

# Automatic spectral coarse spaces for robust finite element tearing and interconnecting and balanced domain decomposition algorithms

N. Spillane<sup>1,2,\*,†</sup> and D. J. Rixen<sup>3</sup>

<sup>1</sup>*Laboratoire Jacques-Louis Lions, CNRS UMR 7598, Université Pierre et Marie Curie, 75005 Paris, France*

<sup>2</sup>*Centre de Technologie de Ladoux, Manufacture des Pneumatiques Michelin, 63040 Clermont-Ferrand, France*

<sup>3</sup>*Institute of Applied Mechanics, Technische Universität München, D-85747 Garching, Germany*

## SUMMARY

We introduce spectral coarse spaces for the balanced domain decomposition and the finite element tearing and interconnecting methods. These coarse spaces are specifically designed for the two-level methods to be scalable and robust with respect to the coefficients in the equation and the choice of the decomposition. We achieve this by solving generalized eigenvalue problems on the interfaces between subdomains to identify the modes that slow down convergence. Theoretical bounds for the condition numbers of the preconditioned operators, which depend only on a chosen threshold, and the maximal number of neighbors of a subdomain are presented and proved. For the finite element tearing and interconnecting method, there are two versions of the two-level method: one based on the full Dirichlet preconditioner and the other on the, cheaper, lumped preconditioner. Some numerical tests confirm these results. Copyright © 2013 John Wiley & Sons, Ltd.

Received 23 November 2012; Revised 9 April 2013; Accepted 10 May 2013

**KEY WORDS:** domain decomposition; FETI; BDD; robustness; scalability; varying coefficients; irregular partitions; interface heterogeneity

## INTRODUCTION

In domain decomposition, it is a real challenge to solve problems with a decomposition given by an automatic partitioner [1, 2], which does not take into account all the difficulties in the problem for the simple reason that there are too many. One well-known challenge for elliptic problems is when the coefficients in the equation are highly heterogeneous. This is often the case in practical applications. Classical coarse spaces are known to give good results when the jumps in the coefficients are across subdomain interfaces (see e.g., [3–6]) or inside the subdomains and not near their boundaries (cf. [7, 8]). However, when the discontinuities are *along* subdomain interfaces, classical results break down, and one may observe very bad convergence of the iterative solvers for the interface problem (see e.g., [9, 10]). In the case of the Darcy equation, the authors in [11–13] prove convergence results, even in some cases where the jumps in the coefficients are along the interfaces. Within a class of coefficient distributions, the theory, which is based on weighted Poincaré inequalities [14], allows to discriminate between jumps in the coefficients, which will affect convergence and those that will not.

It is well known that nonsmooth decompositions (where the interfaces are jagged) [15] or bad aspect ratios of the domains [16] can also lead to poor convergence. This is what we work to improve: we aim to design a method for which the convergence rate does not depend on the choice of the decomposition into subdomains or on any of the coefficients in the equations.

\*Correspondence to: N. Spillane, Laboratoire Jacques-Louis Lions, CNRS UMR 7598, Université Pierre et Marie Curie, 75005 Paris, France.

†E-mail: spillane@ann.jussieu.fr

To achieve this, we will use the strategy introduced in the additive Schwarz framework by [17, 18], and [19]. This strategy is based on the abstract theory of the two-level additive Schwarz method [20]. The strategy is to write the Schwarz theory up to the point where it depends on the set of equations we are dealing with and where assumptions on the coefficient distribution with respect to the decomposition into subdomains are needed to write estimates that do not depend on the parameters. For the Darcy equation ( $-\nabla \cdot \nabla(\alpha u) = b$ ) with the minimal coarse space (the constant functions), the Poincaré inequality and trace theorem are needed to complete the proof, and they require quite strong assumptions. Instead, the authors in [17–19] propose to solve a generalized eigenvalue problem in each subdomain, which selects what modes of the solution satisfy the required estimates for a chosen constant. The other modes, which do not satisfy the estimate, are used to build the coarse space and are basically taken care of with a direct solve in the coarse space. This is what we will refer to as the Schwarz-Generalized Eigenvalues in the Overlaps (GenEO) coarse space. It leads to a two-level method with a convergence rate chosen *a priori* for problems described by symmetric positive definite (spd) matrix.

The idea to use eigenvalue problems to build a coarse space is not new; it was first explored in the algebraic multigrid community. In [21], a strategy to build a coarse space based on spectral information is presented, which allows to achieve any *a priori* chosen target convergence rate. This idea was further developed and implemented in the spectral Algebraic Multigrid based on Element Interpolation method in [22]. More recently, in the framework of two-level overlapping Schwarz, [17–19, 23–26] also build coarse spaces for problems with highly heterogeneous coefficients by solving local eigenproblems. However, compared with the earlier works in the algebraic multigrid context, all of these approaches focus on generalized eigenvalue problems. Within the Schwarz framework, we can distinguish three sets of methods that differ by the choice of the bilinear form in the generalized eigenproblem. First, in the work of [23, 24] for the Darcy equation, the local mass matrix or a ‘homogenized’ version obtained by applying a multiscale partition of unity is used. In [25, 26], still for the Darcy equation, the mass matrix is replaced by an  $L_2$ -product on the subdomain boundary, so that the problem can be reduced to a generalized eigenproblem for the Dirichlet to Neumann operator. This method was analyzed in [27]. Finally in the most recent set of papers, [17–19], which inspired this one, use yet another type of bilinear form that combines the local problem matrix with partition of unity functions. This choice arises from the Schwarz theory and applies to a wide range of problems. Multilevel extensions of some of the previous approaches have been proposed in [28–30]. The approach in [31, 32] within the multigrid framework shares some common ideas.

The purpose of this paper is to extend the GenEO strategy [17–19] to the balancing domain decomposition (BDD) algorithm and the finite element tearing and interconnecting (FETI) algorithm. As it turns out, [33] also proposes to solve generalized eigenvalue problems to deal with heterogeneous coefficients in the BDD by constraints (BDDC) and dual-primal FETI-DP frameworks. More recently, a multilevel extension of this work for BDD by constraints was proposed in [34] with thorough numerical results. Although, at a glance, the generalized eigenvalue problems look similar to the ones in the current paper, they are in fact the result of two different approaches. In [33], the global generalized eigenvalue problem that would need to be solved on the entire domain to achieve a targeted convergence rate is given and then it is made local by restricting it to each of the interfaces between two subdomains. The global condition number *indicator* is chosen to be the maximum over all the set of interfaces of a local estimate. In other words, the global to local *conversion* of the estimate for the condition number is based on heuristics. The approach in this paper is different because it is inspired by previous work in the Schwarz framework. In particular, the global to local *conversion* of the condition number estimate is justified theoretically. It relies very strongly on the abstract Schwarz theory with the result that each local generalized eigenvalue problem is posed on the boundary of a subdomain and not on the interface between two subdomains. With this choice, the fact that the targeted condition number will be achieved is guaranteed theoretically.

The FETI and the BDD are two well-known nonoverlapping domain decomposition methods. Up until now, the GenEO strategy has been applied in the context of overlapping Schwarz, which was first introduced in [35]. The idea of a coarse space correction goes back to [36, 37], and the two-level

overlapping Schwarz preconditioner is due to [38]. As for the BDD method, it is the work of [39] who added a coarse space to the preexisting Neumann–Neumann method [40] to deal with singularities in the local problems. We will refer to the analysis of BDD in [20], which is very closely related to the analysis of the two-level Schwarz preconditioner. Finally, the FETI algorithm was first introduced in [41], and the convergence proof is due to [42, 43]. It is generalized in [44]. Coarse spaces for the FETI method are introduced first in [45] and further developed in [46, 47]. In [48], a two-level FETI method is also introduced for a particular problem, and a convergence result is proved. However, we will follow a very different approach here both for choosing the coarse space and also for writing the proof. In both cases (BDD and FETI), the generalized eigenvalue problem that we solve is used to prove a bound for the largest eigenvalue of the preconditioned operator. As usual, the lower bound for the eigenvalues of the preconditioned operator is one regardless of the coarse space.

The rest of the article is organized as follows. In Section 1, we introduce the notation that will be needed for both algorithms. In Section 2, we introduce the two-level GenEO preconditioner for the BDD algorithm. And in Section 3, we introduce the two-level preconditioner for the FETI algorithm. The definitions of each of the coarse spaces with the corresponding generalized eigenvalue problems can be found in Definitions 2.3 and 3.7, respectively. These generalized eigenvalue problems are chosen specifically to ensure that the so-called stable-splitting properties in Lemmas 2.8 and 3.12 are satisfied. As for the convergence results, they are stated (and proved) in Theorems 2.11 and 3.14. Finally, in Section 4, we give a few numerical results.

## 1. NOTATION FOR FINITE ELEMENT TEARING AND INTERCONNECTING AND BALANCED DOMAIN DECOMPOSITION

For a given domain  $\Omega \in \mathbb{R}^d$  and a finite dimensional Hilbert space  $W_h(\Omega)$ , given a spd bilinear form,

$$\hat{a}(\cdot, \cdot) : W_h(\Omega) \times W_h(\Omega) \rightarrow \mathbb{R}, \quad (1.1)$$

and an element  $\hat{g} \in W_h(\Omega)'$ , we consider the problem of finding  $u \in W_h(\Omega)$ , such that

$$\hat{a}(u, v) = \hat{g}(v), \quad \forall v \in W_h(\Omega). \quad (1.2)$$

To introduce the BDD and the FETI algorithms, we will need to introduce notation for discrete operators at the global and local (on each subdomain) levels.

### 1.1. Problem setting

We begin by rewriting Problem (1.2) in an algebraic framework. As usual in the finite element setting, we start with a triangulation  $\mathcal{T}_h$  of  $\Omega$ :  $\Omega = \bigcup_{\tau \in \mathcal{T}_h} \tau$  and a basis  $\{\phi_k\}_{1 \leq k \leq N}$  for the finite element space  $W_h(\Omega)$ .

#### Assumption 1.1

Given any element  $\tau$  of the mesh  $\mathcal{T}_h$ , let  $W_h(\tau) := \{u|_\tau : u \in W_h(\Omega)\}$ . We assume that for each element  $\tau \in \mathcal{T}_h$ , there exists a symmetric positive semi-definite (spsd) bilinear form  $a_\tau : W_h(\tau) \times W_h(\tau) \rightarrow \mathbb{R}$ , such that

$$\hat{a}(u, v) = \sum_{\tau \in \mathcal{T}_h} a_\tau(u|_\tau, v|_\tau), \quad \forall u, v \in W_h(\Omega),$$

and an element  $g_\tau \in W_h(\tau)'$ , such that

$$\hat{g}(v) = \sum_{\tau \in \mathcal{T}_h} g_\tau(v|_\tau), \quad \forall v \in W_h(\Omega).$$

The stiffness matrix is assembled with the following entries

$$(\hat{K})_{kl} := \hat{a}(\phi_k, \phi_l) \left( = \sum_{\tau \in \mathcal{T}_h} a_\tau(\phi_k|_\tau, \phi_l|_\tau) \right), \quad \forall k, l = 1, \dots, n, \quad (1.3)$$

and the discrete right-hand side  $\hat{f} \in \mathbb{R}^n$  is defined by the entries

$$(\hat{f})_k := \hat{g}(\phi_k) \left( = \sum_{\tau \in \mathcal{T}_h} g_\tau(\phi_k|_\tau) \right), \quad \forall k = 1, \dots, n.$$

As is quite customary, we identify vectors of DOF, which are in some spaces  $\mathbb{R}^m$ , with the associated finite element functions. Operators between the spaces are represented as matrices, and we frequently commit an abuse of notation by using matrices and operators interchangeably. With this abuse of notation, the original problem (1.2) is equivalent to the linear system: find  $u \in W_h(\Omega)$  such that

$$\hat{K}u = \hat{f}, \quad (1.4)$$

with  $\hat{K}$  spd.

## 1.2. Local setting and notation

**1.2.1. Local Setting.** We introduce a partition of the global domain  $\Omega$  into  $N$  nonoverlapping subdomains  $\Omega_i$ , which are resolved by the mesh

$$\bar{\Omega} = \bigcup_{i=1}^N \bar{\Omega}_i \quad \text{and} \quad \Omega_i \cap \Omega_j = \emptyset, \quad i \neq j,$$

and the resulting set of boundaries between subdomains

$$\Gamma := \bigcup_{i \neq i'} \bar{\Omega}_i \cap \bar{\Omega}_{i'}.$$

The reason we have required the information on the nonassembled stiffness matrices is that we want to have access to local matrices for any choice of the partition into subdomains. To do this, we also need to define local finite element spaces and local bilinear forms.

### Assumption 1.2

The basis functions  $\phi_k$  are continuous on  $\Omega$ . In particular for any subset  $D \subset \Omega$ , the restriction  $\phi_k|_D$  of  $\phi_k$  to  $D$  is well defined.

### Definition 1.3 (Local finite element spaces)

For any subset  $D \subset \Omega$ , let the set of DOF in  $D$  be the set

$$dof(D) := \{k = 1, \dots, n; \phi_k|_D \neq 0|_D\}, \quad (1.5)$$

where  $0|_D : D \rightarrow \mathbb{R}$  is identically zero. Then, the finite element space on  $D$  is defined as

$$W_h(D) := \{u|_D; u \in W_h(\Omega)\} = \text{span}\{\phi_k|_D; k \in dof(D)\}. \quad (1.6)$$

The second equality in the definition of  $W_h(D)$  is an immediate consequence.

*Definition 1.4 (Local bilinear forms and local right-hand sides)*

For any open subset  $D \subset \Omega$ , which is resolved by the mesh  $\mathcal{T}_h$ , let the local bilinear form on  $D$  be

$$a_D : W_h(D) \times W_h(D) \rightarrow \mathbb{R}; \quad a_D(v, w) := \sum_{\tau \subset D} a_\tau(v|_\tau, w|_\tau), \quad (1.7)$$

and the local right-hand side be the element

$$g_D \in W'_h(D); \quad g_D(v) := \sum_{\tau \subset D} g_\tau(v|_\tau). \quad (1.8)$$

For any  $i = 1, \dots, N$ , the space of finite element functions on each  $\Omega_i$  follows from (1.6) with  $D = \Omega_i$ :

$$W_h(\Omega_i) = \{u|_{\Omega_i}; u \in W_h(\Omega)\},$$

as well as the trace spaces for  $D = \partial\Omega_i \cap \Gamma$ :

$$W_i := W_h(\Gamma \cap \partial\Omega_i) = \{u|_{\Gamma \cap \partial\Omega_i}; u \in W_h(\Omega)\}.$$

Finally, we define the product space

$$W := \prod_{i=1}^N W_i.$$

We know from (1.6) that  $W_i = \text{span}\{\phi_k|_{\partial\Omega_i \cap \Gamma}; k \in \text{dof}(\partial\Omega_i \cap \Gamma)\}$ ; we make the further assumption that this set of functions is a basis of  $W_i$ .

*Assumption 1.5*

The set  $\{\phi_k|_{\partial\Omega_i \cap \Gamma}; k \in \text{dof}(\partial\Omega_i \cap \Gamma)\}$  is a basis of  $W_i$ .

Throughout the analysis, we will consider elements in the product space  $W$ . Each component  $u_i \in W_i$  is defined on a part  $\Gamma \cap \partial\Omega_i$  of the boundary, and two contributions from two neighboring subdomains do not necessarily match on the shared interface. This is a result of the partition of  $\Omega$  into subdomains. Our finite element approximation of the elliptic problem is, however, based on functions in  $W_h(\Omega)$ , which are defined on the whole of  $\Omega$  with one value per DOF. We denote the space of restrictions of these functions to the set of internal boundaries  $\Gamma$  by  $\hat{W}$ :

$$\hat{W} := W_h(\Gamma) = \{u|_\Gamma; u \in W_h(\Omega)\} \left( = \text{span}\{\phi_k|_\Gamma; k \in \text{dof}(\Gamma)\} \right). \quad (1.9)$$

Next, we introduce interpolation (prolongation) operators  $R_i^\top : W_i \rightarrow \hat{W}$  for  $i = 1, \dots, N$ :

$$\forall u_i = \sum_{k \in \text{dof}(\Gamma \cap \partial\Omega_i)} \alpha_i^k \phi_k|_{\Gamma \cap \partial\Omega_i} \quad (\alpha_i^k \in \mathbb{R}); \quad R_i^\top u_i := \sum_{k \in \text{dof}(\Gamma \cap \partial\Omega_i)} \alpha_i^k \phi_k|_\Gamma.$$

These are the natural interpolation operators represented by boolean matrices: the continuous global function  $R_i^\top u_i \in \hat{W}$  shares the same values as  $u_i$  for DOF in  $\text{dof}(\Gamma \cap \partial\Omega_i)$  and has no contributions from any other DOF. The corresponding restriction operator  $R_i : \hat{W} \rightarrow W_i$  is defined as

$$\forall u = \sum_{k \in \text{dof}(\Gamma)} \alpha^k \phi_k|_\Gamma \quad (\alpha^k \in \mathbb{R}); \quad R_i u := \sum_{k \in \text{dof}(\Gamma \cap \partial\Omega_i)} \alpha^k \phi_k|_{\Gamma \cap \partial\Omega_i}.$$

We note that  $\hat{W} \not\subset W$  and  $\hat{W} = \sum_{i=1}^N R_i^\top W_i$ . It is obvious from the definition of  $R_i^\top$  and Assumption 1.5 that for  $i = 1, \dots, N$  and  $u_i \in W_i$ :

$$u_i = 0_{|\Gamma \cap \partial \Omega_i} \Leftrightarrow R_i^\top u_i = 0_{|\Gamma}. \quad (1.10)$$

**1.2.2. Stiffness matrices.** The local stiffness matrix  $K_i : W_h(\Omega_i) \rightarrow W_h(\Omega_i)$  is the matrix associated with bilinear form  $a_{\Omega_i}$  defined by (1.7) for  $D = \Omega_i$ . From these, the stiffness matrix on the product space is defined as

$$K : W_h(\Omega_1) \times \dots \times W_h(\Omega_N) \rightarrow W_h(\Omega_1) \times \dots \times W_h(\Omega_N); \quad K := \begin{pmatrix} K_1 & 0 & \dots & 0 \\ 0 & K_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & K_N \end{pmatrix} \quad (1.11)$$

so that

$$Ku = (K_1 u_1, \dots, K_N u_N)^\top, \quad \forall u = (u_1, \dots, u_N)^\top \in W_h(\Omega_1) \times \dots \times W_h(\Omega_N). \quad (1.12)$$

**1.2.3. Schur complement matrices.** The DOF,  $dof(\Omega_i)$  in  $W_h(\Omega_i)$ , can be split into the set  $b_i := dof(\Gamma \cap \partial \Omega_i)$  of DOF that are also in the trace space  $W_i$  and the remainder  $I_i := dof(\Omega_i) \setminus dof(\Gamma \cap \partial \Omega_i)$ . This way, we can rewrite the local stiffness matrix in block formulation

$$K_i = \begin{pmatrix} K_i^{b_i b_i} & K_i^{b_i I_i} \\ K_i^{I_i b_i} & K_i^{I_i I_i} \end{pmatrix}.$$

The interior variables of any subdomain are then eliminated in work that can be parallelized across the subdomains. The resulting matrices are the local Schur complements

$$S_i : W_i \rightarrow W_i; \quad S_i := K_i^{b_i b_i} - K_i^{b_i I_i} (K_i^{I_i I_i})^{-1} K_i^{I_i b_i}, \quad i = 1, \dots, N, \quad (1.13)$$

and the Schur complement on the product space is

$$S : \underbrace{W_1 \times \dots \times W_N}_W \rightarrow \underbrace{W_1 \times \dots \times W_N}_W; \quad S := \begin{pmatrix} S_1 & 0 & \dots & 0 \\ 0 & S_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & S_N \end{pmatrix} \quad (1.14)$$

so that

$$Su = (S_1 u_1, \dots, S_N u_N)^\top, \quad \forall u = (u_1, \dots, u_N)^\top \in W. \quad (1.15)$$

The Schur complement  $S$  on the product space  $W$  admits the following counterpart  $\hat{S}$  for functions in  $\hat{W}$ :

$$\hat{S} : \hat{W} \rightarrow \hat{W}; \quad \hat{S}u := \sum_{i=1}^N R_i^\top S_i R_i u. \quad (1.16)$$

We notice that this is the usual Schur complement for the global problem reduced to the set  $\Gamma$  of internal boundaries:

$$\hat{S} = \hat{K}^{bb} - \hat{K}^{bI} (\hat{K}^{II})^{-1} \hat{K}^{Ib}, \quad (1.17)$$

where  $\hat{K}^{bb}$ ,  $\hat{K}^{bI}$ ,  $\hat{K}^{II}$  and  $\hat{K}^{Ib}$  are the components in the bloc formulation of  $\hat{K}$

$$\hat{K} = \begin{pmatrix} \hat{K}^{bb} & \hat{K}^{bI} \\ \hat{K}^{Ib} & \hat{K}^{II} \end{pmatrix}, \quad b := dof(\Gamma) \text{ and } I := dof(\Omega) \setminus dof(\Gamma). \quad (1.18)$$

**1.2.4. Lumped matrices.** In the FETI literature, the lumped version of the stiffness matrix is the extraction of the entries in the stiffness matrix, which correspond to boundary DOF. We have already introduced  $\hat{K}^{bb}$  and  $K_i^{b_i b_i}$ , let  $K^{bb}$  be the counterpart on the product space  $W$ :

$$K^{bb} : \underbrace{W_1 \times \dots \times W_N}_W \rightarrow \underbrace{W_1 \times \dots \times W_N}_W; \quad K^{bb} := \begin{pmatrix} K_1^{b_1 b_1} & 0 & \dots & 0 \\ 0 & K_2^{b_2 b_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & K_N^{b_N b_N} \end{pmatrix}. \quad (1.19)$$

We notice that  $\hat{K}^{bb} = \sum_{i=1}^N R_i^\top K_i^{b_i b_i} R_i$  and the next lemma gives an important relation between lumped matrices and Schur complement matrices.

**Lemma 1.6**

For any  $\hat{u} \in \hat{W}$  and any  $u \in W$ , the following inequalities hold

$$\langle \hat{S}\hat{u}, \hat{u} \rangle \leq \langle \hat{K}^{bb}\hat{u}, \hat{u} \rangle \quad \text{and} \quad \langle Su, u \rangle \leq \langle K^{bb}u, u \rangle.$$

*Proof*

Let  $\hat{u} \in \hat{W}$ . Then by definition of  $\hat{S}$ ,

$$\langle \hat{S}\hat{u}, \hat{u} \rangle = \langle (\hat{K}^{bb} - \hat{K}^{bI}(\hat{K}^{II})^{-1}\hat{K}^{Ib})\hat{u}, \hat{u} \rangle = \langle \hat{K}^{bb}\hat{u}, \hat{u} \rangle - \langle (\hat{K}^{II})^{-1}\hat{K}^{Ib}\hat{u}, \hat{K}^{Ib}\hat{u} \rangle.$$

The first inequality follows by noticing that  $\langle (\hat{K}^{II})^{-1}\hat{K}^{Ib}\hat{u}, \hat{K}^{Ib}\hat{u} \rangle \geq 0$  because  $(\hat{K}^{II})^{-1}$  is spd. For the second, let  $u \in W$ . Then by definition of  $S$ ,

$$\begin{aligned} \langle Su, u \rangle &= \sum_{i=1}^N \langle S_i u_i, u_i \rangle = \sum_{i=1}^N \left\langle \left( K_i^{b_i b_i} - K_i^{b_i I_i} (K_i^{I_i I_i})^{-1} K_i^{I_i b_i} \right) u_i, u_i \right\rangle \\ &= \langle K^{bb}u, u \rangle - \sum_{i=1}^N \left\langle (K_i^{I_i I_i})^{-1} K_i^{I_i b_i} u_i, K_i^{I_i b_i} u_i \right\rangle. \end{aligned}$$

And the second inequality follows by noticing that  $\left\langle (K_i^{I_i I_i})^{-1} K_i^{I_i b_i} u_i, K_i^{I_i b_i} u_i \right\rangle \geq 0$  for any  $i = 1, \dots, N$  because  $(K_i^{I_i I_i})^{-1}$  is spd.  $\square$

**1.2.5. Right-hand sides.** To reduce the problem to the set of interfaces between subdomains, we define the following right-hand side

$$\hat{f}_\Gamma := \hat{f}^b - \hat{K}^{bI}(\hat{K}^{II})^{-1}\hat{f}^I, \quad (1.20)$$

which is the right-hand side of the original problem (1.4) condensed onto the DOF in  $\hat{W}$ . As for the right-hand side on the product space  $W$ , for each subdomain  $i = 1, \dots, N$ : first, let  $f_i$  be the local right-hand side given by (1.8) with  $D = \Omega_i$ . Then condense it onto the interfaces following  $f_{\Gamma,i} := f_i^{b_i} - K_i^{b_i I_i} (K_i^{I_i I_i})^{-1} f_i^{I_i}$ . (We have used the identification between the finite element representation of  $f_i$  and its vector representation.) Finally, the right-hand side for the problem condensed onto the space  $W$  is

$$f_\Gamma = \begin{pmatrix} f_{\Gamma,1} \\ \dots \\ f_{\Gamma,N} \end{pmatrix}. \quad (1.21)$$

Most of this notation is summed up in Table I at the beginning of the article. Some comments are given in Subsection 1.4, along with an important lemma on which of these matrices are positive definite.



Table I. Summary of notations.

Function space	Description	Definition
$W_h(\Omega)$	Global	Solution space for (1.1)
$W_h(\Omega_i)$	Local	$\{u _{\Omega_i}; u \in W_h(\Omega)\} \text{ ((1.6); } D = \Omega_i)$
$W_i$	Local trace	$\{u _{\Gamma \cap \partial \Omega_i}; u \in W_h(\Omega)\} \text{ ((1.6); } D = \Gamma \cap \partial \Omega_i)$
$W$	Product trace	$W_1 \times \dots \times W_N$
$\hat{W}$	Global trace	$\{u _{\Gamma}; u \in W_h(\Omega)\} \text{ ((1.6); } D = \Gamma)$
Stiffness matrices (defined on)	Matrix	Bilinear form
Global ( $W_h(\Omega)$ )	$\hat{K}$ (1.3)	$\hat{a}$ (1.1)
Local ( $W_h(\Omega_i)$ )	$K_i$	$a_{\Omega_i}$ (1.7) for $D = \Omega_i$
Product space ( $\prod_{i=1}^N W_h(\Omega_i)$ )	$K$ (1.11)	None
Lumped global ( $W_h(\Omega)$ )	$\hat{K}^{bb}$ (1.18)	$\hat{a}^{bb}$
Lumped product space ( $\prod_{i=1}^N W_h(\Omega_i)$ )	$K^{bb}$ (1.19)	$a^{bb}$
Schur complement (defined on)	Matrix	Bilinear form
Global ( $\hat{W}$ )	$\hat{S}$ (1.16)	$\hat{s}$
Local ( $W_i$ )	$S_i$ (1.13)	$s_i$ (1.22)
On the product space ( $W$ )	$S$ (1.14)	$s$
Weighted local ( $W_i$ )	$\tilde{S}_i$	$\tilde{s}_i$ (1.23)
Right-hand sides	Notation	
Condensed onto $\Gamma$	$\hat{f}_{\Gamma}$ (1.20)	
Condensed onto $\Gamma \cap \partial \Omega_i$	$f_{\Gamma,i}$ (1.21)	
Condensed on product space $\prod_{i=1}^N \Gamma \cap \partial \Omega_i$	$f_{\Gamma}$ (1.21)	

**Remark 1.7**

Assumption 1.1 is actually stronger than what we really need but enables the use of any partition into subdomains and allowed us to define each component of the algorithm thoroughly. For a given nonoverlapping partition into subdomains, it is enough to have access to the local matrices  $K_i$  on each subdomain, the local right-hand sides  $f_i$ , the local–global interpolation operators  $R_i^{\top}$ , and the information on the boundary of each subdomain  $\Gamma \cap \partial \Omega_i$ .

**1.3. Partition of unity and weighted operators**

An important role in the description of the BDD algorithms is played by a weighting (counting) function on  $W$ . As in the original GenEO algorithm [17, 18], this induces partition of unity operators  $\Xi_i$ , which act directly on the DOF of the finite element functions.

**Definition 1.8 (Partition of unity)**

Let  $\mu = (\mu_1, \dots, \mu_N) \in W$  be a discrete partition of unity

$$\sum_{i=1, \dots, N} R_i^{\top} \mu_i = 1_{|\hat{W}}, \text{ where } 1_{|\hat{W}} \in \hat{W} \text{ and all vector entries are 1.}$$

Then for any function  $u_i \in W_i$  written as

$$u_i = \sum_{k \in \text{dof}(\Gamma \cap \partial \Omega_i)} \alpha_i^k \phi_{k|\Gamma \cap \partial \Omega_i}, \quad \alpha_i^k \in \mathbb{R},$$

the local partition of unity operator  $\Xi_i : W_i \rightarrow W_i$  is defined by



$$\Xi_i(u_i) := \sum_k \mu_i^k \alpha_i^k \phi_{k|\Gamma \cap \partial\Omega_i},$$

where  $\mu_i^k$  is the  $k$ -th entry in  $\mu_i$ . The inverse  $\Xi_i^{-1} : W_i \rightarrow W_i$  is defined by:

$$\Xi_i^{-1}(u_i) := \sum_k \frac{1}{\mu_i^k} \alpha_i^k \phi_{k|\Gamma \cap \partial\Omega_i}.$$

It is clear that the  $\Xi_i$  define a partition of unity from  $\hat{W}$  onto the product space  $W = W_1 \times \cdots \times W_N$  in the sense that

$$u = \sum_{i=1}^N R_i^\top \underbrace{\Xi_i(R_i u)}_{\in W_i}, \quad \forall u \in \hat{W}.$$

It is also clear that  $\Xi_i^{-1}$  is the inverse of  $\Xi_i$  because any  $u_i \in W_i$  satisfies  $\Xi_i^{-1}(\Xi_i(u_i)) = \Xi_i(\Xi_i^{-1}(u_i)) = u_i$ .

*Remark 1.9*

Two common choices for  $\mu$  are the multiplicity scaling where  $\mu_i^k$  is chosen as  $(\#\{i = 1, \dots, N; k \in \text{dof}(\Gamma \cap \partial\Omega_i)\})^{-1}$  and the  $K$ -scaling where  $\mu$  depends on the diagonal entries of the stiffness matrices [44, 49]. In the numerical result section, we mostly use  $K$ -scaling.

We introduce the local bilinear forms that correspond to the local Schur complements  $S_i$  as follows. For  $i = 1, \dots, N$  defines

$$s_i : W_i \times W_i \rightarrow \mathbb{R}, s_i(u_i, v_i) := \langle S_i u_i, v_i \rangle; \quad \forall u_i, v_i \in W_i. \quad (1.22)$$

Next, we use the partition of unity operators to define weighted versions of the Schur complements, which will be instrumental in defining the BDD algorithm.

*Definition 1.10 (Weighted Schur complements)*

For any  $i = 1, \dots, N$ , let  $\tilde{s}_i : W_i \times W_i \rightarrow \mathbb{R}$  be the bilinear form defined by

$$\tilde{s}_i(u_i, v_i) := s_i(\Xi_i^{-1}(u_i), \Xi_i^{-1}(v_i)); \quad \forall u_i, v_i \in W_i, \quad (1.23)$$

where  $s_i$  is the local Schur complement and  $\Xi_i^{-1}$  is the inverse partition of unity operator introduced in Definition 1.8.

Next, let the matrix  $\tilde{S}_i : W_i \rightarrow W_i$  be the matrix counterpart of  $\tilde{s}_i$

$$\langle \tilde{S}_i u_i, v_i \rangle := \tilde{s}_i(u_i, v_i).$$

#### 1.4. Summary of the notation and complements

We have introduced quite a lot of notation. Table I at the beginning of the article sums up most of the notation that will appear in the description of the algorithms and the reference to where it is first introduced. Some of the operators are introduced for the first time ( $\hat{a}^{bb}$ ,  $a^{bb}$ ,  $\hat{s}$ , and  $s$ ) as the bilinear forms associated with a matrix. More precisely, let  $\hat{a}^{bb}$  and  $\hat{s}$  be defined as

$$\hat{a}^{bb} : \hat{W} \times \hat{W} \rightarrow \mathbb{R}; \hat{a}^{bb}(\hat{u}, \hat{v}) := \langle \hat{K}^{bb} \hat{u}, \hat{v} \rangle \quad \text{and} \quad \hat{s} : \hat{W} \times \hat{W} \rightarrow \mathbb{R}; \hat{s}(\hat{u}, \hat{v}) := \langle \hat{S} \hat{u}, \hat{v} \rangle,$$

for any  $\hat{u}$  and  $\hat{v} \in \hat{W}$ , and let  $a^{bb}$  and  $s$  be defined as

$$a^{bb} : W \times W \rightarrow \mathbb{R}; a^{bb}(u, v) := \langle K^{bb} u, v \rangle \quad \text{and} \quad s : W \times W \rightarrow \mathbb{R}; s(u, v) := \langle S u, v \rangle,$$

for any  $u$  and  $v \in W$ .

The operators with a  $\hat{\cdot}$  always correspond to functions defined either on the whole of  $\Omega$  or the whole of  $\Gamma$ . The subscript  $i$  always refers to a local operator defined on a subdomain  $\Omega_i$  or its boundary. Operators without a  $\hat{\cdot}$  or a subscript  $i$  are defined on the product spaces. Finally, operators  $\tilde{S}_i$  are weighted by the inverse partition of unity operators.

In many cases, the local stiffness matrices  $K_i$  are not spd on all floating subdomains. (A floating subdomain is a subdomain that does not *touch* the Dirichlet part of the boundary). For example, in the case of the Darcy equation, the kernel of  $K_i$  for a floating subdomain is the set of constant functions. In the case of linear elasticity, the kernel of  $K_i$  is the set of rigid body motions. It is easy to see that these kernels induce kernels for the corresponding Schur complements  $S_i$  as well as their weighted counterparts  $\tilde{S}_i$  and, possibly, the lumped matrices  $K_i^{b_i b_i}$ . The next lemma makes precise which matrices are positive definite. They are all spd.

*Lemma 1.11*

The stiffness matrix  $K$ , the lumped stiffness matrix  $K^{bb}$ , and the Schur complement  $S$ , which correspond to the product spaces, can be singular. Their respective counterparts,  $\hat{K}$ ,  $\hat{K}^{bb}$ , and  $\hat{S}$ , on the original spaces of functions  $W_h(\Omega)$  and  $\hat{W}$  are spd. Finally, under Assumption 1.5, each of the local matrices  $R_i \hat{K}^{bb} R_i^\top$  and  $R_i \hat{S} R_i^\top$  is also spd.

*Proof*

The fact that  $\hat{K}$  and  $\hat{S}$  are positive definite is clear because the original problem is well posed. The positive definiteness of  $\hat{K}^{bb}$  follows from Lemma 1.6 and the positive definiteness of  $\hat{S}$ : let  $u \in \hat{W}$

$$\langle \hat{K}^{bb} u, u \rangle = 0 \Rightarrow \langle \hat{S} u, u \rangle = 0 \Rightarrow u = 0.$$

The positive definiteness of  $R_i \hat{S} R_i^\top$  and  $R_i \hat{K}^{bb} R_i^\top$  is obvious from the positive definiteness of  $\hat{K}$  and  $\hat{S}$  and (1.10), which is a direct consequence of Assumption 1.5.  $\square$

*Remark 1.12*

Note that in nearly all practical cases,  $K^{bb}$  is also spd.

We are now ready to introduce the BDD preconditioner.

## 2. BALANCING DOMAIN DECOMPOSITION

The problem that we solve is the original problem (1.4) reduced to the set  $\Gamma$  of interfaces between subdomains: find  $u \in \hat{W}$  such that

$$\hat{S}u = \hat{f}_\Gamma. \quad (2.1)$$

### 2.1. One-level balanced domain decomposition preconditioner in the abstract Schwarz framework [20]

The only thing that is needed to define the one-level preconditioner is a solver on each subdomain. Then, we will precondition the global problem (2.1) with a sum of these local solves. The usual BDD strategy is to use the weighted Schur complements  $\tilde{S}_i$  introduced in Definition 1.10 to build local problems. Then each local solve is the solution of a Neumann problem,  $\tilde{S}_i^\dagger$ .

*Definition 2.1 (One-level preconditioner)*

For each  $i = 1, \dots, N$ , let  $\tilde{P}_i$  and  $P_i$  be defined as

$$\tilde{P}_i := \tilde{S}_i^\dagger R_i \hat{S} \quad \text{and} \quad P_i := R_i^\top \tilde{P}_i, \quad (2.2)$$

where  $\tilde{S}_i^\dagger$  is a pseudo inverse of  $\tilde{S}_i$ . Equivalently for any  $u \in \hat{W}$ ,  $\tilde{P}_i u$  is the unique vector in  $\text{range}(\tilde{S}_i^\dagger)$ , which satisfies

$$\tilde{s}_i(\tilde{P}_i u, v_i) = \hat{s}(u, R_i^\top v_i), \quad \forall v_i \in W_i. \quad (2.3)$$

The one-level preconditioner is the sum of local solves  $\sum_{i=1}^N R_i^\top \tilde{S}_i^\dagger R_i$  so the one-level preconditioned operator is  $\sum_{i=1}^N P_i$ .

The next lemma gives a lower bound for the eigenvalues of the one-level preconditioned operator. It does not depend on the specific choice of the pseudo inverse or on any coarse space.

Essentially, what we do is check that a stable-splitting assumption (Assumption 2.2 in [20]) holds on the whole of  $\hat{W}$ . Then we give the result of Lemma 2.5 in [20], which is that this implies a lower bound for the condition number of the one-level preconditioned operator. One of the assumptions in [20] is that the local bilinear forms ( $\tilde{S}_i$  in this case) be positive definite. Here, they are only positive semi-definite but the proof goes through in the exact same way so we do not give it again.

*Lemma 2.2 (Stable-splitting – lower bound for the eigenvalues of the preconditioned operator)*

For any  $u \in \hat{W}$ , there exists a stable-splitting  $(v_1, \dots, v_N)$  of  $u$  onto  $W = W_1 \times \dots \times W_N$

$$u = R_1^\top v_1 + \dots + R_N^\top v_N; v_i \in W_i \text{ and } \sum_{i=1}^N \tilde{s}_i(v_i, v_i) \leq \hat{s}(u, u). \quad (2.4)$$

This implies that the one-level preconditioned operator satisfies

$$\hat{s}(u, u) \leq \hat{s}\left(\sum_{i=1}^N P_i u, u\right) \text{ for any } u \in \hat{W}. \quad (2.5)$$

*Proof*

Let  $u \in \hat{W}$ . The fact that, by definition, the operators  $\Xi_i$  define a partition of unity allows us to write an obvious splitting of  $u$  onto  $W$ :

$$(v_i := \Xi_i(R_i u), \quad \forall i = 1, \dots, N) \hat{\Delta} \Rightarrow u = \sum_{i=1}^N R_i^\top v_i.$$

We prove (2.4) for this splitting using only the definitions of  $\tilde{s}_i$  and  $\hat{s}$

$$\sum_{i=1}^N \tilde{s}_i(v_i, v_i) = \sum_{i=1}^N s_i(\Xi_i^{-1}(\Xi_i(R_i u)), \Xi_i^{-1}(\Xi_i(R_i u))) = \sum_{i=1}^N s_i(R_i u, R_i u) = \hat{s}(u, u).$$

The second part of the lemma is the result of Lemma 2.5 in [20]; we refer the reader there for the proof.  $\square$

The fact that (2.5) provides a lower bound for the eigenvalues of the preconditioned operator  $\sum_{i=1}^N P_i$  is easy to see: suppose  $u$  is an eigenvector associated with eigenvalue  $\lambda$ , then

$$\sum_{i=1}^N P_i u = \lambda u \Rightarrow \hat{S} \sum_{i=1}^N P_i u = \lambda \hat{S} u \Rightarrow \hat{s}\left(\sum_{i=1}^N P_i u, u\right) = \lambda \hat{s}(u, u),$$

and (2.5) implies that  $\lambda \geq 1$ .

In other words, the lower bound for the eigenvalues of the preconditioned operator does not depend on the choice of the coarse space. This is a big difference with the Additive Schwarz method where the proof of a lower bound depends very strongly on the choice of the coarse space and on restrictive assumptions on the coefficient distribution. This is why the Schwarz-GenEO strategy

in [18] is precise to build an enriched coarse space for which the stable-splitting property and thus a lower bound for the spectrum of the preconditioned operator hold, regardless of the partition into subdomains and the coefficient distribution. Luckily, the upper bound for the eigenvalues of the Additive Schwarz operator depends only on the number of neighbors of each subdomain, enabling the proof of a bound for the condition number of the preconditioned operator.

Here, the situation is reversed: Lemma 2.2 gives a lower bound for the eigenvalues of the preconditioned operator that does not depend on the choice of the coarse space thanks to the adequate weighting of the local solvers. However, the upper bound requires more work, and with the usual coarse space, it can only be independent of the coefficients in the equation if some assumptions on the coefficient distribution are satisfied. The GenEO strategy will enable us to waive all of these assumptions.

## 2.2. Generalized Eigenvalues in the Overlaps coarse space for balanced domain decomposition

The abstract Schwarz theory tells us that the upper bound for the eigenvalues of the preconditioned operator is implied by the stability of the local solvers  $\tilde{s}_i$  on the local subspaces once the coarse components have been removed (this is made explicit in Lemma 2.8). This is where the GenEO strategy comes in. We solve a generalized eigenvalue problem that identifies the ‘bad’ modes: in this case, those for which we cannot ensure that the local solver is stable for a constant independent of the coefficients in the equations. These bad modes are then used to span the coarse space, and the local solvers are stable on all remaining local components (the ‘good’ components). More precisely, the next two definitions introduce the generalized eigenvalue problem, the coarse space, and the corresponding two-level BDD-GenEO preconditioners.

*Definition 2.3 (Generalized Eigenvalues in the Overlaps coarse space for balanced domain decomposition)*

For each subdomain  $i = 1, \dots, N$ , find the eigenpairs  $(p_i^k, \lambda_i^k) \in W_i \times \mathbb{R}^+$  of the generalized eigenvalue problem

$$\boxed{\tilde{s}_i(p_i^k, v_i) = \lambda_i^k \hat{a}^{bb}(R_i^\top p_i^k, R_i^\top v_i)} \quad \text{for any } v_i \in W_i. \quad (2.6)$$

Next, given a threshold  $\mathcal{K}_i > 0$  for each subdomain, define the coarse space as

$$W_0 = \text{span} \left\{ R_i^\top p_i^k ; \lambda_i^k < \mathcal{K}_i, i = 1, \dots, N \right\} \quad (\subset \hat{W}). \quad (2.7)$$

Let the interpolation operator  $R_0^\top$  be the matrix whose columns are the coarse basis functions  $\{R_i^\top p_i^k ; \lambda_i^k < \mathcal{K}_i, i = 1, \dots, N\}$ . Finally, let the coarse solver be the exact solver on  $W_0$

$$S_0 := R_0 \hat{S} R_0^\top,$$

and  $P_0$  be the  $\hat{S}$ -orthogonal projection operator defined by

$$P_0 := R_0^\top S_0^\dagger R_0 \hat{S}. \quad (2.8)$$

This definition gives rise to a few immediate remarks.

### Remark 2.4

- (i) The operator  $R_0^\top$  is a mapping between the coordinates of a vector from  $W_0$  in the set of coarse basis functions and its representation in  $\hat{W}$  ( $\text{range}(R_0^\top) \subset \hat{W}$ ). Its transpose  $R_0$  is a restriction operator that maps an element in  $\hat{W}$  to the coordinates of its  $l_2$  projection onto  $W_0$  in the set of coarse basis functions.

- (ii) Eigenvalue 0 for eigenproblem (2.6) is associated with the kernel of  $\tilde{s}_i$  so in some sense the coarse space will take care of the fact that  $\tilde{s}_i$  is not necessarily coercive. Note that if the coarse space would include only the kernel of  $\tilde{s}_i$ , one would obtain the usual coarse grid of the BDD.
- (iii) In the definition of  $P_0$ , we used a pseudo inverse  $S_0^\dagger$  because the columns of  $R_0^\top$  are not necessarily linearly independent. The pseudo inverse is defined up to an element in  $\text{Ker}(R_0^\top)$  and the specific choice of the pseudo inverse makes no difference because the application of  $S_0^\dagger$  is followed by an application of  $R_0^\top$ .
- (iv) The fact that  $P_0$  is an  $\hat{S}$ -orthogonal projection can be proved easily using the definitions of  $P_0$  and  $S_0$ , and it is equivalent to the fact that  $P_0$  is self-adjoint with respect to  $S_0$ .

We are now ready to introduce the BDD-GenEO preconditioner. There are mainly two ways to add the second level once that we have chosen the coarse space: either we use the balanced preconditioner (2.10) with the preconditioned conjugate gradient algorithm or we use the projected preconditioned conjugate gradient (PPCG) algorithm in the space  $\text{range}(I - P_0)$  with the deflated preconditioner (2.9). Both alternatives will lead to essentially identical convergence bounds. In fact, for certain starting vectors in exact arithmetic, they produce the same iterates. A nice comparison (both theoretical and in terms of implementation) between (2.10), (2.9) and some other variants can be found in [50]. The deflated preconditioner is the more natural because it simply restricts the problem to a smaller subspace. However, it suffers from robustness problems when the coarse solves are not sufficiently accurate. The balanced preconditioner is more robust but its application is slightly more expensive.

*Definition 2.5 (Two-level preconditioners)*

Recall that, according to (2.2) and (2.8), we have defined  $P_i = R_i^\top \tilde{S}_i^\dagger R_i \hat{S}$  for any  $i = 1, \dots, N$  and  $P_0 = R_0^\top S_0^\dagger R_0 \hat{S}$ . Then define the deflated preconditioned operator as

$$P_{def} := \sum_{i=1}^N (I - P_0)^\top P_i (I - P_0), \quad (2.9)$$

and the balanced preconditioned operator as

$$P_{bal} := P_0 + \sum_{i=1}^N (I - P_0)^\top P_i (I - P_0). \quad (2.10)$$

In the remainder of this subsection, we show that the BDD-GenEO coarse space leads to an upper bound for the eigenvalues of the preconditioned operators, which does not depend on the number of subdomains or the coefficients in the equations. Instead, it depends on the thresholds  $\mathcal{K}_i$ , which were introduced to select the coarse basis functions. First, we give some properties of the family of generalized eigenvectors (Lemma 2.6). Then, we use these properties to show that the local bilinear forms are stable on the deflated local subspaces (Lemma 2.8) and the upper bound follows from there (Lemma 2.10).

*Lemma 2.6*

For a given subdomain  $i = 1, \dots, N$ , the eigenpairs  $(p_i^k, \lambda_i^k)$  of generalized eigenproblem (2.6) can be chosen so that the set  $\{p_i^k\}_k$  of eigenvectors is an orthonormal basis of  $W_i$  with respect to the inner product induced  $\hat{a}^{bb}(R_i^\top \cdot, R_i^\top \cdot)$ . This writes

$$\hat{a}^{bb}(R_i^\top p_i^k, R_i^\top p_i^k) = 1; \quad \text{and} \quad \hat{a}^{bb}(R_i^\top p_i^k, R_i^\top p_i^{k'}) = 0, \quad k \neq k'.$$

An orthogonality type property with respect to  $\tilde{s}_i$  (which is not necessarily coercive) also holds

$$\tilde{s}_i(p_i^k, p_i^{k'}) = 0, k \neq k'.$$

*Proof*

Lemma 1.11 tells us that  $R_i \hat{K}^{bb} R_i^\top$  is positive definite on  $W_i$  so we may indeed speak of a  $\hat{a}^{bb}(R_i^\top \cdot, R_i^\top \cdot)$  orthonormal basis of  $W_i$ . Then for the proof, see for example [51].  $\square$

*Remark 2.7*

The fact that the generalized eigenproblem (2.6) is equivalent to a nongeneralized eigenproblem implies that all eigenvalues are finite. Because both matrices are spsd, the eigenvalues are also nonnegative: for any  $k$ ,  $0 \leq \lambda_i^k < +\infty$ .

The next lemma states that the local solvers are stable and strongly relies on the definition of the GenEO coarse space. In fact, the purpose of the GenEO strategy is specifically to ensure that Lemma 2.8 holds. This corresponds to Assumption 2.4 in [20].

*Lemma 2.8 (Stability of the local solvers)*

Suppose the pseudo inverse  $\tilde{S}_i^\dagger$  in Definition 2.1 is chosen such that  $\text{range}(\tilde{S}_i^\dagger) = \text{span}\{p_i^k; \lambda_i^k > 0\}$ . Then for any  $i = 1, \dots, N$ , the local solvers are stable in the sense

$$\hat{s}(R_i^\top u_i, R_i^\top u_i) \leq \frac{1}{\mathcal{K}_i} \tilde{s}_i(u_i, u_i), \quad \forall u_i \in \text{range}(\tilde{P}_i(I - P_0)),$$

where the  $\mathcal{K}_i$  are the thresholds that were used to select eigenvectors for the coarse space in Definition 2.3.

*Proof*

We may indeed choose  $\text{range}(\tilde{S}_i^\dagger) = \text{span}\{p_i^k; \lambda_i^k > 0\}$  because the pseudo inverse of an operator is defined up to an element in the kernel of this operator. Precisely, there are an infinity of pseudo inverse, and we may choose the range of  $\tilde{S}_i^\dagger$  among all the spaces that satisfy  $\text{range}(\tilde{S}_i^\dagger) \oplus \text{Ker}(\tilde{S}_i) = W_i$ . Here,  $\text{Ker}(\tilde{S}_i) = \text{span}\{p_i^k; \lambda_i^k = 0\}$  and the set of all  $p_i^k$  is a basis of  $W_i$  so our choice fits this limitation.

Next, we prove that

$$\text{range}(\tilde{P}_i(I - P_0)) \left( = \text{range}(\tilde{S}_i^\dagger R_i \hat{S}(I - P_0)) \right) \subset \text{span}\left\{\left\{p_i^k\right\}_{\mathcal{K}}\right\},$$

where we have introduced the notation  $\{p_i^k\}_{\mathcal{K}}$  for the set of *good* eigenvectors

$$\left\{p_i^k\right\}_{\mathcal{K}} = \left\{p_i^k; \lambda_i^k \geq \mathcal{K}_i\right\}.$$

We will use the following linear algebra identity

$$\text{Ker}\left((I - P_0)^\top \hat{S} R_i^\top\right) \oplus^\perp \text{range}(R_i \hat{S}(I - P_0)) = W_i, \quad (2.11)$$

where the symbol  $\perp$  refers to the  $l_2$  orthogonality between both spaces and  $\oplus$  means that the sum is direct. By definition (2.8) of  $P_0$ ,  $(I - P_0)^\top = I - \hat{S} R_0^\top S_0^\dagger R_0$  so

$$\text{range}(\hat{S} R_0^\top) \subset \text{Ker}((I - P_0)^\top).$$

In particular, for a given  $i = 1, \dots, N$  :  $\text{span} \left\{ \hat{S} R_i^\top p_i^k; \lambda_i^k < \mathcal{K}_i \right\} \subset \text{Ker}((I - P_0)^\top)$ , which implies

$$\text{span} \left\{ \left\{ p_i^k \right\}_{\mathcal{K}} \right\} \subset \text{Ker} \left( (I - P_0)^\top \hat{S} R_i^\top \right). \quad (2.12)$$

Next, we use another linear algebra identity:  $W_i$  is finite dimensional so

$$\text{span} \left\{ \left\{ p_i^k \right\}_{\mathcal{K}} \right\} \oplus^\perp \text{span} \left\{ p_i^k; \lambda_i^k < \mathcal{K}_i \right\}^\perp = W_i. \quad (2.13)$$

According to Lemma 2.6, the  $\{p_i^k\}_{\mathcal{K}}$  form a  $R_i \hat{K}^{bb} R_i^\top$ -orthonormal basis of  $W_i$  so

$$\left\langle p_i^k, R_i \hat{K}^{bb} R_i^\top p_i^{k'} \right\rangle = 0, \quad \forall k \neq k'.$$

This implies that  $\text{span} \left\{ R_i \hat{K}^{bb} R_i^\top p_i^k; \lambda_i^k \geq \mathcal{K}_i \right\} \subset \text{span} \left\{ \left\{ p_i^k \right\}_{\mathcal{K}} \right\}^\perp$ . The equality between these subsets follows by a dimensional argument: the set  $\{p_i^k\}_{\mathcal{K}}$  forms a basis of  $W_i$  and  $R_i \hat{K}^{bb} R_i^\top$  is spd so

$$\text{rank} \left\{ R_i \hat{K}^{bb} R_i^\top p_i^k; \lambda_i^k \geq \mathcal{K}_i \right\} = \text{rank} \left\{ p_i^k; \lambda_i^k \geq \mathcal{K}_i \right\} = \text{rank} \left\{ \left\{ p_i^k \right\}_{\mathcal{K}} \right\}^\perp,$$

and in turn the inclusion becomes an equality

$$\text{span} \left\{ R_i \hat{K}^{bb} R_i^\top p_i^k; \lambda_i^k \geq \mathcal{K}_i \right\} = \text{span} \left\{ \left\{ p_i^k \right\}_{\mathcal{K}} \right\}^\perp.$$

Injecting this into (2.13) implies

$$\text{span} \left\{ \left\{ p_i^k \right\}_{\mathcal{K}} \right\} \oplus^\perp \text{span} \left\{ R_i \hat{K}^{bb} R_i^\top p_i^k; \lambda_i^k \geq \mathcal{K}_i \right\} = W_i. \quad (2.14)$$

Putting (2.11), (2.12), and (2.14), we have

$$\text{range}(R_i \hat{S}(I - P_0)) \subset \text{span} \left\{ R_i \hat{K}^{bb} R_i^\top p_i^k; \lambda_i^k \geq \mathcal{K}_i \right\},$$

where the argument is:

$$(E_1 \oplus^\perp E_2 = E_3 \oplus^\perp E_4 \text{ and } E_1 \subset E_3) \Rightarrow E_4 \subset E_2,$$

for any vector spaces  $E_1, \dots, E_4$ .

By definition of eigenproblem (2.6),  $\lambda_i^k R_i \hat{K}^{bb} R_i^\top p_i^k = \tilde{S}_i p_i^k$  so

$$\text{range}(R_i \hat{S}(I - P_0)) \subset \text{span} \left\{ \tilde{S}_i p_i^k; \lambda_i^k \geq \mathcal{K}_i \right\}.$$

Finally, for the specific choice of the pseudo inverse  $\tilde{S}_i^\dagger$ , it follows that

$$\text{range} \left( \tilde{S}_i^\dagger R_i \hat{S}(I - P_0) \right) (= \text{range}(\tilde{P}_i(I - P_0))) \subset \text{span} \left\{ \left\{ p_i^k \right\}_{\mathcal{K}} \right\}.$$



Now, we prove the inequality in the lemma. Any  $u_i \in \text{range}(\tilde{P}_i(I - P_0))$  writes  $u_i = \sum_{\{k; \lambda_i^k \geq \kappa_i\}} \alpha_i^k p_i^k$  for some coefficients  $\alpha_i^k \in \mathbb{R}$ . From Lemma 1.6, it is obvious that

$$\hat{s}(R_i^\top u_i, R_i^\top u_i) \leq \hat{a}^{bb}(R_i^\top u_i, R_i^\top u_i) = \hat{a}^{bb} \left( R_i^\top \sum_{\{k; \lambda_i^k \geq \kappa_i\}} \alpha_i^k p_i^k, R_i^\top \sum_{\{k; \lambda_i^k \geq \kappa_i\}} \alpha_i^k p_i^k \right).$$

By using successively the first orthogonality property in Lemma 2.6, the definition of the eigenproblem, and the second orthogonality property in Lemma 2.6, we have

$$\begin{aligned} \hat{s}(R_i^\top u_i, R_i^\top u_i) &\leq \sum_{\{k; \lambda_i^k \geq \kappa_i\}} \alpha_i^{k^2} \hat{a}^{bb}(R_i^\top p_i^k, R_i^\top p_i^k) \\ &= \sum_{\{k; \lambda_i^k \geq \kappa_i\}} \frac{1}{\lambda_i^k} \alpha_i^{k^2} \tilde{s}_i(p_i^k, p_i^k) \\ &\leq \frac{1}{\kappa_i} \sum_{\{k; \lambda_i^k \geq \kappa_i\}} \alpha_i^{k^2} \tilde{s}_i(p_i^k, p_i^k) \\ &= \frac{1}{\kappa_i} \tilde{s}_i(u_i, u_i). \end{aligned}$$

□

*Remark 2.9 (Local stability, exact solvers, and choice of the eigenproblem)*

The bilinear form on the left-hand side of the inequality in the lemma is  $\hat{s}(R_i^\top \cdot, R_i^\top \cdot)$ . This is the so-called exact solver on subdomain  $i$  for the global problem given by  $\hat{S}$ . The exact solvers are by definition the solvers that are used to build the Additive Schwarz preconditioner. For the problem  $\hat{S}u = \hat{f}_\Gamma$ , the Additive Schwarz preconditioner would be  $\sum_{i=1}^N R_i^\top \hat{S} R_i$ . If these exact solvers were used instead of  $\tilde{S}_i$ , the upper bound for the eigenvalues of the preconditioned operator would depend only on a constant related to the number of neighbors (introduced in the next lemma). The nice bound that we have for the lowest eigenvalue of the preconditioned operator would no longer hold however. The most straightforward generalized eigenproblem that arises from the theory is

$$\tilde{s}_i(p_i^k, v_i) = \lambda_i^k \hat{s}(R_i^\top p_i^k, R_i^\top v_i) \quad \text{for any } v_i \in W_i, \quad (2.15)$$

so the eigensolve operates some sort of spectral comparison between the exact solver (on the right) and the one that we actually use (on the left). We then isolate the modes for which the chosen preconditioner is not a good enough approximation in the coarse space and use a direct solve on these modes. It is however expensive to assemble and to solve (2.15). This is why in this article, we have chosen to go through only with eigenproblem (2.6) where  $\hat{s}$  is replaced by  $\hat{a}^{bb}$ . For a coarse space based on Eigenproblem (2.15), the theory goes through to the exact same final estimate simply by replacing  $\hat{a}^{bb}$  by  $\hat{s}$  in the proofs.

The following lemma gives a consequence of the stability of the local solvers. It is very narrowly related to Lemma 2.6 in [20].

*Lemma 2.10 (Upper bound for the eigenvalues of the preconditioned operator)*

The stability of each of the local solvers that was proved in Lemma 2.8 implies

$$\hat{s} \left( \sum_{i=1}^N P_i u, u \right) \leq \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\kappa_i} \right) \hat{s}(u, u) \quad \forall u \in \text{range}(I - P_0),$$

where  $\mathcal{N}$  is the maximal number of neighbors of a subdomain (including itself) in the sense

$$\mathcal{N} := \max_{1 \leq i \leq N} (\# \{j; R_j R_i^\top \neq 0\}).$$

*Proof*

This is basically the proof of Lemma 2.6 in [20] but where we have chosen not to rely on strengthened Cauchy Schwarz inequalities. Instead, we make the number of neighbors of a subdomain appear explicitly. Let  $u \in \text{range}(I - P_0)$ , then

$$\begin{aligned}\hat{s}(P_i u, P_i u) &= \hat{s}(R_i^\top \tilde{P}_i u, R_i^\top \tilde{P}_i u) \\ &\leq \frac{1}{\mathcal{K}_i} \tilde{s}_i(\tilde{P}_i u, \tilde{P}_i u) \quad (\text{Lemma 2.8}) \\ &= \frac{1}{\mathcal{K}_i} \hat{s}(u, R_i^\top \tilde{P}_i u) \quad (\text{definition of } \tilde{P}_i \text{ (2.3)}) \\ &= \frac{1}{\mathcal{K}_i} \hat{s}(u, P_i u).\end{aligned}$$

We use the fact that  $P_i = R_i^\top \tilde{P}_i$  and the definition of  $\hat{s}$  to write

$$\hat{s}(P_i u, u) = \sum_{j=1}^N s_j(R_j R_i^\top \tilde{P}_i, R_j u) = \sum_{\{j; R_j R_i^\top \neq 0\}} s_j(R_j R_i^\top \tilde{P}_i, R_j u).$$

We apply the Cauchy Schwarz inequality first for  $s_j$  then for the Euclidean inner product to this and inject the previous result (in the last step)

$$\begin{aligned}\hat{s}(P_i u, u) &\leq \sum_{\{j; R_j R_i^\top \neq 0\}} s_j(R_j R_i^\top \tilde{P}_i, R_j R_i^\top \tilde{P}_i)^{1/2} s_j(R_j u, R_j u)^{1/2} \\ &\leq \left[ \sum_{\{j; R_j R_i^\top \neq 0\}} s_j(R_j R_i^\top \tilde{P}_i, R_j R_i^\top \tilde{P}_i) \right]^{1/2} \left[ \sum_{\{j; R_j R_i^\top \neq 0\}} s_j(R_j u, R_j u) \right]^{1/2} \\ &= \hat{s}(P_i u, P_i u)^{1/2} \left[ \sum_{\{j; R_j R_i^\top \neq 0\}} s_j(R_j u, R_j u) \right]^{1/2} \\ &\leq \left( \frac{1}{\mathcal{K}_i} \hat{s}(u, P_i u) \right)^{1/2} \left[ \sum_{\{j; R_j R_i^\top \neq 0\}} s_j(R_j u, R_j u) \right]^{1/2}.\end{aligned}$$

Raising to the square and simplifying by  $\hat{s}(P_i u, u)$  yields

$$\hat{s}(P_i u, u) \leq \frac{1}{\mathcal{K}_i} \left[ \sum_{\{j; R_j R_i^\top \neq 0\}} s_j(R_j u, R_j u) \right].$$

Finally, summing these inequalities over  $i$  gives the result.  $\square$

### 2.3. Main theorem: convergence bound for balanced domain decomposition with the Generalized Eigenvalues in the Overlaps coarse space

We are now ready to give the estimates for the condition number of BDD with the GenEO coarse space.

*Theorem 2.11 (Main theorem for balanced domain decomposition with the Generalized Eigenvalues in the Overlaps coarse space)*

The condition number for BDD solved in  $\text{range}(I - P_0)$  with the deflated operator (2.9) satisfies

$$\kappa(P_{def}) \leq \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right). \quad (2.16)$$

As for the condition number of the balanced operator (2.10) with the GenEO coarse space, it satisfies

$$\kappa(P_{bal}) \leq \max \left\{ 1, \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \right\}. \quad (2.17)$$

These bounds depend only on the chosen thresholds  $\mathcal{K}_i$ , which we use to select eigenvectors for the coarse space in Definition 2.3 and on the maximal number  $\mathcal{N}$  of neighbors of a subdomain

$$\mathcal{N} = \max_{1 \leq i \leq N} (\# \{j; R_j R_i^\top \neq 0\}).$$

*Proof*

The proof of this theorem is the proof of Theorem 2.13 in [20]. The fact that the local solvers ( $\tilde{S}_i^\dagger$  here) are not spd does not play a role in the proof. The idea is to prove the following bounds

$$\hat{s}(u, u) \leq \hat{s}(P_{def}u, u) \leq \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \hat{s}(u, u); \quad u \in \text{range}(I - P_0), \quad (2.18)$$

and

$$\hat{s}(u, u) \leq \hat{s}(P_{bal}u, u) \leq \max \left\{ 1, \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \right\} \hat{s}(u, u); \quad u \in \hat{W}. \quad (2.19)$$

Following Lemma C.1 in the appendix of [20], these bounds imply the bounds for the condition numbers. They are proved using Lemmas 2.2 and 2.10 combined with the fact that  $P_0$  is an  $\hat{s}$ -orthogonal projection.  $\square$

*Remark 2.12*

The fact that  $\mathcal{K}_i$  can be chosen such that  $\left( \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \right) < 1$  in (2.18) is not a contradiction: in this case the space  $\text{range}(I - P_0)$  is simply empty.

### 3. FINITE ELEMENT TEARING AND INTERCONNECTING

We use the following references to introduce FETI: the book by Toselli and Widlund [20], Tezaur's dissertation [43], and the article by Klawonn and Widlund [44]. A second level was introduced for FETI in [45] and further developed in [46, 47].

#### 3.1. The finite element tearing and interconnecting formulation

In the BDD section, we built the coarse space for problem (2.1), which we simply recall here: find  $\hat{u} \in \hat{W}$  such that  $\hat{S}\hat{u} = \hat{f}_\Gamma$ , where  $\hat{W}$  is the space of functions defined on the interface  $\Gamma$ . Instead, the FETI formulation of the problem is on the product space  $W$  with an additional matching constraint at the interfaces. This constraint is ensured using matrix

$$B = (B_1, B_2, \dots, B_N); \quad Bu = \sum_{i=1, \dots, N} B_i u_i, \quad \forall u \in W, \quad (3.1)$$

which is constructed from entries 0, 1,  $-1$  such that the components  $u_i$  of a vector  $u$  in the product space  $W$  coincide on  $\Gamma$  when  $Bu = 0$ . More precisely, each line in  $B$  corresponds to one continuity constraint for one DOF and two of the subdomains to which it belongs: each line in  $B$  contains one

1 and one  $-1$ , whereas all other entries are zero. Denoting by  $\lambda$  the vector of Lagrange multipliers that is used to enforce the constraint  $Bu = 0$ , we obtain a saddle point formulation of the problem: find  $(u, \lambda) \in W \times U$  such that

$$\begin{pmatrix} S & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} f_\Gamma \\ 0 \end{pmatrix}. \quad (3.2)$$

We note that the solution  $\lambda$  of (3.2) is unique only up to an additive element of  $\text{Ker}(B^\top)$ ; however, the solution  $u$  to our problem does not depend on the choice of  $\lambda$  so this is not an issue in practice. For the theoretical study, we introduce the space

$$U := \text{range}(B) = \text{Ker}(B^\top)^\perp,$$

and will search for  $\lambda \in U$ . Given a basis for  $\text{Ker}(S)$ , which consists of  $n_K$  vectors, an important role is played by the prolongation operator  $\mathcal{R}_N^\top : \mathbb{R}^{n_K} \rightarrow W$  which columns are these basis functions. The transpose  $\mathcal{R}_N$  is a restriction operator that maps an element in  $W$  to the coordinates of its  $l_2$ -orthogonal projection onto  $\text{Ker}(S)$  in the same basis. We have used the subscript  $N$  because  $\text{Ker}(S)$  is often referred to as the *Natural* coarse space for FETI. Going back to the system, the solution of the first equation in (3.2) can be written as

$$u = S^\dagger(f_\Gamma - B^\top \lambda) + \mathcal{R}_N^\top \alpha, \quad \text{for some } \alpha \in \text{range}(\mathcal{R}_N), \quad (3.3)$$

if the right-hand side associated to the operator  $S$  is such that

$$f_\Gamma - B^\top \lambda \perp \text{Ker}(S) \Leftrightarrow \mathcal{R}_N(f_\Gamma - B^\top \lambda) = 0, \quad (3.4)$$

or with notation inspired by the usual FETI notation

$$G_N^\top \lambda = \mathcal{R}_N f_\Gamma, \quad G_N := B \mathcal{R}_N^\top. \quad (3.5)$$

Injecting (3.3) into the second equation in (3.2), we have

$$B S^\dagger B^\top \lambda - G_N \alpha = B S^\dagger f_\Gamma, \quad \text{for some } \alpha \in \text{range}(\mathcal{R}_N).$$

We may again rewrite the problem using a saddle point formulation as

$$\begin{pmatrix} F & -G_N \\ G_N^\top & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \alpha \end{pmatrix} = \begin{pmatrix} d \\ e \end{pmatrix}, \quad (3.6)$$

where

$$F := B S^\dagger B^\top, \quad d := B S^\dagger f_\Gamma, \quad e := \mathcal{R}_N f_\Gamma, \quad \text{and again } G_N = B \mathcal{R}_N^\top. \quad (3.7)$$

To homogenize the second equation and bring the problem down to a single equation, we decompose  $\lambda$  into  $\lambda = \tilde{\lambda} + \lambda_N$ , where  $G_N^\top \tilde{\lambda} = 0$  and  $G_N^\top \lambda_N = e$ . Then, we introduce a projection operator  $\mathcal{P}_N$  as follows: let  $Q : U \rightarrow U$  be a self-adjoint matrix that is positive definite on  $\text{range}(G_N)$ , then define

$$\mathcal{P}_N : U \rightarrow U; \quad \mathcal{P}_N := I - Q G_N (G_N^\top Q G_N)^{-1} G_N^\top. \quad (3.8)$$

### Remark 3.1

It is straightforward to prove that  $\mathcal{P}_N$  is a projection operator from  $U$  onto  $\text{Ker}(G_N^\top)$  and that its transpose  $\mathcal{P}_N^\top = I - G_N (G_N^\top Q G_N)^{-1} G_N^\top Q$  is a  $Q$ -orthogonal projection. It is however less obvious to prove that the inverse  $(G_N^\top Q G_N)^{-1}$  is well defined. This can be derived from the fact that  $Q$  is positive definite on  $\text{range}(G_N)$  so  $G_N^\top Q G_N \beta = 0$  implies  $G_N \beta = 0 \Leftrightarrow B \mathcal{R}_N^\top \beta = 0$ . In other words,  $\mathcal{R}_N^\top \beta \in \text{Ker}(S) \cap \text{Ker}(B)$  and this intersection are zero because the problem is well posed.<sup>‡</sup> Finally,  $\beta = 0$  and  $(G_N^\top Q G_N)^{-1}$  is well defined.

<sup>‡</sup>In case the global operator  $\hat{K}$  is singular, a solution exists for the original problem if  $\hat{f}$  is in the range of  $\hat{K}$ . In that case, the natural coarse grid becomes singular but the FETI approach can still be applied [52].

The system that we solve is the projected system into the space

$$V_N := \text{Ker}(G_N^\top) = \text{range}(\mathcal{P}_N). \quad (3.9)$$

For the choice  $\lambda_N := Q G_N (G_N^\top Q G_N)^{-1} \mathcal{R}_N f_\Gamma$  (which fulfills the condition  $G_N^\top \lambda_N = e$ ), the problem is: find  $\tilde{\lambda} \in V_N$  and  $\alpha \in \text{range}(\mathcal{R}_N)$  such that

$$F \tilde{\lambda} - G_N \alpha = d - F \lambda_N. \quad (3.10)$$

Testing this against elements in  $V_N$  yields the final form of the problem before preconditioning

$$\mathcal{P}_N^\top F \tilde{\lambda} = \mathcal{P}_N^\top (d - F \lambda_N), \quad (3.11)$$

whereas testing against function in  $\text{range}(I - \mathcal{P}_N)$  allows us to define the component  $\alpha$  of the solution completely with respect to  $\tilde{\lambda}$ :

$$(I - \mathcal{P}_N^\top) G_N \alpha = (I - \mathcal{P}_N^\top) (F \tilde{\lambda} - d + F \lambda_N) \Leftrightarrow \alpha = (G_N^\top Q G_N)^{-1} G_N^\top Q (F \lambda - d),$$

where we simply used a multiplication by  $(G_N^\top Q G_N)^{-1} G_N^\top Q$  to write the equivalence. Next, we introduce the two usual FETI preconditioners.

### 3.2. Usual preconditioners for finite element tearing and interconnecting

We first need to introduce diagonal scaling matrices  $D_i : W_i \rightarrow W_i$  for each  $i = 1, \dots, N$ . These are the matrix counterparts of the partition of unity operators  $\Xi_i$  used in the BDD section. Then

let  $D : W \rightarrow W$  be the diagonal scaling matrix  $D := \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & D_N \end{pmatrix}$ , on the product

space. We will consider two different preconditioners for (3.11): the Dirichlet preconditioner with the subscript  $D$  and the lumped preconditioner with the subscript  $L$  [41]. When scaled, those preconditioners can be written as the following operators on  $U$  [44]

$$\mathcal{M}_D^{-1} = \left[ D^{-1} B^\top (B D^{-1} B^\top)^\dagger \right]^\top S \left[ D^{-1} B^\top (B D^{-1} B^\top)^\dagger \right] \quad (3.12)$$

$$\mathcal{M}_L^{-1} = \left[ D^{-1} B^\top (B D^{-1} B^\top)^\dagger \right]^\top K^{bb} \left[ D^{-1} B^\top (B D^{-1} B^\top)^\dagger \right]. \quad (3.13)$$

We use the subscript  $*$  to refer to either of these preconditioners generically: if  $*$  denotes  $D$ , then  $\mathcal{M}_*^{-1} = \mathcal{M}_D^{-1}$  is the Dirichlet preconditioner, and if  $*$  denotes  $L$ , then  $\mathcal{M}_*^{-1} = \mathcal{M}_L^{-1}$  is the Lumped preconditioner. When the diagonal scaling matrix  $D$  is chosen to be the diagonal of the local operator matrix  $K$ , the scaling in the preconditioners (3.12,3.13) are equivalent to the so-called super-Lumped scaling (or  $K$ -scaling) originally proposed in [49].

#### Remark 3.2

In (3.12,3.13), we have used a pseudo inverse where the usual FETI theory uses an inverse. This has no impact on what follows. Indeed,  $(B D^{-1} B^\top)^\dagger$  is defined up to an additive element in  $\text{Ker}(B D^{-1} B^\top)$ , and we have the inclusion  $\text{Ker}(B D^{-1} B^\top) \subset \text{Ker}(B^\top)$  because

$$\lambda \in \text{Ker}(B D^{-1} B^\top) \Rightarrow D^{-1} B^\top \lambda \in \text{Ker}(B) \Rightarrow B^\top \lambda = D v \text{ for some } v \in \text{Ker}(B),$$

and  $\text{Ker}(B) = (\text{range}(B^\top))^\perp$  so  $v^\top B^\top \lambda = v^\top D v = 0 \Rightarrow v = 0 \Rightarrow \lambda \in \text{Ker}(B^\top)$ . The operator  $(B D^{-1} B^\top)^\dagger$  is applied to elements in  $\text{range}(B) = \text{Ker}(B^\top)^\perp$  so this application is well defined. Moreover, the application of  $(B D^{-1} B^\top)^\dagger$  is followed by an application of  $B^\top$  so  $D^{-1} B^\top (B D^{-1} B^\top)^\dagger$  is uniquely defined independently of the choice of the pseudo inverse. This pseudo inverse can be avoided by defining scaling matrices directly on the space of Lagrange multipliers, which is performed for instance in the redundant Lagrange multiplier section

of [44]. For sensible choices, both approaches can lead to identical preconditioners, and in practical implementations, the scaling matrices are actually never computed explicitly as is explained in [49].

By using the subscript  $*$  for either  $D$  or  $L$ , the preconditioned operator is  $\mathcal{M}_*^{-1} \mathcal{P}_N^\top F$ . Because we solve the system using a projected conjugate gradient method, we require that the search directions remain in  $V_N$ . Therefore, we actually solve: find  $\lambda \in V_N$  such that

$$\mathcal{P}_N \mathcal{M}_*^{-1} \mathcal{P}_N^\top F \lambda = \mathcal{P}_N \mathcal{M}_*^{-1} \mathcal{P}_N^\top (d - F \lambda_N). \quad (3.14)$$

Because of the projection step (3.11) and the choice  $\lambda_N := Q G_N (G_N^\top Q G_N)^{-1} \mathcal{R}_N f$ , this is already a two-level preconditioner where the coarse space is  $\text{Ker}(\mathcal{P}_N) = \text{range}(Q G_N) = \text{range}(Q B \mathcal{R}_N^\top)$ . The PPCG solver is initialized with  $\lambda_N$  and the entire solution space is  $\lambda_N + V_N$ . We will refer to  $\mathcal{P}_N$  as the natural coarse space projector.

The theoretical study of the preconditioner is related to operator

$$P_D : W \rightarrow W; \quad P_D := D^{-1} B^\top (B D^{-1} B^\top)^\dagger B, \quad (3.15)$$

where  $D : W \rightarrow W$  is the diagonal scaling matrix already introduced. This is a projection that is orthogonal in the scaled  $l_2$  inner product  $x^\top D y$  ( $x, y \in W$ ). The next two lemmas follow essentially by noticing that  $B P_D u = B u$ . They are Lemmas 4.1 and 4.3 in [44]. We give the proofs for the sake of completeness because they are short.

### Lemma 3.3

For any  $\mu \in U$ , there exists  $\tilde{u} \in \text{range}(P_D)$  such that  $\mu = B \tilde{u}$ .

#### Proof

By definition of  $U$ , there exists  $u \in W$  such that  $\mu = B u$ . Now take  $\tilde{u} = P_D u$ ,  $B \tilde{u} = B u = \mu$ .  $\square$

### Lemma 3.4

Let  $u \in W$ , then

$$P_D u = u - E_D u, \quad (3.16)$$

where  $E_D u : W \rightarrow W$  is an averaging operator defined by its components as:  $(E_D u)_i = R_i \sum_{j=1}^N R_j^\top D_j u_j$ .

#### Proof

We start by noticing that  $B(u - P_D u) = 0$ . This means that  $u - P_D u$  matches at the interfaces and thus, its weighted average satisfies  $E_D(u - P_D u) = u - P_D u$ . A sufficient condition to ensure that the result holds is now  $E_D P_D u = 0$ .

By definition of  $E_D$ ,  $E_D P_D u$  is a  $D$ -weighted average of the values of  $P_D u$ , which correspond to the same global *dof*. One way to compute the averaged value for global *dof*  $k$  is to first compute  $D P_D u = B^\top (B D^{-1} B^\top)^\dagger B u$  and then sum the contributions from the different subdomains for which  $k$  is a DOF. This is the same as computing an  $l_2$  scalar product between  $B^\top (B D^{-1} B^\top)^\dagger B u$  and the function  $e_x \in W$ , which is zero everywhere except at the DOF that correspond to global *dof*  $k$ . By definition,  $B e_x = 0$ . The orthogonality of  $\text{Ker}(B)$  and  $\text{range}(B^\top)$  allows us to conclude that  $\langle B e_x, B^\top (B D^{-1} B^\top)^\dagger B u \rangle = 0$  and thus  $E_D P_D u = 0$ .  $\square$

This last lemma allows us to prove that two suitable choices for  $Q$  in the projection operator  $\mathcal{P}_N$  are  $\mathcal{M}_D^{-1}$  and  $\mathcal{M}_L^{-1}$ .

### Lemma 3.5

Both preconditioners  $\mathcal{M}_D^{-1}$  and  $\mathcal{M}_L^{-1}$  defined by (3.12) and (3.13) are self-adjoint on  $U$  and positive definite on  $\text{range}(G_N)$ . Consequently, they are possible choices for matrix  $Q$  in the natural projection operator defined by (3.8).

*Proof*

We will only prove positive definiteness. Any  $\lambda \in \text{range}(G_N)$  writes  $\lambda = Bz$  for some  $z \in \text{Ker}(S)$ . Moreover, according to Lemma 1.6,  $\lambda \in \text{Ker}(\mathcal{M}_L^{-1})$  implies  $\lambda \in \text{Ker}(\mathcal{M}_D^{-1})$  so whether  $*$  denotes  $D$  or  $L$  we have  $\lambda = Bz \in \text{Ker}(\mathcal{M}_D^{-1})$ . Using the definitions of  $\mathcal{M}_D^{-1}$  and  $P_D$  as well as Lemma 3.4,

$$0 = \langle \mathcal{M}_D^{-1} Bz, Bz \rangle = \langle SP_D z, P_D z \rangle = \langle S(z - E_D z), z - E_D z \rangle.$$

Now we have  $z \in \text{Ker}(S)$  and  $z - E_D z \in \text{Ker}(S)$  so necessarily  $E_D z \in \text{Ker}(S)$ . By definition,  $E_D z \in \text{Ker}(B)$  (it is the  $D$ -weighted average of  $z$ ). The problem is well posed so  $\text{Ker}(S) \cap \text{Ker}(B) = 0$ . Finally,  $z = 0$  and  $\mathcal{M}_*^{-1}$  are positive definite on  $\text{range}(G_N)$ .  $\square$

We have just given two possible choices that complete the definition of the natural coarse space projector and thus the definitions of the spaces  $V_N$  and  $V'_N$ . The main result that we prove holds for these particular choices. For  $*$  denoting either  $D$  or  $L$ , we introduce the notation

$$\mathcal{P}_{*,N} := I - \mathcal{M}_*^{-1} G_N (G_N^\top \mathcal{M}_*^{-1} G_N)^{-1} G_N^\top \quad (3.17)$$

and

$$V_{*,N} = \text{range}(\mathcal{P}_{*,N}), \quad V'_{*,N} = \text{range}(\mathcal{P}_{*,N}^\top). \quad (3.18)$$

The next lemma states a crucial property for the preconditioners, which is that they are positive definite.

*Lemma 3.6*

The preconditioners  $\mathcal{P}_{*,N} \mathcal{M}_*^{-1} : V'_{*,N} \rightarrow V_{*,N}$  are spd for  $*$  denoting either  $D$  or  $L$ .

*Proof*

Again, we only prove positive definiteness. Consider any  $\mu \in V'_{*,N}$  with  $\langle \mathcal{P}_{*,N} \mathcal{M}_*^{-1} \mu, \mu \rangle = \langle \mathcal{M}_*^{-1} \mu, \mu \rangle = 0$ . By Lemma 3.3,  $\mu = B\tilde{u}$  for some  $\tilde{u} \in \text{range}(P_D)$ . Operator  $P_D$  is a projection so  $P_D \tilde{u} = \tilde{u}$ , and we obtain

$$0 = \langle \mathcal{M}_*^{-1} B\tilde{u}, B\tilde{u} \rangle = \begin{cases} |D^{-1} B^\top (BD^{-1} B^\top)^\dagger B\tilde{u}|_S^2 & = |P_D \tilde{u}|_S^2 = |\tilde{u}|_S^2 \text{ if } * = D, \\ |D^{-1} B^\top (BD^{-1} B^\top)^\dagger B\tilde{u}|_{K^{bb}}^2 & = |P_D \tilde{u}|_{K^{bb}}^2 = |\tilde{u}|_{K^{bb}}^2 \text{ if } * = L. \end{cases}$$

According to Lemma 1.6,  $|\tilde{u}|_{K^{bb}}^2 = 0$  implies  $|\tilde{u}|_S^2 = 0$  so, whether  $*$  denotes  $D$  or  $L$ , we have that  $\tilde{u} \in \text{Ker}(S)$ . By definition of  $\mathcal{R}_N$ ,  $\text{Ker}(S) = \text{range}(\mathcal{R}_N^\top)$  and in turn  $\mathcal{M}_*^{-1} B\tilde{u} = \mathcal{M}_*^{-1} \mu \in \text{range}(\mathcal{M}_*^{-1} G_N)$ .

The definition of  $V'_{*,N}$  rewrites

$$V'_{*,N} = \text{range}(\mathcal{P}_{*,N}^\top) = \text{Ker}(G_N^\top \mathcal{M}_*^{-1}) = \text{range}(\mathcal{M}_*^{-1} G_N)^\perp,$$

which together with  $\mu \in V'_{*,N}$  and  $\mathcal{M}_*^{-1} \mu \in \text{range}(\mathcal{M}_*^{-1} G_N)$  implies

$$0 = \langle \mu, \mathcal{M}_*^{-1} \mu \rangle.$$

Finally,  $\tilde{u} \in \text{range}(\mathcal{R}_N^\top)$  implies  $\mu \in \text{range}(G_N)$ , and  $\mathcal{M}_*^{-1}$  is positive definite on  $\text{range}(G_N)$  so  $\mu = 0$ .  $\square$

### 3.3. Two-level finite element tearing and interconnecting preconditioner with the Generalized Eigenvalues in the Overlaps coarse space

The proof of an upper bound for the spectrum of the preconditioned FETI system usually relies on strong assumptions on the set of equations at hand and the coefficient distribution. Once again, we build a coarse space that allows us to waive all of these assumptions. The coarse space is defined next



along with the two-level FETI preconditioners (deflated and balanced). We use again the subscript 0 to refer to the coarse space. To avoid confusion with the BDD case, we use calligraphic notation for the projection operator  $\mathcal{P}_{*,0}$ .

*Definition 3.7 (Generalized Eigenvalues in the Overlaps coarse spaces for finite element tearing and interconnecting)*

Let  $*$  denote either  $D$  (for Dirichlet) or  $L$  (for Lumped). For each subdomain  $i = 1, \dots, N$ , find the eigenpairs  $(q_i^k, \Lambda_i^k) \in W_i \times \mathbb{R}^+$  of the generalized eigenvalue problem:

$$S_i q_i^k = \Lambda_i^k (B_i^\top \mathcal{M}_*^{-1} B_i) q_i^k. \quad (3.19)$$

where  $\mathcal{M}_*^{-1}$  is the preconditioner defined either by (3.12) or (3.13). Next, given a threshold  $\mathcal{K}_i > 0$  for each subdomain, define the coarse space as

$$U_{*,0} = \text{span} \left( \left\{ \mathcal{M}_*^{-1} B_i q_i^k ; 0 < \Lambda_i^k < \mathcal{K}_i, i = 1, \dots, N \right\} \right). \quad (3.20)$$

Let the interpolation operator  $G_{*,0}$  be the matrix whose columns are the coarse basis functions  $\{\mathcal{M}_*^{-1} B_i q_i^k ; 0 < \Lambda_i^k < \mathcal{K}_i, i = 1, \dots, N\}$ . Let the coarse solver be the exact solver on  $U_{*,0}$ :

$$F_{*,0} := G_{*,0}^\top (\mathcal{P}_{*,N}^\top F \mathcal{P}_{*,N}) G_{*,0},$$

and let  $\mathcal{P}_{*,0}$  be the  $(\mathcal{P}_{*,N}^\top F \mathcal{P}_{*,N})$ -orthogonal projection operator defined by

$$\mathcal{P}_{*,0} := I - G_{*,0} F_{*,0}^\dagger G_{*,0}^\top (\mathcal{P}_{*,N}^\top F \mathcal{P}_{*,N}). \quad (3.21)$$

Then the two-level preconditioners (respectively, deflated and balanced) for  $F$  are

$$\mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top \quad \text{and} \quad \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top + \mathcal{P}_{*,N} G_{*,0} F_{*,0}^\dagger G_{*,0}^\top \mathcal{P}_{*,N}^\top. \quad (3.22)$$

The operator  $G_{*,0}$  is a mapping between the coordinates of a vector from  $U_{*,0}$  in the set of coarse basis functions and its representation in  $U$ . Its transpose  $G_{*,0}^\top$  is a restriction operator that maps an element in  $W$  to the coordinates of its  $l_2$  projection onto  $W_{*,0}$  in the set of coarse basis functions. The main difference with the coarse space for BDD is that we have left out the zero eigenvalues that correspond to the kernel of  $S$  because they are already taken care of by the natural coarse space through  $\mathcal{P}_N$ .

*Remark 3.8*

One common point with the BDD-GenEO eigenvalue problem is that one of the operators ( $S_i$ ) is a nonassembled operator on the local space  $W_i$ , whereas the other ( $B_i^\top \mathcal{M}_*^{-1} B_i$ ) is an assembled operator restricted to the local space  $W_i$ . This time, the words assembled and restricted are to be understood in the FETI context and rely on the mappings  $B_i$  between the DOF in  $W_i$  and the Lagrange multipliers in  $U$ . In the same way as for BDD, the role of the GenEO eigenvalue problem for FETI can be interpreted as finding the modes necessary for describing the discrepancy between the interface behavior as seen from a single domain (left-hand side of (3.19)), and the assembled interface operator  $F^{-1}$ , approximated by  $\mathcal{M}_*^{-1}$  (right-hand side of (3.19)). The idea is then to introduce those differences, which will not be well accounted for by the preconditioner, into the coarse space.

Once again in proving our estimate for the condition number, we will take advantage of the orthogonality type properties that result from the generalized eigenvalue problem.

*Lemma 3.9*

Let  $*$  denote either  $D$  or  $L$ . For a given subdomain  $i = 1, \dots, N$ , the eigenpairs  $(q_i^k, \Lambda_i^k)$  of the generalized eigenproblem (3.19) can be chosen so that the set  $\{q_i^k\}_k$  of eigenvectors is an orthonormal basis of  $W_i$  with respect to the inner product induced by  $B_i^\top \mathcal{M}_*^{-1} B_i$ . This writes

$$\langle \mathcal{M}_*^{-1} B_i q_i^k, B_i q_i^k \rangle = 1; \quad \text{and} \quad \langle \mathcal{M}_*^{-1} B_i q_i^k, B_i q_i^{k'} \rangle = 0, \quad k \neq k'.$$

An orthogonality-type property with respect to  $S_i$  (which is not necessarily coercive) also holds

$$\langle S_i q_i^k, q_i^{k'} \rangle = 0, \quad k \neq k'.$$

*Proof*

We proved in Lemma 3.5 that  $\mathcal{M}_*^{-1}$  is spd on  $\text{range}(G_N) = \text{Ker}(\mathcal{P}_N^\top)$ . We also proved in Lemma 3.6 that  $\mathcal{M}_*^{-1}$  is spd on  $V'_N = \text{range}(\mathcal{P}_N^\top)$ . So  $\mathcal{M}_*^{-1}$  is spd on  $\text{Ker}(\mathcal{P}_N^\top) \oplus \text{range}(\mathcal{P}_N^\top) = U$ . Finally, by definition of  $B_i$ ,  $B_i u_i = 0$  implies  $u_i = 0$  so  $B_i^\top \mathcal{M}_*^{-1} B_i$  is spd on  $W_i$  and the result is well known.  $\square$

In the next lemma, we give some useful properties of the projections.

*Lemma 3.10*

- (i)  $\text{range}(\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top) \subset \text{range}(\mathcal{P}_{*,N}^\top)$ .
- (ii)  $\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top = \mathcal{P}_{*,N}^\top \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top$
- (iii)  $\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top$  and  $\mathcal{P}_{*,N}^\top \mathcal{P}_{*,0}^\top$  are projections.

*Proof*

- (i) By definition of  $\mathcal{P}_{*,0}$  (3.21):  $\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top = \mathcal{P}_{*,N}^\top (I - F \mathcal{P}_{*,N} G_0 F_0^\dagger G_0^\top)$ .
- (ii) It follows from (3.10) and the fact that  $\mathcal{P}_{*,N}^\top$  is a projection that  $\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top = \mathcal{P}_{*,N}^\top \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top$ .
- (iii) Then  $\mathcal{P}_{*,0}^\top$  is also a projection so  $\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top = \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top$ .

$\square$

For two spd matrices  $M_1$  and  $M_2$  of same size, the spectrum of  $M_1 M_2$  is identical to the spectrum of  $M_2 M_1$ . Following this idea, we decide to look at the problem in reverse: *Is  $F$  a good preconditioner for  $\mathcal{M}_*^{-1}$ ?* The reason we do this is that then we recognize an abstract Schwarz type preconditioner  $F = \sum_{i=1}^N B_i S_i^\dagger B_i^\top$ . In this framework, the local subspaces are the  $W_i$  and the local solvers are the pseudo inverses  $S_i^\dagger$  of the local bilinear forms  $S_i$ . The prolongation operators are the  $B_i : W_i \rightarrow U$  and the restriction operators are the  $B_i^\top : U \rightarrow W_i$ . Taking advantage of the abstract Schwarz framework, in Lemmas 3.11 and 3.13, we will prove the same estimates as in the BDD subsection for  $F$  viewed as the preconditioner and  $\mathcal{M}_*^{-1}$  viewed as the matrix problem. In the proof of our final theorem, it will become apparent that these estimates allow to prove the condition number of FETI with the two-level preconditioners given by (3.22). In the next Lemma, applying the exact same strategy as in Lemma 2.2, we give an estimate related to a lower bound for the eigenvalues of the preconditioned operator  $F \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1}$ . This bound does not depend on the choice of the coarse space.

*Lemma 3.11 (Stable-splitting – lower bound for the eigenvalues of the preconditioned operator)*

For any  $\mu \in V'_{*,N}$ , there exists a stable-splitting  $(v_1, \dots, v_N) \in W_1 \times \dots \times W_N$  of  $\mu$

$$\mu = B_1 v_1 + \dots + B_N v_N; \quad v_i \in W_i \quad \text{and} \quad \sum_{i=1}^N \langle S_i v_i, v_i \rangle \leq \langle \mathcal{M}_*^{-1} \mu, \mu \rangle. \quad (3.23)$$

This implies

$$\langle \mathcal{M}_*^{-1} \mu, \mu \rangle \leq \langle F \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu, \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu \rangle \text{ for any } \mu \in \text{range}(\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top).$$

*Proof*

Let  $\mu \in V'_{*,N}$  and let  $v_i = D_i^{-1} B_i^\top (B D^{-1} B^\top)^\dagger \mu$  for each  $i = 1, \dots, N$ . This provides a splitting of  $\mu$

$$\sum_{i=1}^N B_i v_i = \sum_{i=1}^N B_i D_i^{-1} B_i^\top (B D^{-1} B^\top)^\dagger \mu = (B D^{-1} B^\top) (B D^{-1} B^\top)^\dagger \mu = \mu,$$

since  $\mu \in \text{range}(B D^{-1} B^\top) = \text{range}(B) = U$ . Moreover, the splitting is stable

$$\begin{aligned} \sum_{i=1}^N \langle S_i v_i, v_i \rangle &= \sum_{i=1}^N \langle S_i D_i^{-1} B_i^\top (B D^{-1} B^\top)^\dagger \mu, D_i^{-1} B_i^\top (B D^{-1} B^\top)^\dagger \mu \rangle \\ &= \langle S D^{-1} B^\top (B D^{-1} B^\top)^\dagger \mu, D^{-1} B^\top (B D^{-1} B^\top)^\dagger \mu \rangle \\ &= \langle \mathcal{M}_D^{-1} \mu, \mu \rangle, \\ &\leq \langle \mathcal{M}_*^{-1} \mu, \mu \rangle, \end{aligned}$$

by Lemma 1.6. This is exactly (3.23). Now let  $\mu \in \text{range}(\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top)$ , then  $\langle \mathcal{M}_*^{-1} \mu, \mu \rangle = \langle \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu, \mu \rangle$ . Moreover, the fact that the  $v_i$  provides a splitting implies

$$\begin{aligned} \langle \mathcal{M}_*^{-1} \mu, \mu \rangle &= \langle \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu, \sum_{i=1}^N B_i v_i \rangle \\ &= \sum_{i=1}^N \langle \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu, B_i (S_i^\dagger S_i) v_i \rangle \\ &= \sum_{i=1}^N \langle S_i v_i, S_i^\dagger B_i^\top \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu \rangle. \end{aligned}$$

Then we apply the Cauchy–Schwarz inequality twice, first in the  $S_i$  inner product and then in the  $l_2$  inner product and finish by using (3.23)

$$\begin{aligned} \langle \mathcal{M}_*^{-1} \mu, \mu \rangle &\leq \sum_{i=1}^N \left[ \langle S_i v_i, v_i \rangle^{1/2} \langle S_i S_i^\dagger B_i^\top \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu, S_i^\dagger B_i^\top \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu \rangle^{1/2} \right] \\ &\leq \left[ \sum_{i=1}^N \langle S_i v_i, v_i \rangle \right]^{1/2} \left[ \sum_{i=1}^N \langle S_i S_i^\dagger B_i^\top \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu, S_i^\dagger B_i^\top \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu \rangle \right]^{1/2} \\ &\leq \langle \mathcal{M}_*^{-1} \mu, \mu \rangle^{1/2} \langle \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu, \sum_{i=1}^N B_i S_i^\dagger B_i^\top \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mu \rangle^{1/2}. \end{aligned}$$

The result follows by raising to the square, simplifying by  $\langle \mathcal{M}_*^{-1} \mu, \mu \rangle$ , and recognizing  $F = \sum_{i=1}^N B_i S_i^\dagger B_i^\top$ .  $\square$

The next lemma is the FETI counterpart of Lemma 2.8 and the proof follows the exact same steps. We prove a crucial result that relies very strongly on the choice of the coarse space. In fact, the coarse space was chosen specifically to ensure that this estimate holds.

*Lemma 3.12 (Stability of the local solvers)*

Let  $*$  denote either  $D$  or  $L$ . For each  $i = 1, \dots, N$ , let the pseudo inverse  $S_i^\dagger$  be chosen such that  $\text{range}(S_i^\dagger) = \text{span}\{q_i^k; \Lambda_i^k > 0\}$ . Then the following estimate for the local solver holds

$$\langle \mathcal{M}_*^{-1} B_i u_i, B_i u_i \rangle \leq \frac{1}{\mathcal{K}_i} \langle S_i u_i, u_i \rangle, \quad \forall u_i \in \text{range}\left(S_i^\dagger B_i^\top \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top\right), \quad (3.24)$$

where the  $\mathcal{K}_i$  are the thresholds that were used to select eigenvectors for the coarse space in Definition 3.7.

*Proof*

First, we prove that  $\text{range}\left(S_i^\dagger B_i^\top \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top\right) \subset \text{span}\{q_i^k; \Lambda_i^k \geq \mathcal{K}_i\}$ . We will use the following linear algebra identity

$$\text{Ker}(\mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} B_i) \oplus^\perp \text{range}(B_i^\top \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top) = W_i, \quad (3.25)$$

where the symbol  $\perp$  refers to the  $l_2$  orthogonality between both spaces and  $\oplus$  means that the sum is direct. According to item (ii) in Lemma 3.10,  $\mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top = \mathcal{P}_{*,N}^\top \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top$ . This implies  $\mathcal{P}_{*,N} \mathcal{P}_{*,0} = \mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{P}_{*,N}$ . So  $\text{Ker}(\mathcal{P}_{*,N}) \subset \text{Ker}(\mathcal{P}_{*,N} \mathcal{P}_{*,0})$ . It is also obvious that  $\text{Ker}(\mathcal{P}_{*,0}) \subset \text{Ker}(\mathcal{P}_{*,N} \mathcal{P}_{*,0})$ . By using the definitions of these projections ((3.17) and (3.21)), this rewrites

$$\text{Ker}(\mathcal{P}_{*,N} \mathcal{P}_{*,0}) \supset (\text{Ker}(\mathcal{P}_{*,N}) \cup \text{Ker}(\mathcal{P}_{*,0})) \supset (\text{range}(G_{*,0}) \cup \text{range}(\mathcal{M}_*^{-1} G_N)).$$

By definition of  $G_{*,0}$  and  $G_N$ , in particular, for each  $i = 1, \dots, N$ ,

$$\text{span}\{\mathcal{M}_*^{-1} B_i q_i^k; \Lambda_i^k < \mathcal{K}_i\} \subset \text{Ker}(\mathcal{P}_{*,N} \mathcal{P}_{*,0}),$$

so

$$\text{span}\{q_i^k; \Lambda_i^k < \mathcal{K}_i\} \subset \text{Ker}(\mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} B_i). \quad (3.26)$$

Following the same procedure as to prove (2.14) in Lemma 2.8, the first orthogonality property in Lemma 3.9 implies that

$$\text{span}\{q_i^k; \Lambda_i^k < \mathcal{K}_i\} \oplus^\perp \text{span}\{B_i^\top \mathcal{M}_*^{-1} B_i q_i^k; \Lambda_i^k \geq \mathcal{K}_i\} = W_i. \quad (3.27)$$

Putting (3.25), (3.26), and (3.27) together tells us that

$$\text{range}(B_i^\top \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top) \subset \text{span}\{B_i^\top \mathcal{M}_*^{-1} B_i q_i^k; \Lambda_i^k \geq \mathcal{K}_i\}.$$

Next, the definition of eigenproblem (3.19),  $S_i q_i^k = \Lambda_i^k (B_i^\top \mathcal{M}_*^{-1} B_i) q_i^k$ , yields

$$\text{range}(B_i^\top \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top) \subset \text{span}\{S_i q_i^k; \Lambda_i^k \geq \mathcal{K}_i\}.$$

Finally, for the specific choice of the pseudo inverse  $S_i^\dagger$ , it is obvious that

$$\text{range}(S_i^\dagger B_i^\top \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top) \subset \text{span}\{q_i^k; \Lambda_i^k \geq \mathcal{K}_i\}.$$

Now it is easy to prove (3.24) using the orthogonality type properties in Lemma 3.9 and the definition of the eigenproblem. Any  $u_i \in \text{range} \left( S_i^\top B_i^\top \mathcal{M}_{*,0}^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top \right)$  writes  $u_i = \sum_{\{k; \Lambda_i^k \geq K_i\}} \alpha_i^k q_i^k$  for some coefficients  $\alpha_i^k \in \mathbb{R}$ , so

$$\begin{aligned} \langle \mathcal{M}_{*}^{-1} B_i u_i, B_i u_i \rangle &= \sum_{\{k; \Lambda_i^k \geq K_i\}} \alpha_i^{k^2} \langle \mathcal{M}_{*}^{-1} B_i q_i^k, B_i q_i^k \rangle \\ &= \sum_{\{k; \Lambda_i^k \geq K_i\}} \frac{1}{\Lambda_i^k} \alpha_i^{k^2} \langle S_i q_i^k, q_i^k \rangle \\ &\leq \frac{1}{K_i} \sum_{\{k; \Lambda_i^k \geq K_i\}} \alpha_i^{k^2} \langle S_i q_i^k, q_i^k \rangle \\ &= \frac{1}{K_i} \langle S_i u_i, u_i \rangle \end{aligned}$$

□

The next lemma is a direct consequence. It is the FETI counterpart of Lemma 2.10 and gives an estimate related to an upper bound for the eigenvalues of the preconditioned operator. The relationship will become apparent in the proof of the final theorem.

*Lemma 3.13 (Upper bound for the eigenvalues of the preconditioned operator)*

The following estimate holds

$$\langle F \mathcal{M}_{*}^{-1} \lambda, \mathcal{M}_{*}^{-1} \lambda \rangle \leq \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{K_i} \right) \langle \mathcal{M}_{*}^{-1} \lambda, \lambda \rangle \text{ for any } \lambda \in \text{range} \left( \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top \right), \quad (3.28)$$

where  $\mathcal{N}$  is the maximal number of neighbors of a subdomain (including itself) in the sense

$$\mathcal{N} = \max_{1 \leq i \leq N} (\#\{j; B_j^\top B_i \neq 0\}).$$

*Proof*

To simplify notation, let us write  $\tilde{\mathcal{P}}_{*,i} := S_i^\top B_i^\top \mathcal{M}_{*}^{-1}$  and  $\mathcal{P}_{*,i} := B_i \tilde{\mathcal{P}}_{*,i}$ . Let  $\lambda \in \text{range} \left( \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top \right)$ , then

$$\begin{aligned} \langle \mathcal{M}_{*}^{-1} \mathcal{P}_{*,i} \lambda, \mathcal{P}_{*,i} \lambda \rangle &= \langle \mathcal{M}_{*}^{-1} B_i \tilde{\mathcal{P}}_{*,i} \lambda, B_i \tilde{\mathcal{P}}_{*,i} \lambda \rangle \\ &\leq \frac{1}{K_i} \langle S_i \tilde{\mathcal{P}}_{*,i} \lambda, \tilde{\mathcal{P}}_{*,i} \lambda \rangle \quad (\text{Lemma 3.12}) \\ &= \frac{1}{K_i} \langle \mathcal{M}_{*}^{-1} \lambda, B_i \tilde{\mathcal{P}}_{*,i} \lambda \rangle \quad (\text{definition of } \tilde{\mathcal{P}}_{*,i}) \\ &= \frac{1}{K_i} \langle \mathcal{M}_{*}^{-1} \lambda, \mathcal{P}_{*,i} \lambda \rangle \end{aligned} \quad (3.29)$$

Taking a close look at the definition of the preconditioners in (3.12) and (3.13), we notice that they can be written as a sum of local contributions:

$$\mathcal{M}_{*}^{-1} = \sum_{j=1}^N \mathcal{M}_{*,j}^{-1}; \quad \mathcal{M}_{*,j}^{-1} := \left[ D_j^{-1} B_j^\top (B D^{-1} B^\top)^\dagger \right]^\top S_j \left[ D_j^{-1} B_j^\top (B D^{-1} B^\top)^\dagger \right],$$

and  $\langle \mathcal{M}_{*,j}^{-1} B_i u_i, u_i \rangle \neq 0$  if and only if  $B_j^\top B_i \neq 0$ . A consequence of this is that

$$\langle \mathcal{M}_{*}^{-1} \lambda, \mathcal{P}_{*,i} \lambda \rangle = \langle \mathcal{M}_{*}^{-1} \lambda, B_i \tilde{\mathcal{P}}_{*,i} \lambda \rangle = \sum_{\{j; B_j B_i^\top \neq 0\}} \langle \mathcal{M}_{*,j}^{-1} \lambda, B_i \tilde{\mathcal{P}}_{*,i} \lambda \rangle.$$

We apply the Cauchy–Schwarz inequality for  $\mathcal{M}_{*,j}^{-1}$  and then for the Euclidean inner product to this and inject the previous result

$$\begin{aligned}
\langle \mathcal{M}_*^{-1} \lambda, \mathcal{P}_{*,i} \lambda \rangle &\leq \sum_{\{j; B_j B_i^\top \neq 0\}} \langle \mathcal{M}_{*,j}^{-1} \lambda, \lambda \rangle^{1/2} \langle \mathcal{M}_{*,j}^{-1} \mathcal{P}_{*,i} \lambda, \mathcal{P}_{*,i} \lambda \rangle^{1/2} \\
&\leq \left[ \sum_{\{j; B_j B_i^\top \neq 0\}} \langle \mathcal{M}_{*,j}^{-1} \lambda, \lambda \rangle \right]^{1/2} \left[ \sum_{\{j; B_j B_i^\top \neq 0\}} \langle \mathcal{M}_{*,j}^{-1} \mathcal{P}_{*,i} \lambda, \mathcal{P}_{*,i} \lambda \rangle \right]^{1/2} \\
&= \left[ \sum_{\{j; B_j B_i^\top \neq 0\}} \langle \mathcal{M}_{*,j}^{-1} \lambda, \lambda \rangle \right]^{1/2} \langle \mathcal{M}_*^{-1} \mathcal{P}_{*,i} \lambda, \mathcal{P}_{*,i} \lambda \rangle^{1/2} \\
&\leq \left[ \sum_{\{j; B_j B_i^\top \neq 0\}} \langle \mathcal{M}_{*,j}^{-1} \lambda, \lambda \rangle \right]^{1/2} \left[ \frac{1}{\mathcal{K}_i} \langle \mathcal{M}_*^{-1} \lambda, \mathcal{P}_{*,i} \lambda \rangle \right]^{1/2} \quad (\text{from (3.29)}).
\end{aligned}$$

Raising to the square and simplifying by  $\langle \mathcal{M}_*^{-1} \lambda, \mathcal{P}_{*,i} \lambda \rangle$  yields

$$\langle \mathcal{M}_*^{-1} \lambda, \mathcal{P}_{*,i} \lambda \rangle \leq \frac{1}{\mathcal{K}_i} \sum_{\{j; B_j B_i^\top \neq 0\}} \langle \mathcal{M}_{*,j}^{-1} \lambda, \lambda \rangle.$$

Finally, summing these inequalities over  $i$  and noticing that  $\sum_{i=1}^N \mathcal{P}_{*,i} = F \mathcal{M}_*^{-1}$  ends the proof.  $\square$

We are now ready to prove the main theorem for the GenEO FETI algorithm, which is similar to Theorem 2.11.

*Theorem 3.14 (Main theorem for finite element tearing and interconnecting with the Generalized Eigenvalues in the Overlaps coarse space)*

Let  $*$  denote either  $L$  for Lumped or  $D$  for Dirichlet. The condition number for FETI solved in  $\text{range}(\mathcal{P}_{*,N} \mathcal{P}_{*,0})$  with the deflated operator satisfies

$$\kappa(\mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top F) \leq \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right). \quad (3.30)$$

As for the balanced two-level preconditioner with the GenEO coarse space in  $\text{range}(\mathcal{P}_{*,N})$ , it satisfies

$$\kappa(\mathcal{P}_{*,N} \mathcal{P}_{*,0} \mathcal{M}_*^{-1} \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top F + \mathcal{P}_{*,N} G_{*,0} F_{*,0}^\dagger G_{*,0}^\top \mathcal{P}_{*,N}^\top F) \leq \max \left\{ 1, \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \right\}. \quad (3.31)$$

These bounds depend only on the chosen thresholds  $\mathcal{K}_i$  we use to select eigenvectors for the coarse space in Definition 3.7 and on the maximal number  $\mathcal{N}$  of neighbors of a subdomain (including itself):

$$\mathcal{N} = \max_{1 \leq i \leq N} (\# \{j; B_j^\top B_i \neq 0\}).$$

*Proof*

From Lemma C.1 in the appendix of [20], to prove (3.30), it is sufficient to show that, for any  $\lambda \in \text{range}(\mathcal{P}_{*,N}\mathcal{P}_{*,0})$ , the following holds:

$$\begin{aligned} \langle (\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mathcal{P}_{*,0}^\top\mathcal{P}_{*,N}^\top)^{-1}\lambda, \lambda \rangle &\leq \langle F\lambda, \lambda \rangle \\ &\leq \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \langle (\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mathcal{P}_{*,0}^\top\mathcal{P}_{*,N}^\top)^{-1}\lambda, \lambda \rangle. \end{aligned} \quad (3.32)$$

Lemma 3.6 tells us that the inverse is well defined. First of all, note that the fact that  $\mathcal{K}_i$  can be chosen such that  $\left( \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \right) < 1$  in (3.32) is not a contradiction: in this case, the space  $\text{range}(\mathcal{P}_{*,N}\mathcal{P}_{*,0})$  is simply empty. Next, we prove (3.32): let  $\mu \in \text{range}(\mathcal{P}_{*,0}^\top\mathcal{P}_{*,N}^\top)$ , Lemma 3.11 tells us that

$$\langle \mathcal{M}_*^{-1}\mu, \mu \rangle \leq \langle F\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu, \mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu \rangle.$$

Then, using the fact that  $\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mathcal{P}_{*,0}^\top\mathcal{P}_{*,N}^\top\mu = \mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu$ , this is equivalent to

$$\begin{aligned} &\langle (\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mathcal{P}_{*,0}^\top\mathcal{P}_{*,N}^\top)^{-1}\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu, \mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu \rangle \\ &\leq \langle F\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu, \mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu \rangle. \end{aligned}$$

In turn,  $\text{range}(\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mathcal{P}_{*,0}^\top\mathcal{P}_{*,N}^\top) = \text{range}(\mathcal{P}_{*,N}\mathcal{P}_{*,0})$  implies

$$\langle (\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mathcal{P}_{*,0}^\top\mathcal{P}_{*,N}^\top)^{-1}\lambda, \lambda \rangle \leq \langle F\lambda, \lambda \rangle, \quad \forall \lambda \in \text{range}(\mathcal{P}_{*,N}\mathcal{P}_{*,0}),$$

which is the lower bound in (3.32).

For the upper bound, we use the result from Lemma 3.13, which is that

$$\langle F\mathcal{M}_*^{-1}\mu, \mathcal{M}_*^{-1}\mu \rangle \leq \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \langle \mathcal{M}_*^{-1}\mu, \mu \rangle, \quad \forall \mu \in \text{range}(\mathcal{P}_{*,0}^\top\mathcal{P}_{*,N}^\top).$$

We know that  $\mathcal{M}_*^{-1}\mu = \mathcal{P}_{*,N}\mathcal{M}_*^{-1}\mu$  and projection  $\mathcal{P}_{*,0}$  is  $(\mathcal{P}_{*,N}^\top F\mathcal{P}_{*,N})$ -orthogonal so

$$\langle F\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu, \mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu \rangle \leq \langle F\mathcal{M}_*^{-1}\mu, \mathcal{M}_*^{-1}\mu \rangle,$$

and in turn

$$\langle F\mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu, \mathcal{P}_{*,N}\mathcal{P}_{*,0}\mathcal{M}_*^{-1}\mu \rangle \leq \mathcal{N} \max_{1 \leq i \leq N} \left( \frac{1}{\mathcal{K}_i} \right) \langle \mathcal{M}_*^{-1}\mu, \mu \rangle.$$

In the same way as for the lower bound, we may then show the upper bound in (3.32). This ends the proof for the condition number of the deflated preconditioned operator (3.30). The proof for the balanced operator (3.31) is similar to the BDD case; it relies simply on the fact that the projection operator  $\mathcal{P}_{*,0}$  is  $(\mathcal{P}_{*,N}^\top F\mathcal{P}_{*,N})$ -orthogonal.  $\square$

#### 4. NUMERICAL RESULTS FOR TWO-DIMENSIONAL ELASTICITY (FINITE ELEMENT TEARING AND INTERCONNECTING)

We give here a few numerical results to confirm the estimate for the condition number in the FETI case. The system of equations that we solve is related to two-dimensional linear elasticity where the domain is clamped on the left-hand side and subject to gravity. An important feature of the methods, which we presented is that, given a FETI code, they do not demand a lot of implementation work: all the mathematical objects that are used to build the coarse space already appear in the algorithms.



All the results that follow were obtained using Freefem++ [53] to build the problem matrices and visualize solutions and Matlab for the solving procedure. The test problems we present here are only small tests that we use to validate our theoretical results. Of course, a full validation of the efficiency of the method would require larger scale tests with an optimized code. Full reorthogonalization at each iteration is used in PPCG. The meshes are regular with quadrilateral elements and the finite element discretization of the two-dimensional elasticity equation uses standard  $\mathbb{P}_1$  (linear) functions. There are two parameters in the linear elasticity system of equations: Young's modulus  $E$  and Poisson's ratio  $\nu$ . Each time an iteration count is given, the stopping criterion is that the relative primal residual at the final iteration  $k$  reach  $10^{-5}$

$$\frac{\|\sum_{i=1}^N R_i^\top S_i D_i^{-1} B_i^\top (B D^{-1} B^\top)^\dagger \mathcal{P}_{*,0}^\top \mathcal{P}_{*,N}^\top (d - F \lambda_k)\|_2}{\|\hat{f}_\Gamma\|_2} < 10^{-4}.$$

The fact that this is indeed the primal residual is explained in [54] and proved for instance in [55].

#### 4.1. Checkerboard coefficient distribution

We discretize a square of size  $1 \times 1$  using  $81 \times 81$  nodes. We use two different decompositions of this unit square: a regular decomposition into  $8 \times 8$  regular subdomains (Figure 1 – left) and a decomposition into 64 subdomains obtained using Metis [1] (Figure 1 – middle). Throughout this subsection, the scaling matrices are chosen to be the  $K$ -scaling matrices [44, 49], meaning that in the definitions of the preconditioners (3.12) and (3.13), we set

$$D_i = \text{diag}(K_i). \quad (4.1)$$

The criterion for selecting which modes are used to build the coarse space is set to

$$\mathcal{K}_i = 0.1; \quad \forall i = 1, \dots, N,$$

so the condition number should satisfy  $\kappa \leq 10 \times \mathcal{N}$  where  $\mathcal{N}$  is the maximal number of neighbors.

**4.1.1. The partition resolves the heterogeneities.** It is well known by now that in the case of a regular decomposition into subdomains, which resolves the jumps in the coefficients and the Dirichlet preconditioner, the use of the  $K$ -scaling matrices (4.1) is sufficient to ensure good convergence. We check here that in these cases, the (automatic) GenEO strategy is to do nothing special, which is to say that no extra modes are selected to build the additional coarse space  $U_0$ . Table II gives the results for the regular partition (Figure 1 – left) into subdomains and a constant coefficient distribution  $(E; \nu) = (10^7; 0.4)$  as well as a *checkerboard* coefficient distribution (Figure 1 – right) where the coefficients take the values  $(E_1; \nu_1) = (10^7; 0.4)$  and  $(E_2; \nu_2) = (10^{12}; 0.3)$ . We have solved each of these problems with the Dirichlet preconditioner and the Lumped preconditioner with and without the GenEO coarse space (we refer to these cases as FETI-GenEO and FETI-1, respectively).

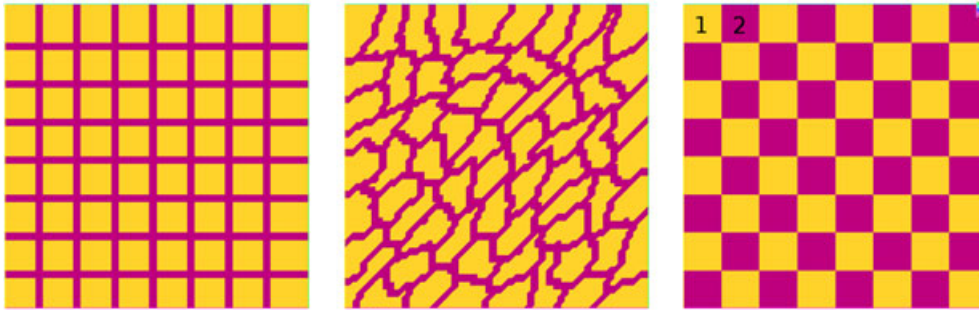


Figure 1. Decomposition of the unit square into 64 regular subdomains (left) – decomposition of the unit square into 64 subdomains using Metis (middle) – checkerboard coefficient distribution (right).

Table II. Checkerboard (64 regular subdomains)  $\kappa$  : condition number;  $\#U_0$ : size of the GenEO coarse space;  $it$ : number of iterations – For the Dirichlet preconditioner, the GenEO coarse space is empty so FETI-GenEO and FETI-1 are identical.

	Dirichlet preconditioner			Lumped preconditioner				
				FETI-GenEO		FETI-1		
Coefficients	$\kappa$	$\#U_0$	$it$	$\kappa$	$\#U_0$	$it$	$\kappa$	$it$
Constant	9.5	0	15	11.1	15	17	86	24
Checkerboard	6.3	0	13	9.7	49	19	93	25

For each test, we give the condition number  $\kappa$  of the preconditioned operator, the size of the GenEO coarse space  $\#U_0$  (if there is one), and the number  $it$  of iterations needed to reach convergence. The first thing that we notice is that in all four cases where the GenEO coarse space is used, the estimate for the condition number is satisfied. In the Dirichlet preconditioner case, no modes were selected to build the coarse space, which is what we expected because the  $K$ -scaling alone is known to be efficient. With the Lumped preconditioner case, only few modes were selected (less than one per subdomain). This test indicates that the GenEO coarse grid circumvents the fact that the Lumped preconditioner does not properly predict the corrections needed on the interface for checkerboard problems.

*4.1.2. The partition does not resolve the heterogeneities.* This time we use the automatic partition into 64 subdomains obtained using METIS [1] (Figure 1 – middle). The coefficient distribution is still the checkerboard distribution shown on the right-hand side of Figure 1 so the subdomain interfaces do not coincide with the jumps in the coefficients. The coefficients are a fixed  $(E_1; v_1) = (10^7; 0.4)$  and a variable  $(E_2, v_2)$  one. Table III gives the results for different values of  $(E_2, v_2)$ . The middle line shows a case where the coefficients are constant throughout the subdomain  $((E_2, v_2) = (E_1; v_1))$ . Once again, we observe that in all cases, the condition number satisfies the estimate and that it hardly varies with the jumps in the coefficients. In the worse case, the number of modes used to build the coarse space is 370 (less than six modes per subdomain on average). Because of bad numerical conditioning, there are a few cases where the FETI-1 residual never reaches  $10^{-4}$ , instead it stagnates. In this case, we report the iteration count before the *plateau* and the corresponding residual. Figure 2 shows a comparison between the convergence curves with and without the additional GenEO coarse space where this phenomenon can be observed. Figure 3 shows the spectrum of the preconditioned operators with and without the additional coarse space. The spectrum is represented in the complex plane but the imaginary part is always almost zero (imaginary parts result from numerical errors in the eigensolver). The zeros in the spectrum correspond to the

Table III. Checkerboard (64 Metis subdomains)  $(E_1; v_1) = (10^7; 0.4)$ ;  $\kappa$  : condition number;  $\#U_0$ : size of the GenEO coarse space;  $it$ : number of iterations. When  $(E_2; v_2) = (10^7; 0.4)$ , there are no jumps in the coefficients.

	Dirichlet preconditioner					Lumped preconditioner				
	FETI-GenEO			FETI-1		FETI-GenEO		FETI-1		
$(E_2; v_2)$	$\kappa$	$\#U_0$	$it$	$\kappa$	$it$	$\kappa$	$\#U_0$	$it$	$\kappa$	$it$
$(10^{12}; 0.3)$	10.4	126	18	$1.5 \cdot 10^6$	142 <sup>(1)</sup>	11.7	186	19	$6.2 \cdot 10^6$	154 <sup>(2)</sup>
$(10^7; 0.4)$	10.5	26	18	447	31	12.2	99	23	$2.1 \cdot 10^3$	58
$(10^2; 0.49)$	12.2	182	21	$5.3 \cdot 10^6$	170 <sup>(3)</sup>	16.3	370	23	$4.0 \cdot 10^7$	198 <sup>(4)</sup>

<sup>(1)</sup> The relative residual reaches a plateau at  $2 \cdot 10^{-4}$  after 142 iterations.

<sup>(2)</sup> The relative residual reaches a plateau at  $3 \cdot 10^{-4}$  after 154 iterations.

<sup>(3)</sup> The relative residual reaches a plateau at  $2 \cdot 10^{-3}$  after 170 iterations.

<sup>(4)</sup> The relative residual reaches a plateau at  $1 \cdot 10^{-3}$  after 198 iterations.

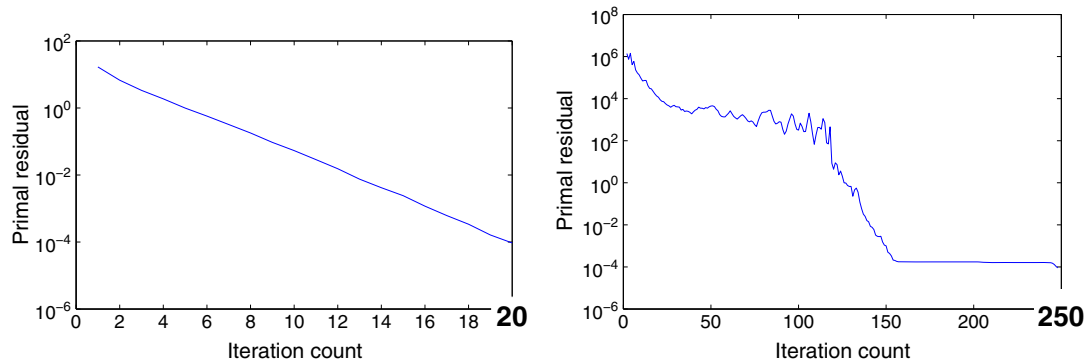


Figure 2. Checkerboard coefficient distribution – convergence curve: primal residual versus iteration count – Left: with GenEO, Right : without GenEO – Lumped preconditioner for the Metis decomposition into 64 subdomains –  $(E_1; v_1) = (10^7; 0.4)$  and  $(E_2; v_2) = (10^{12}; 0.3)$ .

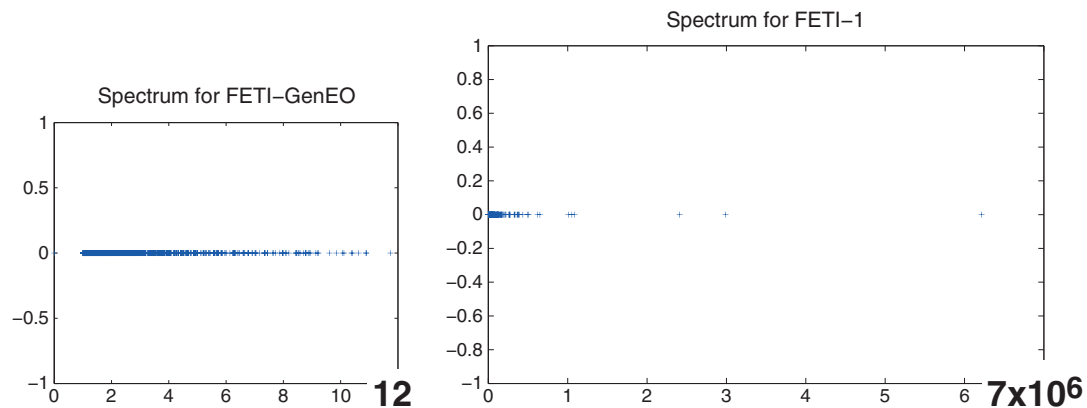


Figure 3. Checkerboard coefficient distribution – spectrum of the preconditioned operator – Left: with GenEO, Right : without GenEO – Lumped preconditioner for the Metis decomposition into 64 subdomains –  $(E_1; v_1) = (10^7; 0.4)$  and  $(E_2; v_2) = (10^{12}; 0.3)$ .

coarse modes (either natural or GenEO) as well as the null space of  $B^\top$ . Whether the GenEO coarse space is used or not, the first nonzero eigenvalue of the preconditioned operator is 1, which is what is expected.

#### 4.2. Discontinuities along the interfaces

In this subsection, we focus only on the GenEO coarse space for the Dirichlet preconditioner, and we conduct a more extensive study. We use a partition into  $N$  regular subdomains of a rectangle of size  $N \times b$  where  $b$  is the aspect ratio of each subdomain (see Figure 4). The discretization of each subdomain is  $n_{el} \times n_{el}$  rectangular elements so that each element has the same aspect ratio as the subdomain to which it belongs. The coefficient distribution consists of a constant value  $v = 0.3$  of Poisson's ratio and seven layers of  $E$  (four soft layers, three hard layers, see again Figure 4). Throughout this subsection, we use again the  $K$ -scaling matrices, (4.1) which is in fact, for this case, equivalent to choosing multiplicity scaling because the coefficient jumps are only along the interfaces.

The parameters are:  $b = 1$  (aspect ratio),  $n_{el} = 21$  (number of elements per direction per subdomain) and  $E_1/E_2 = 10^{-5}$  (jump in the coefficient). The spectrum is plotted in Figure 5 and the corresponding first 11 generalized eigenvectors and corresponding eigenvalues are shown in Figure 6. We observe that there is a gap in the spectrum of the generalized eigenproblem after the



Figure 4. Discontinuities along the interfaces.

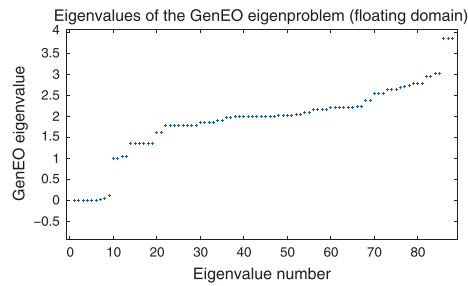


Figure 5. Eigenvalues of the GenEO generalized eigenproblem for the geometry given in Figure 4.

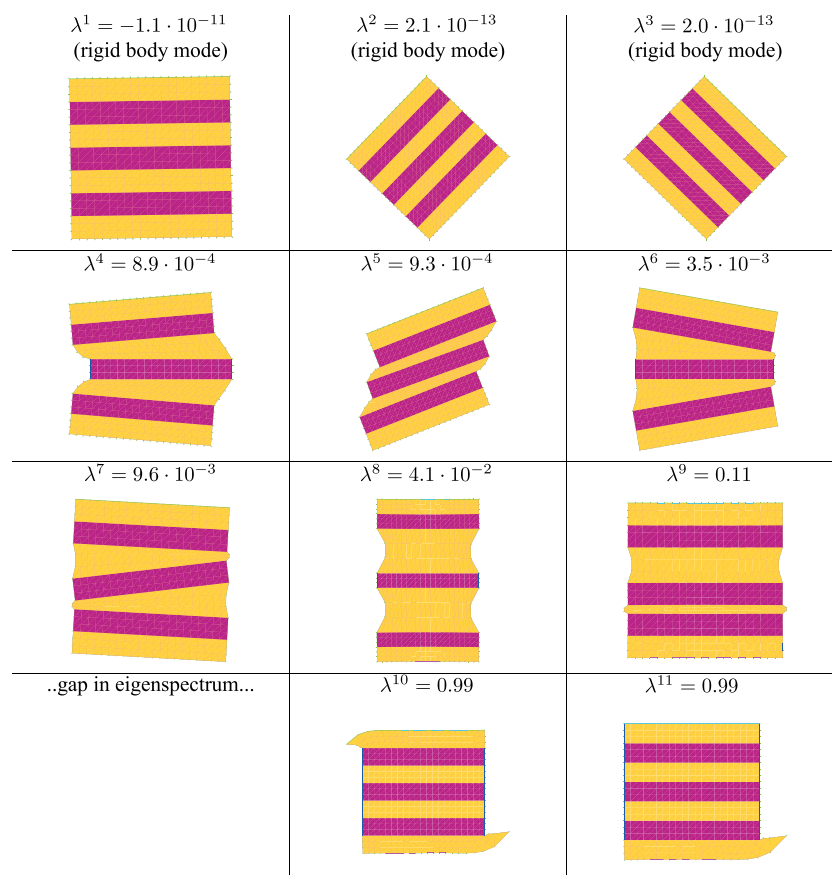


Figure 6. Eigenvalues and eigenmodes of the GenEO generalized eigenproblem for the geometry given in Figure 4 – dark or pink: hard material, light or yellow: soft material – The first eigenmodes (rigid body modes) are part of the natural coarse grid, and the next six are selected for the GenEO coarse space.

ninth generalized eigenvalue because  $\lambda^9 = 0.11$  and  $\lambda^{10} = 0.98$ . For this reason, a judicious choice of the threshold for selecting eigenvectors that are put into the coarse space is for instance

$$\mathcal{K}_i = 0.15,$$

we will use this in all following numerical tests. With this criteria, the GenEO eigenproblem for a floating subdomain will provide nine modes: the first three are rigid body modes included in the usual FETI natural coarse space and six deformation modes that are included in the GenEO coarse grid. As can be seen in Figure 5, those deformation modes represent the behavior of the subdomain when the hard layers deform the soft ones. The nine modes can be seen as a basis to describe the nearly rigid motion of the hard layers (three modes for each of the three layers, amounting to nine modes) and the basis spanned by those modes represent the behavior of the domain as if the hard layers were its backbones. In some sense, the GenEO coarse space can be interpreted in this case as a *skeleton* of the overall problem describing the dominant behavior of the structure according to its hard layers.

Next, we actually solve the problem for different numbers of subdomains, different aspect ratios, and different discretizations. The results are shown in Table IV. The two-level method with the GenEO coarse space is robust throughout all of these tests: the condition number varies between 1.34

Table IV. Three tests for the geometry in Figure 4 –  $\kappa$ : condition number;  $\#U_0$ : size of the GenEO coarse space;  $it$ : number of iterations.

Various number of subdomains ( $N$ ), fixed aspect ratio ( $b = 1$ ), fixed discretization ( $n_{el} = 21$ ), fixed jump in coefficients ( $E_1/E_2 = 10^{-5}$ ), the problem size increases with $N$					
$N$ subdomains	FETI-GenEO			FETI-1	
	$\kappa$	$\#U_0$	$it$	$\kappa$	$it$
4	3	14	5	$1.4 \cdot 10^3$	20
8	1.34	38	5	$1.9 \cdot 10^3$	39
16	1.34	86	4	$2.1 \cdot 10^3$	75
32	1.35	182	4	$2.2 \cdot 10^3$	137
64	1.35	374	4	$2.2 \cdot 10^3$	190
Various aspect ratios ( $b$ ), fixed number of subdomains ( $N = 8$ ), fixed discretization ( $n_{el} = 21$ ), fixed jump in coefficients ( $E_1/E_2 = 10^{-5}$ )					
Aspect ratio $b$	FETI-GenEO			FETI-1	
	$\kappa$	$\#U_0$	$it$	$\kappa$	$it$
5	2.33	43	6	$1.7 \cdot 10^5$	47(*)
2	1.42	40	5	$1.0 \cdot 10^4$	43
1	1.34	38	5	$1.9 \cdot 10^3$	40
1/2	4.51	27	9	446	33
1/5	4.07	14	11	70	22
Various discretizations ( $n_{el}$ ), fixed aspect ratios ( $b = 1$ ), fixed number of subdomains ( $N = 8$ ), fixed jump in coefficients ( $E_1/E_2 = 10^{-5}$ ), the problem size increases with $n_{el}$					
$n_{el}$ elements	FETI-GenEO			FETI-1	
	$\kappa$	$\#U_0$	$it$	$\kappa$	$it$
21	1.34	38	5	$1.92 \cdot 10^3$	39
42	1.42	38	5	$1.93 \cdot 10^3$	40
70	1.46	38	5	$1.94 \cdot 10^3$	40
84	1.47	38	5	$1.94 \cdot 10^3$	40

(\*) The relative residual reaches a plateau at  $2 \cdot 10^{-3}$  after 47 iterations.

and 4.51 only, which is indeed lower than the upper bound given by the theory,  $\mathcal{N}/\mathcal{K}_i = 20$ ,  $\mathcal{N}$  being equal to three in this simple decomposition. Further, the following observations are noteworthy:

- When the number of domains increases, the classical FETI-1 method sees its number of iteration increase significantly, whereas equipped with the GenEO coarse grid, the number of iteration remains small. The dimension of the GenEO coarse grids is roughly proportional to the number of domains in this case.
- The classical FETI method converges very slowly when the height of the domain is large compared with its width ( $b = 5$ ). For that case, the GenEO strategy generates only a small number of modes (43 in total) and converges very fast.
- For this layered structure, the preconditioned interface problem of FETI-1 has a condition number that barely depends on the number of elements per domain, and the number of iterations is nearly invariant with respect to the discretization step. When equipped with the GenEO coarse grid, a small number of modes is included in the coarse space (38 GenEO modes, independent of the discretization step), and the number of iteration is very small.

It is thus remarkable that the GenEO coarse space can handle automatically (once a proper threshold  $\mathcal{K}$  has been chosen) the difficult cases of bad aspect ratios and heterogeneities along the interface.

#### 4.3. Discontinuities along and across interfaces

In this subsection, we consider the case of Figure 7 where the only difference with the previous subsection is that we have added jumps across the interfaces in subdomains 3 and 6 by inverting the soft and the hard layers. The parameters are as follows:  $n_{el} = 21$  elements in each direction and each subdomain,  $N = 8$  subdomains,  $\nu = 0.3$  for Poisson's ratio,  $E_1/E_2 = 10^{-5}$  for the magnitude of the jump in the coefficient,  $b = 1$  for the aspect ratio of the subdomains, and  $\mathcal{K}_i = 0.15$  for the threshold on the GenEO eigenvalues. This is a known hard problem for FETI even with the Dirichlet preconditioner (which we use here again). In this case, we show in Table V that with the  $K$ -scaling matrices (4.1), the number of *bad* eigenmodes is largely reduced compared with the case where multiplicity scaling is used (here, multiplicity scaling reduces to setting all entries of each  $D_i$  to  $1/2$ ). Indeed with  $K$ -scaling, we have selected 46 modes, which is only eight more than for the same case but without the extra jumps across the interfaces (see Table IV – top –  $N = 8$  subdomains). With the multiplicity scaling, the GenEO strategy selects 173 modes. In fact, with



Figure 7. Discontinuities across and along interfaces (subdomains 3 and 6).

Table V. Geometry given in Figure 7 (discontinuities across and along the interfaces),  $n_{el} = 21$ ,  $N = 8$ ,  $E_1/E_2 = 10^{-5} - \kappa$ : condition number;  $\#U_0$ : size of the GenEO coarse space;  $it$ : number of iterations.

Scaling ( $D_i$ )	FETI-GenEO			FETI-1	
	$\kappa$	$it$	$\#U_0$	$\kappa$	$it$
$K$ -scaling	3.71	9	46	$7.0 \cdot 10^4$	55
Multiplicity	3.89	7	173	$4.5 \cdot 10^4$	189(*)

(\*) The relative residual reaches a plateau at  $1.5 \cdot 10^{-3}$  after 189 iterations.



$K$ -scaling, fewer modes are necessary because jumps across the interfaces are already accounted for in the preconditioner. The additional modes are needed to take into account the jumps across the interfaces. This confirms that GenEO compensates for the discrepancy between the preconditioner and the actual inverse of  $F$ ; when inadequate weighting is used, the preconditioner is less effective and hence a larger coarse space is needed. The condition numbers for both types of scaling are almost equal when the GenEO coarse grid is introduced, which confirms the theory.

## 5. CONCLUSION

We have constructed a two-level BDD method and two two-level FETI methods for which the convergence rates depend only on a chosen parameter and the maximal number of neighbors of a subdomain. The choice of this parameter is key in dimensioning the coarse space. Optimizing the choice of the parameter with respect to the efficiency and the size of the coarse space is crucial. Here, it has been set heuristically. For FETI, the result holds for the full preconditioner based on solving Dirichlet problems in the subdomains and also on the Lumped version, which is a lot less expensive to implement. Compared with the Schwarz–GenEO algorithm, these methods have the advantage of being nonoverlapping methods, which means that they do not carry the extra cost of computations in the overlap.

In this paper, the fundamental ideas and proofs underlying the GenEO coarse grid have been explained, and the numerical efficiency has been illustrated on problems hard to solve with classical FETI approaches. Future research will investigate the computational cost incurred by the GenEO coarse grid (computation of the GenEO modes per domain, building, and solving the coarse grid) to assess the overall computational efficiency of the FETI-GenEO when applied to realistic engineering problems.

## ACKNOWLEDGEMENTS

We are very grateful to Victorita Dolean, Patrice Hauret, and Frédéric Nataf for the many fruitful discussions and constructive comments.

## REFERENCES

1. Karypis G, Kumar V. A Fast and Highly Quality Multilevel Scheme for Partitioning Irregular Graphs. *SIAM Journal on Scientific Computing* 1999; **20**(1):359–392.
2. SCOTCH: static mapping, graph partitioning, and sparse matrix block ordering package. Available from: <http://www.labri.fr/~pelegri/scotch/>.
3. Dryja M, Sarkis MV, Widlund OB. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numerische Mathematik* 1996; **72**(3):313–348. DOI: 10.1007/s002110050172. Available from: <http://dx.doi.org/10.1007/s002110050172>.
4. Mandel J, Brezina M. Balancing domain decomposition for problems with large jumps in coefficients. *Mathematics of Computation* 1996; **65**(216):1387–1401. DOI: 10.1090/S0025-5718-96-00757-0. Available from: <http://dx.doi.org/10.1090/S0025-5718-96-00757-0>.
5. Dohrmann CR, Widlund OB. An overlapping Schwarz algorithm for almost incompressible elasticity. *SIAM Journal of Numerical Analysis* 2009; **47**(4):2897–2923. DOI: 10.1137/080724320. Available from: <http://dx.doi.org/10.1137/080724320>.
6. Dohrmann CR, Widlund OB. Hybrid domain decomposition algorithms for compressible and almost incompressible elasticity. *International Journal for Numerical Methods in Engineering* 2010; **82**(2):157–183.
7. Pechstein C, Scheichl R. Scaling up through domain decomposition. *Applied Analysis* 2009; **88**(10–11):1589–1608. DOI: 10.1080/00036810903157204. Available from: <http://dx.doi.org/10.1080/00036810903157204>.
8. Pechstein C, Scheichl R. Analysis of FETI methods for multiscale PDEs. *Numerische Mathematik* 2008; **111**(2):293–333. DOI: 10.1007/s00211-008-0186-2. Available from: <http://dx.doi.org/10.1007/s00211-008-0186-2>.
9. Bhardwaj M, Day D, Farhat C, Lesoinne M, Pierson K, Rixen D. Application of the FETI method to ASCI problems—scalability results on 1000 processors and discussion of highly heterogeneous problems. *International Journal for Numerical Methods in Engineering* 2000; **47**(1–3):513–536.
10. Klawonn A, Rheinbach O. Robust FETI-DP methods for heterogeneous three dimensional elasticity problems. *Journal of Computer Methods in Applied Mechanics and Engineering* 2007; **196**(8):1400–1414.
11. Pechstein C, Scheichl R. Analysis of FETI methods for multiscale PDEs. Part II: interface variation. *Numerische Mathematik* 2011; **118**(3):485–529.



12. Pechstein C, Sarkis M, Scheichl R. New theoretical coefficient robustness results for FETI-DP. In *Domain Decomposition Methods in Science and Engineering XX*, Lecture Notes in Computational Science and Engineering. Springer-Verlag: Berlin Heidelberg, 2011.
13. Pechstein C. *Finite and Boundary Element Tearing and Interconnecting Solvers for Multiscale Problems*. Springer: Berlin Heidelberg, 2012.
14. Pechstein C, Scheichl R. Weighted Poincaré inequalities. *NuMa-Report 2010-10*, Institute of Computational Mathematics, Johannes Kepler University, Linz, 2010. submitted.
15. Klawonn A, Rheinbach O, Widlund OB. An analysis of a FETI-DP algorithm on irregular subdomains in the plane. *SIAM Journal of Numerical Analysis* 2008; **46**(5):2484–2504.
16. Farhat C, Maman N, Brown G. Mesh partitioning for implicit computations via iterative domain decomposition: impact and optimization of the subdomain aspect ratio. *International Journal for Numerical Methods in Engineering* 1995; **38**:989–1000.
17. Spillane N, Dolean V, Hauret P, Nataf F, Pechstein C, Scheichl R. A robust two-level domain decomposition preconditioner for systems of PDEs. *Comptes Rendus Mathématique* 2011; **349**(23-24):1255–1259.
18. Spillane N, Dolean V, Hauret P, Nataf F, Pechstein C, Scheichl R. Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps. *NuMa-Report 2011-07*, Institute of Computational Mathematics, Johannes Kepler University Linz, 2011. submitted, Accepted by Numerische Mathematik, 2013.
19. Efendiev Y, Galvis J, Lazarov R, Willems J. Robust domain decomposition preconditioners for abstract symmetric positive definite bilinear forms. *ESAIM: Mathematical Modelling and Numerical Analysis* 2012; **46**(05):1175–1199.
20. Toselli A, Widlund OB. *Domain Decomposition Methods—Algorithms and Theory*, Springer Series in Computational Mathematics, Vol. 34. Springer-Verlag: Berlin, 2005.
21. Brezina M, Heberton C, Mandel J, Vaněk P. An iterative method with convergence rate chosen a priori. *Technical Report 140*, University of Colorado Denver, CCM, University of Colorado Denver, 1999. Earlier version presented at 1998 Copper Mountain Conference on Iterative Methods, April 1998.
22. Chartier T, Falgout RD, Henson VE, Jones J, Manteuffel T, McCormick S, Ruge J, Vassilevski PS. Spectral AMGe ( $\rho$ AMGe). *SIAM Journal of Scientific Computing* 2003; **25**(1):1–26. DOI: 10.1137/S106482750139892X. Available from: <http://dx.doi.org/10.1137/S106482750139892X>.
23. Galvis J, Efendiev Y. Domain decomposition preconditioners for multiscale flows in high-contrast media. *Multiscale Modelling & Simulation* 2010; **8**(4):1461–1483. DOI: 10.1137/090751190. Available from: <http://dx.doi.org/10.1137/090751190>.
24. Galvis J, Efendiev Y. Domain decomposition preconditioners for multiscale flows in high contrast media: reduced dimension coarse spaces. *Multiscale Modeling & Simulation* 2010; **8**(5):1621–1644. DOI: 10.1137/100790112.
25. Nataf F, Xiang H, Dolean V. A two level domain decomposition preconditioner based on local Dirichlet-to-Neumann maps. *C. R. Mathématique* 2010; **348**(21-22):1163–1167.
26. Dolean V, Nataf F, Spillane N., Xiang H. A coarse space construction based on local Dirichlet-to-Neumann maps. *SIAM Journal on Scientific Computing* 2011; **33**:1623–1642.
27. Dolean V, Nataf F, Scheichl R, Spillane N. Analysis of a two-level Schwarz method with coarse spaces based on local Dirichlet-to-Neumann maps. *Computational Methods in Applied Mathematics* 2012; **12**(4):391–414.
28. Efendiev Y, Galvis J, Vassilevski PS. Spectral element agglomerate algebraic multigrid methods for elliptic problems with high-contrast coefficients. In *Domain Decomposition Methods in Science and Engineering XIX*, Vol. 78, Huang Y, Kornhuber R, Widlund O, Xu J (eds), Lecture Notes in Computational Science and Engineering. Springer: Berlin, 2011; 407–414.
29. Efendiev Y, Galvis J, Vassilevski PS. Multiscale spectral AMGe solvers for high-contrast flow problems. *ISC-Preprint 2012-02*, Inst. Scientific Computation, Texas A&M University, 2012. submitted.
30. Willems J. Robust multilevel methods for general symmetric positive definite operators. *RICAM-Report 2012-06*, Johann Radon Institute for Computational and Applied Mathematics, Linz, 2012. submitted.
31. Napov A, Notay Y. Algebraic analysis of aggregation-based multigrid. *Numerical Linear Algebra with Applications* 2011; **18**(3):539–564.
32. Napov A, Notay Y. An algebraic multigrid method with guaranteed convergence rate. *SIAM Journal on Scientific Computing* 2012; **34**(2):1079–1109.
33. Mandel J, Sousedík B. Adaptive selection of face coarse degrees of freedom in the BDDC and the FETI-DP iterative substructuring methods. *Computer Methods in Applied Mechanics and Engineering* 2007; **196**(8):1389–1399.
34. Šístek J, Sousedík B, Mandel J. Adaptive-multilevel BDDC and its parallel implementation. *Computing* 2013. DOI: 10.1137/080724320. Submitted.
35. Matsokin AM, Nepomnyaschikh SV. A Schwarz alternating method in a subspace. *Soviet Mathematics* 1985; **29**(10):78–84.
36. Nicolaides RA. Deflation of conjugate gradients with applications to boundary value problems. *SIAM Journal of Numerical Analysis* 1987; **24**(2):355–365. DOI: 10.1137/0724027. Available from: <http://dx.doi.org/10.1137/0724027>.
37. Bramble JH, Pasciak JE, Schatz AH. The construction of preconditioners for elliptic problems by substructuring. I. *Mathematics of Computation* 1986; **47**(175):103–134. DOI: 10.2307/2008084. Available from: <http://dx.doi.org/10.2307/2008084>.
38. Dryja M, Widlund OB. Some domain decomposition algorithms for elliptic problems. In *Iterative methods for large linear systems*, Hayes L, Kincaid D (eds). Academic Press: Waltham, MA, 1989; 273–291.

39. Mandel J. Balancing domain decomposition. *Communications of Numerical Methods Engineering* 1993; **9**(3):233–241. DOI: 10.1002/cnm.1640090307. Available from: <http://dx.doi.org/10.1002/cnm.1640090307>.
40. De Roeck Y, Le Tallec P. Analysis and test of a local domain decomposition preconditioner. In *Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Vol. 4, Soc for Industrial & Applied Math: Philadelphia, PA, 1991; 112.
41. Farhat C, Roux FX. A method of finite element tearing and interconnecting and its parallel solution algorithm. *International Journal for Numerical Methods in Engineering* 1991; **32**:1205–1227.
42. Mandel J, Tezaur R. Convergence of a substructuring method with Lagrange multipliers. *Numerische Mathematik* 1996; **73**(4):473–487.
43. Tezaur R. Analysis of Lagrange multiplier based domain decomposition. *PhD Thesis*, University of Colorado at Denver, 1998.
44. Klawonn A, Widlund OB. FETI and Neumann-Neumann iterative substructuring methods: connections and new results. *Communications on Pure and Applied Mathematics* 2001; **54**(1):57–90.
45. Farhat C, Chen PS, Mandel J. A scalable Lagrange multiplier based domain decomposition method for time-dependent problems. *International Journal for Numerical Methods in Engineering* 1995; **38**(22):3831–3853.
46. Farhat C, Mandel J. The two-level FETI method for static and dynamic plate problems - part I: an optimal iterative solver for biharmonic systems. *Journal of Computer Methods in Applied Mechanics and Engineering* 1998; **155**:129–152.
47. Farhat C, Chen PS, Roux FX. The two-level FETI method - part II: extension to shell problems, parallel implementation and performance results. *Journal of Computer Methods in Applied Mechanics and Engineering* 1998; **155**:153–180.
48. Toselli A, Klawonn A. A FETI domain decomposition method for edge element approximations in two dimensions with discontinuous coefficients. *SIAM Journal on Numerical Analysis* 2002; **39**(3):932–956.
49. Rixen D, Farhat C. A simple and efficient extension of a class of substructure based preconditioners to heterogeneous structural mechanics problems. *International Journal for Numerical Methods in Engineering* 1999; **44**(4):489–516.
50. Tang J, Nabben R, Vuik C, Erlangga Y. Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods. *Journal of Scientific Computing* 2009; **39**(3):340–370.
51. Leborgne G. Valeurs et vecteurs propres: Lecture notes for ISIMA: Clermont Ferrand 2008. <http://www.isima.fr/~leborgne/Isimathdiagonalisation/Diagonalisation.pdf>.
52. Rixen D. Dual schur complement method for semi-definite problems. *Contemporary Mathematics* 1998; **218**:341–348. Tenth International Conference on Domain Decomposition Methods, Boulder, CO, August 1997.
53. Hecht F. *Freefem++*, 3.7, Numerical Mathematics and Scientific Computation. Laboratoire J.L. Lions, Université Pierre et Marie Curie, 2010. Available from: <http://www.freefem.org/ff++/>.
54. Rixen D. Extended preconditioners for FETI method applied to constrained problems. *International Journal of Numerical Methods in Engineering* 2002; **54**(1):1–26.
55. Mandel J, Dohrmann C, Tezaur R. An algebraic theory for primal and dual substructuring methods by constraints. *Applied Numerical Mathematics* 2005; **54**(2):167–193.