

## Chapter 2

# One-Level FETI/BETI Methods

This chapter deals with tearing and interconnecting methods based on the finite element method (FEM) and the boundary element method (BEM). Here we allow a mixture of FEM and BEM within a single discretization of a PDE: the computational domain is partitioned (“torn”) into several non-overlapping subdomains, and on each individual subdomain one may choose FEM or BEM as the local discretization. The coupling (“interconnecting”) of these local discretizations is maintained by Lagrange multipliers. Additionally, the tearing and interconnecting framework is used to construct fast solvers for the resulting global system of equations.

The term “one-level” in the title of this chapter refers to a special treatment of the so-called *floating* subdomains that do not touch the Dirichlet boundary. An alternative treatment is used in the dual-primal methods, see Chap. 5.

There are two subclasses of one-level methods: the *classical* formulation, and the *total* or *all-floating* formulation. The *classical* finite element tearing and interconnecting (FETI) method was proposed by Farhat and Roux [FR91, FR92, FR94] as a solver for large-scale finite element systems. Note that some basic ideas can already be found in an earlier work by Glowinsky and Wheeler [GW88] on certain mixed methods. The FETI method was enhanced with the so-called Dirichlet preconditioner by Farhat, Mandel, and Roux [FMR94]. The latter method was first analyzed by Mandel and Tezaur [MT96], who showed that the condition number grows at most as  $C (1 + \log(H/h))^3$  where  $H$  is the subdomain size and  $h$  the element size. See also [Tez98, Bre02, Bre03a] for further analyses. Klawonn and Widlund [KW01] proposed new preconditioners including an earlier algorithm by Rixen and Farhat [RF98a, RF99] and including ideas from balancing Neumann-Neumann methods (see Sect. 2.3). They also generalized the theory in several different respects (e.g., the case of three-dimensions and so-called fully redundant Lagrange multipliers) and could show the improved bound  $C (1 + \log(H/h))^2$ . The boundary element tearing and interconnecting (BETI) and the coupled FETI/BETI methods were proposed and analyzed by Langer and Steinbach [LS03, LS05].

The *total* FETI method and the *all-floating* BETI method were introduced independently by Dostál, Horák, and Kučera [DHK06] and Of [Of06, Of08],

respectively (see also [OS09]), and the corresponding preconditioner was analyzed in [Pec08b]. Note that these methods have been successfully generalized to mechanical contact problems, see e.g. [BDS08, DKV<sup>+</sup>10].

The remainder of this chapter is organized as follows. In Sect. 2.1 we work out a (discrete) skeleton formulation using the (approximate) Steklov-Poincaré operators from Chap. 1. Section 2.2 describes the classical FETI/BETI and the all-floating (total) FETI/BETI method in detail and discusses implementation issues. Section 2.3 briefly introduces the related balancing Neumann-Neumann method. The analysis of one-level FETI/BETI is performed in two steps. In Sect. 2.4 we analyze the unpreconditioned method (which turns out to be sub-optimal) and we begin to analyze the so-called *scaled Dirichlet preconditioner* on an abstract level. After providing a set of technical tools (Sect. 2.5) we will conclude the analysis in Sect. 2.6. Finally, we provide some numerical results in Sect. 2.7, and we discuss generalizations to other equations and/or discretization spaces in Sect. 2.8.

## 2.1 Skeleton Formulations

FETI and BETI are iterative substructuring methods based on a non-overlapping decomposition of the computational domain  $\Omega$ . A good starting point for these methods, especially for BETI, is a discrete skeleton formulation, which is derived from a continuous one.

### 2.1.1 Continuous Skeleton Formulation

Let  $\Omega \subset \mathbb{R}^d$  ( $d = 2$  or  $3$ ) be a bounded Lipschitz domain whose boundary  $\partial\Omega$  consists of a Dirichlet boundary  $\Gamma_D = \overline{\Gamma_D}$  with positive surface measure and a Neumann boundary  $\Gamma_N = \partial\Omega \setminus \Gamma_D$ . The outward unit normal vector to  $\partial\Omega$  is denoted by  $n$ . We consider the weak form of the potential equation (1.21): find  $u \in H^1(\Omega)$ ,  $u|_{\Gamma_D} = g_D$  such that

$$\underbrace{\int_{\Omega} \mathcal{A} \nabla u \cdot \nabla v \, dx}_{=: a(u, v)} = \underbrace{\int_{\Omega} f_{\Omega} v \, dx + \int_{\Gamma_N} g_N v \, ds}_{=: (f, v)} \quad \forall v \in H_D^1(\Omega), \quad (2.1)$$

where  $H_D^1(\Omega) := \{v \in H^1(\Omega) : v|_{\Gamma_D} = 0\}$  and

$$f_{\Omega} \in L^2(\Omega), \quad g_N \in L^2(\Gamma_N), \quad g_D \in H^{1/2}(\Gamma_D) \quad (2.2)$$

are given. We assume that the coefficient  $\mathcal{A}$  fulfills Condition (1.22) from Lemma 1.39 (p. 20), such that we have unique solvability.

**Table 2.1** Geometric sets and quantities associated to the subdomain decomposition

$H_i := \text{diam}(\Omega_i)$	Subdomain diameter
$\Gamma_{ij} := (\partial\Omega_i \cap \partial\Omega_j) \setminus \Gamma_D$	Subdomain interface
$\Gamma := \bigcup_{i \neq j} \Gamma_{ij}$	Interface
$\Gamma_S := \bigcup_i \partial\Omega_i$	Skeleton

Let  $\{\Omega_i\}_{i=1}^s$  be a non-overlapping partition of  $\Omega$  into open Lipschitz domains  $\Omega_i$  (called *subdomains* or *substructures*) such that

$$\overline{\Omega} = \bigcup_{i=1}^s \overline{\Omega}_i, \quad \Omega_i \cap \Omega_j = \emptyset \quad \text{for } i \neq j. \quad (2.3)$$

Furthermore, we introduce a couple of geometric quantities summarized in Table 2.1 (recall that  $\Gamma_D$  is closed!).

Thanks to the assumptions on  $f_\Omega$  and  $g_N$ , we have the splitting property

$$a(u, v) = \sum_{i=1}^s \underbrace{\int_{\Omega_i} \mathcal{A} \nabla u \cdot \nabla v \, dx}_{=: a_i(u|_{\Omega_i}, v|_{\Omega_i})}, \quad \langle f, v \rangle = \sum_{i=1}^s \underbrace{\left( \int_{\Omega_i} f_\Omega v \, dx + \int_{\partial\Omega_i \cap \Gamma_N} g_N v \, ds \right)}_{=: \langle f_i, v|_{\Omega_i} \rangle} \quad (2.4)$$

with  $a_i : H^1(\Omega_i) \times H^1(\Omega_i) \rightarrow \mathbb{R}$  and  $f_i \in H^1(\Omega_i)^*$ .

*Remark 2.1.* We can also allow for a general functional  $f \in H^1(\Omega)^*$  (not necessarily of the form (2.1)), provided that we have a splitting into subdomain functionals  $f_i \in H^1(\Omega_i)^*$  as in (2.4).

For each subdomain  $\Omega_i$ , let  $S_i : H^{1/2}(\partial\Omega_i) \rightarrow H^{-1/2}(\partial\Omega_i)$  denote the Steklov-Poincaré operator corresponding to the bilinear form  $a_i(\cdot, \cdot)$  and  $N_i : H^1(\Omega_i)^* \rightarrow H^{-1/2}(\partial\Omega_i)$  the corresponding Newton potential, see Definition 1.41. Furthermore, we define the skeletal spaces

$$H^{1/2}(\Gamma_S) := \{v \in L^2(\Gamma_S) : \exists \tilde{v} \in H^1(\Omega) : v = \tilde{v}|_{\Gamma_S}\}, \quad (2.5)$$

$$H_D^{1/2}(\Gamma_S) := \{v \in H^{1/2}(\Gamma_S) : v|_{\Gamma_D} = 0\}. \quad (2.6)$$

**Lemma 2.2.** *The variational formulation (2.1) is equivalent to finding  $u \in H^1(\Omega)$  with  $u|_{\Gamma_D} = g_D$  such that*

$$\begin{aligned} \sum_{i=1}^s \langle S_i u|_{\partial\Omega_i}, v|_{\partial\Omega_i} \rangle &= \sum_{i=1}^s \langle N_i f_i, v|_{\partial\Omega_i} \rangle & \forall v \in H_D^{1/2}(\Gamma_S), \\ a_i(u|_{\Omega_i}, v_0) &= \langle f_i, v_0 \rangle_{\Omega_i} & \forall v_0 \in H_0^1(\Omega_i) \quad \forall i = 1, \dots, s. \end{aligned}$$

*Proof.* The equations in the second line follow immediately from (2.1). With these local equations fulfilled, Lemma 1.42(ii) implies that

$$a_i(u_{|\Omega_i}, v_{|\Omega_i}) - \langle f_i, v_{|\Omega_i} \rangle = \langle S_i u_{|\partial\Omega_i}, v_{|\partial\Omega_i} \rangle - \langle N_i f_i, v_{|\partial\Omega_i} \rangle \quad \forall v \in H_D^1(\Omega).$$

This proves the equivalence.  $\square$

The system in Lemma 2.2 can be seen as an algorithm:

1. Determine  $u_S \in H^{1/2}(\Gamma_S)$  with  $u_S|_{\Gamma_D} = g_D$  such that

$$\sum_{i=1}^s \langle S_i u_S|_{\partial\Omega_i}, v_{|\partial\Omega_i} \rangle = \sum_{i=1}^s \langle N_i f_i, v_{|\partial\Omega_i} \rangle \quad \forall v \in H_D^{1/2}(\Gamma_S). \quad (2.7)$$

2. Determine the local functions  $u_i = u_{|\Omega_i}$  as the solution of the variational problems, find  $u_i \in H^1(\Omega_i)$  with  $u_i|_{\partial\Omega_i} = u_S|_{\partial\Omega_i}$  such that

$$a_i(u_i, v_0) = \langle f_i, v_0 \rangle_{\Omega_i} \quad \forall v_0 \in H_0^1(\Omega_i). \quad (2.8)$$

Problem (2.7) is called *skeletal variational formulation*. Under the assumptions of Lemma 1.39, it is straightforward to show (with Theorem 1.1) that Problem (2.7) is well-posed. Should we only be interested in the trace of the solution  $u$  on  $\Gamma_S$ , we can stop after step 1.

Sometimes, the following homogeneous version of (2.7) is convenient. Find  $\tilde{u} \in H_D^{1/2}(\Gamma_S)$  such that

$$\sum_{i=1}^s \langle S_i \tilde{u}|_{\partial\Omega_i}, v_{|\partial\Omega_i} \rangle = \sum_{i=1}^s \langle N_i f_i - S_i \tilde{g}_D|_{\partial\Omega_i}, v_{|\partial\Omega_i} \rangle \quad \forall v \in H_D^{1/2}(\Gamma_S), \quad (2.9)$$

where  $\tilde{g}_D \in H^{1/2}(\Gamma_S)$  fulfills  $\tilde{g}_D|_{\Gamma_D} = g_D$  (see Lemma 1.21). Then  $u = \tilde{g}_D + \tilde{u}$  solves Problem (2.7).

### 2.1.2 Discrete Skeleton Formulations

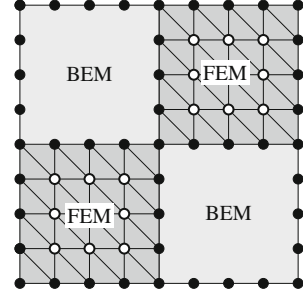
Discrete skeleton formulations are obtained by applying a Galerkin method to (2.7) or (2.9).

**Step 1.** We project the equation to a finite-dimensional space. To this end, we consider a shape regular triangulation  $\mathcal{T}^h(\Gamma_S)$  of the skeleton  $\Gamma_S$  into line segments (if  $d=2$ ) or triangles (if  $d=3$ ). We require that the subdomain boundaries  $\partial\Omega_i$  are unions of elements from  $\mathcal{T}^h(\Gamma_S)$ .

The  $H^{1/2}(\Gamma_S)$ -conforming skeletal finite element space is defined as

$$V^h(\Gamma_S) := \{v \in \mathcal{C}(\Gamma_S) : v|_{\tau} \in \mathcal{P}_1 \quad \forall \tau \in \mathcal{T}^h(\Gamma_S)\}. \quad (2.10)$$

**Fig. 2.1** Example of a skeleton mesh for a decomposition into four subdomains, extended to local meshes in two of the subdomains (● skeleton nodes, ○ interior FE nodes)



For simplicity, we assume that the Dirichlet data  $g_D$  is piecewise linear too.<sup>1</sup> Hence, there is a unique function  $\tilde{g}_D \in V^h(\Gamma_S)$  with  $\tilde{g}_D|_{\Gamma_D} = g_D$  which vanishes on all nodes except those on  $\Gamma_D$ .

**Step 2.** We use an *approximate* bilinear form and right hand side. Let  $\mathcal{I}_{\text{BEM}} \subset \{1, \dots, s\}$  be an index set and assume that

$$\mathcal{A}_{|\Omega_i} = \alpha_i I, \quad \forall i \in \mathcal{I}_{\text{BEM}},$$

with constants  $\alpha_i > 0$ . In each subdomain, the restriction of  $\mathcal{T}^h(\Gamma_S)$  to  $\partial\Omega_i$  is a triangulation of  $\partial\Omega_i$ , simply denoted by  $\mathcal{T}^h(\partial\Omega_i)$ . For  $i \notin \mathcal{I}_{\text{BEM}}$ , we extend  $\mathcal{T}^h(\partial\Omega_i)$  to a shape regular triangulation  $\mathcal{T}^h(\Omega_i)$  of  $\Omega_i$ . For an example see Fig. 2.1. We now replace the local Steklov-Poincaré operators  $S_i$  and Newton potentials  $N_i$  by the FEM- and BEM-approximations from Sects. 1.2.6 and 1.3.8.1. With  $\mathcal{I}_{\text{FEM}} := \{1, \dots, s\} \setminus \mathcal{I}_{\text{BEM}}$ , we set

$$S_{i,h} := \begin{cases} S_{i,\text{BEM}} & \text{if } i \in \mathcal{I}_{\text{BEM}} \\ S_{i,\text{FEM}} & \text{if } i \in \mathcal{I}_{\text{FEM}} \end{cases}, \quad N_{i,h} := \begin{cases} N_{i,\text{BEM}} & \text{if } i \in \mathcal{I}_{\text{BEM}} \\ N_{i,\text{FEM}} & \text{if } i \in \mathcal{I}_{\text{FEM}} \end{cases}.$$

To ensure the invertibility of the discretized single layer potential operators occurring in  $S_{i,\text{BEM}}$ , we assume that in two dimensions,  $\text{diam}(\Omega_i) < 1$  for all  $i \in \mathcal{I}_{\text{BEM}}$ , cf. Assumption 1.72, p. 46. A sufficient and practicable condition for this is of course  $\text{diam}(\Omega) < 1$ , which can be achieved by a simple scaling of the coordinates.

The resulting *discrete skeleton variational problem* reads: find  $u_h \in V^h(\Gamma_S)$  with  $u_h|_{\Gamma_D} = g_D$  such that

$$\sum_{i=1}^s \langle S_{i,h} u_h|_{\partial\Omega_i}, v_h|_{\partial\Omega_i} \rangle = \sum_{i=1}^s \langle N_{i,h} f_i, v_h|_{\partial\Omega_i} \rangle \quad \forall v_h \in V_D^h(\Gamma_S), \quad (2.11)$$

<sup>1</sup>Otherwise, we can use an interpolation (if the data is continuous) or an  $L^2$ -orthogonal projection of the Dirichlet data to  $V^h(\Gamma_D)$ .

where

$$V_D^h(\Gamma_S) := V^h(\Gamma_S) \cap H_D^{1/2}(\Gamma_S). \quad (2.12)$$

Since the above bilinear form is symmetric, Lemma 1.3 implies

$$u_h = \underset{\substack{v_h \in V^h(\Gamma_S) \\ v_h|_{\Gamma_D} = g_D}}{\operatorname{argmin}} \sum_{i=1}^s \left( \frac{1}{2} \langle S_{i,h} v_h|_{\partial\Omega_i}, v_h|_{\partial\Omega_i} \rangle - \langle N_{i,h} f_i, v_h|_{\partial\Omega_i} \rangle \right). \quad (2.13)$$

The homogeneous version reads: find  $\tilde{u}_h \in V_D^h(\Gamma_S)$  such that

$$\sum_{i=1}^s \langle S_{i,h} \tilde{u}_h|_{\partial\Omega_i}, v_h|_{\partial\Omega_i} \rangle = \sum_{i=1}^s \langle N_{i,h} f_i - S_{i,h} \tilde{g}_D|_{\partial\Omega_i}, v_h|_{\partial\Omega_i} \rangle \quad \forall v_h \in V_D^h(\Gamma_S). \quad (2.14)$$

Then  $u_h = \tilde{g}_D + \tilde{u}_h$  solves (2.11). The minimization problem equivalent to (2.14) reads

$$\tilde{u}_h = \underset{v_h \in V_D^h(\Gamma_S)}{\operatorname{argmin}} \sum_{i=1}^s \left( \frac{1}{2} \langle S_{i,h} v_h|_{\partial\Omega_i}, v_h|_{\partial\Omega_i} \rangle - \langle N_{i,h} f_i - S_{i,h} \tilde{g}_D|_{\partial\Omega_i}, v_h|_{\partial\Omega_i} \rangle \right). \quad (2.15)$$

*Remark 2.3.* For a pure FE formulation ( $\mathcal{J}_{\text{FEM}} = \{1, \dots, s\}$ ), system (2.11) is the Schur complement system (cf. Sect. 1.2.6) of the global FE system

$$\text{find } u_h^{\text{FE}} \in V^h(\Omega), u_h^{\text{FE}}|_{\Gamma_D} = g_D : \quad a(u_h^{\text{FE}}, v_h) = \langle f, v_h \rangle \quad \forall v_h \in V^h(\Omega), v_h|_{\Gamma_D} = 0,$$

and  $u_h = u_h^{\text{FE}}|_{\Gamma_S}$ .

*Remark 2.4.* For the case that both FEM and BEM are included, the present form of the skeleton problem is also known as the *symmetric coupling* of FEM and BEM [Cos87]. For the advantages of this coupling and for other types of couplings see e.g. [ZKB77, ZKB79, BJ79, BJN78, JN80, CS90, Lan94, HHKL97, Hip02, KS02, Ste03b, Ste11]. See also the related works [CKL98, HSW00, HW91] on pure boundary element domain decomposition.

*Remark 2.5.* Note that on a FEM subdomain  $\Omega_i$ , the Schur complement matrix may be defined differently by eliminating non-coupling DOFs on the Neumann boundary *together* with the interior DOFs, cf. [TW05, Chap. 4]. The associated operator then maps from  $V^h(\partial\Omega_i \cap (\Gamma \cup \Gamma_D))$  to its dual. Correspondingly, one can define a modified Steklov-Poincaré operator  $S_i : H^{1/2}(\partial\Omega_i \setminus \Gamma_N) \rightarrow H_0^{-1/2}(\partial\Omega_i \setminus \Gamma_N)$ , and show the analogous statements of Sect. 1.2.6 for the two modified operators.

This approach is very natural in implementations, since one only needs to know whether a DOF is a coupling one (or a Dirichlet DOF) rather than if it is on the subdomain boundary. On the contrary, for BEM subdomains this approach is not natural, since all DOFs in  $V^h(\partial\Omega_i)$  are already on the subdomain boundary. Hence, for a unified presentation, we have chosen the separation into “true” boundary DOFs and interior DOFs for all subdomains. However, generalizations of the statements below to the modified operators can be proved without major effort.

### 2.1.3 Error Analysis of the Discrete Skeleton Formulation

The discrete skeleton formulation (2.14) introduces variational crimes when compared to (2.9). In this short section, we provide an a priori error analysis for the simplified case that

$$g_D = 0 \quad \text{and} \quad f_i = 0 \quad \forall i \in \mathcal{J}_{\text{BEM}}.$$

In this case, the solution  $u$  of (2.1) lies in the space

$$V_{S,D} := \{v \in H_D^1(\Omega) : \forall i \in \mathcal{J}_{\text{BEM}} : v|_{\Omega_i} = \mathcal{H}_i(v|_{\partial\Omega_i})\},$$

where  $\mathcal{H}_i$  denotes the harmonic extension from  $H^{1/2}(\partial\Omega_i) \rightarrow H^1(\Omega_i)$ . From the relation

$$a_i(v|_{\Omega_i}, w|_{\Omega_i}) = \langle S_i v|_{\partial\Omega_i}, w|_{\partial\Omega_i} \rangle, \quad \forall v, w \in V_{S,D}, i \in \mathcal{J}_{\text{BEM}},$$

it is straightforward to show that (2.9) is then equivalent to finding  $u \in V_{S,D}$  such that

$$\sum_{i \in \mathcal{J}_{\text{FEM}}} a_i(u|_{\Omega_i}, v|_{\Omega_i}) + \sum_{i \in \mathcal{J}_{\text{BEM}}} \langle S_i u|_{\partial\Omega_i}, v|_{\partial\Omega_i} \rangle = \sum_{i \in \mathcal{J}_{\text{FEM}}} \langle f_i, v|_{\Omega_i} \rangle \quad \forall v \in V_{S,D}. \quad (2.16)$$

We now define the space

$$V_{S,D}^h := \{v \in V_{S,D} : \forall i \in \mathcal{J}_{\text{BEM}} : v|_{\partial\Omega_i} \in V^h(\partial\Omega_i), \forall i \in \mathcal{J}_{\text{FEM}} : v|_{\Omega_i} \in V^h(\Omega_i)\}.$$

The discrete problem (2.14) is equivalent to finding  $u_h \in V_{S,D}^h$  such that

$$\sum_{i \in \mathcal{J}_{\text{FEM}}} a_i(u_h|_{\Omega_i}, v_h|_{\Omega_i}) + \sum_{i \in \mathcal{J}_{\text{BEM}}} \langle S_{i,h} u_h|_{\partial\Omega_i}, v_h|_{\partial\Omega_i} \rangle = \sum_{i \in \mathcal{J}_{\text{FEM}}} \langle f_i, v_h|_{\Omega_i} \rangle \quad \forall v_h \in V_{S,D}^h. \quad (2.17)$$

Comparing (2.16) and (2.17), we see that the variational crime is located in the BEM subdomains only, as we replace  $S_i$  by  $S_{i,h}$  for  $i \in \mathcal{J}_{\text{BEM}}$  (for the case  $\mathcal{J}_{\text{BEM}} = \emptyset$

see also Remark 2.3). Following the proof of [HLP10, Lemma 4.2], which involves a Strang lemma (cf. [Cia87]), we obtain that there exists a constant  $C$  depending on the domain  $\Omega$ , the coefficient  $\mathcal{A}$ , and on the BEM subdomains  $\Omega_i$ ,  $i \in \mathcal{J}_{\text{BEM}}$  such that

$$\begin{aligned} \|u - u_h\|_{H^1(\Omega)} &\leq C \inf_{v_h \in V_{S,D}^h} \|u - v_h\|_{H^1(\Omega)} \\ &+ C \left( \sum_{i \in \mathcal{J}_{\text{BEM}}} \inf_{z_h \in Z_h(\partial\Omega_i)} \|t_i(u) - z_h\|_{H^{-1/2}(\partial\Omega_i)}^2 \right)^{1/2}, \end{aligned}$$

where  $t_i(u) = S_i u|_{\partial\Omega_i}$  is the generalized conormal derivative of the solution  $u$  on  $\partial\Omega_i$  for  $i \in \mathcal{J}_{\text{BEM}}$ . Alternatively, one can employ an error estimate for the associated mixed setting, cf. [Hof11]. Using the minimizing property of the harmonic extension (Lemma 1.54), the FE approximation estimates from Lemma 1.44, and the BE approximation estimates from Lemma 1.89, we finally get the a priori error estimate

$$\|u - u_h\|_{H^1(\Omega)} \leq C h^s \left( |u|_{H^{1+s}(\Omega)}^2 + \sum_{i \in \mathcal{J}_{\text{BEM}}} \|t_i(u)\|_{H_{\text{pw}}^{-1/2+s}(\partial\Omega_i)}^2 \right)^{1/2},$$

provided that  $u \in H^{1+s}(\Omega)$  with  $s \in (0, 1]$ . In terms of  $h$ , this is the same error behavior as for a pure FEM discretization. Note, however, that the constant  $C$  depends on the subdomain partition. Finding an estimate which is explicit in the subdomains is non-trivial, but should be possible under the regularity assumptions that we will introduce in Sect. 2.5.2. For the limit case  $H_i \rightarrow h$  see [HLP10].

### 2.1.4 Conditioning of the Skeleton Problem

**Lemma 2.6.** *Let  $\mathcal{T}^H(\Omega)$  be a shape regular and quasi uniform triangulation of  $\Omega$ , and let each subdomain  $\Omega_i$  be a union of a few coarse elements of  $\mathcal{T}^H(\Omega)$ , such that the number of elements per subdomain is uniformly bounded. Furthermore, let the triangulation  $\mathcal{T}^h(\Gamma_S)$  be quasi-uniform. Then there exist uniform constants  $c_1, c_2 > 0$  such that for all  $v \in V_D^h(\Gamma_S)$ ,*

$$c_1 \alpha_{\min} H \|v\|_{L^2(\Gamma_S)}^2 \leq \underbrace{\sum_{i=1}^s \langle S_{i,h} v|_{\partial\Omega_i}, v|_{\partial\Omega_i} \rangle}_{=: \langle S_h v, v \rangle} \leq c_2 \|\mathcal{A}\|_{L^\infty(\Omega)} h^{-1} \|v\|_{L^2(\Gamma_S)}^2,$$

where  $\alpha_{\min} > 0$  is the constant from (1.22) (p. 21). Hence, the conditioning of the skeleton problems (2.11) and (2.14) is given by

$$\kappa(S_h) = \mathcal{O}\left(\frac{\|\mathcal{A}\|_{L^\infty(\Omega)}}{\alpha_{\min}} H^{-1} h^{-1}\right).$$



*Proof.* The following proof is identical to [TW05, Lemma 4.11]. To simplify its presentation, we assume that  $\text{diam}(\Omega) = 1$ , and we extend  $v$  from  $\Gamma_S$  to  $\Omega$  by setting  $v|_{\Omega_i} := \mathcal{H}_i v$ . By Theorem 1.23 and a scaling argument we get that

$$\begin{aligned} H_i \|v\|_{L^2(\partial\Omega_i)}^2 &\leq H_i^2 \left( |v|_{H^{1/2}(\partial\Omega_i)}^2 + H_i^{-1} \|v\|_{L^2(\partial\Omega_i)}^2 \right) \\ &\lesssim H_i^2 \left( |v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|v\|_{L^2(\Omega_i)}^2 \right). \end{aligned}$$

Above,  $\lesssim$  stands for  $\leq C$ , where  $C$  is a generic, uniform constant. Summing over  $i = 1, \dots, s$ , using that  $H_i \approx H \leq 1$ , Friedrichs' inequality, as well as the spectral equivalence relations from Sect. 1.3.8.3, we obtain

$$\begin{aligned} \alpha_{\min} H \|v\|_{L^2(\Gamma_S)}^2 &\lesssim \alpha_{\min} \|v\|_{H^1(\Omega)}^2 \lesssim \alpha_{\min} |v|_{H^1(\Omega)}^2 \\ &\leq \sum_{i=1}^s \langle S_i v, v \rangle_{\partial\Omega_i} \lesssim \sum_{i=1}^s \langle S_{i,h} v, v \rangle_{\partial\Omega_i}. \end{aligned}$$

This shows the lower bound. For the upper bound, note first that with the same notation as above,

$$\sum_{i=1}^s \langle S_{i,h} v, v \rangle_{\partial\Omega_i} \lesssim \sum_{i=1}^s \langle S_i v, v \rangle_{\partial\Omega_i} \lesssim \|\mathcal{A}\|_{L^\infty(\Omega)}^2 \sum_{i=1}^s |v|_{H^{1/2}(\partial\Omega_i)}^2.$$

To conclude the proof, we apply the inverse inequality

$$|w|_{H^{1/2}(\partial\Omega_i)} \lesssim h^{-1/2} \|w\|_{L^2(\Omega_i)} \quad \forall w \in V^h(\partial\Omega_i),$$

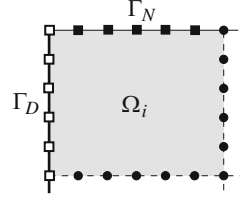
which can be derived using interpolation theory (cf. [TW05, Lemma B.27]).  $\square$

*Remark 2.7.* Preconditioners for the skeleton problem are also called Schur complement preconditioners. Examples are the BPS type and wirebasket preconditioners (see [TW05, Sect. 5] and the pioneering papers [BPS86, BPS87, BPS88, BPS89]), the Neumann-Neumann preconditioners (Sect. 2.3) and the BDDC preconditioner (Sect. 5.1.4). See also the related papers [HL92, HLM91a, HLM91b, CKL98, HS01, Ste03a, KL04] as well as the work of Nepomnyaschikh [Nep91b, Nep07].

## 2.2 Formulation of One-Level FETI/BETI Methods

In the following Sect. 2.2.1, we derive the classical FETI/BETI methods starting from the homogenized minimization problem (2.14). Section 2.2.2 deals with the all-floating (total) FETI/BETI formulation starting from the non-homogenized

**Fig. 2.2** Nodes of the set  $\partial\Omega_i^h$  for a single subdomain  $\Omega_i$  (● interface nodes in  $\Gamma^h \cap \partial\Omega_i^h$ , □ Dirichlet nodes, ■ remaining Neumann nodes)



minimization problem (2.11). In Sect. 2.2.3, we give an interpretation of the Lagrange multipliers involved in both formulations. Section 2.2.4 introduces various preconditioners, and Sect. 2.2.5 discusses implementation issues.

Throughout the remainder of this chapter, we work exclusively in the discrete setting. To simplify the notation, we drop the subscript  $h$  in the operators  $S_{i,h}$  and  $N_{i,h}$ . That means from now on,  $S_i$ ,  $N_i$  refer to discrete operators

$$S_i : V^h(\partial\Omega_i) \rightarrow V^h(\partial\Omega_i)^*, \quad N_i : H^1(\Omega_i)^* \rightarrow V^h(\partial\Omega_i)^*. \quad (2.18)$$

Moreover, we need a few definitions concerning the skeletal triangulation. Let  $\Gamma_S^h$  be the set of nodes on  $\Gamma_S$ . Analogously,  $\partial\Omega_i^h$ ,  $\Gamma_D^h$ ,  $\Gamma^h$ ,  $\Gamma_{ij}^h$  are the sets of nodes on the respective parts of the skeleton, see also Fig. 2.2. A typical node will be denoted by  $x^h$ .

*Remark 2.8.* Note that, for all FEM subdomains, we can define the discrete operators  $S_i$ ,  $N_i$  differently by eliminating non-coupling Neumann DOFs *together* with the interior DOFs, which is much more practicable in implementations, cf. Remark 2.5 and [TW05, Chap. 4]. This results in operators

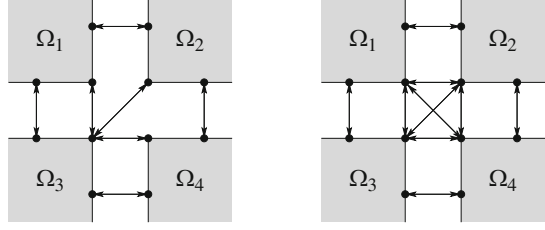
$$S_i : V^h(\partial\Omega_i \setminus \Gamma_N) \rightarrow V^h(\partial\Omega_i \setminus \Gamma_N)^*, \quad N_i : H^1(\Omega_i) \rightarrow V^h(\partial\Omega_i \setminus \Gamma_N)^*. \quad (2.19)$$

This is in fact done in the original FETI method, but for a unified presentation of FETI/BETI, we have chosen the setting in (2.18). The theory below, however, can be generalized without major difficulties to the case of (2.19).

### 2.2.1 Formulation of Classical FETI/BETI

In this section we derive the classical formulation of the FETI method (as introduced in [FR94]) as well as the BETI and coupled FETI/BETI methods (as introduced in [LS03, LS05]). Our presentation mainly follows [KW01, LS05], and [TW05, Sect. 6.3]. Our starting point is the homogeneous skeleton problem (2.14).

**Fig. 2.3** Sketch of non-redundant constraints (*left*) and fully redundant constraints (*right*) for a node that is shared by four subdomains



### 2.2.1.1 Tearing and Interconnecting

The idea of “tearing” is to introduce separate unknowns for  $\tilde{u}|_{\partial\Omega_i} \in V_D^h(\partial\Omega_i)$  on the subdomain boundaries. To this end we define the spaces<sup>2</sup>

$$W_i := V_D^h(\partial\Omega_i) \quad \text{and} \quad W := \prod_{i=1}^s W_i. \quad (2.20)$$

We denote the components of  $w \in W$  by  $w_i$  and write

$$w := [w_i]_{i=1}^s \in W. \quad (2.21)$$

**A Jump Operator.** Functions in the product space  $W$  are typically discontinuous across subdomain interfaces. Continuity (“interconnecting”) is enforced by constraints of the form

$$w_i(x^h) - w_j(x^h) = 0 \quad \text{for } x^h \in \Gamma_{ij}^h, \quad i > j. \quad (2.22)$$

In this book, we restrict ourselves to *fully redundant* constraints, i.e., we impose *all* constraints of the form (2.22), see [TW05, Sect. 6.3.3] and Fig. 2.3, right. They turn out to be advantageous in implementations due to the full symmetry. In the non-redundant case (see [TW05, Sect. 6.3.2] and Fig. 2.3, left) a minimal number of necessary constraints is used. For other variants, such as orthogonal constraints, see e.g. [FP03, FP04].

**Definition 2.9.** Let us assume a numbering of the (fully redundant) constraints (2.22) with  $M$  being the total number of constraints. The entry of a vector  $\mu \in \mathbb{R}^M$  corresponding to the constraint (2.22) at the node  $x^h \in \Gamma_{ij}^h$  is denoted by  $\mu_{ij}(x^h)$ .

Introducing the (linear) jump operator  $B : W \rightarrow \mathbb{R}^M$ , given by

$$(Bw)_{ij}(x^h) = w_i(x^h) - w_j(x^h) \quad \text{for } x^h \in \Gamma_{ij}^h, \quad i > j, \quad (2.23)$$

<sup>2</sup>For the setting (2.19),  $W_i = V_D^h(\partial\Omega_i \setminus \Gamma_N)$ .

we can write all constraints compactly as

$$B w = 0.$$

The essential property of  $B$  is (and should be) that the subspace

$$\widehat{W} := \{w \in W : B w = 0\} = \ker(B) \quad (2.24)$$

can be identified with  $V_D^h(\Gamma_S)$ , in short:  $V_D^h(\Gamma_S) \equiv \widehat{W}$ .

*Remark 2.10.* With respect to the standard nodal basis, the operator  $B$  is represented by signed Boolean matrix (with entries 0, 1 or  $-1$ ). For other discretization spaces we refer to Sect. 2.8.

**Saddle point formulation.** In the sequel, to avoid cumbersome notation, we will regard the approximate Steklov-Poincaré approximants  $S_i$  as operators mapping from  $W_i$  to  $W_i^*$ , with the only exception of  $S_i \tilde{g}_{D|\partial\Omega_i}$ . In addition, we define  $S : W \rightarrow W^*$  by

$$\langle S v, w \rangle := \sum_{i=1}^s \langle S_i v_i, w_i \rangle \quad \text{for } v, w \in W, \quad (2.25)$$

in short  $S := \text{diag}(S_i)_{i=1}^s$ , and the linear functional  $g \in W^*$  by

$$\langle g, w \rangle := \sum_{i=1}^s \langle g_i, w_i \rangle := \sum_{i=1}^s \langle N_i f_i - S_i \tilde{g}_{D|\partial\Omega_i}, w_i \rangle \quad \text{for } w \in W, \quad (2.26)$$

in short  $g := [g_i]_{i=1}^s = [N_i f_i - S_i \tilde{g}_{D|\partial\Omega_i}]_{i=1}^s$ . Using this notation we can write the minimization problem (2.15) equivalently as the constrained minimization problem

$$\tilde{u} = \underset{\substack{w \in W \\ B w = 0}}{\text{argmin}} \frac{1}{2} \langle S w, w \rangle - \langle g, w \rangle. \quad (2.27)$$

For simplicity, we use the symbol  $\tilde{u}$  simultaneously for the solution of (2.14) and the solution of the above problem. The equivalent saddle point problem reads as follows (see also the following Sect. 2.2.1.2). Find  $(\tilde{u}, \lambda) \in W \times U$ :

$$\begin{bmatrix} S & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} \tilde{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} g \\ 0 \end{bmatrix}, \quad (2.28)$$

where  $U = \mathbb{R}^M$  is the space of Lagrange multipliers. The second equation ( $B \tilde{u} = 0$ ) ensures that  $\tilde{u} \in \widehat{W} \equiv V_D^h(\Gamma_S)$ .

**Lemma 2.11.** *Under the assumptions made in Sect. 2.1,  $\ker(S) \cap \ker(B) = \{0\}$ . Therefore, Problem (2.28) is uniquely solvable up to adding elements from  $\ker B^\top$  to  $\lambda$ . The solution  $\tilde{u} \in \widehat{W} \equiv V_D^h(\Gamma_S)$  is the unique solution of (2.14).*

*Proof.* Since  $\ker(B) = \widehat{W} \equiv V_D^h(\Gamma_S)$  and the bilinear form  $\langle S \cdot, \cdot \rangle$  is coercive on  $V_D^h(\Gamma_S)$ , it follows that  $\ker(S) \cap \ker(B) = \{0\}$ . The classical Brezzi theory (cf. Lemma 1.16) implies the solvability of (2.28). Finally,  $\tilde{u}$  solves the constrained minimization problem (2.27), which is equivalent to (2.15) and (2.14).  $\square$

For reasons that will become clear later on, we agree that the Lagrange multipliers are in  $U$ , but  $B w$  is in its dual:

$$B : W \rightarrow U^*, \quad B^\top : U \rightarrow W^*.$$

Of course,  $U^* = \mathbb{R}^M$  too, and the duality pairing in  $U^* \times U$  is nothing else than the Euclidean inner product, i.e.,  $\langle \mu, \lambda \rangle = (\mu, \lambda)_{\ell^2}$  for  $\lambda \in U$ ,  $\mu \in U^*$ . Nevertheless, we will keep track of the subtle difference between  $U$  and  $U^*$  for this has several advantages.

The jump operator  $B$  can be decomposed into local operators  $B_i : W_i \rightarrow U^*$ ,  $i = 1, \dots, s$  such that

$$\langle B w, \lambda \rangle = \sum_{i=1}^s \langle B_i w_i, \lambda \rangle \quad \forall \lambda \in U.$$

### 2.2.1.2 Interpretation of the Lagrange Multipliers

This subsection may be skipped by readers who are mainly interested in the derivation of the FETI/BETI method. It contains an interpretation of the Lagrange multipliers as normal fluxes, and it discusses the adjoint  $B^\top$  in more detail.

The saddle point formulation (2.28) can also be derived directly from the discrete skeleton formulation (2.14) without a detour via the minimization problem.<sup>3</sup> We show this alternative derivation because it provides more insight on the Lagrange multipliers. Let  $\tilde{u} \in V_D^h(\Gamma_S)$  be the solution of (2.14) and let

$$t_i := S_i \tilde{u}|_{\partial\Omega_i} - g_i$$

denote the discrete (generalized) conormal derivative<sup>4</sup> and set  $t = [t_i]_{i=1}^s \in W^*$ . Substituting this formula into (2.14) and using that  $V_D^h(\Gamma_S) \equiv \widehat{W} = \ker(B)$ , we obtain

$$\langle t, w \rangle = 0 \quad \forall w \in \ker(B). \quad (2.29)$$

<sup>3</sup>This way, FETI formulations can be derived for non-symmetric or indefinite problems.

<sup>4</sup>Note that the given Neumann data  $g_N$  is already incorporated in  $g_i$ ; if  $\tilde{u}$  is the solution,  $t_i$  vanishes on all the interior nodes of the local Neumann boundary.

This is equivalent to

$$t \in \ker(B)^\circ = \text{range}(B^\top),$$

cf. Lemma 1.8. We now

- Parametrize the solution  $\tilde{u} \in V_D^h(\Gamma_S)$  by  $\tilde{u} \in W$  with the side condition  $B \tilde{u} = 0$ ,
- Parametrize the conormal derivative  $t \in W^*$  by  $-B^\top \lambda$  with  $\lambda \in U$ , therefore fulfilling condition (2.29) automatically.

From the definition of  $t$  and  $g$ , we get the equation  $-B^\top \lambda = S \tilde{u} - g$ . Together with  $B \tilde{u} = 0$ , this yields exactly (2.28). Under this perspective, the Lagrange multipliers themselves can be interpreted as normal fluxes. More precisely, they *parametrize* the normal fluxes of the solution  $\tilde{g}_D + \tilde{u}$  of (2.11):

$$-B_i^\top \lambda = S_i(\tilde{g}_D|_{\partial\Omega_i} + \tilde{u}_i) - N_i f_i.$$

For an interpretation in a mechanical context see [RF98a, RF99].

**Definition 2.12.** For  $x^h \in \Gamma_S^h$  we define

$$\mathcal{N}_{x^h} := \{k = 1, \dots, s : x^h \in \partial\Omega_k^h\},$$

i.e., the index set of the subdomains sharing the node  $x^h$ . Furthermore, we set

$$\mathcal{N}_i := \{k = 1, \dots, s : \partial\Omega_i \cap \partial\Omega_k \neq \emptyset\}.$$

**Definition 2.13.** For  $i = 1, \dots, s$  and  $x^h \in \partial\Omega_i^h$ , let  $\varphi_{i,x^h} \in V^h(\partial\Omega_i)$  be the nodal basis function corresponding to node  $x^h$ . For  $t \in W^*$ , we set

$$t_{i,x^h} := \langle t_i, \varphi_{i,x^h} \rangle_{\partial\Omega_i}.$$

**Lemma 2.14.** For  $\mu \in U^*$  and  $\lambda \in U$ ,

$$\langle \mu, \lambda \rangle = \sum_{x^h \in \Gamma^h} \sum_{\substack{i, j \in \mathcal{N}_{x^h} \\ i > j}} \mu_{ij}(x^h) \lambda_{ij}(x^h).$$

The adjoints  $B^\top : U \rightarrow W^*$  and  $B_i^\top : U \rightarrow W_i^*$  fulfill

$$(B^\top \lambda)_{i,x^h} = \langle B_i^\top \lambda, \varphi_{i,x^h} \rangle_{\partial\Omega_i} = \sum_{j \in \mathcal{N}_{x^h} \setminus \{i\}} \text{sign}(i-j) \lambda_{ij}(x^h) \quad \text{for } x^h \in \partial\Omega_i^h \cap \Gamma^h.$$

*Proof.* The proof is straightforward.

### 2.2.1.3 Eliminating Local Variables

The first set of equations in (2.28) read

$$S_i \tilde{u}_i = g_i - B_i^\top \lambda \quad \forall i = 1, \dots, s.$$

The goal of this subsection is to obtain an explicit formula for  $\tilde{u}_i$ , which is not straightforward as some of the operators  $S_i$  are singular.

In the usual theory of iterative substructuring methods, a *floating subdomain* is defined as a subdomain  $\Omega_i$  whose boundary  $\partial\Omega_i$  does not intersect the Dirichlet boundary  $\Gamma_D$ . Here, we use a more abstract form.

**Definition 2.15 (floating subdomain).** A subdomain  $\Omega_i$  is called a *floating subdomain* if  $S_i : W_i \rightarrow W_i^*$  is singular, otherwise it is a *non-floating subdomain*. The index set corresponding to the floating subdomains is denoted by  $\mathcal{J}_{\text{float}}$ .

In the case of the potential equation, the solution of the pure Neumann problem on a floating subdomain is only unique up to an additive constant, and so

$$\ker(S_i) = \begin{cases} \text{span}\{1_{\partial\Omega_i}\} & \text{if } i \in \mathcal{J}_{\text{float}}, \\ \{0\} & \text{else.} \end{cases}$$

In a non-floating subdomain,

$$\tilde{u}_i = S_i^{-1}(g_i - B_i^\top \lambda) \quad \forall i \notin \mathcal{J}_{\text{float}}.$$

For the remaining subdomains, we need the solvability conditions

$$g_i - B_i^\top \lambda \in \text{range}(S_i) \quad \forall i \in \mathcal{J}_{\text{float}}.$$

Choosing an injective operator

$$R_i : \mathbb{R}^{\dim(\ker(S_i))} \rightarrow W_i, \quad \text{range}(R_i) = \ker(S_i), \quad (2.30)$$

the local solution  $u_i$  can be represented by

$$\tilde{u}_i = S_i^\dagger(g_i - B_i^\top \lambda) + R_i \xi_i, \quad (2.31)$$

where  $S_i^\dagger$  is a pseudo inverse of  $S_i$  (see Sect. 1.1.3.4), and  $\xi_i \in \mathbb{R}^{\dim(\ker(S_i))}$ . Actually, formula (2.31) is valid for all  $i = 1, \dots, s$ . In our setting, we choose  $R_i \xi_i = \xi_i$  if  $i \in \mathcal{J}_{\text{float}}$  and 0 otherwise. Since  $\text{range}(S_i) = \ker(S_i)^\circ = \text{range}(R_i)^\circ = \ker(R_i^\top)$  (see Lemma 1.8), the compatibility conditions rewrite as

$$R_i^\top(g_i - B_i^\top \lambda) = 0 \quad \forall i = 1, \dots, s, \quad (2.32)$$

where for  $i \notin \mathcal{I}_{\text{float}}$ , this condition is trivial. Introducing

$$Z := \prod_{i=1}^s \mathbb{R}^{\dim(\ker(S_i))}$$

and the operators  $R = \text{diag}(R_i)_{i=1}^s : Z \rightarrow W$ , and  $S^\dagger := \text{diag}(S_i^\dagger) : W^* \rightarrow W$ , we see from (2.31) and (2.32) that

$$\tilde{u} = S^\dagger(g - B^\top \lambda) + R \xi \quad (2.33)$$

for some  $\xi \in Z$ , provided that

$$R^\top B^\top \lambda = R^\top g. \quad (2.34)$$

Eliminating  $\tilde{u}$  from (2.28) using (2.33) yields

$$B S^\dagger(g - B^\top \lambda) + B R \xi = 0.$$

After reordering the terms above and adding (2.34) to the set of equations, we obtain the system

$$\begin{aligned} B S^\dagger B^\top \lambda - B R \xi &= B S^\dagger g, \\ R^\top B^\top \lambda &= R^\top g. \end{aligned}$$

By defining the operators

$$F := B S^\dagger B^\top, \quad G := B R, \quad (2.35)$$

we see that the above system has saddle point structure. In the following, we briefly discuss the solvability of this problem and its relation to (2.28).

**Lemma 2.16.** *The problem of finding  $(\lambda, \xi) \in U \times Z$  such that*

$$\begin{bmatrix} F & -G \\ G^\top & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \xi \end{bmatrix} = \begin{bmatrix} B S^\dagger g \\ R^\top g \end{bmatrix} \quad (2.36)$$

*is uniquely solvable up to adding an element from  $\ker(B^\top)$  to  $\lambda$ . With*

$$\tilde{u} = S^\dagger(g - B^\top \lambda) + R \xi,$$

*$(\tilde{u}, \lambda)$  solves (2.28), and  $\tilde{u} \in \widehat{W} \equiv V_D^h(\Gamma_S)$  solves (2.14).*

*Proof.* First note that the operators  $F, G^\top$  vanish on  $\ker(B^\top)$ . Therefore,  $F$  and  $G^\top$  are well-defined on the factor space  $U_{/\ker(B^\top)}$ . We apply Lemma 1.16 (with  $V \mapsto U_{/\ker(B^\top)}$ ,  $Q \mapsto Z$ ). The assumptions hold because  $\ker(F) \cap \ker(G^\top) = \ker(B^\top)$ .



One easily shows that  $\ker(G) = \ker(B|_{\text{range}(R)}) = \ker(B) \cap \ker(S) = \{0\}$ , which implies uniqueness. The rest of the proof is straightforward.  $\square$

*Remark 2.17.* For a pure FEM, i.e.,  $\mathcal{J}_{\text{BEM}} = \emptyset$ , the use of Schur complement operators in the derivation can be circumvented. For each  $i = 1, \dots, s$ , let  $X_i := V_D^h(\Omega_i)$  be the FE space on  $\Omega_i$  with respect to  $\mathcal{T}^h(\Omega_i)$  and set  $X := \prod_{i=1}^s X_i$ . Moreover, let  $A_i : X_i \rightarrow X_i^*$  and  $\ell_i \in X_i^*$  be the operator and functional corresponding to the local stiffness matrix and load vector (after the homogenization). Let  $\bar{B}_i : X_i \rightarrow U^*$  be the jump operator defined by  $\bar{B}_i v_i := B_i v_i|_{\partial\Omega_i}$ , then the global problem is identical to find  $(u, \lambda) \in V \times U$ :

$$\begin{bmatrix} A & \bar{B}^\top \\ \bar{B} & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} \ell \\ 0 \end{bmatrix},$$

where  $A = \text{diag}(A_i)_{i=1}^s$ , and  $\ell = [\ell_i]_{i=1}^s \in V^*$ . Having at hand pseudo inverses  $A_i^\dagger$  and injective operators  $\bar{R}_i : \mathbb{R}^{\dim(\ker(A_i))} \rightarrow X_i$  such that  $\ker(A_i) = \text{range}(\bar{R}_i)$ , we can reduce this system as in Sect. 2.2.1.3. The resulting system is identical to (2.36), i.e.,

$$F = B S^\dagger B^\top = \bar{B} A^\dagger \bar{B}^\top, \quad B S^\dagger g = \bar{B} A^\dagger \ell, \quad G = B R = \bar{B} \bar{R}, \quad R^\top g = \bar{R}^\top \ell,$$

cf. [RFTM99, Sect. 2.1.3].

*Remark 2.18.* The result of Remark 2.17 also holds for arbitrary combinations of FEM and BEM if we reformulate the equations in a mixed setting. For  $i \in \mathcal{J}_{\text{BEM}}$ , we use the local space  $X_i := V_D^h(\partial\Omega_i) \times Z^h(\partial\Omega_i)$  and define  $A_i : X_i \rightarrow X_i^*$ ,  $\bar{B}_i : X_i \rightarrow U^*$ , and  $\ell_i \in X_i^*$  by

$$A_i := \begin{bmatrix} D_i & \frac{1}{2}I + K_i^\top \\ \frac{1}{2}I + K_i & -V_i \end{bmatrix}, \quad \bar{B}_i \begin{bmatrix} v_i \\ t_i \end{bmatrix} := B_i v_i, \quad \ell_i := \begin{bmatrix} N_i f_i - D_i \tilde{g}_{D|\partial\Omega_i} \\ -(\frac{1}{2}I + K_i) \tilde{g}_{D|\partial\Omega_i} \end{bmatrix}.$$

### 2.2.1.4 A Projection Method

Note that the saddle point problem (2.36) is different in its structure from the saddle point problem (2.28) because the variables  $\xi$  lives in the space  $Z$  of small dimension:

$$\dim(Z) < s \ll \dim(U).$$

Therefore, an inversion of a sparse system on  $Z$  is acceptable. Since the lower right block of System (2.36) is zero, we can use a *projection method*, cf. [FR91, FR94].

Following [FCM95], we define the space of *admissible Lagrange increments*

$$U_{\text{ad}} := \ker(G^\top) = \{\lambda \in U : B^\top \lambda \in \text{range}(S)\}. \quad (2.37)$$

**The Case  $R^\top g = 0$ .** Assume for a moment that  $R^\top g = 0$ . Then the solution  $\lambda$  of (2.36) lies in  $U_{\text{ad}}$ . We can always reach this subspace by an orthogonal projection (see Sect. 1.1.3.3). Let the self-adjoint linear operator

$$Q : U^* \rightarrow U$$

be positive definite on  $\text{range}(G)$ ; specific choices of  $Q$  will be given later on. The operator

$$P := I - Q G (G^\top Q G)^{-1} G^\top \quad (2.38)$$

is a projection from  $U$  onto  $U_{\text{ad}}$ . Note that  $G^\top Q G$  is SPD because of the assumptions on  $Q$  and because  $\ker(G) = \{0\}$ , see the proof of Lemma 2.16. Furthermore, for a suitable choice of  $Q$ , the matrix representing  $G^\top Q G$  is sparse, cf. Sect. 2.2.5.

Hence, if  $R^\top g = 0$ , it suffices to test the first line of (2.36) (rewritten as a variational problem) with test functions from  $U_{\text{ad}} = \text{range}(P)$ , i.e.,

$$\text{find } \lambda \in U_{\text{ad}} : \quad P^\top F \lambda = P^\top B S^\dagger g.$$

**The Case  $R^\top g \neq 0$ .** In the general case, we can decompose  $\lambda = \lambda_0 + \tilde{\lambda}$ , where  $G^\top \lambda_0 = R^\top g$  and  $\tilde{\lambda} \in U_{\text{ad}}$ . Apparently, the choice

$$\lambda_0 = Q G (G^\top Q G)^{-1} R^\top g \quad (2.39)$$

fulfills these requirements, and so we can homogenize the equation. Summarizing, we obtain the problem

$$\text{find } \tilde{\lambda} \in U_{\text{ad}} : \quad P^\top F \tilde{\lambda} = \underbrace{P^\top (B S^\dagger g - F \lambda_0)}_{= B S^\dagger (g - B^\top \lambda_0)}. \quad (2.40)$$

We will discuss this equation in detail in Sect. 2.2.1.5. Before, we need to see how to recover the variable  $\xi$  from  $\lambda_0$  and  $\tilde{\lambda}$ . Testing the first line of (2.36) (rewritten as a variational problem) with test functions from  $\text{range}(I - P)$ , we obtain

$$\begin{aligned} (I - P^\top) F \lambda - \underbrace{(I - P^\top) G \xi}_{= G \xi} &= (I - P^\top) B S^\dagger g \\ \iff G \xi &= \underbrace{(I - P^\top)}_{G(G^\top Q G)^{-1} G^\top Q} \underbrace{(F \lambda - B S^\dagger g)}_{B S^\dagger (B^\top \lambda - g)}. \end{aligned}$$

Applying  $(G^\top Q G)^{-1} G^\top Q$  to the last equation, we obtain the formula

$$\xi = -(G^\top Q G)^{-1} G^\top Q B S^\dagger (g - B^\top \lambda). \quad (2.41)$$

**Algorithm 2:** Classical FETI/BETI method based on PCG

---


$$g = [N_i f_i - S_i \tilde{g}_D|_{\partial\Omega_i}]_{i=1}^s$$

$$\lambda_0 = Q G (G^\top Q G)^{-1} R^\top g$$

$$d = P^\top B S^\dagger (g - B^\top \lambda_0)$$

solve  $P^\top F \tilde{\lambda} = d$  with PCG and initial value  $\tilde{\lambda}^{(0)} = 0$  and stop after  $k$  iterations

$$\lambda^{(k)} = \lambda_0 + \tilde{\lambda}^{(k)}$$

$$\xi^{(k)} = -(G^\top Q G)^{-1} G^\top Q B S^\dagger (g - B^\top \lambda^{(k)})$$

$$\tilde{u}^{(k)} = S^\dagger (g - B^\top \lambda^{(k)}) + R \xi^{(k)}$$


---

**Table 2.2** Overview on the spaces and operators involved in FETI/BETI

Spaces	
$U = \mathbb{R}^M$	Space of Lagrange multipliers
$U_{\text{ad}} = \ker(G^\top) \subset U$	Space of admissible Lagrange increments
$\tilde{U}_{\text{ad}} = U_{\text{ad}} / \ker(B^\top)$	Factor space modulo redundancies
$V_D^h(\Gamma_S)$	Skeletal FE space with homogeneous Dirichlet conditions
$W_i = \begin{cases} V_D^h(\partial\Omega_i) & \text{(classical)} \\ V^h(\partial\Omega_i) & \text{(all-floating)} \end{cases}$	Local spaces
$W = \prod_{i=1}^s W_i$	Product space (“discontinuous” functions)
$\widehat{W} = \ker(B) \subset W$	Subspace of continuous functions in $W$ , identifiable with $V_D^h(\Gamma_S)$
$Z = \prod_{i=1}^s \mathbb{R}^{\dim(\ker(S_i))}$	Parameter space of $\ker(S)$
Operators	
$B : W \rightarrow U^*$	Jump operator
$F : U \rightarrow U^*$	$F = B S^\dagger B^\top$
$G : Z \rightarrow U^*$	$G = B R$
$P : U \rightarrow U_{\text{ad}}$	Projection, $P = I - Q G (G^\top Q G)^{-1} G^\top$
$P^\top : U^* \rightarrow U_{\text{ad}}^*$	Projection, $P^\top = I - G (G^\top Q G)^{-1} G^\top Q$
$Q : U^* \rightarrow U$	Self-adjoint operator, SPD on $\text{range}(G)$
$R : Z \rightarrow W$	Injective operator, $\text{range}(R) = \ker(S)$
$S : W \rightarrow W^*$	Block operator of local approximate Steklov-Poincaré operators
$S^\dagger : W^* \rightarrow W$	Pseudo inverse of $S$

---

The complete method is summarized in Algorithm 2, p. 81. An overview on the spaces and operators involved is given in Table 2.2, p. 81.

### 2.2.1.5 The Central FETI/BETI Equation

Herein, we discuss the solvability and a solution algorithm of Problem (2.40). Since the Lagrange multipliers are only unique modulo  $\ker(B^\top)$ , we introduce the factor space

$$\widetilde{U}_{\text{ad}} := U_{\text{ad}/\ker(B^\top)}. \quad (2.42)$$

As mentioned in the proof of Lemma 2.16,  $F$  is well-defined on  $\widetilde{U}_{\text{ad}}$  because it vanishes on  $\ker(B^\top)$ . Recall from Sect. 1.1.3.3 that  $\text{range}(P^\top)$  is a realization of the dual of  $U_{\text{ad}}$ . Moreover, one can show that  $\text{range}(P^\top) \cap \text{range}(B)$  is a realization of the dual of  $\widetilde{U}_{\text{ad}}$ . In short,

$$\widetilde{U}_{\text{ad}}^* = \text{range}(P^\top) \cap \text{range}(B) = \{\mu \in \text{range}(B) : \langle Bz, Q\mu \rangle = 0 \quad \forall z \in \ker(S)\}. \quad (2.43)$$

**Lemma 2.19.** *The operator  $P^\top F$  maps  $\widetilde{U}_{\text{ad}}$  to  $\widetilde{U}_{\text{ad}}^*$  and  $P^\top F|_{\widetilde{U}_{\text{ad}}}$  is SPD.*

*Proof.* By definition,  $F$  maps  $U_{\text{ad}}$  to  $\text{range}(B)$ . From the definition of  $P^\top$  we see that  $P^\top(\text{range}(B)) \subset \text{range}(B) \cap \text{range}(P^\top) = \widetilde{U}_{\text{ad}}^*$ .

Since  $F$  is self-adjoint and positive semi-definite and  $P$  is a projection onto  $U_{\text{ad}}$ , it follows immediately that  $P^\top F|_{\widetilde{U}_{\text{ad}}}$  is self-adjoint and positive semi-definite. It remains to show the definiteness. Assume that for  $\lambda \in \widetilde{U}_{\text{ad}}$ ,

$$0 = \langle P^\top F \lambda, \lambda \rangle = \langle S^\dagger B^\top \lambda, B^\top \lambda \rangle.$$

Since  $\lambda \in \widetilde{U}_{\text{ad}}$  implies  $B^\top \lambda \in \text{range}(S)$ , we conclude from the properties of the pseudo inverse  $S^\dagger$  that  $B^\top \lambda = 0$ . Since  $\lambda$  is in the factor space modulo  $\ker(B^\top)$ , this means  $\lambda = 0$  and concludes the proof.  $\square$

As a consequence of Lemma 2.19, Problem (2.40) can be solved using a PCG method, see Sect. 1.2.4.2 and Corollary 1.50. Recall that we have left the choice of  $Q : U^* \rightarrow U$  open yet, and that we need a preconditioner (at least a formal one) mapping from  $\widetilde{U}_{\text{ad}}^*$  back to  $\widetilde{U}_{\text{ad}}$ .

**The unpreconditioned case.** We choose  $Q = I$  (recall that  $U = \mathbb{R}^M = U^*$ ). Then  $P^\top = P$  and  $U_{\text{ad}}^* = U_{\text{ad}}$ . Hence,  $P^\top F|_{\widetilde{U}_{\text{ad}}}$  maps to  $U_{\text{ad}} \cap \text{range}(B)$ , which is naturally embedded in the factor space  $\widetilde{U}_{\text{ad}}$ . Summarizing, the formal preconditioner can be chosen as the identity, i.e., we take *no* preconditioner.

**The preconditioned case.** All preconditioners under our consideration have the form

$$PM^{-1} \quad (2.44)$$

where  $M^{-1} : U^* \rightarrow U$ . The projection  $P$  makes sure that the preconditioner maps back to the space  $U_{\text{ad}}$ , which is embedded in  $\widetilde{U}_{\text{ad}}$ . Of course,  $(PM^{-1})|_{\widetilde{U}_{\text{ad}}^*}$  must be

SPD in order to make PCG applicable. The choice  $M^{-1} = I$  and  $Q = I$  gives the unpreconditioned case.

The complete FETI/BETI method for the classical formulation is summarized in Algorithm 2, p. 81. Note that in that algorithm, we can substitute  $g$ ,  $B$ ,  $R$ ,  $S$ , and  $S^\dagger$  by  $\ell$ ,  $\bar{B}$ ,  $\bar{R}$ ,  $A$ , and  $A^\dagger$  from Remarks 2.17 and 2.18, respectively (and omit the first line), see also Sect. 2.2.5.

*Remark 2.20.* By simple linear algebra, one shows the residual identity

$$d - P^\top F \tilde{\lambda}^{(k)} = B \tilde{u}^{(k)},$$

(see e.g. [FR94, FCRR98]) i.e., the residual in the CG algorithm controls the jump of the approximant  $\tilde{u}^{(k)}$ , and equivalently the jump of  $u^{(k)} = [\tilde{g}_D]_{\partial\Omega_i}^s_{i=1} + \tilde{u}^{(k)}$ . We note that the entire method can be rewritten in terms of the variables  $\iota^{(k)} := -B^\top(\lambda_0 + \tilde{\lambda}^{(k)})$  and  $\tilde{u}^{(k)}$ , see Sects. 2.2.1.2 and 2.2.3.

*Remark 2.21.* In Algorithm 2, we have chosen  $\tilde{\lambda}^{(0)} = 0$ , but any value in  $U_{\text{ad}}$  would be suitable. Formally, the PCG runs in the factor space  $\tilde{U}_{\text{ad}}$ , but in a standard implementation, we just use vectors in  $U$  (projected to  $U_{\text{ad}}$ ). Working in factor spaces might often be dangerous in practice, but not in the current case: the components in  $\ker(B^\top)$  of the iterates cannot blow up. If an iterate  $\tilde{\lambda}^{(k)}$  should have a non-zero contribution from  $\ker(B^\top)$ , the next iterate  $\tilde{\lambda}^{(k+1)}$  does not depend on this contribution, see also Sect. 2.2.3.

*Remark 2.22.* If a preconditioner  $M^{-1}$  is SPD on the whole of  $U^*$ , one can use  $Q = M^{-1}$ . In that case,  $PQ = QP^\top$  (cf. formula (1.5)) and so

$$PM^{-1}P^\top F = M^{-1}P^\top F,$$

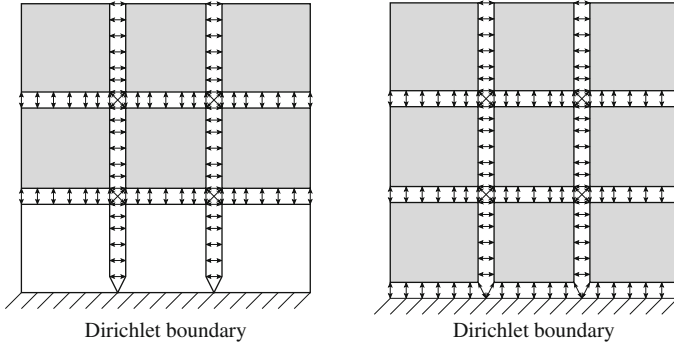
i.e., we can leave out  $P$  in (2.44), see also Sect. 2.2.5.

We leave it up to the reader to follow the subsequent presentation linearly or not. Here is a guide for “nonlinear” readers.

**Conditioning.** For the case  $Q = M^{-1} = I$ , the convergence of the CG method is determined by the condition number  $\kappa(P^\top F|_{\tilde{U}_{\text{ad}}})$ , cf. Lemma 1.49, which we will analyze in Sect. 2.4.1. In case of a global quasi-uniform mesh  $\mathcal{T}^h(\Gamma_S)$  with mesh parameter  $h$  and suitable assumptions on the subdomains, it can be shown that

$$\kappa(P^\top F|_{\tilde{U}_{\text{ad}}}) \leq C \frac{\|\mathcal{A}\|_{L^\infty(\Omega)} H}{\alpha_{\min} h},$$

where  $H$  denotes the maximal subdomain diameter (see Theorem 2.38). The constant  $C$  depends on the shape of the subdomains but is independent of  $\mathcal{A}$ ,  $h$ , and  $H$ , in particular independent of the number of subdomains. Note that the robustness with respect to the number of subdomains comes from the projection  $P$  which plays the role of a coarse solve, cf. [FMR94].



**Fig. 2.4** Illustration of a classical formulation (*left*) and the all-floating (or total) formulation (*right*). Floating subdomains are *dark-shaded*

**Preconditioning** is (in general) required if

- $H/h$  gets large (there is a large number of local unknowns in each subdomain), or
- $\|\mathcal{A}\|_{L^\infty(\Omega)}/\alpha_{\min}$  gets large (the coefficient  $\mathcal{A}$  varies over several orders of magnitude).

For the definition of preconditioners for piecewise constant coefficients see Sect. 2.2.4, for their analysis see Sect. 2.6. The case of highly varying (multiscale) coefficients is subject of Chap. 3.

**All-floating (Total) FETI/BETI.** An important variant of the classical FETI/BETI method introduced in the current subsection is the *all-floating (total) FETI/BETI method*, see Sect. 2.2.2, where additional Lagrange multipliers are used to enforce the Dirichlet boundary conditions. This simplifies the method in a certain sense.

**Implementation and Parallelization** of the classical and the all-floating FETI/BETI method are discussed in Sect. 2.2.5.

### 2.2.2 All-Floating (Total) FETI/BETI

The all-floating method is a variant of the classical FETI/BETI method, where one introduces additional Lagrange multipliers that enforce Dirichlet conditions. For an illustration see Fig. 2.4.

**Saddle Point Formulation.** We start with the inhomogeneous skeleton formulation (2.11). In contrast to Sect. 2.2.1, the working spaces are chosen as<sup>5</sup>

<sup>5</sup>For the setting (2.19),  $W_i = V^h(\partial\Omega_i \setminus \Gamma_N)$ .

$$W_i := V^h(\partial\Omega_i) \quad \text{and} \quad W := \prod_{i=1}^N W_i, \quad (2.45)$$

i.e., we do not incorporate any Dirichlet boundary conditions. In this subsection, we regard  $S_i$  as a mapping from  $W_i$  to  $W_i^*$  (opposed to Sect. 2.2.1).

In addition to the (fully redundant) interface constraints

$$w_i(x^h) - w_j(x^h) = 0 \quad \text{for } x^h \in \Gamma_{ij}^h, \quad i > j, \quad (2.46)$$

(cf. (2.22)) we also require

$$w_i(x^h) = \tilde{g}_D(x^h) \quad \text{for } x^h \in \Omega_i^h \cap \Gamma_D. \quad (2.47)$$

**Definition 2.23.** Let us assume a numbering of the constraints (2.46) and (2.47) with  $M$  being the total number of constraints. The entry of a vector  $\mu \in \mathbb{R}^M$  corresponding to constraint (2.46) is denoted by  $\mu_{ij}(x^h)$ , the entry corresponding to constraint (2.47) is denoted by  $\mu_{iD}(x^h)$ .

We define  $B : W \rightarrow U^*$  (where  $U = \mathbb{R}^M$ ) by

$$\left. \begin{aligned} (B w)_{ij}(x^h) &= w_i(x^h) - w_j(x^h) \text{ for } x^h \in \Gamma_{ij}^h, \quad i > j, \\ (B w)_{iD}(x^h) &= w_i(x^h) \quad \text{for } x^h \in \partial\Omega_i^h \cap \Gamma_D. \end{aligned} \right\} \quad (2.48)$$

Recall that we have chosen  $\tilde{g}_D \in V^h(\Gamma_S)$  such that it vanishes on all nodes in  $\Gamma_S^h \setminus \Gamma_D$ . Of course, we can identify  $\tilde{g}_D$  with a function in  $W$ , and (for simplicity) denote it again by  $\tilde{g}_D$ . With these considerations, the condition

$$B w = B \tilde{g}_D$$

requires that  $w$  is continuous across the subdomain interfaces, and that  $w|_{\Gamma_D} = g_D$ . Hence, the space

$$\widehat{W} := \ker(B) \equiv V_D^h(\Gamma_S)$$

coincides with that from Sect. 2.2.1. Analogously to Sect. 2.2.1.2, we derive a saddle point formulation from (2.11). We define

$$t_i := S_i u|_{\partial\Omega_i} - N_i f_i, \quad t = [t_i]_{i=1}^s \in W^*.$$

Equation (2.11) implies that

$$\langle t, v \rangle = 0 \quad \forall v \in \widehat{W},$$

and so  $t \in \ker(B)^\circ = \text{range}(B^\top)$ . Parametrizing  $t$  by  $-B^\top \lambda$  for  $\lambda \in U$  and representing the solution as  $u \in W$  with  $B u = B \tilde{g}_D$  yields the saddle point problem:

---

**Algorithm 3:** All-floating FETI/BETI method based on PCG
 

---

$g = [N_i f_i]_{i=1}^s$   
 $\lambda_0 = Q G (G^\top Q G)^{-1} R^\top g$   
 $d = P^\top B [S^\dagger (g - B^\top \lambda_0) - \tilde{g}_D]$   
 solve  $P^\top F \tilde{\lambda} = d$  with PCG and initial value  $\tilde{\lambda}^{(0)} = 0$  and stop after  $k$  iterations  
 $\lambda^{(k)} = \lambda_0 + \tilde{\lambda}^{(k)}$   
 $\xi^{(k)} = -(G^\top Q G)^{-1} G^\top Q B [S^\dagger (g - B^\top \lambda^{(k)}) - \tilde{g}_D]$   
 $u^{(k)} = S^\dagger (g - B^\top \lambda^{(k)}) + R \xi^{(k)}$

---

find  $(u, \lambda) \in W \times U$  such that

$$\begin{bmatrix} S & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} [N_i f_i]_{i=1}^s \\ B \tilde{g}_D \end{bmatrix}. \quad (2.49)$$

This problem has the same properties as (2.28): the solution  $u$  is unique, and  $\lambda$  is unique modulo  $\ker(B^\top)$ , cf. Lemma 2.11. Note that  $-B^\top \lambda$  coincides with the discrete (generalized) conormal derivative of the solution  $u$  (also on the Dirichlet boundary).

**Dual Saddle Point Formulation.** Going through the same steps as in Sect. 2.2.1.3, we derive the dual saddle point formulation. The spaces and operators are slightly different because (in our setting)

$$\ker(S_i) = \text{span}\{1_{\partial\Omega_i}\} \quad \forall i = 1, \dots, s,$$

and so *all subdomains are floating subdomains* in the sense of Definition 2.15. Hence,  $Z = \mathbb{R}^s$  and  $R : Z \rightarrow \ker(S)$  with  $(R\xi)_i = \xi_i$ . With the notations from Table 2.2, p. 81, the dual saddle point problem reads as follows. Find  $(\lambda, \xi) \in U \times Z$  such that

$$\begin{bmatrix} F & -G \\ G^\top & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \xi \end{bmatrix} = \begin{bmatrix} B(S^\dagger [N_i f_i]_{i=1}^s - \tilde{g}_D) \\ R^\top [N_i f_i]_{i=1}^s \end{bmatrix}.$$

We see that the structure of this system is identical to (2.40). The projection method is applied analogously to Sect. 2.2.1.4 and leads to an SPD problem for the variable  $\tilde{\lambda}$ , see Lemma 2.24 below. The complete all-floating FETI/BETI method is summarized in Algorithm 3. Also here, we can substitute  $g$ ,  $B$ ,  $R$ ,  $S$ , and  $S^\dagger$  by  $\bar{\ell}$ ,  $\bar{B}$ ,  $\bar{R}$ ,  $A$ , and  $A^\dagger$  from Remarks 2.17 and 2.18, respectively (and omit the first line), see also Sect. 2.2.5.

**Lemma 2.24.** *In the all-floating formulation, the operator  $P^\top F$  maps  $\tilde{U}_{\text{ad}}$  to  $\tilde{U}_{\text{ad}}^*$  and  $P^\top F|_{\tilde{U}_{\text{ad}}}$  is SPD.*

*Proof.* The proof is identical to that of Lemma 2.19.



*Remark 2.25.* Similarly to Remark 2.20, one can show the residual identity

$$d - P^\top F \widetilde{\lambda}^{(k)} = B(u^{(k)} - \tilde{g}_D),$$

where above,  $\tilde{g}_D$  is interpreted as an element in  $W$ . Hence, in the all-floating method, the residual in the CG method controls the jump of the approximant  $u^{(k)}$  as well as the error in the Dirichlet conditions.

**Lemma 2.26.** *In the all-floating formulation, for  $\mu \in U^*$  and  $\lambda \in U$ ,*

$$\langle \mu, \lambda \rangle = \sum_{x^h \in \Gamma^h} \sum_{\substack{i, j \in \mathcal{N}_{x^h} \\ i > j}} \mu_{ij}(x^h) \lambda_{ij}(x^h) + \sum_{i=1}^s \sum_{x^h \in \partial\Omega_i^h \cap \Gamma_D} \mu_{iD}(x^h) \lambda_{iD}(x^h).$$

The adjoints  $B^\top : U \rightarrow W^*$  and  $B_i^\top : U \rightarrow W_i^*$  fulfill

$$(B^\top \lambda)_{i, x^h} = \langle B_i^\top \lambda, \varphi_{i, x^h} \rangle_{\partial\Omega_i} = \begin{cases} \sum_{j \in \mathcal{N}_{x^h} \setminus \{i\}} \text{sign}(i - j) \lambda_{ij}(x^h) & \text{if } x^h \in \partial\Omega_i^h \cap \Gamma^h, \\ \lambda_{iD}(x^h) & \text{if } x^h \in \partial\Omega_i^h \cap \Gamma_D, \\ 0 & \text{else.} \end{cases}$$

*Proof.* The proof is straightforward.

### 2.2.3 FETI/BETI in Terms of Fluxes and Traces

Algorithm 4 displays the FETI/BETI method (both for the classical and the all-floating formulation) including explicitly the PCG method. Note that in order to unify the two formulations, we have used that in the classical formulation,  $B \tilde{g}_D = 0$  and

$$u^{(k)} = \tilde{u}^{(k)} + [\tilde{g}_D]_{\partial\Omega_i}^s_{i=1} = P_Z S^\dagger ([N_i f_i]_{i=1}^s - B^\top \lambda^{(k)}),$$

where

$$P_Z := I - R(G^\top Q G)^{-1} G^\top Q B. \quad (2.50)$$

We note that this particular projection is also treated in [FP03, SM08]. One can now show that for *both* formulations,

$$\tilde{u}^{(k)} = u^{(k)} - \tilde{g}_D = P_Z (S^\dagger (g - B^\top \lambda^{(k)}) - \tilde{g}_D).$$

We have seen in Sect. 2.2.1.2 that the Lagrange multipliers are temporary variables that parametrize fluxes. Also, due to Remarks 2.20 and 2.25,

$$d^{(k)} = B \tilde{u}^{(k)}.$$

---

**Algorithm 4:** FETI/BETI method in terms of Lagrange multipliers  $(\lambda^{(k)})$  and jumps  $(d^{(k)})$  including PCG

---


$$g = \begin{cases} [N_i f_i - S_i \tilde{g}_D]_{i=1}^s & \text{in classical formulation} \\ [N_i f_i]_{i=1}^s & \text{in all-floating formulation} \end{cases}$$

$$\lambda^{(0)} = Q G (G^\top Q G)^{-1} R^\top g$$

$$d^{(0)} = P^\top B (S^\dagger (g - B^\top \lambda^{(0)}) - \tilde{g}_D)$$

$$k = 0$$

**repeat**

$$z^{(k)} = P M^{-1} d^{(k)} \quad (\text{in the unpreconditioned case: } z^{(k)} = d^{(k)})$$

$$q^{(k)} = z^{(k)} + \beta_{k-1} q^{(k-1)} \quad \text{where } \beta_{-1} = 0, \beta_{k-1} = \frac{(d^{(k)}, z^{(k)})_{\ell^2}}{(d^{(k-1)}, z^{(k-1)})_{\ell^2}}$$

$$\lambda^{(k+1)} = \lambda^{(k)} + \alpha^{(k)} q^{(k)} \quad \text{where } \alpha_k = \frac{(d^{(k)}, z^{(k)})_{\ell^2}}{(P^\top F q^{(k)}, q^{(k)})_{\ell^2}}$$

$$d^{(k+1)} = d^{(k)} - \alpha_k P^\top F q^{(k)} = B(u^{(k)} - \tilde{g}_D)$$

$$k = k + 1$$

**until** *stopping criterion fulfilled for*  $d^{(k)}$

$$u^{(k)} = \tilde{g}_D + P_Z (S^\dagger (g - B^\top \lambda^{(k)}) - \tilde{g}_D)$$


---

Introducing the additional variables

$$\begin{aligned} t^{(k)} &:= -B^\top \lambda^{(k)}, \\ s^{(k)} &:= B^\top z^{(k)} = B^\top P M^{-1} d^{(k)} = B^\top P M^{-1} B \tilde{u}^{(k)}, \\ p^{(k)} &:= B^\top q^{(k)}, \end{aligned}$$

we can rewrite the whole algorithm in terms of the fluxes  $t^{(k)}$ , the (discontinuous) Dirichlet traces  $\tilde{u}^{(k)}$ , the preconditioned “residual”  $s^{(k)}$ , and the search directions  $p^{(k)}$  in the flux space, see Algorithm 5. In that algorithm we have to use the projection operators  $P_Z$  and  $P_Z^\top$  rather than  $P$  and  $P^\top$ . We have

$$\begin{aligned} \text{range}(P_Z) &= \ker(S)^\perp_{B^\top Q B}, & P^\top B &= B P_Z \\ \text{range}(P_Z^\top) &= \text{range}(S) & B^\top P &= P_Z^\top B^\top. \end{aligned}$$

We notice that Algorithm 5 is formally a PCG algorithm (cf. Algorithm 1) with the operator

$$P_Z S^\dagger : Y \rightarrow Y^*$$

and the preconditioner

$$P_Z^\top B^\top M^{-1} B : Y^* \rightarrow Y,$$

where the flux space  $Y$  and the Dirichlet trace space  $Y^*$  are given by

---

**Algorithm 5:** FETI/BETI method including PCG rewritten in terms of fluxes  $(t^{(k)})$  and traces  $(\tilde{u}^{(k)})$

---

$g = \begin{cases} [N_i f_i - S_i \tilde{g}_D|_{\partial\Omega_i}]_{i=1}^s & \text{in classical formulation} \\ [N_i f_i]_{i=1}^s & \text{in all-floating formulation} \end{cases}$   
 $t^{(0)} = -(I - P_Z^\top)g$   
 $\tilde{u}^{(0)} = P_Z(S^\dagger(g + t^{(0)}) - \tilde{g}_D)$   
 $k = 0$   
**repeat**  
      $s^{(k)} = P_Z^\top B^\top M^{-1} B \tilde{u}^{(k)}$  (in the unpreconditioned case:  $s^{(k)} = B^\top B \tilde{u}^{(k)}$ )  
      $p^{(k)} = s^{(k)} + \beta_{k-1} p^{(k-1)}$  where  $\beta_{-1} = 0$ ,  $\beta_{k-1} = \frac{(\tilde{u}^{(k)}, s^{(k)})_{\ell^2}}{(\tilde{u}^{(k-1)}, s^{(k-1)})_{\ell^2}}$   
      $t^{(k+1)} = t^{(k)} - \alpha_k p^{(k)}$  where  $\alpha_k = \frac{(\tilde{u}^{(k)}, s^{(k)})_{\ell^2}}{(P_Z S^\dagger p^{(k)}, p^{(k)})_{\ell^2}}$   
      $\tilde{u}^{(k+1)} = \tilde{u}^{(k)} - \alpha_k P_Z S^\dagger p^{(k)} = P_Z(S^\dagger(g + t^{(k)}) - \tilde{g}_D)$   
      $k = k + 1$   
**until** stopping criterion fulfilled for  $B \tilde{u}^{(k)}$   
 $u^{(k)} = \tilde{g}_D + \tilde{u}^{(k)}$

---

$$Y := \text{range}(B^\top) \cap \text{range}(S),$$

$$Y^* := \{w \in \ker(S)^{\perp_{B^\top Q B}} : S w \in \text{range}(B^\top)\} = \{w \in \ker(S)^{\perp_{B^\top Q B}} \cap \ker(B)^{\perp_S}\}$$

(it can be argued that  $Y^*$  above is a realization of the dual of  $Y$ ). If  $Q = M^{-1} = I$ , then the operator  $P_Z^\top$  above can be left out (see Remark 2.22). These spaces make perfect sense because the solution  $\tilde{u}$  must fulfill  $B \tilde{u} = 0$ , thus it must lie in  $\ker(S)^{\perp_{B^\top Q B}}$ . On the other hand, its flux must vanish on  $\ker(B)$ , i.e., the flux must lie in  $\ker(B)^\circ = \text{range}(B^\top)$ . During the algorithm, the flux iterates are indeed in equilibrium:

$$t^{(k)} - t^{(0)} \in \ker(B)^\circ.$$

## 2.2.4 Preconditioning

As announced in Sect. 2.2.1.5, all preconditioners for  $P^\top F|_{\tilde{U}_{\text{ad}}}$  under our consideration have the form

$$P M^{-1} : \tilde{U}_{\text{ad}}^* \rightarrow \tilde{U}_{\text{ad}},$$

where  $M^{-1} : U^* \rightarrow U$ .

### 2.2.4.1 The Dirichlet Preconditioner

The *Dirichlet preconditioner* proposed by Farhat, Mandel, and Roux [FMR94] is given by

$$M_D^{-1} = B S B^\top. \quad (2.51)$$

It was first analyzed by Mandel and Tezaur [MT96] who showed that in two dimensions, the condition number of FETI with the Dirichlet preconditioner and with  $Q = I$  is bounded by

$$C \max_{i=1}^s (1 + \log(H_i/h_i))^\beta,$$

where the constant  $C$  is independent of  $h_i$  (which is the local mesh size of subdomain  $\Omega_i$ ),  $H_i$ , and the number of subdomains. In general,  $\beta = 3$ . In some situations,  $\beta = 2$ , see [MT96, Lemma 3.8 and Remark 3.9]. With their pioneering article, Mandel and Tezaur paved the ground for all the refined FETI type analyses that appeared later. Note also that Tezaur [Tez98] showed that a method by Park, Justino, and Felippa [PJF97] is equivalent to the method in [FR91], see also [RFTM99].

Note that in general, the constant  $C$  above depends on the coefficient  $\mathcal{A}$  including possible jumps. Also, classical primal substructuring methods (see [BPS86] and [TW05, Chap. 5]) are known to have a condition number involving just two powers of the logarithmic term. To get rid of the third power in FETI and to address coefficient jumps, one has to use a scaling of the jump operator in the preconditioner.

### 2.2.4.2 The Scaled Dirichlet Preconditioner

The scaled Dirichlet preconditioner has its roots in the following works. Rixen and Farhat [RF98a] provided a derivation using an energy-minimizing smoothing procedure. Klawonn and Widlund [KW01] used the so-called *weighted counting functions*, which are a basic ingredient of balancing Neumann-Neumann methods [DL91, Man93, MB96] (see also Sect. 2.3) and of related methods, see [DSW94, DW95, DSW96, Sar93, Sar94, Sar97]. A special choice of scalings leads to the method in [RF98a]. Moreover, Klawonn and Widlund [KW01] gave a rigorous analysis (covering also the three-dimensional case), showing that FETI with the scaled Dirichlet preconditioner results in a condition number of

$$C \max_{i=1}^s (1 + \log(H_i/h_i))^2,$$

where the constant  $C$  is independent of the local mesh sizes  $h_i$ , the diameters  $H_i$ , the number of subdomains, and of jumps in the coefficients (provided that  $\mathcal{A}$  is isotropic and piecewise constant on each of the subdomains). We will present this analysis in detail in Sect. 2.6 below.

**Scalings and Weighted Counting Functions.** Firstly, we need scalar weights

$$\rho_i(x^h) > 0 \quad (2.52)$$

for each  $i = 1, \dots, s$  and for each node  $x^h \in \partial\Omega_i^h$ . We will discuss several choices below. Secondly, for each  $j \in \{1, \dots, s\}$  and  $x^h \in \Gamma_S^h$ , we define the weight

$$\delta_j^\dagger(x^h) := \begin{cases} \frac{\rho_j(x^h)^\gamma}{\sum_{k \in \mathcal{N}_{x^h}} \rho_k(x^h)^\gamma} & \text{for } x^h \in \partial\Omega_j^h, \\ 0 & \text{for } x^h \in \Gamma_S^h \setminus \partial\Omega_j^h, \end{cases} \quad (2.53)$$

where  $\mathcal{N}_{x^h} = \{i = 1, \dots, s : x^h \in \partial\Omega_i^h\}$ , cf. Definition 2.12, and  $\gamma \in [1/2, \infty)$  is a fixed exponent. A default choice for  $\gamma$  is one. The resulting piecewise linear functions  $\delta_j^\dagger \in V^h(\Gamma_S)$  for  $j = 1, \dots, s$ , are called *weighted counting functions*. The union of these functions forms a partition of unity on the skeleton, i.e.,

$$\sum_{j=1}^s \delta_j^\dagger(x^h) = 1 \quad \forall x^h \in \Gamma_S^h.$$

Note finally that  $\delta_j^\dagger(x^h) = 1$  for all  $x^h \in \partial\Omega_j \setminus \Gamma^h$ .

*Remark 2.27.* The limit case  $\gamma \rightarrow \infty$  corresponds to the choice

$$\delta_j^\dagger(x^h) := \begin{cases} 1/m & \text{if } x^h \in \partial\Omega_j^h \text{ and } \rho_j(x^h) = \max_{k \in \mathcal{N}_{x^h}} \rho_k(x^h), \\ 0 & \text{else,} \end{cases} \quad (2.54)$$

where  $m = \#\{k \in \mathcal{N}_{x^h} : \rho_k(x^h) = \max_{\ell \in \mathcal{N}_{x^h}} \rho_\ell(x^h)\}$  is the number of times the maximal coefficient is attained. For an early domain decomposition method with this scaling see [WX94].

The following scalings are common.

*Multiplicity Scaling.* If there is no (or only very little) variation in the coefficient  $\mathcal{A}$ , one usually chooses

$$\rho_i(x^h) = 1.$$

Then  $\delta_j^\dagger(x^h)$  simply equals the reciprocal of the multiplicity of the node  $x^h$ .

*Coefficient Scaling.* In case of coefficient variation, the weight  $\rho_i(x^h)$  should in a way resemble the coefficient  $\mathcal{A}$  in subdomain  $\Omega_i$  around  $x^h$ . If  $\mathcal{A}|_{\Omega_i} = \alpha_i I$ , we can set  $\rho_i(x^h) = \alpha_i$ . The case of a varying coefficient is treated in Chap. 3. If  $\mathcal{A}$  is isotropic and globally constant, coefficient and multiplicity scaling are equivalent.

**Stiffness Scaling.** For a pure FETI method,  $\rho_i(x^h)$  is often chosen as the diagonal entry of the subdomain stiffness matrix corresponding to the node  $x^h$ , see also Sects. 2.6.4 and 3.3.2.

Note that the coefficient scaling is frequently called  $\rho$ -scaling in the literature (where the scalar coefficient itself is denoted by  $\rho$ ), see e.g. [KRW08]. The stiffness scaling was first proposed by Rixen and Farhat [RF98a, RF99], there called *superlumped smoothing*. For more scalings see Sect. 2.6.4 and the recent paper [DW12b].

**A Weighted Jump Operator.** Recall from Definition 2.13 that

$$t_{i,x^h} = \langle t_i, \varphi_{i,x^h} \rangle_{\partial\Omega_i} \quad \text{for } t \in W,$$

where  $\varphi_{i,x^h} \in V^h(\partial\Omega_i)$  is the nodal basis function associated to node  $x^h$ . The weighted jump operator  $B_D : W^* \rightarrow U$  is given by

$$(B_D t)_{ij}(x^h) = \delta_j^\dagger(x^h) t_{i,x^h} - \delta_i^\dagger(x^h) t_{j,x^h} \quad \text{for } x^h \in \Gamma_{ij}^h, i > j. \quad (2.55)$$

In the all-floating formulation, we set

$$(B_D t)_{iD}(x^h) = t_{i,x^h} \quad \text{for } x^h \in \partial\Omega_i^h \cap \Gamma_D. \quad (2.56)$$

**Lemma 2.28.** *The adjoint  $B_D^\top : U^* \rightarrow W$  is given by*

$$(B_D^\top \mu)_i(x^h) = \begin{cases} \sum_{j \in \mathcal{N}_{x^h} \setminus \{i\}} \delta_j^\dagger(x^h) \text{sign}(i-j) \mu_{ij}(x^h) & \text{if } x^h \in \Gamma^h \cap \partial\Omega_i, \\ \mu_{iD}(x^h) & \text{if } x^h \in \partial\Omega_i^h \cap \Gamma_D \text{ (in the} \\ & \text{all-floating formulation),} \\ 0 & \text{else.} \end{cases}$$

*Proof.* The proof is straightforward.

**The Scaled Dirichlet Preconditioner** is finally given by

$$M_{\text{SD}}^{-1} = B_D S B_D^\top. \quad (2.57)$$

**Remark 2.29.** Each block  $S_i$  in  $S$  appearing in the preconditioner (2.57) may be replaced by the hypersingular operator on  $\partial\Omega_i$  (then called *scaled hypersingular BETI preconditioner*, cf. [LS03, LS05]), or by any other operator that is spectrally equivalent to  $S_i$ .

**Remark 2.30.** For non-redundant Lagrange multipliers, the preconditioner takes the same form but with

$$B_D = (B D^{-1} B^\top)^{-1} B D^{-1},$$

where  $D : W^* \rightarrow W$  is a diagonal scaling operator where the entry corresponding to node  $x^h$  in subdomain  $\Omega_i$  is  $\delta_i^\dagger(x^h)$ , cf. [KW01] and [TW05, Sect. 6.3.2]. For an efficient algorithmic construction of  $B_D$  see also [Of06, Sect. 5.5.2].

### 2.2.4.3 Lumped Preconditioners

For a pure FETI method, “lumped” preconditioners (see [FR91, Sect. 5] and [FMR94, Sect. 6]) are constructed by replacing the FE Schur complements  $S_i$  in (2.51) or (2.57) by the block  $\mathbf{K}_{i,BB}$  of the stiffness matrix corresponding to the boundary/interface unknowns. In other words, instead of solving for the PDE-harmonic extension in each subdomain, we simply extend by zero.

The application of a lumped preconditioner is more economic because one does not need the factorizations and solves appearing in the operators  $S_i$ . In some situations, the overall CPU time of a FETI method with a lumped preconditioner can be shorter than with a Dirichlet preconditioner, even though the condition number of the preconditioned system might be larger. This is also due to a superconvergence effect, cf. [FMR94, Sect. 7].

*Remark 2.31.* The hypersingular BETI preconditioner could be seen as a BEM equivalent of the lumped preconditioner, because there is no additional solving involved. However, this operator has a corresponding dense matrix and thus needs an effective implementation. Note also that the hypersingular preconditioner is still quasi-optimal.

For the remainder of this book, we will only treat the scaled Dirichlet preconditioner, because it turns out to be quasi-optimal with respect to the condition number.

### 2.2.4.4 The Operator $Q$

If there is no (or only little) coefficient variation, one commonly chooses

$$Q = I,$$

which is supported by the analysis, see Remark 2.111. In case of large coefficient variation, one often chooses

$$Q = M_{\text{SD}}^{-1}.$$

where  $M_{\text{SD}}^{-1}$  is the scaled Dirichlet preconditioner, cf. [BDF<sup>+</sup>00]. Another practically successful option is to set  $Q$  to the lumped Dirichlet preconditioner. These choices originate from [FR94] and were further investigated in [Rix97]. There is also a diagonal choice

$$Q = Q_{\text{diag}}$$

(first suggested and analyzed in [KW01]) which mimics the choice  $Q = M_{\text{SD}}^{-1}$  under certain regularity assumptions and which simplifies the implementation, see

Sects. 2.6.3 and 3.3.5.4. Note all these different choices of  $Q$  lead to a sparse coarse matrix  $\mathbf{G}^\top \mathbf{Q} \mathbf{G}$ , but to different sparsity patterns, cf. [FR94].

### 2.2.5 Implementation Issues

In this subsection, we discuss implementation issues of the classical and the all-floating FETI/BETI methods described above (cf. also [LP98, RF98b, Kam00]). In particular, we address the issue of parallelization.

#### 2.2.5.1 The Basic Input Data

To make an implementation of FETI/BETI possible, one needs the vector representations for the local spaces  $V^h(\Omega_i)$  or  $V^h(\Gamma_i)$  with respect to the standard nodal basis ( $V_D^h(\Omega_i)$  or  $V_D^h(\Gamma_i)$  in the classical formulation) and a local numbering of the degrees of freedom (DOFs). Correspondingly, for each FEM subdomain we need the local stiffness matrix (for a floating subdomain the “Neumann” matrix) and the load vector associated to each of these spaces. For each BEM subdomain we need (data-sparse) matrix approximations of the boundary integral operators (for matrix-free methods see also Sect. 2.2.6). For the classical formulation, we assume that the system is homogenized, i.e., the prescribed Dirichlet values are already contained in the load vector, and there are either no Dirichlet DOFs, or these are decoupled from the remaining DOFs.

#### 2.2.5.2 Interconnecting

In order to do interconnecting, each DOF on the interface must have a local and a global index. With this information, one can set up the Lagrange multipliers on the interface: each multiplier (numbered by an index) is described by two subdomain indices and two local DOF indices, cf. (2.22). The global index is used for identification only. In the all-floating formulation, one additionally needs to know which DOFs are on the Dirichlet boundary in order to set up the additional multipliers there, cf. (2.47).

If the subdomain decomposition is generated from a mesh partitioner such as METIS [KK98], the local to global mappings come together with the output (at least for scalar  $P^1$ -elements).

#### 2.2.5.3 Additional Input Data

For the scaled Dirichlet preconditioner, one needs the scalings  $\rho_i(x^h)$  corresponding to each interface DOF. For the coefficient scaling (cf. Sect. 2.2.4.2), the coefficient is



required, whereas for the stiffness scaling and the multiplicity scaling no additional information is needed. See however Sect. 2.6.4.2 for undesired effects that can occur with the stiffness scaling and/or with varying coefficients.

Most importantly, one needs a description of the local kernels, i.e., one needs matrices  $\mathbf{R}_i$  corresponding to  $R_i$  from (2.30). If  $i \in \mathcal{J}_{\text{FEM}}$ , it is advantageous to set up  $\mathbf{R}_i$  such that the columns of  $\mathbf{R}_i$  span the kernel of the local stiffness matrix  $\mathbf{K}_i$  (rather than that of the Schur complement  $\mathbf{S}_i$  which should never be formed).

If the kernel is not known a priori, one can run a singular value decomposition or try to get low-frequent eigenpairs of  $\mathbf{K}_i$  by an inverse power method (see e.g. [GV96]) or the LOBPCG method [Kny01], but this is in general expensive. Here lies a true advantage of the all-floating formulation: for the most widely used types of PDEs, especially for the potential equation or linear elasticity (see also Sect. 2.8), the kernel is known a priori in all the subdomains and the matrix  $\mathbf{R}_i$  can be given explicitly.

For the diagonal choice  $Q = Q_{\text{diag}}$  due to Klawonn and Widlund, which we will expose in Sect. 2.6.3, additional mesh information ( $H_i$  and  $h_i$ ) is required. Furthermore, one needs to know whether a coupling DOF is associated to a subdomain vertex, or if it is in the interior of a subdomain edge or face. However, this can be figured out from the Lagrange multipliers by combinatorial means (cf. Sect. 2.48).

### 2.2.5.4 Implementation of the Underlying Operators

Having  $R$  and  $R^\top$  at our disposal, and assuming that  $Q \in \{I, M_{\text{SD}}^{-1}, Q_{\text{diag}}\}$ , we can reduce Algorithms 2 and 3 to the applications of  $B, B^\top, B_D, B_D^\top, S, S^\top$ , and  $(G^\top Q G)^{-1}$ , which are discussed in the sequel.

**Jump Operators.** The operators  $B, B^\top, B_D$  and  $B_D^\top$  need not be stored but are encoded as routines which perform their application to vectors. These routines mainly use the description of the Lagrange multipliers and the scalings  $\rho_i(x^h)$ .

**Local FEM Neumann Problems.** For each  $i \in \mathcal{J}_{\text{FEM}}$ , the action  $v_i = S_i^\dagger f_i$  for a given  $f_i \in \text{range}(S_i)$  is performed as follows. Let  $\mathbf{v}_B$  and  $\mathbf{f}_B$  denote the vector representations of  $v_i$  and  $f_i$ , respectively. Then, with the analogous notation as in Sect. 1.2.6, the equation  $\mathbf{S}_i \mathbf{v}_B = \mathbf{f}_B$  is equivalent to

$$\begin{bmatrix} \mathbf{K}_{i,BB} & \mathbf{K}_{i,BI} \\ \mathbf{K}_{i,IB} & \mathbf{K}_{i,II} \end{bmatrix} \begin{bmatrix} \mathbf{v}_B \\ \mathbf{v}_I \end{bmatrix} = \begin{bmatrix} \mathbf{f}_B \\ 0 \end{bmatrix},$$

i.e., we need to solve a local Neumann problem, cf. Sect. 1.2.5. The most convenient way is to regularize the matrix  $\mathbf{K}_i$  (if  $\Omega_i$  is floating) and store its factorization. For the setting (2.19) (see also Remark 2.5, p. 68), the set “ $I$ ” above includes interior DOFs as well as non-coupling Neumann DOFs.

**Local BEM Neumann Problems.** For each  $i \in \mathcal{J}_{\text{BEM}}$ , the action  $v_i = S_i^\dagger f_i$  for a given  $f_i \in \text{range}(S_i)$  is performed as follows. Let  $\mathbf{v}$  and  $\mathbf{f}$  denote the

vector representations of  $v_i$  and  $f_i$ , respectively. Let  $\mathbf{D}_i$ ,  $\mathbf{K}_i$ , and  $\mathbf{V}_i$  denote the matrix representations of the hypersingular operator, the double layer potential, and the single layer potential, respectively, and let  $\mathbf{M}_i$  be the mass matrix from Sect. 1.3.7. Then instead of solving  $\mathbf{S}_i \mathbf{v} = \mathbf{f}$  with  $\mathbf{S}_i = \mathbf{D}_i + (\frac{1}{2}\mathbf{M}_i^\top + \mathbf{K}_i^\top)\mathbf{V}_i^{-1}(\frac{1}{2}\mathbf{M}_i + \mathbf{K}_i)$ , we solve the equivalent saddle point problem

$$\begin{bmatrix} \mathbf{D}_i & \frac{1}{2}\mathbf{M}_i^\top + \mathbf{K}_i^\top \\ \frac{1}{2}\mathbf{M}_i + \mathbf{K}_i & -\mathbf{V}_i \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{t} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix},$$

which is a standard BEM problem corresponding to the local Neumann problem. Recall that for  $d = 2$ , we should ensure that  $\text{diam}(\Omega_i) < 1$  (e.g. by scaling all the coordinates of  $\Omega$ ) in order to ensure the invertibility of  $\mathbf{V}_i$ .

Again, if  $\Omega_i$  is a floating subdomain, we can regularize this problem by regularizing the hypersingular operator  $\mathbf{D}_i$  similar as in Sect. 1.2.5. As briefly described in Sect. 1.3.7, the matrices  $\mathbf{D}_i$ ,  $\mathbf{K}_i$ , and  $\mathbf{V}_i$  can be approximated in data-sparse form using  $\mathcal{H}$ -matrices. Thus, also the matrix corresponding to the above (possibly regularized) saddle point problem can be represented by an  $\mathcal{H}$ -matrix, and for each BEM subdomain its  $\mathcal{H}$ -LU factorization can be built and stored in quasi-optimal time and memory complexity in the preprocessing phase. For matrix-free fast BEM, such as the fast multipole methods, in connection with FETI/BETI methods see Sect. 2.2.6.

**Local Dirichlet Problems.** The action of  $S_i$  for  $i \in \mathcal{I}_{\text{FEM}}$ , corresponds essentially to solving a system of the form

$$\mathbf{K}_{i,II} \mathbf{v}_I = -\mathbf{K}_{i,IB} \mathbf{v}_B.$$

For the input  $\mathbf{v}_B$ , the output is given by  $\mathbf{S}_i \mathbf{v}_B = \mathbf{K}_{i,BB} \mathbf{v}_B + \mathbf{K}_{i,BI} \mathbf{v}_I$ . Hence, it is most convenient to build and store a factorization of  $\mathbf{K}_{i,II}$  in the preprocessing phase. Again, for the setting (2.19) (see also Remark 2.5, p. 68), the set “ $I$ ” above includes interior DOFs as well as non-coupling Neumann DOFs. For  $i \in \mathcal{I}_{\text{BEM}}$ , the action of  $S_i$  (see Sect. 1.3.8.1) involves the inverse of the single layer potential, which can be realized by  $\mathcal{H}$ -LU factorization.

**The Coarse Problem.** Let  $\mathbf{G}$  and  $\mathbf{Q}$  denote the matrix representations of  $G$  and  $Q$ . As discussed in Remark 2.108 below, the matrix  $\mathbf{G}^\top \mathbf{Q} \mathbf{G}$  is sparse, and its sparsity pattern is determined by the connectivity graph of the subdomain partition where each floating subdomain is a node of that graph. Once  $\mathbf{G}^\top \mathbf{Q} \mathbf{G}$  is assembled (cf. [RF98b, Sect. 3.2]), its factorization can be computed efficiently during the preprocessing phase, as long as the number of subdomains is not very large. If we set  $\mathbf{Q} = M_{\text{SD}}^{-1}$ , an efficient assembly of  $\mathbf{G}^\top \mathbf{Q} \mathbf{G}$  is possible but tricky (see Remark 2.36). Note that the extra cost of applying  $\mathbf{Q}$  during the FETI/BETI algorithm (see Algorithm 2 and (2.38)) involves the solution of additional local Dirichlet problems. However, within each step of the iterative solver, this is compensated by the fact that we can leave out the action of  $P$  (cf. Remark 2.22). Nevertheless, from the implementation point of view, it is more attractive to use a diagonal choice of  $\mathbf{Q}$  if the context allows to do so; see Sects. 2.6.3 and 3.3.5.4 below.

*Remark 2.32.* We warn the reader that if the coefficient  $\alpha$  varies extremely within a single FEM subdomain (see also Chap. 3), the matrices  $\mathbf{K}_{i,II}$  and  $\mathbf{K}_i$  (the latter possibly regularized) can become very ill-conditioned. For example if the coefficient varies of between 1 and  $10^9$ , even direct solvers may run into stability problems. The same can happen with the coarse matrix  $\mathbf{G}^\top \mathbf{Q} \mathbf{G}$  if the coefficient varies a lot throughout the global domain  $\Omega$ .

### 2.2.5.5 Parallelization

The FETI/BETI algorithm is very suitable for multiprocessor machines with shared and especially with distributed memory. The coding should follow the MIMD (multiple instruction multiple data) paradigm, cf. [DHL03, Haa99, SBG96]. In the following, let the processors be numbered from 1 to  $p$ . For software supporting parallelization (on different levels), we refer e.g. to the following frameworks.

- MPI (message-passing interface) standard [MPI09]
- PETSc: <http://www-unix.mcs.anl.gov/petsc/petsc-as/>
- HyPre: <http://acts.nersc.gov/hypre/>
- DUNE: <http://www.dune-project.org/>
- Parallel toolbox: <http://paralleltoolbox.sourceforge.net/>

For further literature see also [Bas96, SBG96, Zum03].

Since the main work are subdomain solves that are independent of each other, we assign each subdomain  $\Omega_i$  to a processor  $p_i \in \{1, \dots, p\}$ . Hence, each processor handles one or several subdomains. There are two kinds of global objects involved in the algorithm that need to be parallelized:

- (i) Lagrange multipliers, i.e., vectors from  $U, U^*$ ,
- (ii) Coarse vectors representing elements from  $Z, Z^*$ .

The remaining variables are elements from the spaces  $W = \prod_{i=1}^s W_i$  and  $W^*$  which are parallel by construction.

We will first show how to deal with parallelizing the Lagrange multipliers (kind (i)) using the concept of accumulated and distributed vectors, see e.g. [DHL03, Haa99].

#### Accumulated and Distributed Vectors

**Definition 2.33.** For  $\lambda \in \mathbb{R}^M$  and  $q = 1, \dots, p$ , we denote by  $\lambda_{\text{acc}}^{(q)} \in \mathbb{R}^{M_q}$  the vector of those entries  $\lambda_{ij}(x^h)$  where

$$p_i = q \text{ or } p_j = q \quad \text{and} \quad x^h \in \Gamma_{ij}^h,$$

i.e., the entries *shared* by processor  $q$ . As for  $\lambda$  itself, the local numbering within  $\lambda_{\text{acc}}^{(q)}$  can be arbitrary. To be general, we denote the entries of  $\lambda_{\text{acc}}^{(q)}$  again by  $\lambda_{\text{acc},ij}^{(q)}(x^h)$ . For the all-floating formulation, the vector  $\lambda_{\text{acc}}^{(q)}$  additionally contains those entries  $\lambda_{iD}(x^h)$  where  $p_i = q$  and  $x^h \in \partial\Omega_i^h \cap \Gamma_D^h$ . The parallel vector  $\lambda_{\text{acc}} := [\lambda_{\text{acc}}^{(q)}]_{q=1}^p$  is called *accumulated* realization of  $\lambda$ . It has the property that  $\lambda_{\text{acc},ij}^{(q)}(x^h) = \lambda_{ij}(x^h)$  for all processors  $q$ .

**Definition 2.34.** A parallel vector  $\lambda_{\text{dist}} = [\lambda_{\text{dist}}^{(q)}]_{q=1}^p$  with  $\lambda_{\text{dist}}^{(q)} \in \mathbb{R}^{M_q}$  is called *distributed* realization of  $\lambda \in \mathbb{R}^M$ , if

$$\lambda_{ij}(x^h) = \sum_{q \in \{p_i, p_j\}} \lambda_{\text{dist},ij}^{(q)}(x^h).$$

In the implementation, we will use both representations, but the vectors  $\lambda_{\text{acc}}, \lambda_{\text{dist}}$  are never formed, but only their local components  $\lambda_{\text{acc}}^{(q)}, \lambda_{\text{dist}}^{(q)}$  on each processor  $q$ .

It is immediate, that for scalars  $\alpha, \beta \in \mathbb{R}$  and vectors  $\lambda, \mu \in U$  or  $U^*$ , the accumulated (or distributed) representation of  $\alpha\lambda + \beta\mu$  is simply  $\alpha\lambda_{\text{acc}} + \beta\mu_{\text{acc}}$  (or  $\alpha\lambda_{\text{dist}} + \beta\mu_{\text{dist}}$ , respectively), i.e., these operations can be performed fully in parallel.

The next lemma shows that computing the scalar product between a distributed and an accumulated vector can be performed by computing the local scalar products and then just communicate the (scalar) results between all processors and add them up. In the MPI standard, this is done by the `allreduce` command, see [MPI09].

**Lemma 2.35.** For  $\lambda \in U$  and  $\mu \in U^*$ ,

$$\langle \mu, \lambda \rangle = (\mu, \lambda)_{\ell^2} = \sum_{q=1}^p (\mu_{\text{dist}}^{(q)}, \lambda_{\text{acc}}^{(q)})_{\ell^2}.$$

*Proof.* The proof is straightforward.

With a slight abuse of notation, we will write  $\langle \mu_{\text{dist}}, \lambda_{\text{acc}} \rangle = \langle \mu, \lambda \rangle$  in the sequel, which indicates that the calculation of the inner product is based on the result of Lemma 2.35.

Let  $U_{\text{acc}}, U_{\text{dist}}$  and  $U_{\text{acc}}^*, U_{\text{dist}}^*$  denote the (formal) spaces of accumulated and distributed vectors representing elements in  $U, U^*$ , respectively. Let  $\mathbf{A}$  denote the *accumulation operator* such that for  $\lambda \in U$  or in  $U^*$ ,  $\mathbf{A}\lambda_{\text{dist}} = \lambda_{\text{acc}}$ . Moreover, let  $\mathbf{D}$  denote the *distribution operator* under whose action entries will be distributed by simply dividing them by their multiplicity. We have  $\mathbf{A}\mathbf{D} = \mathbf{I}$  but in general  $\mathbf{D}\mathbf{A} \neq \mathbf{I}$ . Note also that  $\mathbf{D} \neq \mathbf{A}^\top$ .

When we want to compute the (global) Euclidean norm of a vector  $\lambda$  from its accumulated representation  $\lambda_{\text{acc}}$ , we can calculate  $\sqrt{(\mathbf{D}\lambda_{\text{acc}}, \lambda_{\text{acc}})_{\ell^2}}$ , without major communication. If only  $\lambda_{\text{dist}}$  is available, we can calculate  $\sqrt{(\lambda_{\text{dist}}, \mathbf{A}\lambda_{\text{dist}})_{\ell^2}}$ , which involves more communication due to the accumulation operator.

### Setting Up Parallel Lagrange Multipliers

The most convenient starting point is a setting where each local DOF (in a subdomain) is associated to a global DOF. This can be easily achieved for continuous  $P^1$  elements, if one has global indices for each vertex of the mesh (for high order elements, one additionally needs global indices for the edges/facets of the mesh). For simplicity, we continue with the  $P^1$  elements for the scalar elliptic equation, but the concept can be generalized straightforwardly.

For each DOF, it should be known which subdomains are formally sharing it. If this information is not known a priori, it can be got from the associated global indices by sorting and communicating (see, e.g., [Lie06]). The communication can be reduced if the next neighbors of each processor are known.

Each multiplier  $\lambda_{ij}(x^h)$  is identified by the triple  $(i, j, g_{x^h})$  where  $g_{x^h}$  is the global index of node  $x^h$  and  $i > j$ . When it comes to sending/receiving entries to/from another processor, we order the corresponding entries with respect to the global index triple. This ensures that the entries are sent/received in the correct order.

### Parallelization of $Z$ and $Z^*$

Since each entry of a vector  $\xi \in Z$  or in  $Z^*$  corresponds to a *unique* subdomain, parallelization is easier. For  $\xi \in Z$  and  $q = 1, \dots, p$ , let  $\xi_{\text{loc}}^{(q)}$  be the local vector of entries  $\xi_i$  such that  $p_i = q$ , and let  $\xi_{\text{loc}} := [\xi_{\text{loc}}^{(q)}]_{q=1}^p$  be the corresponding parallel vector. Moreover, let  $Z_{\text{loc}}, Z_{\text{loc}}^*$  be the spaces of such parallel vectors and let

$$\mathbf{A} : Z_{\text{loc}}^* \rightarrow Z^*, \quad \mathbf{D} : Z \rightarrow Z_{\text{loc}}$$

denote the accumulation and distribution operator, respectively. Opposed to the global vectors of Lagrange multipliers, here the operator  $\mathbf{A}$  creates a global vector *on each processor* by collecting the vectors from all processors. Note that the size of this global vector is of the same order as the number of subdomains. For the distribution operator, we have the identity  $\mathbf{D} = \mathbf{A}^\top$ .

### Parallelization of the FETI/BETI Operators

We will now replace the operators occurring in Algorithm 4 by operators involving the parallel spaces defined above. The following operators arise naturally.

If  $\mathbf{M}^{-1} = I$ , we have to set  $\mathbf{M}^{-1} := \mathbf{A}$  as an operator mapping  $U_{\text{dist}}^*$  to  $U_{\text{acc}}$ .

The implementation of  $\mathbf{Q} : U_{\text{dist}}^* \rightarrow U_{\text{acc}}$  depends on the choice of  $Q$ :

- If  $Q = I$  or  $Q = Q_{\text{diag}}$ , we first implement local diagonal operators  $\mathbf{Q}^{(q)}$ . Applying them in parallel leads to the operator  $\mathbf{Q}_{\text{loc}} : U_{\text{dist}}^* \rightarrow U_{\text{dist}}$ . Finally, we set  $\mathbf{Q} := \mathbf{A} \mathbf{Q}_{\text{loc}}$ .
- If  $Q = M_{\text{SD}}^{-1}$ , we set  $\mathbf{Q} := \mathbf{M}_{\text{SD}}^{-1}$ .

Operator		Adjoint	
$\mathbf{B}$	$: W \rightarrow U_{\text{dist}}^*$	$\mathbf{B}^\top$	$: U_{\text{acc}} \rightarrow W^*$
$\mathbf{B}_D$	$: W^* \rightarrow U_{\text{dist}}$	$\mathbf{B}_D^\top$	$: U_{\text{acc}}^* \rightarrow W$
$\mathbf{R}$	$: Z_{\text{loc}} \rightarrow W$	$\mathbf{R}^\top$	$: W^* \rightarrow Z_{\text{loc}}^*$
$\mathbf{G} := \mathbf{B} \mathbf{R} \mathbf{D}$	$: Z \rightarrow U_{\text{dist}}^*$	$\mathbf{G}^\top := \mathbf{A} \mathbf{R}^\top \mathbf{B}^\top$	$: U_{\text{acc}} \rightarrow Z^*$
$\mathbf{F} := \mathbf{B} \mathbf{S}^\dagger \mathbf{B}^\top$	$: U_{\text{acc}} \rightarrow U_{\text{dist}}^*$		–
$\mathbf{M}_{\text{sD}}^{-1} := \mathbf{A} \mathbf{B}_D \mathbf{S} \mathbf{B}_D^\top \mathbf{A}$	$: U_{\text{dist}}^* \rightarrow U_{\text{acc}}$		–

The parallel operators corresponding to the projections  $\mathbf{P}$ ,  $\mathbf{P}^\top$  are as follows:

$$\begin{aligned} \mathbf{P} &:= \mathbf{I} - \mathbf{Q} \mathbf{G} (\mathbf{G}^\top \mathbf{Q} \mathbf{G})^{-1} \mathbf{G}^\top : U_{\text{acc}} \rightarrow U_{\text{acc}}, \\ \mathbf{P}^\top &:= \mathbf{I} - \mathbf{G} (\mathbf{G}^\top \mathbf{Q} \mathbf{G})^{-1} \mathbf{G}^\top \mathbf{Q} : U_{\text{dist}}^* \rightarrow U_{\text{dist}}^*. \end{aligned}$$

Note that the matrix corresponding to the coarse operator  $(\mathbf{G}^\top \mathbf{Q} \mathbf{G}) : Z \rightarrow Z^*$  can be assembled and made available on each processor, and its factorization can be built and stored. Recall from Remark 2.22 that if  $\mathbf{Q} = \mathbf{M}_{\text{sD}}^{-1}$ , we can leave out  $\mathbf{P}$ .

*Remark 2.36.* A parallelly efficient assembly of  $\mathbf{G}^\top \mathbf{Q} \mathbf{G}$  for the case  $\mathbf{Q} = \mathbf{M}_{\text{sD}}^{-1}$  is tricky but possible. One possible way of implementation is described in [FR94], see also [BDF<sup>+</sup>00] for a parallel speed-up of the factorization. Another recipe is the following. In Sect. 2.4.2.1, we will see that there is an averaging operator  $\mathbf{E}_D : W \rightarrow \widehat{W}$  such that  $\mathbf{G}^\top \mathbf{M}_{\text{sD}}^{-1} \mathbf{G} = \mathbf{R}^\top (\mathbf{I} - \mathbf{E}_D)^\top \mathbf{S} (\mathbf{I} - \mathbf{E}_D) \mathbf{R} = \mathbf{R}^\top \mathbf{E}_D^\top \mathbf{S} \mathbf{E}_D \mathbf{R}$ , where the latter identity holds true because  $\mathbf{S} \mathbf{R} = 0$ . From this formula, we see that averaging each kernel function (by  $\mathbf{E}_D$ ) and employing suitable neighbor communication, we can figure out the corresponding row of the matrix  $\mathbf{G}^\top \mathbf{M}_{\text{sD}}^{-1} \mathbf{G}$ .

The final FETI/BETI algorithm (including PCG) is displayed in Algorithm 6.

## 2.2.6 Inexact FETI/BETI Methods

Recall that the action of  $\mathbf{S}_i^\dagger$  requires to solve a system on subdomain  $\Omega_i$ . For a FEM subdomain, the system matrix is the (regularized) stiffness matrix, for a BEM subdomain, the system matrix has the form

$$\begin{bmatrix} \widetilde{\mathbf{D}}_i & \frac{1}{2} \mathbf{M}_i^\top + \mathbf{K}_i^\top \\ \frac{1}{2} \mathbf{M}_i + \mathbf{K}_i & -\mathbf{V}_i \end{bmatrix},$$

where  $\widetilde{\mathbf{D}}_i$  is the matrix corresponding to the hypersingular operator, regularized if  $\Omega_i$  is floating, see Sect. 2.2.5.4. We already mentioned that this matrix can be factorized using  $\mathcal{H}$ -LU (or Cholesky) factorization. If one uses the fast multipole method, only the fast *application* of the above matrix to a vector is available. In [LOSZ07b, LOSZ07a], (see also [KW00] for the original idea in a FETI setting), an

**Algorithm 6:** Parallel FETI/BETI method including PCG

---


$$\begin{aligned}
g &= \begin{cases} [N_i f_i - S_i \tilde{g}_D|_{\partial\Omega_i}]_{i=1}^s & \text{in classical formulation} \\ [N_i f_i]_{i=1}^s & \text{in all-floating formulation} \end{cases} \\
\lambda_{\text{acc}}^{(0)} &= \mathbf{Q} \mathbf{G} (\mathbf{G}^\top \mathbf{Q} \mathbf{G})^{-1} \mathbf{A} \mathbf{R}^\top g \\
d_{\text{dist}}^{(0)} &= \mathbf{P}^\top \mathbf{B} (S^\dagger (g - \mathbf{B}^\top \lambda_{\text{dist}}^{(0)}) - \tilde{g}_D) \\
k &= 0 \\
\text{repeat} \\
&\quad \begin{cases} z_{\text{acc}}^{(k)} = \mathbf{P} \mathbf{M}^{-1} d_{\text{dist}}^{(k)} & \text{(in the unpreconditioned case: } z_{\text{acc}}^{(k)} = \mathbf{A} d_{\text{dist}}^{(k)}) \\ q_{\text{acc}}^{(k)} = z_{\text{acc}}^{(k)} + \beta_{k-1} q_{\text{acc}}^{(k-1)} & \text{where } \beta_{-1} = 0, \\ \beta_{k-1} = \frac{\langle d_{\text{dist}}^{(k)}, z_{\text{acc}}^{(k)} \rangle}{\langle d_{\text{dist}}^{(k-1)}, z_{\text{acc}}^{(k-1)} \rangle} \\ \lambda_{\text{acc}}^{(k+1)} = \lambda_{\text{acc}}^{(k)} + \alpha_k q_{\text{acc}}^{(k)} & \text{where } \alpha_k = \frac{\langle d_{\text{dist}}^{(k)}, z_{\text{acc}}^{(k)} \rangle}{\langle \mathbf{P}^\top \mathbf{F} q_{\text{acc}}^{(k)}, q_{\text{acc}}^{(k)} \rangle} \\ d_{\text{dist}}^{(k+1)} = d_{\text{dist}}^{(k)} - \alpha_k \mathbf{P}^\top \mathbf{F} q_{\text{acc}}^{(k)} & = \mathbf{B}(u^{(k)} - \tilde{g}_D) \\ k = k + 1 \end{cases} \\
\text{until stopping criterion fulfilled for } d_{\text{dist}}^{(k)} \\
u^{(k)} &= \tilde{g}_D + (\mathbf{I} - \mathbf{R} \mathbf{D} (\mathbf{G}^\top \mathbf{Q} \mathbf{G})^{-1} \mathbf{G}^\top \mathbf{Q} \mathbf{B}) (S^\dagger (g - \mathbf{B}^\top \lambda_{\text{acc}}^{(k)}) - \tilde{g}_D)
\end{aligned}$$


---

inexact BETI method has been developed which circumvents this factorization. The BETI system is rewritten as the threefold saddle point system

$$\begin{bmatrix} -V & \frac{1}{2}I + K & 0 \\ \frac{1}{2}I + K^\top & D & B^\top \\ 0 & B & \end{bmatrix} \begin{bmatrix} t \\ u \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ f \\ 0 \end{bmatrix}, \quad (2.58)$$

where  $V = \text{diag}(V_i)$ ,  $D = \text{diag}(D_i)$ , and  $K = \text{diag}(K_i)$ , see also Remark 2.18. The authors use and extend the classical theory of saddle point preconditioners, see [BP88, Zul02, Zul11]. Their full preconditioner is based on the scaled hypersingular BETI preconditioner ( $M_{\text{SD}}^{-1}$  with  $S_i$  replaced by  $D_i$ ) and further preconditioners for the local single layer potential operators  $V_i$ . For the latter, several choices have been proposed in the literature, see e.g. [FS97, Ste03a]. The same technique applies to hybrid FETI/BETI methods, as outlined in [LP08].

## 2.3 Balancing Neumann-Neumann Methods

In this section, we define the balancing Neumann-Neumann method. Its close connection to FETI will be subject of Sect. 2.4.2.4. The Neumann-Neumann methods were first developed by Bourgat, Glowinski, Le Tallec, and Vidrascu, as

well as De Roeck, see [BGLV89, De 91, DL91]. They were considerably improved by adding a second coarse level, see Mandel and Brezina [Man93, MB96], Le Tallec [Le 94], as well as Dryja and Widlund [DW95]. See also [TW05, Sect. 6.2] for more details on the method. For an alternative with a different coarse space see [BS02, Sect. 7.7]. The connection between balancing Neumann-Neumann and FETI methods was demonstrated in [KW01], and supported with an underpinning theory in [SM08]. Note also that the Neumann-Neumann methods have been further developed by Dohrmann and Mandel, leading to the BDDC methods, see Sect. 5.1.4.

Assume for simplicity that  $g_D = 0$  and let  $W, \widehat{W}, S : W \rightarrow W^*$ , and  $g \in W^*$  be defined either as in the classical or as in the all-floating formulation. Let  $V_D^h(\Gamma_S)$  be as in (2.12) and recall that  $V_D^h(\Gamma_S) \equiv \widehat{W}$ . Let  $A^\top : V_D^h(\Gamma_S) \rightarrow W$  denote the natural embedding and  $A : W^* \rightarrow V_D^h(\Gamma_S)^*$  its adjoint (the latter can be seen as an assembling operator). Furthermore, we define  $\widehat{S} := A S A^\top$  and  $\widehat{g} := A g$ . The equation

$$\text{find } u \in V_D^h(\Gamma_S) : \quad \widehat{S} u = \widehat{g} \quad (2.59)$$

is then identical to the skeleton formulation (2.11).

The balancing Neumann-Neumann preconditioner can be classified as a symmetric, hybrid two-level Schwarz preconditioner for  $\widehat{S}$  (cf. [TW05, Sect. 2.2]; here *hybrid* indicates the combination of Schwarz projectors in both an additive and multiplicative way). First, we define the “coarse” level. Let the averaging operator  $\widehat{E}_D : W \rightarrow V_D^h(\Gamma_S)$  be defined by

$$(\widehat{E}_D)(x^h) := \begin{cases} \sum_{j \in \mathcal{N}_{x^h}} \delta_j^\dagger(x^h) w_j(x^h) & \text{for } x^h \in \partial\Omega_i^h \setminus \Gamma_D, \\ 0 & \text{for } x^h \in \partial\Omega_i^h \cap \Gamma_D, \end{cases} \quad (2.60)$$

where  $\delta_j^\dagger$  are the weighted counting functions from (2.53) (for a fixed choice of scalings  $\rho_j$ ). Let  $r^{(i)} \in W$  be such that  $\text{span}\{r_i^{(i)}\} = \ker(S_i)$  and  $r_j^{(i)} = 0$  for all  $j \neq i$ . If  $\Omega_i$  is non-floating, then  $r^{(i)} = 0$ . Moreover, we define the subspace

$$\widehat{W}_0 := \text{span}\{\widehat{E}_D r^{(i)} : i \in \mathcal{I}_{\text{float}}\} \subset V_D^h(\Gamma_S),$$

with the basis  $\{\widehat{E}_D r^{(i)}\}_{i \in \mathcal{I}_{\text{float}}}$ . Let  $\widehat{R}_0^\top : \widehat{W}_0 \rightarrow V_D^h(\Gamma_S)$  denote the natural embedding and  $\widehat{R}_0$  its adjoint ( $\widehat{R}_0$  is a restriction and  $\widehat{R}_0^\top$  the corresponding prolongation). Let  $\widehat{S}_0 := \widehat{R}_0^\top \widehat{S} \widehat{R}_0$  be the corresponding projection of  $\widehat{S}$  (which is still SPD). Note that the matrix  $\widehat{\mathbf{S}}_0$  corresponding to  $\widehat{S}_0$  with respect to the basis  $\{\widehat{E}_D r^{(i)}\}_{i \in \mathcal{I}_{\text{float}}}$  is sparse. Moreover, it can be shown that for the choice  $Q = M_{\text{SD}}^{-1}$ , the matrix  $\widehat{\mathbf{S}}_0$  is identical to  $\mathbf{G}^\top \mathbf{Q} \mathbf{G}$  from the FETI method, cf. Remark 2.36.

Let the Schwarz projector  $\widehat{P}_0 : V_D^h(\Gamma_S) \rightarrow \widehat{W}_0$  be defined by

$$\widehat{P}_0 := \widehat{R}_0^\top \widehat{S}_0^{-1} \widehat{R}_0 \widehat{S}, \quad (2.61)$$



see also Sect. 1.1.3.3. Finally, the balancing Neumann-Neumann preconditioner  $M_{\text{BNN}}^{-1} : V_D^h(\Gamma_S)^* \rightarrow V_D^h(\Gamma_S)$  is given by

$$M_{\text{BNN}}^{-1} := (I - \widehat{P}_0) \widehat{E}_D S^\dagger \widehat{E}_D^\top (I - \widehat{P}_0^\top) + \widehat{R}_0^\top \widehat{S}_0^{-1} \widehat{R}_0. \quad (2.62)$$

Note that the operator  $\widehat{E}_D^\top : V_D^h(\Gamma_S)^* \rightarrow W^*$  distributes (global) residuals to the (local) subdomains using the weights  $\delta_j^\dagger$ . The process of replacing a residual  $r \in \widehat{S}$  by  $(I - \widehat{P}_0^\top)r$  when applying the first part of the preconditioner is called *balancing*, cf. [Man93, Sect. 2].

Let  $P_Z : W \rightarrow W$  be defined as in (2.50). With Lemma 2.39 in Sect. 2.4.2.1 below, one can show that

$$(I - \widehat{P}_0) \widehat{E}_D = \widehat{E}_D P_Z. \quad (2.63)$$

Since  $\text{range}(P_Z^\top) = \text{range}(S)$ , this identity shows that the application of the pseudo inverse  $S^\dagger$  in (2.62) is valid and the output is always consistent, cf. Definition 1.13.

We can apply the preconditioner (2.62) either to the original Schur complement problem (2.59) or to an auxiliary problem (2.64) that we will introduce below. In the latter case, many simplifications can be made in the algorithm. Furthermore, it can be shown that the two algorithms are equivalent.

In a first step, we use  $\widehat{P}_0$  to project Eq. (2.59) to a subspace. Each element  $w \in V_D^h(\Gamma_S)$  can be decomposed as

$$w = w_0 + \tilde{w}, \quad \text{where } w_0 \in \widehat{W}_0, \tilde{w} \in \text{range}(I - \widehat{P}_0), \langle \widehat{S} w_0, \tilde{w} \rangle = 0.$$

Thanks to the  $\widehat{S}$ -orthogonality of  $\widehat{W}_0$  and  $\text{range}(I - \widehat{P}_0)$ , the part  $u_0$  of the solution  $u$  of (2.59) is given by

$$u_0 = \widehat{R}_0^\top \widehat{S}_0^{-1} \widehat{R}_0 \hat{g}.$$

It is easily seen that  $\hat{g} - \widehat{S} u_0 = (I - \widehat{P}_0^\top) \hat{g}$ . The orthogonal part  $\tilde{u}$  is given as the solution of the following problem:

$$\text{find } \tilde{u} \in \text{range}(I - \widehat{P}_0) : (I - \widehat{P}_0^\top) \widehat{S} \tilde{u} = (I - \widehat{P}_0^\top) \hat{g}. \quad (2.64)$$

Then  $u = u_0 + \tilde{u}$  is the solution of (2.59). We now solve (2.64) using PCG with preconditioner  $M_{\text{BNN}}^{-1}$ . Two simplifications can then be made. Firstly, when applying  $M_{\text{BNN}}^{-1}$  to this equation, the second term  $\widehat{R}_0^\top \widehat{S}_0^{-1} \widehat{R}_0$  in  $M_{\text{BNN}}^{-1}$  can be left out, as well as the term  $(I - \widehat{P}_0^\top)$ . To explain the second simplification, we first rewrite the whole procedure as a method in the original variables  $u = u_0 + \tilde{u}$  (assuming that we have chosen the initial value 0 for  $\tilde{u}^{(0)}$ ). The resulting algorithm is displayed in Algorithm 7. Due to the properties of  $u^{(0)}$ , all the residuals are in  $\text{range}(I - \widehat{P}_0^\top)$ .

---

**Algorithm 7:** Balancing Neumann-Neumann algorithm based on PCG
 

---

```

 $u^{(0)} = \widehat{R}_0^\top \widehat{S}_0^{-1} \widehat{R}_0 \hat{g}$ 
 $r^{(0)} = \hat{g} - \widehat{S} u^{(0)}$ 
 $k = 0$ 
repeat
   $z^{(k)} = \widehat{E}_D S^\dagger \widehat{E}_D^\top r^{(k)}$ 
   $s^{(k)} = (I - \widehat{P}_0) z^{(k)}$ 
   $p^{(k)} = s^{(k)} + \beta_{k-1} p^{k-1}$    where  $\beta_{-1} = 0, \quad \beta_{k-1} = \frac{\langle r^{(k)}, s^{(k)} \rangle}{\langle r^{(k-1)}, s^{(k-1)} \rangle}$ 
  for  $k > 0$ 
     $u^{(k)} = u^{(k)} + \alpha_k p^{(k)}$    where  $\alpha_k = \frac{\langle r^{(k)}, s^{(k)} \rangle}{\langle \widehat{S} p^{(k)}, p^{(k)} \rangle}$ 
     $r^{(k+1)} = r^{(k)} - \alpha_k \widehat{S} p^{(k)}$ 
   $k = k + 1$ 
until stopping criterion fulfilled for  $r^{(k)}$ 

```

---

Hence, the projection step  $(I - \widehat{P}_0^\top)$  in the operator  $(I - \widehat{P}_0^\top) \widehat{S}$  can be omitted as well (and is not included in Algorithm 7).

A brief analysis of the balancing Neumann-Neumann preconditioner will be given in Sect. 2.4.2.4, where we will also see a close connection to FETI/BETI.

## 2.4 Introduction to the Analysis of FETI/BETI

In this section, we first investigate the conditioning of the FETI/BETI operator itself (Sect. 2.4.1). Second, we introduce an abstract framework for analyzing the condition number of the preconditioned operator (Sect. 2.4.2).

### 2.4.1 The Unpreconditioned Case

Following [FMR94, Sect. 3], we analyze the convergence of Algorithms 2 and 3 for the case  $M^{-1} = Q = I$ . Thanks to Lemma 1.49, it suffices to find an upper bound for the condition number of the corresponding operator. Since  $Q = I$ , the operator  $P^\top F|_{\widetilde{U}_{\text{ad}}} \mapsto U_{\text{ad}}^* = U_{\text{ad}}$ , and so, as a formal preconditioner we can choose the natural embedding of  $U_{\text{ad}}$  into the factor space  $\widetilde{U}_{\text{ad}}$ . Since each element in the factor space has a unique representative in  $\text{range}(B)$ , we can use  $U_{\text{ad}} \cap \text{range}(B)$  instead of  $\widetilde{U}_{\text{ad}}$  and obtain

$$\kappa(P^\top F|_{\tilde{U}_{\text{ad}}}) = \frac{\sup_{\lambda \in U_{\text{ad}} \cap \text{range}(B)} \frac{\langle F\lambda, \lambda \rangle}{\|\lambda\|_{\ell^2}^2}}{\inf_{\lambda \in U_{\text{ad}} \cap \text{range}(B)} \frac{\langle F\lambda, \lambda \rangle}{\|\lambda\|_{\ell^2}^2}}. \quad (2.65)$$

In the following let  $a \lesssim b$  be a short hand for  $a \leq C b$  and  $a \approx b$  for  $a \lesssim b$  and  $b \lesssim a$ , where  $C$  is a generic constant.

**Lemma 2.37.** *On  $U_{\text{ad}} \cap \text{range}(B)$ , the operator  $B^\top$  is injective. Let  $\mathbf{B}_i \in \mathbb{R}^{M \times \dim(W_i)}$  be the matrix representation of  $B_i$  with respect to the nodal basis of  $W_i$ . Then*

$$\sum_{i=1}^s \|\mathbf{B}_i^\top \lambda\|_{\ell^2}^2 \approx \|\lambda\|_{\ell^2}^2 \quad \forall \lambda \in U_{\text{ad}} \cap \text{range}(B),$$

where the equivalence constants only depend on the maximal number of subdomains that share a single node  $x^h \in \Gamma_S^h$ .

*Proof.* In the classical formulation, we have

$$\sum_{i=1}^s \|\mathbf{B}_i^\top \lambda\|_{\ell^2}^2 = \sum_{x^h \in \Gamma^h} \sum_{i \in \mathcal{N}_{x^h}} \underbrace{\left| \sum_{j \in \mathcal{N}_{x^h} \setminus \{i\}} \text{sign}(i-j) \lambda_{ij}(x^h) \right|^2}_{=: b_1(\lambda, x^h)}$$

(see Lemma 2.14) and

$$\|\lambda\|_{\ell^2}^2 = \sum_{x^h \in \Gamma^h} \underbrace{\sum_{i, j \in \mathcal{N}_{x^h}, i > j} |\lambda_{ij}(x^h)|^2}_{=: b_2(\lambda, x^h)}.$$

For  $\lambda \in \text{range}(B)$ , the expressions  $\sqrt{b_1(\lambda, x^h)}$  and  $\sqrt{b_2(\lambda, x^h)}$  are norms on a space of dimension  $\#(\mathcal{N}_{x^h})$  and as such equivalent. Since there are only a bounded number of topologically different configurations of a shared node  $x^h$ , the equivalence in the statement of the lemma is indeed uniform. The treatment of the additional terms in the all-floating formulation is straightforward.  $\square$

**Theorem 2.38.** *Assume that*

- (i) *The triangulation  $\mathcal{T}^h(\Gamma_S)$  is shape regular,*
- (ii) *The local triangulations  $\mathcal{T}^h(\partial\Omega_i)$  are quasi-uniform with mesh parameter  $h_i$ ,*
- (iii) *The intersection  $\partial\Omega_i \cap \Gamma_D$  is either empty or has positive surface measure,*
- (iv) *The number of subdomains sharing a node  $x^h \in \Gamma_S^h$  is uniformly bounded.*

*Then*

$$\kappa(P^\top F|_{\tilde{U}_{\text{ad}}}) \leq C \frac{\|\mathcal{A}\|_{L^\infty(\Omega)}}{\alpha_{\min}} \frac{\max_{i=1}^s H_i h_i^{1-d}}{\min_{i=1}^s h_i^{2-d}},$$

where the constant  $C$  is independent of  $H_i$ ,  $h_i$ ,  $\mathcal{A}$ , and the number of subdomains, but it depends on the subdomain shapes. If  $\mathcal{T}^h(\Gamma_S)$  is quasi-uniform, then

$$\kappa(P^\top F|_{\tilde{U}_{\text{ad}}}) = \mathcal{O}\left(\frac{H}{h}\right) \quad \text{as } h \rightarrow \infty,$$

where  $H := \max_{i=1}^s H_i$ .

*Proof.* We define the subspaces

$$W_i^\perp := \begin{cases} \{w_i \in W_i : (\mathbf{w}_i, \mathbf{1})_{\ell^2} = 0\} & \text{if } \Omega_i \text{ floating,} \\ W_i & \text{else,} \end{cases}$$

where  $\mathbf{w}_i \leftrightarrow w_i$  and  $\mathbf{1}$  is the vector of ones. Furthermore, we set

$$|w_i|_{S_i} := \langle S_i w_i, w_i \rangle \quad \text{for } w_i \in W_i.$$

Throughout the proof, let  $\lambda \in U_{\text{ad}} \cap \text{range}(B)$  be arbitrary but fixed. Due to Lemma 1.14,

$$\langle F\lambda, \lambda \rangle = \langle S^\dagger B^\top \lambda, B^\top \lambda \rangle = \sum_{i=1}^s \sup_{w_i \in W_i^\perp} \frac{\langle B_i^\top \lambda, w_i \rangle^2}{|w_i|_{S_i}^2}.$$

Assume first that  $\mathcal{J}_{\text{BEM}} = \emptyset$ . Corollary 1.61 implies that

$$\alpha_{\min} H_i^{-1} h_i^{d-1} \|\mathbf{w}_i\|_{\ell^2}^2 \lesssim |w_i|_{S_i}^2 \lesssim \|\mathcal{A}\|_{L^\infty(\Omega_i)} h_i^{d-2} \|\mathbf{w}_i\|_{\ell^2}^2 \quad \forall \mathbf{w}_i \leftrightarrow w_i \in W_i^\perp.$$

If  $\mathcal{J}_{\text{BEM}} \neq \emptyset$  we can use the spectral equivalence in Corollary 1.94, then the equivalence constants depend additionally on the constant  $c_0(\partial\Omega_i)$  defined as in Lemma 1.77. Moreover, since  $B_i^\top \lambda \in \text{range}(S_i)$  implies  $(\mathbf{B}_i^\top \lambda, \mathbf{1})_{\ell^2} = 0$ , we obtain

$$\sup_{\mathbf{w}_i \leftrightarrow w_i \in W_i^\perp} \frac{\langle B_i^\top \lambda, w_i \rangle^2}{\|\mathbf{w}_i\|_{\ell^2}^2} = \sup_{\mathbf{w}_i \leftrightarrow w_i \in W_i^\perp} \frac{(\mathbf{B}_i^\top \lambda, \mathbf{w}_i)_{\ell^2}^2}{\|\mathbf{w}_i\|_{\ell^2}^2} = \|\mathbf{B}_i^\top \lambda\|_{\ell^2}^2.$$

Combining the above estimate with Lemma 2.37 yields

$$\begin{aligned} \langle F \lambda, \lambda \rangle &\lesssim \sum_{i=1}^s \alpha_{\min}^{-1} H_i h_i^{1-d} \|\mathbf{B}_i^\top \lambda\|_{\ell^2}^2 \lesssim \alpha_{\min}^{-1} \max_{i=1}^s (H_i h_i^{1-d}) \|\lambda\|_{\ell^2}^2, \\ \langle F \lambda, \lambda \rangle &\gtrsim \sum_{i=1}^s \|\mathcal{A}\|_{L^\infty(\Omega)}^{-1} h_i^{2-d} \|\mathbf{B}_i^\top \lambda\|_{\ell^2}^2 \gtrsim \|\mathcal{A}\|_{L^\infty(\Omega)}^{-1} \min_{i=1}^s h_i^{2-d} \|\lambda\|_{\ell^2}^2. \end{aligned}$$

Together with (2.65) this concludes the proof.  $\square$

### 2.4.2 Abstract Framework for the Preconditioned Case

In this subsection, we collect abstract results, whose proofs can be performed on the operator level. Later on, we will have to make rather strong assumptions on the coefficient and the subdomains and use technical tools for finite element functions.

#### 2.4.2.1 The Projection Operators $P_D$ and $E_D$

An important role in the analysis of the scaled Dirichlet preconditioner (2.57) is played by the projection operator

$$P_D := B_D^\top B, \quad (2.66)$$

whose properties are summarized in the following lemma. The result for classical FETI methods was proved by Klawonn and Widlund [KW01], see also [TW05].

**Lemma 2.39.** *The operator  $P_D : W \rightarrow W$  defined in (2.66) satisfies the identities*

$$B M_{\text{SD}}^{-1} B^\top = P_D^\top S P_D, \quad (2.67)$$

$$B P_D = B, \quad (2.68)$$

for both the classical and the all-floating formulation. Furthermore,  $E_D := I - P_D$  is a projection onto the subspace  $\widehat{W}$ , and can be evaluated by

$$(E_D w)_i(x^h) = \begin{cases} \sum_{j \in \mathcal{N}_{x^h}} \delta_j^\dagger(x^h) w_j(x^h) & \text{for } x^h \in \partial\Omega_i^h \setminus \Gamma_D, \\ 0 & \text{for } x^h \in \partial\Omega_i^h \cap \Gamma_D. \end{cases} \quad (2.69)$$

*Proof.* Identity (2.67) follows from the definitions of  $M_{\text{SD}}^{-1}$  and  $P_D$ . Recall that

$$\begin{aligned} (B w)_{ij}(x^h) &= \text{sign}(i - j) (w_i(x^h) - w_j(x^h)) & \text{for } x^h \in \Gamma_{ij}^h, \\ (B w)_{iD}(x^h) &= w_i(x^h) & \text{for } x^h \in \partial\Omega_i^h \cap \Gamma_D, \end{aligned}$$

where the last formula holds only in the all-floating formulation. From Lemma 2.28, we immediately obtain that

$$(P_D w)_i(x^h) = \begin{cases} \sum_{j \in \mathcal{N}_{x^h}} \delta_j^\dagger(x^h) (w_i(x^h) - w_j(x^h)) & \text{for } x \in \partial\Omega_i^h \cap \Gamma^h, \\ w_i(x^h) & \text{for } x^h \in \partial\Omega_i^h \cap \Gamma_D, \\ 0 & \text{else.} \end{cases} \quad (2.70)$$

This implies formula (2.69), and we see that  $E_D w \in \widehat{W}$ . Hence  $B E_D = 0$ , which implies (2.68). Therefore,  $P_D$  is a projection, and so is  $E_D$ .  $\square$

**Corollary 2.40.** *For each  $\mu \in \text{range}(B)$  we can find a function  $w \in \text{range}(P_D)$  such that  $\mu = B w$ .*

*Proof.* Lemma 2.39 implies that  $\text{range}(B) = \text{range}(B P_D)$ .  $\square$

**Remark 2.41.** The operator  $E_D$  is a weighted averaging operator and equals the operator  $\widehat{E}_D$  from Sect. 2.3, up to identification of  $\widehat{W}$  and  $V_D^h(\Gamma_S)$ , cf. [KW01]. For the multiplicity scaling  $\rho_i(x^h) = 1$ ,  $(E_D w)_i(x^h)$  is the algebraic mean value of  $\{w_j(x^h)\}_{j \in \mathcal{N}_{x^h}}$ .

### 2.4.2.2 Positivity of the Preconditioner

For a diagonal choice of  $Q$  (see [KW01]), the operator  $Q$  is SPD per definition, and so the projections  $P$  and  $P^\top$  are well-defined. As the following lemma shows,  $P$  and  $P^\top$  are also well-defined if  $Q = M_{\text{SD}}^{-1}$ .

**Lemma 2.42.** *The scaled Dirichlet preconditioner  $M_{\text{SD}}^{-1}$  is SPD on  $\text{range}(G)$ . Consequently, if  $Q = M_{\text{SD}}^{-1}$ , then the projections  $P$  and  $P^\top$  are well-defined.*

*Proof.* From (2.57) it is immediate that  $M_{\text{SD}}^{-1}$  is positive semi-definite. To show the definiteness on  $\text{range}(G)$ , assume that  $\langle B z, M_{\text{SD}}^{-1} B z \rangle = 0$  for some  $z \in \ker(S)$ . Due to identity (2.67) we obtain  $|P_D z|_S^2 = 0$  which implies that  $P_D z = z - E_D z \in \ker(S)$  and consequently,  $E_D z \in \ker(S)$ . However,  $E_D z \in \ker(B)$  and  $\ker(S) \cap \ker(B) = \{0\}$  imply that  $E_D z = 0$ . This means that the function  $z$ , which is piecewise constant on the subdomains, is continuous across the subdomain interfaces and vanishes on the Dirichlet boundary. Since the domain  $\Omega$  is connected, there is no other possibility than  $z = 0$ , which shows the definiteness.  $\square$

The next lemma discusses the positivity of  $P M_{\text{SD}}^{-1}$  on  $\widetilde{U}_{\text{ad}}^*$ .

**Lemma 2.43.** *If  $Q$  is SPD on  $\text{range}(G)$ , then  $P M_{\text{SD}}^{-1}$  is SPD on  $\widetilde{U}_{\text{ad}}^*$ .*

*Proof.* From (2.57) it is immediate that  $M_{\text{SD}}^{-1}$  is positive semi-definite. To show the definiteness on  $\widetilde{U}_{\text{ad}}^*$ , assume that  $\langle \mu, M_{\text{SD}}^{-1} \mu \rangle = 0$  for some  $\mu \in \widetilde{U}_{\text{ad}}^*$ . Due to

Corollary 2.40, there exists a function  $w \in \text{range}(P_D)$  with  $\mu = B w$ , and so

$$0 = \langle \mu, M_{\text{sD}}^{-1} \mu \rangle = |P_D w|_S^2 = |w|_S^2 = 0.$$

Hence,  $w \in \ker(S)$  and  $\mu = B w \in \text{range}(G)$ . The definition (2.43) of  $\tilde{U}_{\text{ad}}^*$  yields

$$\langle \mu, Q \mu \rangle = \langle \mu, Q B w \rangle = 0.$$

Since  $Q$  is SPD on  $\text{range}(G)$ , it follows that  $\mu = 0$ .  $\square$

### 2.4.2.3 An Abstract Condition Number Estimate for FETI/BETI

With the following two lemmas, we can reduce the condition number estimate of the preconditioned FETI/BETI method to a single stability estimate of the  $P_D$  operator. For the original proofs see [KW01] and also [MT96].

**Lemma 2.44.** *Let  $Q$  be SPD on  $\text{range}(G)$ . Then, for any  $w \in W$ , then there exists a unique element  $z_w \in \ker(S)$  such that  $B(w + z_w) \in \tilde{U}_{\text{ad}}^*$ , given by*

$$z_w = -R(G^\top Q G)^{-1} G^\top Q B w = -(I - P_Z)w,$$

where  $P_Z$  is defined in (2.50). Moreover, if  $Q$  is SPD on  $\tilde{U}^*$ , then

$$z_w = \underset{z \in \ker S}{\operatorname{argmin}} \|B(w + z)\|_Q, \quad \text{and} \quad \|B z_w\|_Q \leq \|B w\|_Q,$$

where  $\|\mu\|_Q := \langle \mu, Q \mu \rangle^{1/2}$ .

*Proof.* The statements follow from Sect. 1.1.3.3 and Lemma 1.3.  $\square$

**Lemma 2.45.** *Assume that  $P M_{\text{sD}}^{-1}$  is SPD on  $\tilde{U}_{\text{ad}}^*$  and let  $\omega$  be a parameter such that*

$$|P_D(w + z_w)|_S^2 \leq \omega |w|_S^2 \quad \forall w \in W,$$

where  $z_w$  is the unique element from Lemma 2.44. Then

$$\kappa(P M_{\text{sD}}^{-1} P^\top F|_{\tilde{U}_{\text{ad}}}) \leq \omega.$$

*Proof.* Let  $M_{\text{sD}} : \tilde{U}_{\text{ad}} \rightarrow U_{\text{ad}}^*$  be the inverse of  $(P M_{\text{sD}}^{-1})|_{\tilde{U}_{\text{ad}}^*}$ . From Corollary 1.50, we see that

$$\kappa(P M_{\text{sD}}^{-1} P^\top F|_{\tilde{U}_{\text{ad}}}) = \frac{\sup_{\lambda \in \tilde{U}_{\text{ad}}} \frac{\langle F \lambda, \lambda \rangle}{\langle M_{\text{sD}} \lambda, \lambda \rangle}}{\inf_{\lambda \in \tilde{U}_{\text{ad}}} \frac{\langle F \lambda, \lambda \rangle}{\langle M_{\text{sD}} \lambda, \lambda \rangle}}. \quad (2.71)$$

In order to estimate the numerator in (2.71), we bound  $F$  in terms of  $M_{\text{SD}}$ . Let  $\lambda \in \widetilde{U}_{\text{ad}}$  arbitrary but fixed. Lemma 1.14 and our assumptions imply

$$\langle F\lambda, \lambda \rangle = \sup_{w \in W} \frac{\langle Bw, \lambda \rangle^2}{|w|_S^2} \leq \omega \sup_{w \in W} \frac{\langle Bw, \lambda \rangle^2}{|P_D(w + z_w)|_S^2}.$$

From the definition (2.37) of  $U_{\text{ad}}$  and from  $\widetilde{U}_{\text{ad}} = U_{\text{ad}/\ker(B^\top)}$ , we can conclude that

$$\langle Bw, \lambda \rangle = \langle B(w + z_w), \lambda \rangle.$$

Together with the above it follows from (2.67) and Lemma 2.44 that

$$\langle F\lambda, \lambda \rangle \leq \omega \sup_{w \in W} \frac{\langle B(w + z_w), \lambda \rangle^2}{\langle B(w + z_w), M_{\text{SD}}^{-1} B(w + z_w) \rangle} \leq \omega \sup_{\mu \in \widetilde{U}_{\text{ad}}^*} \frac{\langle \mu, \lambda \rangle^2}{\langle \mu, M_{\text{SD}}^{-1} \mu \rangle} = \omega \langle M_{\text{SD}} \lambda, \lambda \rangle.$$

We now turn to the denominator in (2.71). Let  $\lambda \in \widetilde{U}_{\text{ad}}$  be fixed and  $\mu \in \widetilde{U}_{\text{ad}}^*$  arbitrary. Thanks to Corollary 2.40, there exists  $w \in \text{range}(P_D)$  such  $Bw = \mu$ . Hence,

$$\langle F\lambda, \lambda \rangle \geq \frac{\langle Bw, \lambda \rangle^2}{|w|_S^2} = \frac{\langle Bw, \lambda \rangle^2}{|P_D w|_S^2} = \frac{\langle Bw, \lambda \rangle^2}{\langle Bw, M_{\text{SD}}^{-1} Bw \rangle} = \frac{\langle \mu, \lambda \rangle^2}{\langle \mu, M_{\text{SD}}^{-1} \mu \rangle}.$$

Since  $\mu \in \widetilde{U}_{\text{ad}}^*$  was arbitrary, we get that

$$\langle F\lambda, \lambda \rangle \geq \langle M_{\text{SD}} \lambda, \lambda \rangle \quad \forall \lambda \in \widetilde{U}_{\text{ad}}.$$

Combining the two bounds concludes the proof.  $\square$

In Sect. 2.6 we will work out in detail the missing bound

$$|P_D(w + z_w)|_S^2 \leq \omega |w|_S^2 \quad \forall w \in W,$$

which requires a series of assumptions and technical tools that we present in Sect. 2.5.

#### 2.4.2.4 An Estimate for the Balancing Neumann-Neumann Method

For the original analysis of the balancing Neumann-Neumann preconditioner we refer to [Man93, MB96], see also [DW95] and [TW05, Sect. 6.2.3]. Here, we provide an analysis which makes use of an abstract framework evolved from the above works on Neumann-Neumann as well as from the analysis of the more recent BDDC methods (see Sect. 5.1.4). We note that the following theorem (cf. [MS07a, Theorem 2] and [Sou10, Theorem 2.5]) has certain relations to the *fictitious space lemma* by Nepomnyaschikh, cf. [Nep91a, Nep92, Nep07].



**Theorem 2.46.** Let  $\widehat{X} \subset X$  be two Hilbert spaces,  $a(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$  a symmetric positive definite bilinear form, and let  $A : X \rightarrow X^*$  and  $\widehat{A} : \widehat{X} \rightarrow \widehat{X}^*$  denote the corresponding operator and its restriction to  $\widehat{X}$ . Furthermore, let  $\widehat{Q} : X \rightarrow \widehat{X}$  be a projector. The abstract BDDC preconditioner  $\widehat{B}^{-1} : \widehat{X}^* \rightarrow \widehat{X}$  is defined by

$$\widehat{B}^{-1} : \widehat{Q} A^{-1} \widehat{Q}^\top.$$

Let  $X^M := \{v \in X : \forall z \in X : \widehat{Q} v = \widehat{Q} z \implies \|v\|_a \leq \|z\|_a\}$ , where  $\|\cdot\|_a$  is the norm associated to  $a(\cdot, \cdot)$ , and let  $\omega$  be the minimal constant such that

$$\|\widehat{Q} v\|_a^2 \leq \omega \|v\|_a^2 \quad \forall v \in X^M.$$

Then

$$\kappa(\widehat{B}^{-1} \widehat{A}) \leq \omega = \sup_{v \in X} \frac{\|\widehat{Q} v\|_a^2}{\|v\|_a^2}.$$

*Proof.* For completeness we display the proof from [MS07a, Sou10]. Let  $\widehat{G} : \widehat{X} \rightarrow X$  be defined by

$$\widehat{G} v := \operatorname{argmin}_{w \in X^M : v = \widehat{Q} w} a(w, w).$$

A short computation reveals that

$$\widehat{G} = A^{-1} \widehat{Q}^\top \widehat{B}, \quad \widehat{Q} \widehat{G} = I, \quad \widehat{G}^\top A \widehat{G}^\top = \widehat{B}, \quad (2.72)$$

where  $\widehat{B} : \widehat{X} \rightarrow \widehat{X}^*$  is the inverse of  $\widehat{B}^{-1}$ . Using (2.72), the definition of  $\widehat{G}$ , and the fact that  $\widehat{Q}$  is a projection, we obtain

$$\langle \widehat{B} v, v \rangle = a(\widehat{G} v, \widehat{G} v) \leq a(v, v) \quad \forall v \in \widehat{X},$$

which shows that  $\lambda_{\min}(\widehat{B}^{-1} \widehat{A}) \geq 1$ . Now, we conclude from (2.72) and the defining property of  $\omega$  that for any  $v \in \widehat{W}$ ,

$$\|v\|_a^2 = \|\widehat{Q} \widehat{G} v\|_a^2 \leq \omega \|\widehat{G} v\|_a^2 = \omega \langle \widehat{B} v, v \rangle.$$

The last estimate implies that  $\lambda_{\max}(\widehat{B}^{-1} \widehat{A}) \leq \omega$ . The alternative characterization of  $\omega$  is trivial.  $\square$

The following lemma makes use of the above abstract theorem and shows that a similar estimate as in Lemma 2.45 implies a bound for the condition number of the balancing Neumann-Neumann preconditioner.

**Lemma 2.47.** For each  $i = 1, \dots, s$ , let  $W_i^\perp$  be a subspace of  $W_i$  such that the sum  $W_i = \ker(S_i) \oplus W_i^\perp$  is direct, and let  $W^\perp := \prod_{i=1}^s W_i^\perp$ . Then a bound of the form

$$|E_D w|_S^2 \leq \omega |w|_S^2 \quad \forall w \in W^\perp \quad (2.73)$$

implies that

$$\kappa(M_{\text{BNN}}^{-1} \widehat{S}_{|\text{range}(I-\widehat{P}_0)}) \leq \omega.$$

*Proof.* Firstly, we convince ourselves that

$$M_{\text{BNN}}^{-1} \widehat{S}_{|\text{range}(I-\widehat{P}_0)} = (I - \widehat{P}_0) \widehat{E}_D S^\dagger \widehat{E}_D^\top (I - \widehat{P}_0^\top) \widehat{S}_{|\text{range}(I-\widehat{P}_0)}.$$

Using identity (2.63) and the fact that  $P_Z S^\dagger|_{\text{range}(S)}$  is the inverse of the restriction of  $S$  to  $\text{range}(P_Z)$ , it can be shown that the balancing Neumann-Neumann preconditioner fits into the framework of Theorem 2.46 with

$$X := \text{range}(P_Z), \quad \widehat{X} := \text{range}(I - \widehat{P}_0), \quad \widehat{Q} := (I - \widehat{P}_0) \widehat{E}_D = \widehat{E}_D P_Z.$$

Hence, Theorem 2.46 implies that

$$\kappa(M_{\text{BNN}}^{-1} \widehat{S}_{|\text{range}(I-\widehat{P}_0)}) \leq \sup_{w \in \text{range}(P_Z)} \frac{|E_D w|_S^2}{|w|_S^2} = \sup_{w \in W^\perp} \frac{|E_D P_Z w|_S^2}{|w|_S^2},$$

where in the last step, we have used the same argumentation as in Lemma 2.103 below. Finally, the statement of the lemma follows since  $\widehat{E}_D P_Z = (I - \widehat{P}_0) E_D$  and  $(I - \widehat{P}_0)$  is an  $\widehat{S}$ -orthogonal projector with its  $\widehat{S}$ -norm less or equal than one.  $\square$

The connection between FETI and balancing Neumann-Neumann methods can firstly be seen from the ingredients of the two methods, which are related by the two operators  $S$  and  $S^\dagger$ , by  $B_D^\top B = P_D = I - E_D$ , by the fact that the coarse matrices coincide for the case  $Q = M_{\text{SD}}^{-1}$ , cf. Sect. 2.3, and by many common parts of their analyses (see also [KW01]). Secondly, since the norm of a non-trivial projection in a finite-dimensional Hilbert space depends only on the angle between its kernel and range [IM95], the bound (2.73) implies

$$|P_D w|_S^2 \leq \omega |w|_S^2 \quad \forall w \in W^\perp, \quad (2.74)$$

and vice versa. Moreover, Sousedik and Mandel [SM08] proved that the spectra of the FETI preconditioner (with  $Q = M_{\text{SD}}^{-1}$ ) and the balancing Neumann-Neumann preconditioner (with corresponding scalings) are identical except for the eigenvalue of one, i.e.,  $\sigma(M_{\text{BNN}}^{-1} \widehat{S}_{|\text{range}(I-\widehat{P}_0)}) \setminus \{1\} = \sigma(M_{\text{SD}}^{-1} F_{|\tilde{U}_{\text{ad}}}) \setminus \{1\}$ .

## 2.5 Technical Tools

In this section, we first define a partition of the skeleton into so-called *glob*s and introduce associated cut-off functions. These will be used to split the contributions in the operator  $P_D$  that involve different neighbors. To estimate the effect of the cut-off, we need regularity assumptions of the subdomains (Sect. 2.5.2) and a series of technical estimates (Sects. 2.5.3–2.5.6). Readers interested in the main line of the analysis may initially bypass the technical estimates, continue with Sect. 2.6, and return to the necessary lemmas at a later stage.

### 2.5.1 Globs and Cut-Off Functions

When looking to formula (2.70), we see that the operator  $P_D$  involves different subsets of the functions  $\{w_i\}_{i=1}^N$  on different parts of the skeleton  $\Gamma_S$ . This is the main reason for the following definition.

**Definition 2.48.** (i) For  $x \in \Gamma_S$  we define the set of subdomains that share  $x$ :

$$\mathcal{N}_x := \{i = 1, \dots, s : x \in \partial\Omega_i\}.$$

(ii) The interface  $\Gamma$  decomposes into equivalence classes of the relation

$$x \sim y \iff \mathcal{N}_x = \mathcal{N}_y.$$

Each *connected* component of such an equivalence class is called a *glob* (cf. [Mat08]). For a glob  $\mathcal{G}$ , we simply denote by  $\mathcal{N}_{\mathcal{G}}$  the set of sharing subdomains ( $\mathcal{N}_{\mathcal{G}} = \mathcal{N}_x$  for all  $x \in \mathcal{G}$ ).

(iii) In three dimensions, a glob  $\mathcal{G}$  on the interface  $\Gamma$  is called

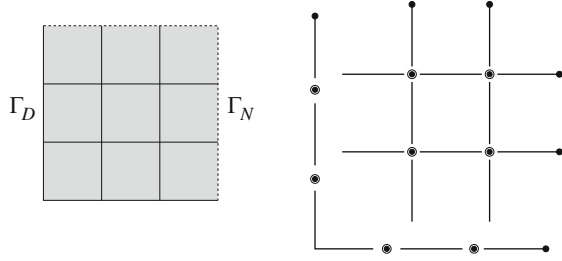
- *Subdomain vertex* if it consists of a single point,
- *Subdomain face* or *subdomain facet* if it is shared by exactly two subdomains,
- *Subdomain edge* otherwise.

In two dimensions, a glob  $\mathcal{G}$  on the interface  $\Gamma$  is called

- *Subdomain edge* or *subdomain facet* if it is shared by exactly two subdomains,
- *Subdomain vertex* otherwise (being a single point).

Note that subdomain edges may include one of their endpoints (see the definition of the interface  $\Gamma$ , Table 2.1, p. 65). Similarly, there are subdomain faces which include a part of the Neumann boundary.

**Fig. 2.5** Example of a decomposition of a two-dimensional domain with the corresponding globs on  $\Gamma \cup \Gamma_D$ , subdomain vertices indicated by  $\bullet$ , subdomain edges indicated by  $\text{---}$  and  $\text{---}\bullet$



- (iv) We decompose the Dirichlet boundary  $\Gamma_D$  into equivalence classes of the same type as above, and each connected component of such an equivalence class is called a *Dirichlet glob*.
- (v) In three dimensions, a Dirichlet glob  $\mathcal{G}$  is called

- *Subdomain vertex* if it consists of a single point,
- *Subdomain face* or *subdomain facet* if it belongs to exactly one subdomain,
- *Subdomain edge* otherwise.

In two dimensions, we have again subdomain vertices consisting of a single point, all other globs are called subdomain edges or also subdomain facets.

Figure 2.5 illustrates the decomposition of  $\Gamma$  and  $\Gamma_D$  into globs for a two-dimensional example.

*Remark 2.49.* Our definition of subdomain faces, edges, and vertices slightly differs from the “common” one in [TW05, Definition 4.1] because we will need to treat all-floating methods that operate on the Dirichlet boundary as well.

*Notation.* We denote subdomain facets, edges, and vertices on  $\Gamma$  that are shared by (at least) two subdomains  $\Omega_i$  and  $\Omega_j$  generically by  $\mathcal{F}_{ij}$ ,  $\mathcal{E}_{ij}$ ,  $\mathcal{V}_{ij}$ , respectively. We would like to point out that an index pair does not necessarily specify a subdomain facet/edge/vertex uniquely. Subdomain facets, edges, and vertices on  $\partial\Omega_i$  (possibly shared by  $\Omega_i$  and the outer boundary) are denoted generically by  $\mathcal{F}_i$ ,  $\mathcal{E}_i$ , and  $\mathcal{V}_i$ , respectively.

**Definition 2.50.** For a subdomain face  $\mathcal{F}$ , let  $\mathcal{F}^h$  be the set of nodes contained in  $\mathcal{F}$ , and for a subdomain edge  $\mathcal{E}$ , let  $\mathcal{E}^h$  be the set of nodes contained in  $\mathcal{E}$ . For a subdomain vertex  $\mathcal{V}$  we set  $\mathcal{V}^h := \{\mathcal{V}\}$ .

In order to separate the contributions of  $P_D$  on the different globs, we define the following cut-off functions, according to [TW05, Sect. 4.6]. Let  $V^h(\partial\Omega_i)$  denote the restriction of  $V^h(\Omega_i)$  to the boundary, cf. Sect. 1.2.6.2.

**Definition 2.51 (Finite element cut-off functions).**

- For a subdomain vertex  $\mathcal{V}_i$ , the piecewise linear function  $\theta_{\mathcal{V}_i} \in V^h(\partial\Omega_i)$  equals one at the vertex  $\mathcal{V}_i$ , and zero on all other nodes.
- For a subdomain edge  $\mathcal{E}_i$ , the piecewise linear function  $\theta_{\mathcal{E}_i} \in V^h(\partial\Omega_i)$  equals one at all nodes contained in  $\mathcal{E}_i$ , and zero at all other nodes.

- For a subdomain face  $\mathcal{F}_i$ , the piecewise linear function  $\theta_{\mathcal{F}_i} \in V^h(\partial\Omega_i)$  equals one at the nodes contained in  $\mathcal{F}_i$ , and zero at all other nodes.

Extending these functions by zero, we have  $\theta_{\mathcal{G}_i} \in V^h(\Gamma_S)$ .

**Definition 2.52.** Let  $I^h$  denote the nodal interpolator that interpolates continuous functions to  $V^h(\Omega_i)$  or  $V^h(\partial\Omega_i)$ .

*Notation.* By writing

$$\sum_{\mathcal{G}_i} I^h(\theta_{\mathcal{G}_i} w_i), \quad \sum_{\mathcal{G}_i \subset \Gamma} I^h(\theta_{\mathcal{G}_i} w_i), \quad \text{and} \quad \sum_{\mathcal{G}_i \subset \Gamma_D} I^h(\theta_{\mathcal{G}_i} w_i),$$

we mean that we sum over all globs on  $\partial\Omega_i$ ,  $\partial\Omega_i \cap \Gamma$ , and  $\partial\Omega_i \cap \Gamma_D$ , respectively.

The cut-off functions from Definition 2.51 provide a partition of unity in the sense that

$$\sum_{\mathcal{G}_i} I^h(\theta_{\mathcal{G}_i} v) = v \quad \forall v \in V^h(\partial\Omega_i), \quad v|_{\partial\Omega_i \cap \Gamma_N^h} = 0, \quad (2.75)$$

where

$$\Gamma_N^h := \Gamma_S^h \cap (\Gamma_N \setminus \Gamma) \quad (2.76)$$

denotes the set of non-coupling Neumann nodes.

Within the following sections, we work out tools in order to estimate the effect of a single cut-off in the  $H^1$ -energy norm.

## 2.5.2 Regularity Assumptions on the Subdomains

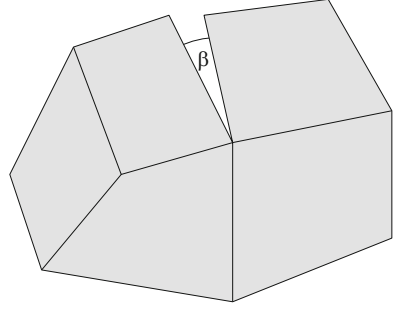
In this subsection we collect all the regularity assumptions that we need for the technical tools exposed in subsequent sections. Most of the assumptions can be relaxed (while of course complicating the theory) and we will indicate that at the corresponding places.

**Assumption 2.53.** The subdomain triangulations  $\mathcal{T}^h(\Omega_i)$ ,  $i \in \mathcal{I}_{\text{FEM}}$  and  $\mathcal{T}^h(\partial\Omega_i)$ ,  $i \in \mathcal{I}_{\text{BEM}}$  are quasi-uniform with mesh parameter  $h_i$ .

**Assumption 2.54.** There is a shape regular simplicial coarse triangulation  $\mathcal{T}^H(\Omega)$  of  $\Omega$ , such that each subdomain  $\Omega_i$  is the union of coarse elements and the number of coarse elements per subdomain is uniformly bounded.

Under Assumption 2.54, the subdomain diameter  $H_i$  is equivalent to the local mesh parameter of  $\mathcal{T}^H(\Omega_i)$ . Furthermore, we can extend the triangulations  $\mathcal{T}^h(\partial\Omega_i)$ ,  $i \in \mathcal{I}_{\text{BEM}}$  to auxiliary triangulations  $\mathcal{T}^h(\Omega_i)$  that are quasi-uniform with mesh parameter  $h_i$ . We will use these auxiliary triangulations in the analysis.

**Fig. 2.6** Non-convex domain with arbitrarily small exterior angle  $\beta$ . The displayed subdomain decomposition satisfies Assumptions 2.54 and 2.55 with shape regularity constants uniform in  $\beta$



The next assumption essentially states that the exterior angles of the BEM subdomains must be bounded from below. Note that even for a pure BETI method, this does not necessarily prohibit the global domain  $\Omega$  from having bad exterior angles, cf. Fig. 2.6.

**Assumption 2.55.** For each subdomain  $\Omega_i$ ,  $i \in \mathcal{I}_{\text{BEM}}$ , there exists a neighborhood  $\Omega'_i \supset \overline{\Omega_i}$  and a shape regular simplicial coarse triangulation  $\mathcal{T}^H(\Omega'_i)$ , such that the shape regularity constants and the number of coarse elements in  $\mathcal{T}^H(\Omega'_i)$  are uniformly bounded.

**Assumption 2.56.** The Dirichlet boundary  $\Gamma_D$  is the union of facets (faces/edges) of the coarse triangulation  $\mathcal{T}^H(\Omega)$  from Assumption 2.54.

Thanks to Assumptions 2.54 and 2.56, each glob is the union of a few vertices, edges, and faces of the coarse triangulation  $\mathcal{T}^H(\Omega)$ .

*Notation.* We write  $a \lesssim b$  if there is a constant  $C$  depending only on the shape regularity and quasi-uniformity constants from Assumptions 2.53–2.55 such that  $a \leq C b$ . The notation  $a \approx b$  is a short hand for  $a \lesssim b$  and  $b \lesssim a$ .

### 2.5.3 An Explicit Poincaré Inequality

In this subsection, we provide a uniform bound of the Poincaré constant  $C_P(\Omega_i)$  (see Lemma 1.27) under Assumption 2.54. The following lemma is taken from [PS12b, Lemma A.1]. For similar approaches see also [BH70, DS80, DW06, VV09].

**Lemma 2.57.** *Let  $T$  be a triangle ( $d = 2$ ) or tetrahedron ( $d = 3$ ) and let  $F$  be one of its facets (edge if  $d = 2$  and triangular face if  $d = 3$ ). Then*

$$\|u - \bar{u}^F\|_{L^2(T)} \leq \text{diam}(T) |u|_{H^1(T)} \quad \forall u \in H^1(T).$$

*Proof.* Due to Veeder and Verfürth (see [VV09, Sect. 4, Remark 4.6, formula (2.3), and Corollary 4.5]), for all  $v \in H^1(T)$ :

$$\frac{1}{\text{meas}_{d-1}(F)} \|v\|_{L^2(F)}^2 \leq \frac{1}{\text{meas}_d(T)} \|v\|_{L^2(T)}^2 + \frac{2 \text{diam}(T)}{d \text{meas}_d(T)} \|v\|_{L^2(T)} |v|_{H^1(T)}, \quad (2.77)$$

Due to [PW60, Beb03],

$$\|u - \bar{u}^T\|_{L^2(T)} \leq \frac{\text{diam}(T)}{\pi} |u|_{H^1(T)} \quad \forall u \in H^1(T), \quad (2.78)$$

because  $T$  is convex. With the triangle inequality and Cauchy's inequality,

$$\begin{aligned} \|u - \bar{u}^F\|_{L^2(T)} &\leq \|u - \bar{u}^T\|_{L^2(T)} + \sqrt{\text{meas}_d(T)} |\bar{u}^T - \bar{u}^F| \\ &\leq \|u - \bar{u}^T\|_{L^2(T)} + \frac{\sqrt{\text{meas}_d(T)}}{\sqrt{\text{meas}_{d-1}(F)}} \|u - \bar{u}^T\|_{L^2(F)} \end{aligned}$$

Using (2.77) and (2.78) in the estimate above yields

$$\begin{aligned} \|u - \bar{u}^F\|_{L^2(T)} &\leq \frac{\text{diam}(T)}{\pi} |u|_{H^1(T)} + \sqrt{\|u - \bar{u}^T\|_{L^2(T)}^2 + \frac{2 \text{diam}(T)}{d} \|u - \bar{u}^T\|_{L^2(T)} |u|_{H^1(T)}} \\ &\leq \frac{\text{diam}(T)}{\pi} |u|_{H^1(T)} + \sqrt{\frac{\text{diam}(T)^2}{\pi^2} |u|_{H^1(T)}^2 + \frac{2 \text{diam}(T)}{d} \frac{\text{diam}(T)}{\pi} |u|_{H^1(T)}^2} \\ &= \left( \frac{1}{\pi} + \sqrt{\frac{1}{\pi^2} + \frac{2}{d\pi}} \right) \text{diam}(T) |u|_{H^1(T)} \end{aligned}$$

Since  $d \geq 2$ , the factor in the parentheses is  $\leq 0.96609936 \leq 1$ .  $\square$

**Definition 2.58.** Let  $\mathcal{T}^H(\Omega)$  be a simplicial triangulation. An open set  $Y \subset \Omega$  is called ( $d$ -dimensional) *agglomerate* of  $\mathcal{T}^H(\Omega)$  (in short  $\mathcal{T}^H$ -agglomerate) if  $Y$  is a connected union of elements from  $\mathcal{T}^H(\Omega)$ .

**Definition 2.59.** Let  $Y$  be an agglomerate of  $\mathcal{T}^H(\Omega)$  with  $\bar{Y} = \bigcup_{\ell=1}^L \bar{T}_\ell$ . We call the region  $P_{\ell_1, \ell_s} = \text{interior}(\bar{T}_{\ell_1} \cup \bar{T}_{\ell_2} \cup \dots \cup \bar{T}_{\ell_s})$ ,  $1 \leq \ell_1, \dots, \ell_s \leq L$ , a *path* from  $T_{\ell_1}$  to  $T_{\ell_s}$  of *length*  $s$  if for each  $i = 1, \dots, s-1$ , the elements  $\bar{T}_{\ell_i}$  and  $\bar{T}_{\ell_{i+1}}$  share a common facet  $F_i$ .

The following two lemmas lift the result of Lemma 2.57 from a single simplex to an agglomerate, cf. [PS11c, Lemmas 2 and 4]. We will revisit the same approach in Sect. 3.4.4.

**Lemma 2.60.** *Let  $Y$  be an agglomerate of  $\mathcal{T}^H(\Omega)$ , let  $P_{\ell_1, \ell_s} = \text{interior}(\overline{T_{\ell_1}} \cup \overline{T_{\ell_2}} \cup \dots \cup \overline{T_{\ell_s}})$  be a path and let  $F_i$  be the facet shared by  $\overline{T_{\ell_i}}$  and  $\overline{T_{\ell_{i+1}}}$  for  $i = 1, \dots, s-1$ . Furthermore, let  $F_s$  be any facet of  $T_{\ell_s}$ . Then*

$$\|u - \bar{u}^{F_s}\|_{L^2(T_{\ell_1})}^2 \leq 4 \left( \sum_{i=1}^s \frac{\text{meas}_d(T_{\ell_1}) \text{diam}(T_{\ell_i})^2}{\text{meas}_d(T_{\ell_i})} \right) |u|_{H^1(P_{\ell_1, \ell_s})}^2 \quad \forall u \in H^1(Y).$$

*Proof.* By a telescoping argument we have

$$\|u - \bar{u}^{F_s}\|_{L^2(T_{\ell_1})} \leq \|u - \bar{u}^{F_1}\|_{L^2(T_{\ell_1})} + \sum_{i=2}^s \sqrt{\text{meas}_d(T_{\ell_1})} |\bar{u}^{F_{i-1}} - \bar{u}^{F_i}|.$$

With Lemma 2.57,

$$\|u - \bar{u}^{F_1}\|_{L^2(T_{\ell_1})} \leq \text{diam}(T_{\ell_1}) |u|_{H^1(T_{\ell_1})}.$$

For fixed  $2 \leq i \leq s$  the same lemma implies that

$$\begin{aligned} |\bar{u}^{F_{i-1}} - \bar{u}^{F_i}|^2 &\leq \frac{2}{\text{meas}_d(T_{\ell_i})} \left( \|\bar{u}^{F_{i-1}} - u\|_{L^2(T_{\ell_i})}^2 + \|u - \bar{u}^{F_i}\|_{L^2(T_{\ell_i})}^2 \right) \\ &\leq \frac{4}{\text{meas}_d(T_{\ell_i})} \text{diam}(T_{\ell_i})^2 |u|_{H^1(T_{\ell_i})}^2. \end{aligned}$$

Combining the three estimates above, Cauchy's inequality (in  $\mathbb{R}^s$ ) yields

$$\|u - \bar{u}^{F_s}\|_{L^2(T_{\ell_1})}^2 \leq \left[ \sum_{i=1}^s \frac{4 \text{meas}_d(T_{\ell_1}) \text{diam}(T_{\ell_i})^2}{\text{meas}_d(T_{\ell_i})} \right] |u|_{H^1(P_{\ell_1, \ell_s})}^2,$$

which completes the proof.  $\square$

**Lemma 2.61.** *Let  $Y$  an agglomerate of  $\mathcal{T}^H(\Omega)$  consisting of at most  $L$  elements. Then*

$$C_P(Y) \leq C,$$

where  $C$  only depends on  $L$  and on the shape regularity constant of  $\mathcal{T}^H(\Omega)$ .

*Proof.* One easily shows that

$$\text{meas}_d(T) \geq c \text{meas}_d(Y), \quad \text{diam}(T) \geq c \text{diam}(Y)$$

for all elements  $T$  contained in  $Y$ , where the constant  $c$  only depends on  $L$  and on the shape regularity constant of  $\mathcal{T}^H(\Omega)$ . Let  $F$  be an arbitrary facet of  $Y$ . For each element  $T$  in  $Y$  there exists a path from  $T$  to the element containing  $F$  of maximal length  $L$ . Summing the estimate in Lemma 2.60 over all elements in  $Y$  yields

$$\inf_{c \in \mathbb{R}} \|u - c\|_{L^2(Y)}^2 \leq \|u - \bar{u}^F\|_{L^2(Y)}^2 \leq C \text{diam}(Y)^2 |u|_{H^1(Y)}^2.$$

for all  $u \in H^1(Y)$ .  $\square$



**Corollary 2.62.** *Under Assumption 2.54, the Poincaré constants of all the subdomains are uniformly bounded.*

*Remark 2.63.* Without Assumption 2.54, one can instead assume that the isoperimetric constants of the subdomains are uniformly bounded, see [DKW08b, DKW08a, KRW08]. Note that ragged boundaries alone do not make these constants blow up.

### 2.5.4 Trace and Poincaré Inequalities for FE Functions

In this subsection, we elaborate on inequalities for finite element functions, which do (in general) not hold in the Sobolev space  $H^1$ . Therefore, we obtain expressions depending on the mesh parameter  $h$ . All the estimates can e.g. be found in [TW05, Sect. 4.6] as well as in early papers [BPS86, BPS87, BPS88, BPS89]. Here, we work out the common pattern and the dependence of the constants on the geometry, using the assumptions of Sect. 2.5.2.

**Definition 2.64.** For  $0 \leq m \leq d - 1$ , the  $m$ -facets of a simplicial (coarse) triangulation  $\mathcal{T}^H(\Omega)$  are

- The vertices of  $\mathcal{T}^H(\Omega)$  if  $m = 0$ ,
- The edges if  $m = 1$ ,
- The triangular faces if  $m = 2$ .

**Definition 2.65.** For  $0 \leq m \leq d - 1$ , a set  $X \subset \overline{\Omega}$  is called  $m$ -dimensional *agglomerate* of  $\mathcal{T}^H(\Omega)$  (in short  $\mathcal{T}^H$ -agglomerate) if  $X$  is a union of  $m$ -facets. The dimension of  $X$  is denoted by  $d_X := m$ .

Apparently, a zero-dimensional agglomerate consists of finitely many (isolated) vertices. We see that under Assumption 2.54, the globs from Definition 2.48 are  $m$ -agglomerates of the coarse triangulation  $\mathcal{T}^H(\Omega)$  with  $0 \leq m \leq d - 1$ .

If  $X$  is a  $\mathcal{T}^H$ -agglomerate of dimension  $d_X > 0$ , the  $d_X$ -dimensional Lebesgue measure properly defines the space  $L^2(X)$  and the average

$$\bar{v}^X := \frac{1}{\text{meas}_{d_X}(X)} \int_X v \, ds \quad \text{for } v \in L^2(X).$$

The case  $d_X = 0$  is treated in the following definition.

**Definition 2.66.** Let  $X = \{p_j : j = 1, \dots, J\}$  be a zero-dimensional agglomerate of  $\mathcal{T}^H(\Omega)$ , consisting of  $J$  different points. For a function  $v : X \rightarrow \mathbb{R}$ , we define

$$\text{meas}_0(X) := J, \quad \int_X v \, ds := \sum_{j=1}^J v(p_j), \quad \|v\|_{L^2(X)} := \left( \int_X |v|^2 ds \right)^{1/2}$$

and the average

$$\bar{v}^X := \frac{1}{\text{meas}_0(X)} \int_X v \, ds = J^{-1} \sum_{j=1}^J v(p_j).$$

Let  $\{\mathcal{T}^h(\Omega)\}_h$  be a family of triangulations that are refinements of  $\mathcal{T}^H(\Omega)$  and let  $V^h(\Omega)$  denote the finite element space of continuous and piecewise linear functions with respect to  $\mathcal{T}^h(\Omega)$ . For any  $d$ -dimensional agglomerate  $Y$  of  $\mathcal{T}^H(\Omega)$ , we denote by  $\mathcal{T}^h(Y)$  and  $V^h(Y)$  the restrictions of  $\mathcal{T}^h(\Omega)$  and  $V^h(\Omega)$  to  $Y$ .

**Definition 2.67.** Let  $Y$  be a  $d$ -dimensional  $\mathcal{T}^H$ -agglomerate.

- (i) For a  $\mathcal{T}^H$ -agglomerate  $X \subset \bar{Y}$  of dimension  $0 \leq d_X \leq d$ , let  $C_P(Y, X; h)$  denote the smallest parameter such that

$$\|v - \bar{v}^X\|_{L^2(Y)} \leq C_P(Y, X; h) \text{diam}(Y) |v|_{H^1(Y)} \quad \forall v \in V^h(Y).$$

For  $d_X \geq d - 1$  let  $C_P(Y, X)$  denote the smallest constant such that the above inequality holds in  $H^1(Y)$ .

- (ii) For  $\mathcal{T}^H$ -agglomerates  $X, W \subset \bar{Y}$  of dimensions  $d_X, d_W \in [0, d]$ , let  $C_P(Y, X, W; h)$  denote the smallest parameter such that

$$\|v - \bar{v}^X\|_{L^2(W)} \leq C_P(Y, X, W; h) \text{diam}(Y) \sqrt{\frac{\text{meas}_{d_W}(W)}{\text{meas}_d(Y)}} |v|_{H^1(Y)} \quad \forall v \in V^h(Y).$$

For  $d_X \geq d - 1$  and  $d_W \geq d - 1$  let  $C_P(Y, X, W)$  denote the smallest constant such that the above inequality holds in  $H^1(Y)$ .

By definition, for  $d_W = d$  and  $d_X \geq d - 1$ , we have

$$C_P(D) \leq C_P(D, X), \quad C_P(D, X, W) \leq C_P(D, X).$$

**Definition 2.68.** For an integer  $j \geq 0$ , we define the indicator function

$$\sigma^j(x) := \begin{cases} 1 & \text{if } j \leq 1, \\ 1 + \log(x) & \text{if } j = 2, \\ x^{j-2} & \text{if } j \geq 3. \end{cases}$$

The following lemma states a discrete trace inequality with explicit dependence on the geometric parameters.

**Lemma 2.69 (discrete trace inequality).** *Let Assumption 2.54 hold, let  $Y$  and  $W$  be  $\mathcal{T}^H$ -agglomerates of dimensions  $d_Y = d$  and  $0 \leq d_W \leq d$  with  $W \subset \bar{Y}$ . Furthermore, let  $\mathcal{T}^h(Y)$  be quasi-uniform. Then*

$$\|v\|_{L^2(W)}^2 \leq C \sigma^{d-d_W} \left(\frac{H_Y}{h}\right) \frac{\text{meas}_{d_W}(W)}{\text{meas}_d(Y)} \left(H_Y^2 |v|_{H^1(Y)}^2 + \|v\|_{L^2(Y)}^2\right) \quad \forall v \in V^h(Y),$$

where  $H_Y = \text{diam}(Y)$ . The constant  $C$  only depends on the shape regularity and quasi-uniformity constants of  $\mathcal{T}^H(\Omega)$  and  $\mathcal{T}^h(Y)$ , and on the number of coarse elements contained in  $Y$ .

*Proof.* For  $d_W = d - 1$ , the estimate follows from the trace theorem and a scaling argument. For the case  $d = 2, d_W = 0$ , see [TW05, Lemma 4.15] and also [BX91]. For  $d = 3, d_W = 0$  we combine the embedding of  $H^1(D)$  in  $L^{2d/(d-2)}(D)$  and the inverse inequality  $\|v\|_{L^\infty(D)} \lesssim h^{-(d-2)/2} \|v\|_{H^1(D)}$  on a domain of unit diameter. Finally, the case  $d = 3, d_W = 1$  is treated by integrating the two-dimensional  $L^\infty$ -estimate along the edges forming  $W$ , see also [TW05] and [PS12b, Lemma A.4].  $\square$

The statement of the following lemma can be found (in different form) in [TW05, Sect. 4.6], see also the references therein.

**Lemma 2.70 (discrete Poincaré type inequality).** *Let Assumption 2.54 hold, let  $Y$  be an agglomerate and let  $X, W$  be agglomerates of dimensions  $d_X, d_W \in [0, d]$ . Furthermore, let  $\mathcal{T}^h(Y)$  be quasi-uniform. Then*

$$\|v - \bar{v}^X\|_{L^2(W)}^2 \leq C \sigma^{d-\min(d_X, d_W)} \left(\frac{H}{h}\right) \frac{\text{meas}_{d_W}(W)}{\text{meas}_d(Y)} H^2 |v|_{H^1(Y)}^2 \quad \forall v \in V^h(Y),$$

where  $H = \text{diam}(Y)$ . The constant  $C$  only depends on the shape regularity and quasi-uniformity constants of  $\mathcal{T}^H(\Omega)$  and  $\mathcal{T}^h(Y)$ , respectively, and on the number of elements that  $Y$  contains.

*Proof.* We have

$$\begin{aligned} \frac{1}{2} \|v - \bar{v}^X\|_{L^2(W)}^2 &\leq \|v - \bar{v}^Y\|_{L^2(W)}^2 + \text{meas}_{d_W}(W) |\bar{v}^Y - \bar{v}^X|^2 \\ &\leq \|v - \bar{v}^Y\|_{L^2(W)}^2 + 2 \frac{\text{meas}_{d_W}(W)}{\text{meas}_{d_X}(X)} \left( \|v - \bar{v}^Y\|_{L^2(X)}^2 + \|v - \bar{v}^X\|_{L^2(X)}^2 \right). \end{aligned}$$

Due to the Ritz minimum principle (cf. Lemma 1.3),

$$\|v - \bar{v}^X\|_{L^2(X)}^2 = \inf_{c \in \mathbb{R}} \|v - c\|_{L^2(X)}^2 \leq \|v - \bar{v}^Y\|_{L^2(X)}^2.$$

Hence, applying Lemma 2.69 for the manifolds  $X$  and  $W$ , we obtain

$$\begin{aligned} \frac{1}{2} \|v - \bar{v}^X\|_{L^2(W)}^2 &\leq \|v - \bar{v}^Y\|_{L^2(W)}^2 + 4 \frac{\text{meas}_{d_W}(W)}{\text{meas}_{d_X}(X)} \|v - \bar{v}^Y\|_{L^2(X)}^2 \\ &\leq \left( \sigma^{d-d_W} \left(\frac{H}{h}\right) + 4 \sigma^{d-d_X} \left(\frac{H}{h}\right) \right) \frac{\text{meas}_{d_W}(W)}{\text{meas}_d(Y)} \left( |v|_{H^1(Y)}^2 + \text{diam}(Y)^{-2} \|v - \bar{v}^Y\|_{L^2(Y)}^2 \right). \end{aligned}$$

Applying Poincaré's inequality on  $Y$  and taking Lemma 2.61 into account concludes the proof.  $\square$

**Corollary 2.71.** *Under the assumptions of Lemma 2.70,*

$$C_P(Y, X, W; h) \leq C \sqrt{\sigma^{d-\min(d_X, d_W)} \left(\frac{H}{h}\right)},$$

$$C_P(Y, X; h) \leq C \sqrt{\sigma^{d-d_X} \left(\frac{H}{h}\right)}.$$

### 2.5.5 Cut-Off Estimates

In this subsection, we estimate the effect of a cut-off in the energy norm. As for Sect. 2.5.4, the following estimates can be found in [TW05]. They mainly stem from the early works [BPS86, BX91, Dry87, DW94, DW95] on iterative substructuring.

**Definition 2.72.** Let  $\mathcal{F}_i$  be a subdomain facet (i.e., a subdomain face if  $d = 3$ , a subdomain edge if  $d = 2$ ) on  $\partial\Omega_i$ . We set

$$\vartheta_{\mathcal{F}_i} := \mathcal{H}_i^h(\theta_{\mathcal{F}_i}) \in V^h(\Omega_i).$$

Here, the space  $V^h(\Omega_i)$  and the discrete harmonic extension  $\mathcal{H}_i^h$  are defined with respect to the (possibly auxiliary) triangulation  $\mathcal{T}^h(\Omega_i)$ . For any  $\mathcal{G}_i$  on  $\partial\Omega_i$  that is not a subdomain facet, we define  $\vartheta_{\mathcal{G}_i} \in V^h(\Omega_i)$  as the extension of  $\theta_{\mathcal{G}_i} \in V^h(\partial\Omega_i)$  that vanishes at all interior nodes of  $\Omega_i$ .

*Remark 2.73.* For simplicity, we did not indicate the subdomain index in the definition above. For any interface glob  $\mathcal{G}$  we may regard  $\vartheta_{\mathcal{G}}$  as a global function in  $V^h(\Omega)$  only supported in the subdomains that share the glob  $\mathcal{G}$ .

For any  $v \in V^h(\Omega_i)$ , the function  $I^h(\vartheta_{\mathcal{G}_i} v)$  is an extension of  $I^h(\theta_{\mathcal{G}_i} v|_{\partial\Omega})$  and so by the minimizing property of  $\mathcal{H}_i^h$  (cf. Definition 1.55),

$$|\mathcal{H}_i^h(\theta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)} \leq |I^h(\vartheta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)}.$$

*Notation.* Above and in the following, with a small abuse of notation, whenever  $v$  is a function with well-defined values at the nodes contained in  $\mathcal{G}_i$ , by

$$\mathcal{H}_i^h(\theta_{\mathcal{G}_i} v) \tag{2.79}$$

we mean that we first interpret the expression  $\theta_{\mathcal{G}_i} v$  as a function in  $V^h(\partial\Omega_i)$  (by interpolating at the nodes contained in  $\mathcal{G}_i$ , and choosing zero at all other nodes) and then apply  $\mathcal{H}_i^h$  to it.

The following lemmas give further bounds in terms of the original function  $v$ . First, we discuss the case of lower-dimensional globs.

**Lemma 2.74.** *Let  $\mathcal{T}^h(\Omega_i)$  be quasi-uniform. Then for any glob  $\mathcal{G}_i$  of dimension  $0 \leq m < d - 1$  (a subdomain vertex or edge if  $d = 3$ , a subdomain vertex if  $d = 2$ ),*

$$|I^h(\vartheta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)}^2 \leq C h_i^{d-2-m} \|v\|_{L^2(\mathcal{G}_i)}^2 \quad \forall v \in V^h(\Omega_i),$$

where  $\|v\|_{L^2(\mathcal{V}_i)} = |v(\mathcal{V}_i)|$ . In particular,

$$|\vartheta_{\mathcal{V}_i}|_{H^1(\Omega_i)}^2 \leq C h_i^{d-2}.$$

For  $d = 3$ , let  $\mathcal{E}_i$  be a subdomain edge such that  $\mathcal{V}_i \in \overline{\mathcal{E}_i}$ . Then

$$|I^h(\vartheta_{\mathcal{V}_i} v)|_{H^1(\Omega_i)}^2 \leq C \|v\|_{L^2(\mathcal{E}_i)}^2 \quad \forall v \in V^h(\Omega_i).$$

If  $d = 3$  and if Assumption 2.54 holds then

$$|\vartheta_{\mathcal{E}_i}|_{H^1(\Omega_i)}^2 \leq C H_i.$$

The constant  $C$  only depends on the shape regularity constants of  $\mathcal{T}^h(\Omega_i)$  (and on the shape regularity of  $\mathcal{T}^H(\Omega)$  in the last estimate).

*Proof.* Let  $v \in V^h(\Omega_i)$  be arbitrary but fixed and note that  $I^h(\vartheta_{\mathcal{G}_i} v)$  vanishes on all nodes of  $\mathcal{T}^h(\Omega_i)$  except those on  $\mathcal{G}_i$ . Then, by an inverse inequality,

$$|I^h(\vartheta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)}^2 \leq C \sum_{x^h \in \mathcal{G}_i} h_i^{d-2} |v(x^h)|^2 \leq C h_i^{d-2-m} \sum_{x^h \in \overline{\mathcal{G}_i}} h_i^m |v(x^h)|^2.$$

The sum in the last expression is equivalent to  $\|v\|_{L^2(\mathcal{G}_i)}^2$ . For  $d = 3$ , we obtain from the above that  $|I^h(\vartheta_{\mathcal{V}_i} v)|_{H^1(\Omega_i)}^2 \leq C h_i |v(\mathcal{V}_i)|^2 \leq C \|v\|_{L^2(\mathcal{E}_i)}^2$ . The last estimate follows from  $\|1\|_{L^2(\mathcal{E}_i)}^2 = \text{meas}_1(\mathcal{E}_i) \leq C H_i$ .  $\square$

Combining Lemma 2.74 with the trace inequality from Lemma 2.69, we obtain the following result.

**Lemma 2.75.** *Let Assumptions 2.53 and 2.54 hold. Then for any glob  $\mathcal{G}_i$  of dimension  $0 \leq m < d - 1$  (a subdomain vertex or edge if  $d = 3$ , a subdomain vertex if  $d = 2$ ),*

$$|I^h(\vartheta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)}^2 \lesssim \left(1 + \log\left(\frac{H_i}{h_i}\right)\right) \left(|v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|v\|_{L^2(\Omega_i)}^2\right) \quad \forall v \in V^h(\Omega_i).$$

*Proof.* First, we select a glob  $\mathcal{X}_i$  of dimension at least  $d - 2$ . If  $d = 3$  and  $m = 0$ , we choose  $\mathcal{X}_i$  as a subdomain edge touching  $\mathcal{G}_i$ , otherwise we set  $\mathcal{X}_i = \mathcal{G}_i$ . Thanks to Lemma 2.74,

$$|I^h(\vartheta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)}^2 \lesssim \|v\|_{L^2(\mathcal{X}_i)}^2.$$

Lemma 2.69 yields

$$\|v\|_{L^2(\mathcal{X}_i)}^2 \lesssim \sigma^2\left(\frac{H_i}{h_i}\right) \frac{\text{meas}_{d-2}(\mathcal{X}_i) H_i^2}{\text{meas}_d(\Omega_i)} \left( |v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|v\|_{L^2(\Omega_i)}^2 \right).$$

Since  $\text{meas}_{d-2}(\mathcal{X}_i) \approx H_i^{d-2}$  and  $\text{meas}_d(\Omega_i) \approx H_i^d$ , this concludes the proof.  $\square$

We now turn to the case of a subdomain facet (i.e., a subdomain face if  $d = 3$ , a subdomain edge if  $d = 2$ ).

**Lemma 2.76.** *Let Assumptions 2.53 and 2.54 hold. Then for any subdomain facet  $\mathcal{F}_i$ ,*

$$|I^h(\vartheta_{\mathcal{F}_i} v)|_{H^1(\Omega_i)}^2 \lesssim \left(1 + \log\left(\frac{H_i}{h_i}\right)\right)^2 \left( |v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|v\|_{L^2(\Omega_i)}^2 \right) \quad \forall v \in V^h(\Omega_i),$$

and

$$|\vartheta_{\mathcal{F}_i}|_{H^1(\Omega_i)}^2 \lesssim \left(1 + \log\left(\frac{H_i}{h_i}\right)\right) H_i^{d-2}.$$

*Proof.* For a proof in three dimensions, we refer to [TW05, Lemmas 4.24 and 4.25]. For the two-dimensional case, see [MB96] as well as the pioneering paper [BPS86]. Sharpness of the estimates is shown in [BS00].

*Remark 2.77.* On a subdomain facet  $\mathcal{F}_i$ , the term  $|\mathcal{H}_i^h(\theta_{\mathcal{F}_i} v)|_{H^1(\Omega_i)}$  realizes the  $H_{00}^{1/2}(\mathcal{F}_i)$ -norm of  $v \in V^h(\mathcal{F}_i)$ , see also [TW05, Lemma 4.26] and Sect. 2.5.7.

Combining the cut-off lemmas and the discrete Poincaré inequality straightforwardly may create several powers of the term  $(1 + \log(H_i/h_i))$ . The following corollaries help to keep a power of two.

**Corollary 2.78.** *Let Assumptions 2.53 and 2.54 hold, and let  $\mathcal{G}_i$  be a glob. Then*

$$|I^h(\vartheta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)}^2 \lesssim \left(1 + \log\left(\frac{H_i}{h_i}\right)\right)^2 |v|_{H^1(\Omega_i)}^2 + \left(1 + \log\left(\frac{H_i}{h_i}\right)\right) H_i^{-2} \|v\|_{L^2(\Omega_i)}^2$$

for all  $v \in V^h(\Omega_i)$ .

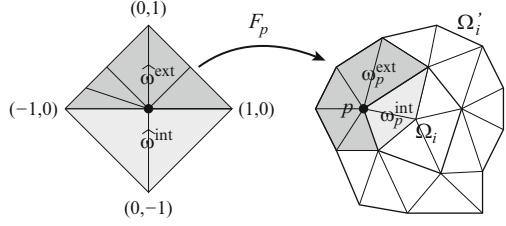
*Proof.* If the dimension of  $\mathcal{G}_i$  is less than  $d - 1$ , the statement follows from Lemma 2.74. If  $\mathcal{G}_i = \mathcal{F}_i$  is a subdomain facet, we have

$$|I^h(\vartheta_{\mathcal{F}_i} v)|_{H^1(\Omega_i)}^2 \lesssim |I^h(\vartheta_{\mathcal{F}_i} (v - \bar{v}^{\Omega_i}))|_{H^1(\Omega_i)}^2 + |\vartheta_{\mathcal{F}_i}|_{H^1(\Omega_i)}^2 |\bar{v}^{\Omega_i}|^2.$$

By Cauchy's inequality (cf. (1.13)) and the estimates from Lemma 2.76, we get

$$\begin{aligned} |I^h(\vartheta_{\mathcal{F}_i} v)|_{H^1(\Omega_i)}^2 &\lesssim \left(1 + \log\left(\frac{H_i}{h_i}\right)\right)^2 \left( |v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|v - \bar{v}^{\Omega_i}\|_{L^2(\Omega_i)}^2 \right) \\ &\quad + \left(1 + \log\left(\frac{H_i}{h_i}\right)\right) H_i^{d-2} |\Omega_i|^{-1} \|v\|_{L^2(\Omega_i)}^2 \end{aligned}$$

**Fig. 2.7** Mapping of a node patch  $\omega_p$  in two dimensions



Poincaré's inequality (Lemma 2.61) and using that  $|\Omega_i| \approx H_i^d$  concludes the proof.  $\square$

**Corollary 2.79.** *Let Assumptions 2.53 and 2.54 hold, and let  $\mathcal{G}_i$ ,  $\mathcal{X}_i$  be two globs of  $\Omega_i$ , such that the dimension of  $\mathcal{X}_i$  is at least  $d - 2$ . Then*

$$|I^h(\vartheta_{\mathcal{G}_i}(v - \bar{v}^{\mathcal{X}_i}))|_{H^1(\Omega_i)}^2 \lesssim (1 + \log(\frac{H_i}{h_i}))^2 |v|_{H^1(\Omega_i)}^2 \quad \forall v \in V^h(\Omega_i).$$

*Proof.* The statement follows immediately from Corollary 2.78 and Lemma 2.70.  $\square$

Summing the estimate from Corollary 2.78 over all globs of a subdomain, we obtain the following stable decomposition (recall our convention (2.79) from p. 122),

$$\begin{aligned} \sum_{\mathcal{G}_i} |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)}^2 &\lesssim (1 + \log(\frac{H_i}{h_i}))^2 |v|_{H^1(\Omega_i)}^2 \\ &+ (1 + \log(\frac{H_i}{h_i})) H_i^{-2} \|v\|_{L^2(\Omega_i)}^2 \quad \forall v \in V^h(\Omega_i). \end{aligned} \quad (2.80)$$

The following lemma is kind of inverse to (2.80).

**Lemma 2.80.** *Let Assumption 2.54 hold. Then*

$$|\mathcal{H}_i^h v|_{H^1(\Omega_i)}^2 \leq C \sum_{\mathcal{G}_i} |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} v)|_{H^1(\Omega_i)}^2 \quad \forall v \in V^h(\partial\Omega_i), \quad v|_{\partial\Omega_i \cap \Gamma_N^h} = 0,$$

where the constant  $C$  only depends on the shape regularity constant of Assumption 2.54.

*Proof.* Since the functions  $\{\theta_{\mathcal{G}_i}\}$  provide a partition of unity on  $\partial\Omega_i \setminus \Gamma_N$ , the function  $\sum_{\mathcal{G}_i} \mathcal{H}_i^h(\theta_{\mathcal{G}_i} v)$  is a discrete extension of  $v$  from  $\partial\Omega_i$  to  $\Omega_i$ . Therefore, the result follows from Lemma 1.54, the triangle inequality, and the fact that the number of globs of  $\Omega_i$  is uniformly bounded (see Assumption 2.54). See also Lemma 3.21, p. 169.  $\square$

**Remark 2.81.** Almost all proofs of the cut-off estimates in this section rely on Assumption 2.54 and so does Lemma 2.69. An extension to less regular subdomains

(so-called John domains that can have ragged boundaries) for two dimensions can be found in [KRW08]. The three-dimensional case seems to be open yet.

### 2.5.6 An Explicit Sobolev Extension

In this subsection, we define a Sobolev extension operator for Lipschitz polytopes in the spirit of Stein [Ste70] and provide an explicit estimate in terms of shape regularity constants only.

For a Lipschitz polytope  $\Omega_i$ , let  $\Omega'_i \supset \overline{\Omega}_i$  be a neighborhood such that both domains are resolved by a shape regular coarse triangulation  $\mathcal{T}^H(\Omega'_i)$  consisting of a bounded number of elements, cf. Assumption 2.55. For an illustration see Fig. 2.7 (right). Let  $\partial\Omega_i^H$  denote the set of coarse vertices of  $\mathcal{T}^H(\Omega'_i)$  that lie on  $\partial\Omega_i$ . For each vertex  $p \in \partial\Omega_i^H$ , we define the open vertex patch  $\omega_p$  by

$$\overline{\omega}_p = \bigcup \{ \overline{T} : T \in \mathcal{T}^H(\Omega'), p \in \overline{T} \},$$

and

$$\omega_p^{\text{int}} := \omega_p \cap \Omega_i, \quad \omega_p^{\text{ext}} := \omega_p \cap (\Omega'_i \setminus \overline{\Omega}_i),$$

cf. Fig. 2.7 (right). Without loss of generality, we assume that  $\omega_p^{\text{int}}$  and  $\omega_p^{\text{ext}}$  each contain at least one coarse vertex that does not lie on  $\partial\Omega_i$ . This condition can always be fulfilled by formally subdividing some of the coarse elements.

We define the open reference patch  $\widehat{\omega}$  by

$$\widehat{\omega} = \begin{cases} \text{conv}(\{(-1, 0), (1, 0), (0, 1), (0, -1)\}) & \text{if } d = 2, \\ \text{conv}(\{(-1, 0, 0), (1, 1, 0), (1, -1, 0), (0, 0, 1), (0, 0, -1)\}) & \text{if } d = 3, \end{cases}$$

where  $\text{conv}(S)$  denotes the convex hull of the set  $S$ . Furthermore, we define the subsets

$$\widehat{\omega}^{\text{int}} := \widehat{\omega} \cap \{x : x_d < 0\}, \quad \widehat{\omega}^{\text{ext}} := \widehat{\omega} \cap \{x : x_d > 0\},$$

where  $x_d$  refers to the  $d$ -th component of  $x$ .

Let  $\mathcal{T}_p(\widehat{\omega})$  be a shape regular simplicial triangulation of  $\widehat{\omega}$  such that there exists a bijective continuous mapping  $F_p : \widehat{\omega} \rightarrow \omega_p$  with the following properties.

- For each element  $T \in \mathcal{T}_p(\widehat{\omega})$ , the restricted mapping  $F_p|_T$  is affine linear,
- $F_p(0) = p$ ,
- $F_p(\widehat{\omega} \cap \{x : x_d = 0\}) = \omega_p \cap \partial\Omega_i$ ,
- $F_p(\widehat{\omega}^{\text{int}}) = \omega_p^{\text{int}}$  and  $F_p(\widehat{\omega}^{\text{ext}}) = \omega_p^{\text{ext}}$ ,
- For each element  $T \in \mathcal{T}_p(\widehat{\omega})$ ,



$$c_1 H_i^d \leq \det(F'_{p|T}) \leq c_2 H_i^d,$$

$$\|F'_{p|T}\|_{\ell^2} \leq c_3 H_i, \quad \|(F'_{p|T})^{-1}\|_{\ell^2} \leq c_4 H_i^{-1},$$

where the constants  $c_1$ ,  $c_2$ ,  $c_3$ , and  $c_4$  only depend on the shape regularity constants of  $\mathcal{T}^H(\Omega_i^H)$ .

For an illustration in two dimensions, see Fig. 2.7. Under the conditions on  $\mathcal{T}^H(\Omega_i)$  stated in Assumption 2.55, such a triangulation and mapping exists for every coarse vertex  $p \in \partial\Omega_i^H$ .

On the reference patch, we define the linear operator

$$\widehat{E} : \mathcal{C}^\infty(\widehat{\omega}^{\text{int}}) \rightarrow \mathcal{C}^\infty(\widehat{\omega}^{\text{ext}}), \quad (\widehat{E}w)(x_1, \dots, x_d) := w(x_1, \dots, x_{d-1}, -x_d),$$

i.e., the reflection of  $w$  across the hyperplane  $\{x : x_d = 0\}$ . By construction we have  $\|\widehat{E}w\|_{H^1(\widehat{\omega}^{\text{ext}})} = \|w\|_{H^1(\widehat{\omega}^{\text{int}})}$ . Since the  $\mathcal{C}^\infty$ -functions are dense in  $H^1$ , it follows that  $\widehat{E} : H^1(\widehat{\omega}^{\text{int}}) \rightarrow H^1(\widehat{\omega}^{\text{ext}})$  is linear and continuous. For each coarse node  $p \in \partial\Omega_i^H$ , we define the linear operator

$$E^{(p)} : H^1(\omega_p^{\text{int}}) \rightarrow H^1(\omega_p^{\text{ext}}), \quad E^{(p)}v := (\widehat{E}(v \circ F_p)) \circ F_p^{-1}.$$

Since  $F_p$  is continuous and piecewise affine linear,  $E^{(p)}$  maps indeed into  $H^1$  and is linear and continuous. Furthermore, one easily shows that

$$(E^{(p)}v)|_{\omega_p \cap \partial\Omega_i} = v|_{\omega_p \cap \partial\Omega_i}.$$

Finally, we define the extension operator

$$E_i : H^1(\Omega_i) \rightarrow H^1(\Omega_i'), \quad E_i v := \begin{cases} v & \text{in } \Omega_i, \\ \sum_{p \in \partial\Omega_i^H} \varphi_p E^{(p)}v & \text{in } \Omega_i' \setminus \Omega_i, \\ 0 & \text{else.} \end{cases} \quad (2.81)$$

where  $\varphi_p$  is the nodal finite element basis function on  $\mathcal{T}^H(\Omega_i')$  associated with the coarse node  $p$ .

**Lemma 2.82.** *Let  $\Omega_i$  be a subdomain and  $\Omega_i' \supset \overline{\Omega_i}$  a Lipschitz domain that fully contains  $\Omega_i$  and that is resolved by a shape regular coarse triangulation  $\mathcal{T}^H(\Omega_i')$ . Then the extension operator  $E_i$  as defined above maps into  $H^1(\mathbb{R}^d)$ . Furthermore, there exists a constant  $C_{E_i}$  depending only on the number of coarse elements in  $\Omega_i$  and on the shape regularity constants of  $\mathcal{T}^H(\Omega_i')$  such that*

$$|E_i v|_{H^1(\mathbb{R}^d)}^2 + H_i^{-2} \|E_i v\|_{L^2(\mathbb{R}^d)}^2 \leq C_{E_i} \left( |v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|v\|_{L^2(\Omega_i)}^2 \right) \quad \forall v \in H^1(\Omega_i).$$

*Proof.* Let  $v \in H^1(\Omega_i)$  be arbitrary but fixed. For each coarse vertex  $p_j$ ,  $j = 1, \dots, J$ , the function  $\varphi_{p_j} E^{(p_j)} v$  vanishes on  $\mathbb{R}^d \setminus (\overline{\Omega}_i \cup \omega_p^{\text{ext}})$ . Hence,

$$(E_i v)|_{\mathbb{R}^d \setminus \overline{\Omega}_i} \in H^1(\mathbb{R}^d \setminus \overline{\Omega}_i).$$

Thanks to the partition of unity property

$$\sum_{j=1}^J \varphi_{p_j}(x) = 1 \quad \forall x \in \partial\Omega_i,$$

we can conclude that  $(E_i v)|_{\partial\Omega_i} = v|_{\partial\Omega_i}$ . Since in addition  $(E_i v)|_{\partial\Omega'_i} = 0$ , it follows that  $E_i v \in H^1(\mathbb{R}^d)$ . With standard finite element techniques (see e.g. [BS02, Cia87]), one shows that for each coarse node  $p$  on  $\partial\Omega_j$ ,

$$|E^{(p)} v|_{H^1(\omega_p^{\text{ext}})} \leq C |v|_{H^1(\omega_p^{\text{int}})}, \quad \|E^{(p)} v\|_{L^2(\omega_p^{\text{ext}})} \leq C \|v\|_{L^2(\omega_p^{\text{int}})}.$$

The constant  $C$  is uniform because there are only a small number of different triangulations  $\mathcal{T}_p(\widehat{\omega})$ . Since  $\|\varphi_p\|_{L^\infty} = 1$ , it follows from the above that

$$\|\varphi_p E^{(p)} v\|_{L^2(\omega_p^{\text{ext}})} \leq C \|v\|_{L^2(\omega_p^{\text{int}})}.$$

Summing over  $p \in \partial\Omega_i^H$ , we obtain

$$\|E_i v\|_{L^2(\Omega'_i \setminus \Omega_i)} \leq \sum_{p \in \partial\Omega_i^H} C \|v\|_{L^2(\omega_p^{\text{int}})} \leq C \|v\|_{L^2(\Omega_i)}, \quad (2.82)$$

where we have used that the number of coarse nodes on  $\partial\Omega_i$  is uniformly bounded.

We now turn to the  $H^1$ -seminorm. Since  $\|\nabla \varphi_p\|_{L^\infty} \leq C H_i^{-1}$ , we can conclude from the product rule and the local  $L^2$ -estimate from above that

$$\begin{aligned} |\varphi_p(E^{(p)} v)|_{H^1(\omega_p^{\text{ext}})}^2 &\leq C \left( |E^{(p)} v|_{H^1(\omega_p^{\text{ext}})}^2 + H_i^{-2} \|E^{(p)} v\|_{L^2(\omega_p^{\text{ext}})}^2 \right) \\ &\leq C \left( |v|_{H^1(\omega_p^{\text{int}})}^2 + H_i^{-2} \|v\|_{L^2(\omega_p^{\text{int}})}^2 \right). \end{aligned}$$

Summing over all coarse nodes  $p \in \partial\Omega_i^H$ , we get

$$|E_i v|_{H^1(\Omega'_i \setminus \Omega_i)}^2 \leq C \left( |v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|v\|_{L^2(\Omega_i)}^2 \right) \quad (2.83)$$

Using the definition of  $E_i$  and combining (2.82) and (2.83) concludes the proof.  $\square$

*Remark 2.83.* The relaxation of Assumptions 2.54 and 2.55 is possible by using Jones' extension operator introduced in [Jon81], which is defined for the so-called  $\varepsilon$ - $\delta$  domains or uniform domains, and which employs a bound of the same form as

in Lemma 2.82. The constant can be made explicit in a geometric parameter linked to the domain, see also [DKW08b, DKW08a, KRW08]. As for Poincaré's constant, ragged boundaries alone do not make the extension constant blow up. For a degree-independent extension operator see also [Rog06].

*Remark 2.84.* The extension operator  $E_i$  is also well-defined and bounded from  $L^2(\Omega_i)$  to  $L^2(\mathbb{R}^d)$ , and by classical interpolation theory (see [AF03, Sect. 7.22f]) it follows that it is bounded from  $H^s(\Omega_i)$  to  $H^s(\mathbb{R}^d)$  for  $s \in [0, 1]$ , cf. [Ste70].

**Lemma 2.85.** *The operator  $\tilde{E}_i : H^1(\Omega_i) \rightarrow H^1(\Omega'_i)$  defined by*

$$\tilde{E}_i v := \begin{cases} v & \text{in } \Omega_i, \\ \bar{v}^{\Omega_i} + \sum_{p \in \partial\Omega_i^H} \varphi_p E^{(p)}(v - \bar{v}^{\Omega_i}) & \text{in } \Omega'_i \setminus \Omega_i \end{cases}$$

(cf. (2.81)) is linear and continuous, and it preserves constants, i.e.,

$$\tilde{E}_i c = c \quad \forall c \in \mathbb{R}. \quad (2.84)$$

Furthermore, we have the separate stability estimates

$$\left. \begin{aligned} \|\tilde{E}_i v\|_{L^2(\Omega'_i)} &\leq C \|v\|_{L^2(\Omega_i)} \\ |\tilde{E}_i v|_{H^1(\Omega'_i)} &\leq C |v|_{H^1(\Omega_i)} \end{aligned} \right\} \quad \forall v \in H^1(\Omega_i),$$

where the constant  $C$  depends only on the number of coarse elements in  $\Omega_i$  and on the shape regularity constants of  $\mathcal{T}^H(\Omega'_i)$ .

*Proof.* The proof of the  $L^2$ -stability follows from (2.82), from Cauchy's inequality, and the fact that  $\text{meas}_d(\Omega'_i \setminus \Omega_i) \approx \text{meas}_d(\Omega_i)$ . The  $H^1$ -stability immediately follows from Poincaré's inequality (Lemma 2.61).  $\square$

### 2.5.7 A Subdomain Transfer Operator for FE Functions

In this subsection, we construct an extension operator that transfers finite element functions from a subdomain to its neighboring subdomains. This technique was first used in [KRW08, Lemma 4.5], see also the references therein. Note that the original analysis of FETI [KW01] was performed using trace norms, which we can circumvent by the transfer operators. Such, one can make the dependence on the shapes of the subdomains more explicit.

**Definition 2.86.** For any glob  $\mathcal{G}$ , the glob patch  $\mathcal{U}_{\mathcal{G}}$  is given by

$$\overline{\mathcal{U}_{\mathcal{G}}} := \bigcup_{k \in \mathcal{N}_{\mathcal{G}}} \overline{\Omega_k}.$$

**Lemma 2.87.** *Let Assumption 2.54 hold. Then, for any glob  $\mathcal{G}_i$  there exists a discrete extension operator  $E_{i,\mathcal{G}_i}^h : V^h(\Omega_i) \rightarrow V^h(\mathcal{U}_{\mathcal{G}_i})$  such that  $(E_{i,\mathcal{G}_i}^h v)|_{\Omega_i} = v$  and*

$$\left. \begin{aligned} |E_{i,\mathcal{G}_i}^h v|_{H^1(\mathcal{U}_{\mathcal{G}_i})} &\leq C |v|_{H^1(\Omega_i)} \\ \|E_{i,\mathcal{G}_i}^h v\|_{L^2(\mathcal{U}_{\mathcal{G}_i})} &\leq C \|v\|_{L^2(\Omega_i)} \end{aligned} \right\} \quad \forall v \in V^h(\Omega_i),$$

where the constant  $C$  only depends on the shape regularity constant of  $\mathcal{T}^H(\Omega)$ . In particular,  $E_{i,\mathcal{G}_i}^h$  preserves the (fine) nodal values on  $\overline{\mathcal{G}_i}$  and it preserves constants.

*Proof.* Assume that  $\mathcal{G}_i \subset \Gamma$  (otherwise  $\mathcal{U}_{\mathcal{G}_i} = \Omega_i$  and the extension is trivial). Therefore,  $\mathcal{U}_{\mathcal{G}_i} \setminus \Omega_i$  contains at least one coarse element and  $\text{meas}_d(\mathcal{U}_{\mathcal{G}_i} \setminus \Omega_i) \approx \text{meas}_d(\Omega_i)$ . Let  $\widetilde{E}_{i,\mathcal{G}_i} : H^1(\Omega_i) \rightarrow H^1(\mathcal{U}_{\mathcal{G}_i})$  be defined analogously to Lemma 2.85, with the only modification that we work on the coarse mesh on  $\mathcal{U}_{\mathcal{G}_i}$  instead of  $\Omega_i'$  and that we sum only the reflections corresponding to the coarse nodes on  $\overline{\mathcal{G}_i}$ . Doing so, we obtain

$$\left. \begin{aligned} \|\widetilde{E}_{i,\mathcal{G}_i} v\|_{L^2(\mathcal{U}_{\mathcal{G}_i})} &\leq C \|v\|_{L^2(\Omega_i)} \\ |\widetilde{E}_{i,\mathcal{G}_i} v|_{H^1(\mathcal{U}_{\mathcal{G}_i})} &\leq C |v|_{H^1(\Omega_i)} \end{aligned} \right\} \quad v \in H^1(\Omega_i). \quad (2.85)$$

Note that, again, the extension preserves constants:  $\widetilde{E}_{i,\mathcal{G}_i} c = c$  for all  $c \in \mathbb{R}$ . Let  $\Pi^h : H^1(\mathcal{U}_{\mathcal{G}_i} \setminus \Omega_i) \rightarrow V^h(\mathcal{U}_{\mathcal{G}_i} \setminus \Omega_i)$  denote the Scott-Zhang operator, which preserves piecewise linear data on the boundary. We set

$$E_{i,\mathcal{G}_i}^h v := \begin{cases} v & \text{in } \Omega_i, \\ \Pi^h \widetilde{E}_{i,\mathcal{G}_i} v & \text{in } \mathcal{U}_{\mathcal{G}_i} \setminus \Omega_i. \end{cases}$$

Indeed,  $E_{i,\mathcal{G}_i}^h v$  is continuous across  $\partial\Omega_i \cap \overline{\mathcal{U}_{\mathcal{G}_i}}$  and the stability estimates follow immediately from (2.85) and Lemma 1.45.  $\square$

*Remark 2.88.* We note that the essential assumption in Lemma 2.87 is that the target domain  $\mathcal{U}_{\mathcal{G}_i}$  obeys a shape regular coarse triangulation consisting of a bounded number of coarse elements.

## 2.5.8 Uniform Bounds for the Constants $c_0$ and $c_K$

Let  $V_i$  denote the single layer potential operator on  $\partial\Omega_i$  and  $D_i$  the corresponding hypersingular integral operator. Moreover, let  $c_{0,i}$  and  $c_{K,i}$  denote the respective constants from Lemma 1.77. In this subsection, we show that under the assumptions made in Sect. 2.5.2, at least in three dimensions, we can bound  $c_{0,i}$  from below in

terms of the shape regularity constants of  $\mathcal{T}^H(\Omega)$  and  $\mathcal{T}^H(\Omega'_i)$ . The following presentation is based on [Pec12].

To get a bound for  $c_0$ , we introduce special trace norms.

**Definition 2.89.** For  $v \in H^{1/2}(\partial\Omega_i)$  we define

$$\|v\|_{\star, H^{1/2}(\partial\Omega_i)} := \left( |\mathcal{H}_i v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|\mathcal{H}_i v\|_{L^2(\Omega_i)}^2 \right)^{1/2}$$

and

$$\|w\|_{\star, H^{-1/2}(\partial\Omega_i)} := \sup_{v \in H^{1/2}(\partial\Omega_i)} \frac{\langle w, v \rangle}{\|v\|_{\star, H^{1/2}(\partial\Omega_i)}}.$$

According to Definition 1.41 and Sect. 1.2.1.6, these norms are equivalent to  $\|\cdot\|_{H^{1/2}(\partial\Omega_i)}$  and  $\|\cdot\|_{H^{-1/2}(\partial\Omega_i)}$ .

**Definition 2.90.** Assume that  $V_i$  is coercive and let  $c_{V_i}^*$  and  $c_{D_i}^*$  be the largest constants and  $C_{P,i}^*$  the smallest constant such that

$$\begin{aligned} \langle w, V_i w \rangle &\geq c_{V_i}^* \|w\|_{\star, H^{-1/2}(\partial\Omega_i)}^2 & \forall w \in H^{-1/2}(\partial\Omega_i), \\ \langle D_i v, v \rangle &\geq c_{D_i}^* |\mathcal{H}_i v|_{H^1(\Omega_i)}^2 & \forall v \in H^{1/2}(\partial\Omega_i), \\ \|\mathcal{H}_i v\|_{L^2(\Omega_i)} &\leq C_{P,i}^* H_i |\mathcal{H}_i v|_{H^1(\Omega_i)} & \forall v \in H_*^{1/2}(\partial\Omega_i). \end{aligned}$$

**Lemma 2.91.** We have that

$$c_{0,i} \geq \frac{c_{V_i}^* c_{D_i}^*}{1 + (C_{P,i}^*)^2}.$$

*Proof.* With a standard duality argument, it follows that

$$\langle V_i^{-1} v, v \rangle \leq (c_{V_i}^*)^{-1} \|v\|_{\star, H^{1/2}(\partial\Omega_i)}^2 \quad \forall v \in H^{1/2}(\partial\Omega_i).$$

Using the definition of  $c_{0,i}$ , Definition 2.90, and the above boundedness result, we obtain

$$c_{0,i} = \inf_{v \in H_*^{1/2}(\partial\Omega_i)} \frac{\langle D_i v, v \rangle}{\langle V_i^{-1} v, v \rangle} \leq \inf_{v \in H_*^{1/2}(\partial\Omega_i)} \frac{c_{D_i}^* |\mathcal{H}_i v|_{H^1(\Omega_i)}^2}{(c_{V_i}^*)^{-1} (|\mathcal{H}_i v|_{H^1(\Omega_i)}^2 + H_i^{-2} \|\mathcal{H}_i v\|_{L^2(\Omega_i)}^2)}.$$

Using the definition of  $C_{P,i}^*$  concludes the proof.  $\square$

**Lemma 2.92.** Let Assumption 2.55 holds. Then in three dimensions,

$$c_{V_i}^* \geq \frac{1}{2 C_{E_i}^2},$$

where  $C_{E_i}$  is the constant from Lemma 2.82. Hence  $c_{D_i}^*$  can be bounded from below in terms of the shape regularity constants of  $\mathcal{T}^H(\Omega'_i)$ .

*Proof.* For a proof see [Pec12, Corollary 6.2]. It follows basically the line of the standard coercivity proof (see e.g. [Ste08, Sect. 6.6.1]), but uses the carefully chosen norms and the extension operator.

To bound the constant for the hypersingular integral operator, we need another extension operator which extends functions in the annulus  $\Omega'_i \setminus \overline{\Omega}_i$  back to  $\Omega_i$ . By changing the roles of  $\Omega_i$  and the annulus in the construction of Lemma 2.85, we get the extension operator

$$\widetilde{E}'_i : H^1(\Omega'_i \setminus \overline{\Omega}_i) \rightarrow H^1(\Omega'_i)$$

which preserves constants and obeys the following stability estimates.

**Lemma 2.93.** *There exists a constant  $C_{\widetilde{E}'_i}$  depending only on the maximal number of coarse elements per subdomain and on the shape regularity constants of  $\mathcal{T}^H(\Omega'_i)$  such that for each subdomain  $\Omega_i$ ,*

$$\left. \begin{aligned} |\widetilde{E}'_i v|_{H^1(\Omega'_i)} &\leq C_{\widetilde{E}'_i} |v|_{H^1(\Omega'_i \setminus \overline{\Omega}_i)} \\ \|\widetilde{E}'_i v\|_{L^2(\Omega'_i)} &\leq C_{\widetilde{E}'_i} \|v\|_{L^2(\Omega'_i \setminus \overline{\Omega}_i)} \end{aligned} \right\} \quad \forall v \in H^1(\Omega'_i \setminus \overline{\Omega}_i).$$

**Lemma 2.94.** *Let Assumption 2.55 hold. Then*

$$c_{D_i}^* \geq \frac{1}{2 C_{\widetilde{E}'_i}^2},$$

where  $C_{\widetilde{E}'_i}^2$  is the constant from Lemma 2.93. Hence  $c_{D_i}^*$  can be bounded from below in terms of the shape regularity constants of  $\mathcal{T}^H(\Omega'_i)$ .

*Proof.* For a proof see [Pec12, Lemma 6.4].

**Lemma 2.95.** *Let Assumption 2.55 holds. Then in three dimensions,*

$$C_{P,i}^* \leq \left[ 2 C_P(\Omega_i)^2 + \frac{1 + C_P(\Omega_i)^2}{C_{\widetilde{E}_i}^2} \right]^{1/2}.$$

Hence,  $C_{P,i}^*$  can be bounded from above in terms of the shape regularity constants from  $\mathcal{T}^H(\Omega'_i)$ .

*Proof.* For a proof see [Pec12, Lemma 6.7].

**Corollary 2.96.** *Let Assumption 2.55 holds. Then in three dimensions, the constant  $c_{0,i}$  can be bounded from below and  $c_{K,i}$  from above in terms of the shape regularity constants of  $\mathcal{T}^H(\Omega_i^l)$ .*

*Remark 2.97.* The generalization of these uniformity results to the case of two dimensions is not yet known, mainly due to the particularity of the two-dimensional exterior problem, see Sect. 1.3.3, or one might say due to the logarithm in the fundamental solution. Ideas towards such estimates are given in [Pec12, Remark 4].

### 2.5.9 An Elementary Inequality

The following lemma provides an important inequality involving the weighted counting functions from (2.53). See e.g. [DW95, MB96, Sar94] for early works where this result has been used.

**Lemma 2.98.** *For  $x^h \in \Gamma_S^h$ , let  $\{\rho_k(x^h)\}_{k \in \mathcal{N}_{x^h}}$  be arbitrary positive weights and let  $\{\delta_k^\dagger(x^h)\}_{k \in \mathcal{N}_{x^h}}$  be defined as in (2.53). Then*

$$\rho_i(x^h) (\delta_j^\dagger(x^h))^2 \leq \min(\rho_i(x^h), \rho_j(x^h)) \quad \forall i, j \in \mathcal{N}_{x^h}$$

for any choice of the exponent  $\gamma \in [1/2, \infty)$ . The same estimate holds for the choice (2.54), which corresponds to  $\gamma \rightarrow \infty$ .

*Proof.* For simplicity, we drop the dependence on  $x^h$  during the proof. Recall that

$$\delta_j^\dagger = \frac{\rho_j^\gamma}{\sum_{k \in \mathcal{N}} \rho_k^\gamma}.$$

Assume that  $\#(\mathcal{N}) \geq 2$  (otherwise the inequality is trivial). Since  $0 < \delta_j^\dagger \leq 1$ , we immediately get that  $\rho_i (\delta_j^\dagger)^2 \leq \rho_i$ . Secondly,

$$\rho_i (\delta_j^\dagger)^2 \leq \frac{\rho_i \rho_j^{2\gamma}}{(\rho_i + \rho_j)^{2\gamma}} = \frac{\rho_i}{(\rho_i + \rho_j)} \frac{\rho_j^{2\gamma-1}}{(\rho_i + \rho_j)^{2\gamma-1}} \rho_j.$$

The first factor on the right-hand side is less than one. Since the function  $y \mapsto y^{2\gamma-1}$  is monotonically non-decreasing for  $\gamma \in [1/2, \infty)$ , the second factor is less than one as well. This implies  $\rho_i (\delta_j^\dagger)^2 \leq \rho_j$ . The proof of the estimate for the choice (2.54) (corresponding to  $\gamma \rightarrow \infty$ ) is trivial.  $\square$

## 2.6 Preconditioner Analysis for Subdomain Resolved Coefficients

In this section, we give the complete convergence analysis of FETI/BETI (both classical and all-floating) for the case that the diffusion coefficient  $\mathcal{A}$  is isotropic and piecewise constant in each subdomain, and with two further assumptions stated below.

**Assumption 2.99.** In each subdomain  $\Omega_i$ , there exists a constant  $\alpha_i > 0$  such that

$$\mathcal{A}|_{\Omega_i} = \alpha_i I.$$

**Assumption 2.100.** The *coefficient scaling* is used, i.e.,  $\rho_i(x^h) = \alpha_i$ , cf. Sect. 2.2.4.2.

*Remark 2.101.* Note that the theory below can be carried over to matrix-valued coefficients with mild anisotropy that are piecewise constant in each subdomain. In such a case, the value of  $\rho_i(x^h)$  should be chosen as the maximal eigenvalue of the matrix coefficient on  $\Omega_i$ .

**Assumption 2.102.** For the classical formulation of FETI/BETI in three dimensions only, we assume that  $\Gamma_D \cap \partial\Omega_i$  is either empty or contains at least a subdomain edge (i.e., it should not collapse to a subdomain vertex).

Recall the missing estimate of the operator  $P_D$  from Lemma 2.45,

$$|P_D(w + z_w)|_S^2 \leq \omega |w|_S^2 \quad \forall w \in W.$$

**Lemma 2.103.** For each  $i = 1, \dots, s$ , let  $W_i^\perp$  be a subspace of  $W_i$  such that the sum

$$W_i = \ker(S_i) \oplus W_i^\perp$$

is direct. Let  $W^\perp := \prod_{i=1}^s W_i^\perp$ . Then

$$|P_D(w + z_w)|_S^2 \leq \omega |w|_S^2 \quad \forall w \in W^\perp$$

implies

$$|P_D(w + z_w)|_S^2 \leq \omega |w|_S^2 \quad \forall w \in W.$$

*Proof.* Let  $w \in W^\perp$  be arbitrary but fixed. First, Lemma 2.44 states that the mapping  $v \mapsto z_v$  (for  $v \in W$ ) is linear and that  $z_y = -y$  for  $y \in \ker(S)$ . Therefore, we have the invariants

$$\left. \begin{aligned} w + z_w &= (w + y) + (z_{w+y}) \\ |w + y|_S &= |w|_S \end{aligned} \right\} \quad \forall y \in \ker(S).$$



Since  $W = W^\perp \oplus \ker(S)$  the second inequality follows immediately.  $\square$

Throughout this section, we choose

$$W_i^\perp := \left\{ w_i \in W_i : \overline{\mathcal{H}_i^h w_i}^{\Omega_i} = 0 \right\} \quad \text{if } i \in \mathcal{I}_{\text{float}}, \quad (2.86)$$

and  $W_i^\perp := W_i$  else. The strategy is to show bounds for  $P_D w$  and  $P_D z_w$  separately, for  $w \in W^\perp$ .

### 2.6.1 An Energy Estimate of $P_D$

The following lemma is essential for the condition number estimate and goes back to [KW01, Lemma 4.7], see also [MT96, Sect. 3.3] and [TW05, Lemma 6.3].

**Lemma 2.104.** *Let Assumptions 2.53–2.56 and 2.99–2.102 hold. Then*

$$|P_D w|_S^2 \lesssim \max_{i=1,\dots,s} (1 + \log(H_i/h_i))^2 |w|_S^2 \quad \forall w \in W^\perp.$$

*Proof.* Let  $w \in W^\perp$  and  $i = 1, \dots, s$  be fixed. Recall the characterization (2.70) of the  $P_D$  operator (p. 108), which reveals that there is no contribution from  $\Gamma_N^h$  (the non-coupling Neumann nodes). Recall also that  $\mathcal{N}_{x^h} = \mathcal{N}_{\mathcal{G}_i}$  for all  $x^h \in \mathcal{G}_i$ . Using the BEM-FEM spectral equivalence from Corollary 1.94 and the cut-off result from Lemma 2.80 (with convention (2.79) from p. 122) we obtain

$$\begin{aligned} |(P_D w)_i|_{S_i}^2 &\lesssim \alpha_i |\mathcal{H}_i^h(P_D w)_i|_{H^1(\Omega_i)}^2 \\ &\lesssim \underbrace{\alpha_i \sum_{\mathcal{G}_i \subset \Gamma} \left| \mathcal{H}_i^h \left( \theta_{\mathcal{G}_i} \sum_{j \in \mathcal{N}_{\mathcal{G}_i}} \delta_j^\dagger (w_i - w_j) \right) \right|_{H^1(\Omega_i)}^2}_{=: \gamma_i} + \alpha_i \sum_{\mathcal{G}_i \subset \Gamma_D} |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} w_i)|_{H^1(\Omega_i)}^2. \end{aligned} \quad (2.87)$$

Since  $\rho_i(x^h) = \alpha_i$ , the functions  $\delta_j^\dagger$  are constant on each glob  $\mathcal{G}_i$ . Using the fact that each glob is shared by a uniformly bounded number of subdomains and the elementary inequality from Lemma 2.98, we obtain

$$\begin{aligned} \gamma_i &\lesssim \sum_{\mathcal{G}_i \subset \Gamma} \sum_{j \in \mathcal{N}_{\mathcal{G}_i}} \alpha_i (\delta_j^\dagger)_{|\mathcal{G}_i}^2 |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} (w_i - w_j))|_{H^1(\Omega_i)}^2 \\ &\leq \sum_{\mathcal{G}_i \subset \Gamma} \sum_{j \in \mathcal{N}_{\mathcal{G}_i}} \min(\alpha_i, \alpha_j) |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} (w_i - w_j))|_{H^1(\Omega_i)}^2 \end{aligned}$$

Inserting the estimate for  $\mathcal{V}_i$  into (2.87) and using the triangle inequality yields

$$|(P_D w)_i|_{S_i}^2 \quad (2.88)$$

$$\begin{aligned} &\lesssim \sum_{\mathcal{G}_i \subset \Gamma} \sum_{j \in \mathcal{N}_{\mathcal{G}_i}} \min(\alpha_i, \alpha_j) |\mathcal{H}_i^h(\theta_{\mathcal{G}_i}(w_i - w_j))|_{H^1(\Omega_i)}^2 + \sum_{\mathcal{G}_i \subset \Gamma_D} \alpha_i |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} w_i)|_{H^1(\Omega_i)}^2 \\ &\lesssim \sum_{\mathcal{G}_i \subset \Gamma \cup \Gamma_D} \sum_{j \in \mathcal{N}_{\mathcal{G}_i}} \alpha_j |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} w_j)|_{H^1(\Omega_i)}^2. \end{aligned} \quad (2.89)$$

Note that the function  $w_j$  from  $\Omega_j$  is cut down to  $\mathcal{G}_{ij}$  and harmonically extended to the (possibly different) subdomain  $\Omega_i$ . In order to estimate the energy of the extension in terms of the energy of the original function, we use the transfer operator from Sect. 2.5.7. Let  $\tilde{w}_j = \mathcal{H}_{ij}^h w_j \in V^h(\Omega_j)$ . Then the function

$$I^h(\vartheta_{\mathcal{G}_i} E_{j, \mathcal{G}_i}^h \tilde{w}_j)$$

extends  $\theta_{\mathcal{G}_i} w_j$  to  $\Omega_i$ . Using Corollary 2.78 and Lemma 2.87, we get

$$\begin{aligned} |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} w_j)|_{H^1(\Omega_i)}^2 &\leq |I^h(\vartheta_{\mathcal{G}_i} E_{j, \mathcal{G}_i}^h \tilde{w}_j)|_{H^1(\Omega_i)}^2 \\ &\lesssim (1 + \log(H_i/h_i))^2 |E_{j, \mathcal{G}_i}^h \tilde{w}_j|_{H^1(\Omega_i)}^2 + H_i^{-2} (1 + \log(H_i/h_i)) \|E_{j, \mathcal{G}_i}^h \tilde{w}_j\|_{L^2(\Omega_i)}^2 \\ &\lesssim (1 + \log(H_i/h_i))^2 |\mathcal{H}_j^h w_j|_{H^1(\Omega_j)}^2 + H_i^{-2} (1 + \log(H_i/h_i)) \|\mathcal{H}_j^h w_j\|_{L^2(\Omega_j)}^2. \end{aligned}$$

If  $\Omega_j$  is floating,  $w_j \in W_j^\perp$ , and we can eliminate the  $L^2$ -term using Poincaré's inequality. If  $\Omega_j$  is non-floating, we know that  $w_j$  vanishes at least on a glob of dimension  $d - 2$  (see Assumption 2.102). Hence, we can eliminate the  $L^2$ -term at the cost of another factor of  $(1 + \log(H_i/h_i))$  using Lemma 2.70. In either case, what we obtain is

$$|\mathcal{H}_i^h(\theta_{\mathcal{G}_i} w_j)|_{H^1(\Omega_i)}^2 \lesssim (1 + \log(H_i/h_i))^2 |\mathcal{H}_j^h w_j|_{H^1(\Omega_j)}^2. \quad (2.90)$$

Combining (2.89) with (2.90) and using that the number of globs and neighbors per subdomain is uniformly bounded yields

$$\sum_{i=1}^s |(P_D w)_i|_{S_i}^2 \lesssim \max_{i=1}^s (1 + \log(H_i/h_i))^2 \sum_{j=1}^s \alpha_j |\mathcal{H}_j^h w_j|_{H^1(\Omega_j)}^2.$$

The proof is concluded by the fact that  $\alpha_j |\mathcal{H}_j^h w_j|_{H^1(\Omega_j)}^2 = |w_j|_{S_j}^2$  for  $j \in \mathcal{J}_{\text{FEM}}$  and  $\alpha_j |\mathcal{H}_j^h w_j|_{H^1(\Omega_j)}^2 \lesssim |w_j|_{S_j}^2$  for  $j \in \mathcal{J}_{\text{BEM}}$  (by Corollary 1.94).  $\square$

### 2.6.2 The Case $Q = M_{\text{SD}}^{-1}$

To complete the estimate from Lemma 2.45, we still need to bound  $|P_D z_w|_S$ . The next lemma stems from [KW01, Lemma 4.8].

**Lemma 2.105.** *Let  $z_w$  be defined as in Lemma 2.44 and let  $Q = M_{\text{SD}}^{-1}$ . Then*

$$|P_D z_w|_S \leq |P_D w|_S \quad \forall w \in W.$$

*Proof.* Using identity (2.67), i.e.,  $P_D^\top S P_D = B^\top M_{\text{SD}}^{-1} B = B^\top Q B$ , and Lemma 2.44 we have  $|P_D z_w|_S^2 = \|B z_w\|_Q^2 \leq \|B w\|_Q^2 = |P_D w|_S^2$ .  $\square$

The next theorem estimates the condition number of classical an all-floating FETI/BETI.

**Theorem 2.106.** *Let Assumptions 2.53–2.56 and 2.99–2.102 hold. Then, for the classical or the all-floating FETI/BETI method with the scaled Dirichlet preconditioner  $M_{\text{SD}}^{-1}$  and with the choice  $Q = M_{\text{SD}}^{-1}$ ,*

$$\kappa(P M_{\text{SD}}^{-1} P^\top F_{|\tilde{U}_{\text{ad}}}) \leq C \max_{i=1}^s (1 + \log(H_i / h_i))^2,$$

where the constant  $C$  depends only on the uniform constants from the mentioned assumptions. If the subdomain meshes are not quasi-uniform (but still shape-regular), the analogous bound holds but then  $h_i$  has to be replaced by the minimal element diameter of  $\mathcal{T}^h(\Omega_i)$ .

*Proof.* The estimate follows immediately by combining Lemmas 2.45 and 2.103–2.105.  $\square$

### 2.6.3 Diagonal Choice of $Q$

An implementation of the FETI/BETI method with  $Q$  a diagonal matrix is of course much easier. The following choice, proposed and analyzed by Klawonn and Widlund [KW01], still gives a robust method with respect to coefficient jumps.

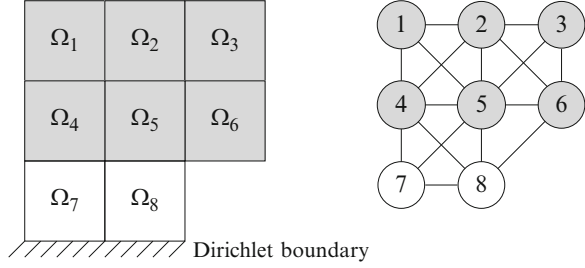
**Definition 2.107.** For each node  $x^h \subset \Gamma^h \cup \Gamma_D^h$  we define

$$q_i(x^h) := \begin{cases} (1 + \log(H_i / h_i)) \frac{h_i^{d-1}}{H_i} & \text{if } x^h \text{ lies on a subdomain facet,} \\ h_i^{d-2} & \text{else.} \end{cases}$$

Furthermore, we define the operator  $Q_{\text{diag}} : U^* \rightarrow U$  by

$$(Q_{\text{diag}} \mu)_{ij}(x^h) := \min(\rho_i(x^h), \rho_j(x^h)) q_{ij}(x^h) \mu_{ij}(x^h) \quad \text{for } \mu \in U^*, \quad (2.91)$$

**Fig. 2.8** *Left:* Subdomains with Dirichlet boundary.  
*Right:* Corresponding connectivity graph



where  $q_{ij}(x^h) = \min(q_i(x^h), q_j(x^h))$  and (in case of the all-floating formulation)

$$(Q_{\text{diag}} \mu)_{iD}(x^h) := \rho_i(x^h) q_i(x^h) \mu_{iD}(x^h) \quad \text{for } \mu \in U^*, \quad (2.92)$$

cf. [KW01, (4.14)] and [Of06, Pec08b].

Note that if  $H_i \approx H_j$  and  $h_i \approx h_j$  for neighboring subdomains  $\Omega_i$  and  $\Omega_j$ , then we have also  $q_i(x^h) \approx q_j(x^h)$ . The operator  $Q_{\text{diag}}$  mimics the action of  $M_{sD}^{-1}$  when restricted to  $\text{range}(G)$ , and it will be better understood in the proof of Lemma 2.109 below, where we analyze FETI/BETI with  $Q = Q_{\text{diag}}$ . Note that if the coefficient is globally constant, we may also choose  $Q = I$ , see Remark 2.111.

*Remark 2.108.* For  $Q = Q_{\text{diag}}$ , let us investigate the structure of the matrix  $G^\top Q G$  that appears in the projections  $P$  and  $P^\top$ . Consider the connectivity graph whose nodes correspond to the subdomains  $\Omega_i$  with an edge between two nodes whenever the corresponding subdomains are neighboring, cf. Fig. 2.8. Recall that  $G = B R$  and that  $R : Z \rightarrow \ker S$  with

$$Z = \prod_{i=1}^s \mathbb{R}^{\dim(\ker(S_i))}.$$

We can think of elements from  $Z$  as discrete functions on the nodes of the connectivity graph which satisfy homogeneous boundary conditions at the nodes which correspond to the non-floating subdomains, cf. Fig. 2.8. Using the definition of the jump operator  $B$ , we find that

$$\langle G^\top Q G y, z \rangle = \sum_{\substack{i > j \\ \Gamma_{ij} \neq \emptyset}} (y_i - y_j) \underbrace{\left( \min(\alpha_i, \alpha_j) \sum_{x^h \in \Gamma_{ij}^h} q_{ij}(x^h) \right)}_{=: \beta_{ij}} (z_i - z_j) \quad \forall y, z \in Z.$$

In the all-floating formulation, we have to add

$$\sum_{i=1}^s y_i \underbrace{\left( \alpha_i \sum_{x^h \in \partial \Omega_i^h \cap \Gamma_D} q_i(x^h) \right)}_{=:\beta_{iD}} z_i .$$

This bilinear form corresponds to the (sparse) matrix induced by the graph Laplacian (see e.g., [Fie73]) where we assign each edge between node  $i$  and  $j$  in  $\mathcal{G}$  the weight  $\beta_{ij}$ . As the subsequent analysis will show, the operator  $(G^\top Q G)^{-1}$ , as such solving a discrete Laplace problem on the connectivity graph, acts as a coarse problem for the FETI/BETI algorithm. In three dimensions, under Assumption 2.53, each subdomain face  $\mathcal{F}_i$  contains  $\mathcal{O}((H_i/h_i)^2)$  nodes and each subdomain edge  $\mathcal{E}_i$  contains  $\mathcal{O}(H_i/h_i)$  nodes. Hence,

$$\beta_{ij} \approx \begin{cases} \min(\alpha_i, \alpha_j) (1 + \log(H_i/h_i)) H_i & \text{if } \Omega_i \text{ and } \Omega_j \text{ share a subdomain face,} \\ \min(\alpha_i, \alpha_j) H_i & \text{if } \Omega_i \text{ and } \Omega_j \text{ share only a subd. edge,} \\ \min(\alpha_i, \alpha_j) h_i & \text{if } \Omega_i \text{ and } \Omega_j \text{ share only a subd. vertex.} \end{cases}$$

In two dimensions,

$$\beta_{ij} \approx \begin{cases} \min(\alpha_i, \alpha_j) (1 + \log(H_i/h_i)) & \text{if } \Omega_i \text{ and } \Omega_j \text{ share a subdomain edge,} \\ \min(\alpha_i, \alpha_j) & \text{if } \Omega_i \text{ and } \Omega_j \text{ share only a subd. vertex.} \end{cases}$$

We observe that vertex connections in three dimensions are weighted weaker than others, and that connections between subdomains with large coefficients are in general weighted stronger than others.

The following lemma is essentially [KW01, Lemma 4.10] (there stated for the classical FETI method).

**Lemma 2.109.** *Let Assumptions 2.53–2.56 and 2.99–2.102 hold. Then, for  $Q = Q_{\text{diag}}$ ,*

$$|P_D z_w|_S^2 \lesssim \max_{i=1}^s (1 + \log(H_i/h_i))^2 |w|_S^2 \quad \forall w \in W^\perp .$$

*Proof.* Note that  $z_w$  is constant on each subdomain and vanishes on the non-floating subdomains. We denote the components by  $z_i$ . Using inequality (2.88) from the proof of Lemma 2.104 we obtain

$$\begin{aligned} |P_D z_w|_S^2 &\lesssim \sum_{\mathcal{G}_i \subset \Gamma} \sum_{j \in \mathcal{N}_{\mathcal{G}_i}} \min(\alpha_i, \alpha_j) |\mathcal{H}_i^h(\theta_{\mathcal{G}_i}(z_i - z_j))|_{H^1(\Omega_i)}^2 + \sum_{\mathcal{G}_i \subset \Gamma_D} \alpha_i |\mathcal{H}_i^h(\theta_{\mathcal{G}_i} z_i)|_{H^1(\Omega_i)}^2 \\ &\lesssim \sum_{\mathcal{G}_i \subset \Gamma} \sum_{j \in \mathcal{N}_{\mathcal{G}_i}} \min(\alpha_i, \alpha_j) |\mathcal{H}_i^h \theta_{\mathcal{G}_i}|_{H^1(\Omega_i)}^2 |z_i - z_j|^2 + \sum_{\mathcal{G}_i \subset \Gamma_D} \alpha_i |\mathcal{H}_i^h \theta_{\mathcal{G}_i}|_{H^1(\Omega_i)}^2 |z_i|^2 . \end{aligned}$$

By the subdomain facet estimate from Lemma 2.76, Definition 2.107, and the fact that a subdomain facet contains  $\mathcal{O}((H_i/h_i)^{d-1})$  nodes, we can conclude that

$$|\mathcal{H}_i^h \theta_{\mathcal{F}_i}|_{H^1(\Omega_i)}^2 \lesssim (1 + \log(H_i/h_i)) H_i^{d-2} \lesssim \sum_{x^h \in \mathcal{F}_i} q_i(x^h).$$

For the remaining globs, we can conclude from Lemma 2.74, Definition 2.107, and the fact that a subdomain edge contains  $\mathcal{O}(H_i/h_i)$  nodes, we can conclude that

$$\begin{aligned} |\mathcal{H}_i^h \theta_{\mathcal{V}_i}|_{H^1(\Omega_i)}^2 &\lesssim h_i^{d-2} \lesssim \sum_{x^h \in \mathcal{V}_i} q_i(x^h) \\ |\mathcal{H}_i^h \theta_{\mathcal{E}_i}|_{H^1(\Omega_i)}^2 &\lesssim H_i \lesssim \sum_{x^h \in \mathcal{E}_i} q_i(x^h) \quad \text{if } d = 3. \end{aligned}$$

Since  $q_i(x^h) \approx q_j(x^h)$  for  $x^h \in \mathcal{G}_{ij}$ , and since

$$(B z_w)_{ij}(x^h) = \pm |z_i - z_j|, \quad (B z_w)_{iD}(x^h) = z_i,$$

we obtain (comparing with Definition 2.107) that

$$|P_D z_w|_S^2 \lesssim \|B z_w\|_{Q_{\text{diag}}}^2 \leq \|B w\|_{Q_{\text{diag}}}^2, \quad (2.93)$$

where in the last step we have used Lemma 2.44. The particular choice of  $Q_{\text{diag}}$  actually stems from the estimates above.

In order to bound  $\|B w\|_{Q_{\text{diag}}}^2$  in terms of  $|w|_S^2$ , we sort the contributions with respect to the globs. Using the definition of  $Q_{\text{diag}}$  and the quasi-uniformity of  $\mathcal{T}^h(\Omega_i)$  we obtain

$$\begin{aligned} &\|B w\|_{Q_{\text{diag}}}^2 \\ &\lesssim \sum_{\mathcal{G}_i \subset \Gamma} \sum_{j \in \mathcal{N}_{\mathcal{G}_i}} \min(\alpha_i, \alpha_j) \sum_{x^h \in \mathcal{G}_{ij}^h} q_{ij}(x^h) |w_i(x^h) - w_j(x^h)|^2 + \sum_{\mathcal{G}_i \subset \Gamma_D} \alpha_i \sum_{x^h \in \mathcal{G}_i^h} q_i(x^h) |w_i(x^h)|^2 \\ &\lesssim \sum_{\mathcal{G}_i \subset \Gamma \cup \Gamma_D} \alpha_i \sum_{x^h \in \mathcal{G}_i^h} q_i(x^h) |w_i(x^h)|^2 \\ &\lesssim \sum_{\mathcal{G}_i \subset \Gamma \cup \Gamma_D} \alpha_i q_i|_{\mathcal{G}_i^h} h_i^{-d_{\mathcal{G}_i}} \|w\|_{L^2(\mathcal{G}_i)}^2, \end{aligned} \quad (2.94)$$

where  $\mathcal{G}_i^h$  is as in Definition 2.50 and  $q_i|_{\mathcal{G}_i^h}$  denotes the constant value of  $q_i$  on the nodes on  $\mathcal{G}_i$ . Note that in the classical formulation, the globs on  $\Gamma_D$  can be left out. Combining this estimate with the discrete trace inequality from Lemma 2.69 yields

$$\|B w\|_{Q_{\text{diag}}}^2 \lesssim \sum_{\mathcal{G}_i \subset \Gamma \cup \Gamma_D} \alpha_i r_{\mathcal{G}_i}(H_i, h_i) \left( |\mathcal{H}_i^h w_i|_{H^1(\Omega_i)}^2 + H_i^{-2} \|\mathcal{H}_i^h w_i\|_{L^2(\Omega_i)}^2 \right), \quad (2.95)$$

where

$$r_{\mathcal{G}_i}(H_i, h_i) := q_i|_{\mathcal{G}_i^h} h_i^{-d_{\mathcal{G}_i}} \sigma^{d-d_{\mathcal{G}_i}} \left(\frac{H_i}{h_i}\right) H_i^{2+d_{\mathcal{G}_i}-d}.$$

In the all-floating formulation, we obtain the desired bound

$$|P_D w|_S^2 \lesssim \|B w\|_{Q_{\text{diag}}}^2 \lesssim \max_{i=1}^s (1 + \log(H_i/h_i))^2 |w|_S^2 \quad \forall w \in W^\perp$$

(by Poincaré's inequality from Lemma 2.61) if

$$r_{\mathcal{G}_i}(H_i, h_i) \lesssim (1 + \log(H_i/h_i))^2. \quad (2.96)$$

For quasi-uniform fine and coarse triangulations, condition (2.96) is even necessary. A short computation reveals that (2.96) holds if and only if

$$\begin{aligned} (1 + \log(\frac{H_i}{h_i})) \frac{h_i^{d-1}}{H_i} &\lesssim q_i(x^h) \lesssim (1 + \log(\frac{H_i}{h_i}))^2 \frac{h_i^{d-1}}{H_i} && \text{if } x^h \in \mathcal{F}_i, \\ h_i &\lesssim q_i(x^h) \lesssim (1 + \log(\frac{H_i}{h_i})) h_i && \text{if } x^h \in \mathcal{E}_i, d = 3, \\ h_i^{d-2} &\lesssim q_i(x^h) \lesssim (1 + \log(\frac{H_i}{h_i}))^{d-1} h_i^{d-2} && \text{if } x^h = \mathcal{V}_i, \end{aligned}$$

where we have included the lower bounds for  $q_i(x^h)$  that we have used in the beginning of the proof. In the classical formulation, we need to be able to apply a discrete Poincaré inequality in each non-floating subdomain (cf. Assumption 2.56) at the cost of a factor of  $(1 + \log(H_i/h_i))$ . Hence, this factor must be taken away from the upper bounds above unless all non-floating subdomains have a subdomain face in common with the Dirichlet boundary. In any case, we see that the respective bounds for  $q_i(x^h)$  hold for the choice made in Definition 2.107, which concludes the proof.  $\square$

**Theorem 2.110.** *Let Assumptions 2.53–2.56 and 2.99–2.102 hold. Then, for the classical and the all-floating FETI/BETI method with the scaled Dirichlet preconditioner  $M_{\text{SD}}^{-1}$  and with the choice  $Q = Q_{\text{diag}}$  (cf. Definition 2.107),*

$$\kappa(P M_{\text{SD}}^{-1} P^\top F|_{\tilde{U}_{\text{ad}}}) \leq C \max_{i=1}^s (1 + \log(H_i/h_i))^2.$$

where the constant  $C$  depends only on the uniform constants from the mentioned assumptions.

*Proof.* The estimate follows immediately by combining Lemmas 2.45, 2.103, 2.104, and 2.109.  $\square$

Opposed to Theorem 2.106, the statement of Theorem 2.110 does not generalize to the case of non quasi-uniform meshes, as we heavily used the quasi-uniformity in the proof of Lemma 2.109.

*Remark 2.111.* Let us suppose that the coefficients are locally quasi-monotone, i.e., for each glob  $\mathcal{G}_{ij}$  we can find an admissible face path  $\mathcal{F}_{k_1,k_2}, \mathcal{F}_{k_2,k_3}, \dots, \mathcal{F}_{k_{m-1},k_m}$  with

$$\min(\alpha_i, \alpha_j) = \alpha_{k_1} \leq \alpha_{k_2} \leq \dots \leq \alpha_{k_m} = \max(\alpha_i, \alpha_j).$$

In that situation, the weights  $q_{ij}(x^h)$  for subdomain vertices (and edges if  $d = 3$ ) on the interface  $\Gamma$  can be decreased arbitrarily, see also the remarks after Theorems 4.11 and 5.7 in [KW01]. As a consequence, if the coefficient is globally constant (or at least  $\alpha_i \approx \alpha_j$  for all neighboring subdomains  $\Omega_i, \Omega_j$ ), and if the subdomain decomposition and the global mesh are quasi-uniform ( $h_i \approx h$  and  $H_i \approx H$ ), then the matrix  $Q$  can be chosen as the identity matrix or any multiple thereof.

## 2.6.4 Alternative Scalings

In this short section we discuss other scalings than the coefficient scaling, which is assumed in the theory above (cf. Assumption 2.100). Note a further scalings (not discussed below) based on eigensolves can be found in [DW12b].

### 2.6.4.1 Multiplicity Scaling

If we use the multiplicity scaling (which also effects  $Q_{\text{diag}}$ , cf. Definition 2.107), the statements of Theorems 2.106 and 2.110 remain true under the stated assumptions, but the condition number bounds have to be multiplied by a factor of

$$\max_{i=1}^s \max_{j \in \mathcal{N}_i} \frac{\alpha_i}{\alpha_j},$$

i.e., the maximal jump *between* subdomains, where  $\mathcal{N}_i$  contains the indices of the subdomains neighboring  $\Omega_i$ , cf. Definition 2.12. This is seen from the proofs of Lemmas 2.104 and 2.109 by using that  $\delta_j^\dagger(x^h) \leq 1$  as well as the simple fact that

$$\alpha_i \leq \left( \max_{i=1}^s \max_{k \in \mathcal{N}_i} \frac{\alpha_i}{\alpha_k} \right) \alpha_j$$

whenever  $j \in \mathcal{N}_i$ .

### 2.6.4.2 On the Stiffness Scaling and on Effects of Varying Coefficients

If one has no a priori knowledge on the coefficient  $\alpha_i$ , a common choice in FETI is the stiffness scaling, where  $\rho_i(x^h)$  is set to the diagonal entry of the stiffness matrix  $\mathbf{K}_i$  corresponding to the node  $x^h$ . Assuming an isotropic coefficient  $\mathcal{A} = \alpha I$ , the



stiffness scaling is given by

$$\rho_i(x^h) = \int_{\omega_{i,x^h}} \alpha |\nabla \varphi_{i,x^h}|^2 dx,$$

where  $\omega_{i,x^h}$  is the union of all elements  $\tau$  from  $\mathcal{T}^h(\Omega_i)$  such that  $x^h \in \bar{\tau}$ . Similar choices are the local average

$$\rho_i(x^h) = \frac{1}{\text{meas}_{\omega_{i,x^h}}} \int_{\omega_{i,x^h}} \alpha dx$$

and the local maximum

$$\rho_i(x^h) = \text{ess.sup}_{x \in \omega_{i,x^h}} \alpha(x).$$

If  $\alpha|_{\Omega_i} = \alpha_i$  is constant, the latter two choices reproduce the coefficient scaling  $\rho_i(x^h) = \alpha_i$ . For the stiffness scaling,  $\rho_i(x^h) \approx h_i^{d-2} \alpha_i$  (if  $\mathcal{T}^h(\Omega_i)$  is quasi-uniform) which may look promising. However, in the presence of rough (or ragged) interfaces as they appear in METIS partitions, the stiffness scaling leads to extremely poor convergence of FETI type methods as it has been demonstrated in [KRW08, Sect. 5] (the term  $\rho$ -scaling therein corresponds to the *coefficient scaling* in this book).

In the analysis, we used that (for the coefficient scaling)  $\delta_j^\dagger$  is constant at the nodes of each glob, which is not anymore true for the stiffness scaling. In the case of rough interfaces, the function  $\rho_i \in V^h(\partial\Omega_i)$  becomes *oscillatory*, i.e.,

$$\frac{\rho_i(x^h) - \rho_i(y^h)}{\rho_i(x^h)} \gg \frac{|x^h - y^h|}{H_i}.$$

Typically, the weighted counting functions  $\delta_j^\dagger$  have the same property, which is at high probability the reason behind the poor convergence.

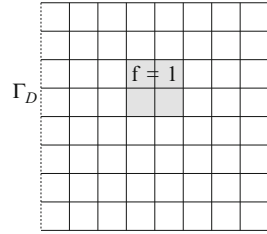
Even if the interfaces are smooth, the phenomenon of an oscillatory function  $\delta_j^\dagger$  can occur when the coefficient  $\mathcal{A} = \alpha I$  is mildly varying in  $\Omega_i$ . By this we mean that  $\text{ess.sup}_{x,y \in \Omega_i} \frac{\alpha(x)}{\alpha(y)}$  is relatively small, but the coefficient may change from element to element.

Summarizing, in case of rough interfaces or a mildly varying (isotropic) coefficient, provided that the subdomain meshes are quasi-uniform, a good choice is

$$\rho_i(x^h) = \max_{x^h \in \partial\Omega_i^h} k_{i,x^h} \quad \text{or} \quad \rho_i(x^h) = \max_{x^h \in \Omega_i^h} k_{i,x^h},$$

where  $k_{i,x^h}$  denotes the diagonal entry of the stiffness matrix  $\mathbf{K}_i$  at node  $x^h$ . For an analysis with mildly varying coefficients (but smooth interfaces) see Sect. 3.2. For a sound theory of FETI-DP methods for the case of rough interfaces in two dimensions (using the coefficient scaling) see [KRW08].

**Fig. 2.9** Setup for the experiments on the unit square with a homogeneous coefficient



## 2.7 Numerical Results

In this section, we present some numerical results for one-level FETI/BETI methods for a two-dimensional model problem with and without coefficient jumps. This is mainly to give the reader an impression on how small the condition numbers and the number of PCG iterations actually are. The main implementation was done in  $C^{++}$ . The FEM stiffness matrices and the coarse matrix were factorized using PARDISO [PAR05, SG04, SG06]. For the boundary element method we have used Olaf Steinbach's Fortran package OSTBEM [Ste00]. The condition numbers are estimated using the Lanczos method, see Remark 1.51. Mainly interested in verifying the theoretical results of this chapter, we have not used any data-sparse approximation of the boundary element matrices.

Note that computational results for FETI methods (including the case of linear elasticity) can be found in [FR91, FR91, LP98, RFTM99, Rhe02]; for all-floating BETI methods see [Of06, Of08].

**Unit Square – Homogeneous Coefficient.** We consider the unit square  $\Omega = (0, 1)^2$ , subdivided into 64 equally-sized square-shaped subdomains, with homogeneous Dirichlet boundary conditions on the left side  $\Gamma_D$ , and homogeneous Neumann boundary conditions on the rest of  $\partial\Omega$ . The source term  $f$  is chosen to be zero except for the four shaded subdomains in Fig. 2.9, and the coefficient  $\alpha$  is set uniformly to one.

Tables 2.3 and 2.4 show the results for FETI and FETI/BETI, respectively. There, the column entitled “Lagr. mult.” indicates number of Lagrange multipliers (additional multipliers enforcing the Dirichlet boundary conditions in the all-floating method are not counted). For simplicity,  $H$  denotes the height/width of the subdomain. In the columns entitled “PCG” we give the number of PCG steps needed to get a reduction of  $\varepsilon = 10^{-8}$  in the residual, and the columns entitled “cond.” show the estimated condition number using the Lanczos method. We see that the condition numbers of the preconditioned systems behave as predicted by the theory. From the first column in the two tables one can observe the reduction in the global DOFs when using the boundary element method. In Table 2.5 we demonstrate the scalability, i.e., the robustness with respect to the number of subdomains.

**Table 2.3** Unit square with homogeneous coefficient; classical one-level vs. all-floating FETI method; 64 subdomains

Global DOFs	Lagr. mult.	Local DOFs	$H/h$	Std. one-level		All-floating	
				PCG	Cond.	PCG	Cond.
289	406	9	2	9	1.67	8	1.40
1,089	630	25	4	11	2.20	10	1.88
4,225	1,078	81	8	13	2.97	12	2.43
16,641	1,974	289	16	16	3.92	14	3.15
66,049	3,766	1,089	32	18	5.05	16	4.05
263,169	7,350	4,225	64	21	6.33	18	5.12
1,050,625	14,518	16,641	128	23	7.77	19	6.36
4,198,403	28,854	66,049	256	24	9.38	21	7.76
16,785,409	57,526	263,169	512	25	11.15	23	9.33

**Table 2.4** Unit square with homogeneous coefficient; classical one-level vs. all-floating FETI/BETI method; 64 subdomains (60 BEM, 4 FEM)

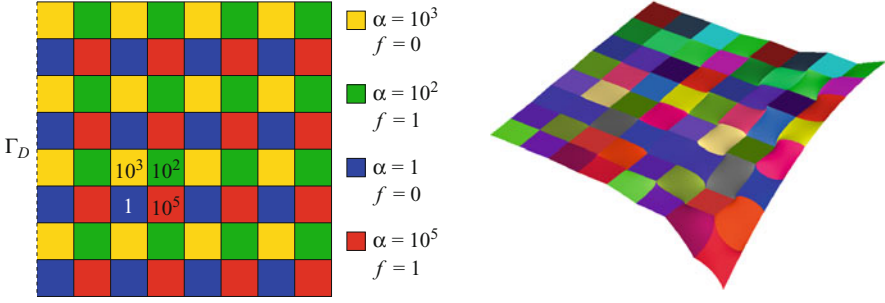
Global DOFs	Lagr. mult.	FEM loc. DOFs	BEM loc. DOFs	$H/h$	Std. one-level		All-floating	
					PCG	Cond.	PCG	Cond.
229	406	9	8	2	9	1.65	9	1.64
549	630	25	16	4	10	1.91	9	1.67
1,285	1,078	81	32	8	13	2.58	11	2.08
3,141	1,974	289	64	16	15	3.44	13	2.72
8,389	3,766	1,089	128	32	18	4.48	16	3.54
25,029	7,350	4,225	256	64	20	5.68	18	4.54
82,885	14,518	16,641	512	128	23	7.03	20	5.71
296,901	28,854	66,049	1,024	256	24	8.55	22	7.05
1,118,149	57,526	263,169	2,048	512	25	10.24	23	8.55

**Table 2.5** Unit square with homogeneous coefficient; classical vs. all-floating FETI method; fixed ratio  $H/h = 32$ ; fixed number 1,089 of local FEM DOFs, varying number of subdomains

Number of subdomains	Global DOFs	Std. one-level		All-floating	
		PCG	Cond.	PCG	Cond.
64	66,049	18	5.049	16	4.045
256	263,169	18	5.055	16	4.064
1,024	591,361	18	5.055	16	4.064
4,096	1,050,625	18	5.053	15	4.057

**Unit Square – Heterogeneous Coefficient.** In this example we consider the unit square  $(0, 1)^2$  with the same partitioning as in Fig. 2.9, but we choose the coefficient  $\alpha$  and the source  $f$  according to Fig. 2.10, left. The Dirichlet boundary conditions read

$$u(x_1, x_2) = 8x_2(1 - 8x_2) \quad \text{for } (x_1, x_2) \in \Gamma_D = \{(0, x_2) : x_2 \in (0, 1)\}.$$



**Fig. 2.10** *Left:* Setup of unit square with heterogeneous coefficient. *Right:* Visualization of the solution  $u$  via the graph  $(x_1, x_2, u(x_1, x_2))$ ; different colors indicate different subdomains

On the remainder of  $\partial\Omega$  we impose homogeneous Neumann boundary conditions.

Tables 2.6 and 2.7 show the number of PCG steps and the estimated condition number for the classical one-level and all-floating FETI and FETI/BETI method, respectively. In the second case, the BEM subdomains are exactly those where  $f = 0$ . The numbers in the tables demonstrate the robustness with respect to the heterogeneous coefficient, which would not be the case without the careful scalings in  $B_D$  and  $Q$ . Figure 2.10, right displays the solution  $u$  to the problem. We see that in the areas with large coefficients the solution is relatively flat.

## 2.8 Other PDEs and Other Discretization Spaces

The theory of the preceding sections carries over immediately from  $P^1$  to  $Q^1$  elements, i.e., bilinear quadrilateral or trilinear hexahedral elements. In this section, however, we give a brief overview (mostly in form of references) on FETI/BETI and related methods for PDEs other than the potential equation, and for discretizations other than piecewise linear FEM/BEM. For DD methods other than substructuring methods, we refer to [Mat08, TW05]. In Sect. 2.8.1 we discuss the influence of the discretization on the coupling (interconnecting), whereas Sect. 2.8.2 treats the change of the PDE (possibly implying specific discretizations).

More developments than mentioned below are and will be documented in the proceedings of the international conferences on domain decomposition methods, see <http://www.ddm.org/>.

### 2.8.1 Other Discretizations Spaces for $H^1$ -Problems

Let  $\Omega$  be the computational domain, and  $V(\Omega)$  a space of functions on  $\Omega$  related to the problem, e.g.,  $V = H^1(\Omega)$ . First, we treat the case of a conforming finite

**Table 2.6** Unit square with heterogeneous coefficient; classical one-level vs. all-floating FETI, 64 subdomains

Global DOFs	Local DOFs	$H/h$	Std. one-level		All-floating	
			PCG	Cond.	PCG	Cond.
289	9	2	10	2.23	8	1.54
1,089	25	4	11	2.69	11	2.12
4,225	81	8	13	3.18	13	2.90
16,641	289	16	15	3.84	15	3.84
66,049	1,089	32	17	4.91	17	4.91
263,169	4,255	64	19	6.10	19	6.10
1,050,625	16,641	128	21	7.42	22	7.43
4,198,401	66,049	256	23	8.91	24	8.88
16,785,409	263,169	512	26	10.60	25	10.47

**Table 2.7** Unit square with heterogeneous coefficient; classical one-level vs. all-floating FETI/BETI, 32 FEM, 32 BEM subdomains

Global DOFs	FEM loc. DOFs	BEM loc. DOFs	$H/h$	Std. one-level		All-floating	
				PCG	Cond.	PCG	Cond.
257	9	8	2	7	1.68	7	1.64
801	25	16	4	10	2.24	9	2.24
2,657	81	32	8	11	3.09	12	3.09
9,441	289	64	16	13	4.08	13	4.08
35,297	1,089	128	32	14	5.21	15	5.21
136,161	4,255	256	64	16	6.46	17	6.46
534,497	16,641	512	128	17	7.83	18	7.83
2,117,601	66,049	1,024	256	18	9.33	20	9.33
8,429,537	263,169	2,048	512	19	10.96	21	10.96

element space  $V^h(\Omega) \subset V$  that fulfills Assumption 2.112 below. Other types of discretization spaces will be treated in Sects. 2.8.1.3–2.8.1.5.

- Assumption 2.112.** (i) The FE space  $V^h(\Omega)$  is based on a triangulation  $\mathcal{T}^h(\Omega)$  of  $\Omega$ , i.e., to each mesh element  $T \in \mathcal{T}^h$  there is an associated space  $V_T$  of shape functions and a set  $\mathcal{N}_T \subset V_T^*$  of DOFs (nodal variables), cf. [BS02, Definition 3.1.1] and [Cia87].
- (ii) Each DOF is associated to a (fine) vertex, edge, face, or element of the mesh.
- (iii) The global space  $V^h(\Omega)$  is composed of the local spaces  $V_T$ , where corresponding DOFs on the same entity (vertex, edge, face) are globally identified.

Assumption 2.112 holds for the case for high order  $H^1$ -conforming spaces (either of hierarchical nature as in Sect. 1.2.3.7, or for spectral elements; see e.g., [BM97, KS99, Sch98b]). Note that Assumption 2.112 holds as well for the  $H(\text{curl})$ -conforming Nédélec edge elements, the  $H(\text{div})$ -conforming Raviart-Thomas (Nédélec face) elements of any order, see, e.g., [RT77, Néd80, Néd86, Mon03, Zag06], and the mixed elements in [Sin08, PS11a, PS12a].

With Assumption 2.112 being fulfilled, we can define the interface (or skeleton) DOFs as those DOFs associated to (fine) vertices, edges, or faces lying on the (geometrical) interface  $\Gamma$  (or the skeleton  $\Gamma_S$ , respectively) and eliminate the remaining (interior) DOFs. This yields a discrete skeleton formulation. Furthermore, we can define the local restrictions to  $\Omega_i$  or  $\partial\Omega_i$ , which leads to the local spaces  $W_i$  and the (discontinuous) product space  $W$ . Reinstalling the original continuity follows the finite element construction above. Let  $\psi_{i,k}$  and  $\psi_{j,\ell}$  be two DOFs on subdomain  $\Omega_i$  and  $\Omega_j$ , respectively, that are globally identified. Then the correct jump condition simply reads

$$\psi_{i,k}(w_i) - \psi_{j,\ell}(w_j) = 0,$$

and all these conditions together define the jump operator  $B : W \rightarrow U^* = \mathbb{R}^M$ . If we represent  $B$  as a matrix  $\mathbf{B}$  with respect to the nodal FE basis, then  $\mathbf{B}$  is again signed Boolean (see, e.g., [TK01, TV03]).

Corresponding boundary element spaces on the skeleton  $\Gamma_S$  can be easily derived by restricting  $V^h(\Omega)$  to the skeleton. This restricted space is parameterized by the DOFs associated to  $\Gamma_S$ , and the coupling procedure is identical to the above one.

The following two sections deal with specific  $H^1$ -conforming discretizations.

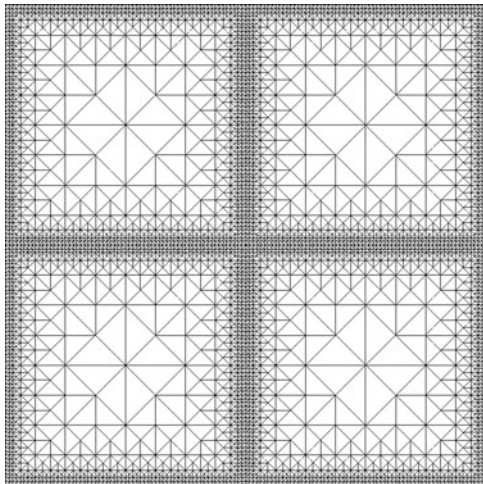
### 2.8.1.1 $H^1$ -Conforming Interface Concentrated FETI/BETI

The interface concentrated FETI method was introduced by Beuchler, Eibner, and Langer [BEL08] as a solver for the scalar potential equation in  $H^1$ , see also [LP08] for a generalization to interface concentrated FETI/BETI. The main idea is to employ a *boundary concentrated* FEM (cf. [KM03] and Sect. 1.2.3.8) in each FEM subdomain, such that the polynomial degree equals one on each subdomain boundary. See Fig. 2.11 for an illustration. Obviously, the coupling is the same as for the low-order FETI/BETI methods. Since it can be shown that the corresponding FEM Schur complement in each subdomain is spectrally equivalent to the Steklov-Poincaré operator (cf. [BEL08, Theorem 3.13]), provided that the boundary mesh is quasi uniform, all the theoretical results of Sect. 2.6 carry over immediately to the interface concentrated FETI/BETI. Numerical results can be found in [BEL08, LP08], and also in [Pec08b, Sect. 2.3].

### 2.8.1.2 High Order $H^1$ -Conforming Spaces

There are two classes of high order elements. Spectral elements [BM97, KS99] have uniform polynomial degree throughout the domain. The associated DOFs are point evaluations at Gauss-Lobatto-Legendre points. Hierarchical high order elements (see e.g. [Dem07, Zag06]) allow for variable polynomial degree, but the DOFs are associated to vertices, edges, faces, or elements and not necessarily point evaluations, but typically averages.

**Fig. 2.11** Example for an interface concentrated mesh for a partition of the unit square into four subdomains



For spectral elements of degree  $k$ , it was shown that Neumann-Neumann and FETI type methods (in two and three dimensions) lead to condition number bounds of the form

$$\kappa \leq C (1 + \log(k))^2,$$

cf. [TW05, Sect. 7.2] as well as [Pav97, Pav07, KPR08]. However, in three dimensions these bounds do not carry over to the hierarchical high order elements, even if spanning the same polynomial space, see [TW05, Sect. 7.5] (this is because the coupling relies on the underlying DOFs, which are different for the two approaches). Indeed, the convergence can slow down for the hierarchical case. For numerical studies of Neumann-Neumann and FETI methods see also [TV03, TV04, TV06].

*Remark 2.113.* Instead of coupling directly the DOFs, one could also introduce a suitable set of interpolation points and couple function values there. This is possible for any basis of a high order  $H^1$ -conforming FE space. In that case the structure of  $B$  is more involved.

### 2.8.1.3 Mortar Discretizations

*Mortar* discretizations (see e.g. [BMP94]) are FEM or BEM on non-conforming meshes. The coupling is done by integrating the jumps on the coupling interface against test functions that act as Lagrange multipliers. Thus, mortar methods contain already the needed interconnecting. A FETI like jump operator can be derived using the mass matrices on the coupling interface. See [SK98, Ste01, KL05, Kim07, KDW08, Kim08a, Kim08c, Kim08b, KT09] for FETI and balancing type methods for such mortar discretizations. Moreover, see [Rou09] for a related method called FETI-2LM.

### 2.8.1.4 Discontinuous Galerkin Method

Discontinuous Galerkin methods (see e.g. [DE12, HW08a, Riv08]) typically work with spaces of discontinuous functions across elements. The coupling is performed by suitable jump terms that are added to the bilinear form. Balancing type methods for DG discretizations on conforming and non-conforming meshes have been thoroughly investigated by Dryja, Galvis, and Sarkis [DGS07, DGS08, DGS11, DGS12].

### 2.8.1.5 Isogeometric Analysis

For the discretizations in *isogeometric analysis* (cf. [HCB05, BBC<sup>+</sup>06, CHB09]), Assumption 2.112 is not necessarily fulfilled. This is because for  $C^k$ -elements, the NURBS basis functions cannot be localized to individual elements, and thus it is difficult to associate the DOFs to a *geometric* interface. See however [BSEH11] for a localization by a change to the Bernstein basis. The recently introduced IETI method, cf. [KPJT12], covers the case of  $C^0$ -continuity across subdomain interfaces, which makes the coupling more natural. The case of  $C^k$ -continuity with  $k > 0$  has been studied recently in [BCPS12], using the concept of *fat interfaces*.

## 2.8.2 FETI and Balancing Type Methods for General PDEs

This section consists mainly of an (incomplete) list of references for FETI and balancing type methods for a variety of different PDEs. Note that many of the references are on FETI-DP and BDDC methods, which will be dealt with later on in Chap. 5. In the following, we briefly touch advection-diffusion problems, and problems in continuum mechanics, acoustics, and electromagnetics.

### 2.8.2.1 Dynamic Problems and Advection-Diffusion Problems

In this section, we consider the modified PDE

$$-\operatorname{div}(\alpha \nabla u) + \beta \cdot \nabla u + \gamma u = f, \quad (2.97)$$

**The Case  $\beta = 0$  and  $\gamma > 0$ .** Assume that  $\beta = 0$  and that  $\gamma(x) \geq \gamma_0 > 0$  a.e. in  $\Omega$ . This kind of problem occurs particularly in implicit time stepping for the dynamic problem

$$\frac{\partial u}{\partial t} - \operatorname{div}(\alpha \nabla u) = f.$$



In each time step, one has to solve a problem of the form (2.97) with  $\beta = 0$  and  $\gamma = 1/\Delta t$ , where  $\Delta t$  is the time step.

Let  $W_i$  and  $W$  be the spaces from either the classical or the all-floating formulation and let  $S_i : W_i \rightarrow W_i^*$  denote the corresponding (discrete) Steklov-Poincaré operators. On a FEM subdomain,  $S_i$  corresponds to the Schur complement of the matrix

$$\mathbf{K}_i + \mathbf{M}_i$$

where  $\mathbf{K}_i$  is the stiffness matrix corresponding to  $\int_{\Omega} \alpha \nabla u \cdot \nabla v \, dx$  and  $\mathbf{M}_i$  the mass matrix corresponding to  $\int_{\Omega} \gamma u v \, dx$ . For a BEM subdomain, the operator  $S_i$  is given analogously to Sect. 1.3 but the fundamental solution in the boundary integral operators has to be modified accordingly (see, e.g., [SS11]). Using  $S := \text{diag}(S_i)_{i=1}^s$ , one derives the FETI/BETI saddle point formulation analogously to Sect. 2.2.1.1: find  $(u, \lambda) \in W \times U$  such that

$$\begin{bmatrix} S & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} g \\ 0 \end{bmatrix},$$

see also [FCM95, FM98, Tos01]. For convenience we split the local Steklov-Poincaré operators into the part corresponding to the stationary term  $-\text{div}(\alpha \nabla u)$  and a remainder,

$$S_i = S_i^K + S_i^M.$$

Observe that, even if  $S_i^K$  might have a non-trivial kernel, the operator  $S_i$  is always SPD, and so

$$\ker(S_i) = \{0\}.$$

That is, according to Definition 2.15, all subdomains are non-floating. Therefore, no projection is needed to eliminate the unknown  $u$  from the saddle point problem, and the resulting dual system simply takes the form

$$\text{find } \lambda \in U : \underbrace{BS^{-1}B^\top}_{=:F} \lambda = \underbrace{BS^{-1}g}_{=:d},$$

where  $S^{-1} = \text{diag}(S_i^{-1})_{i=1}^s$ . The solution  $\lambda$  is unique up to  $\ker(B^\top)$ . Hence, let  $\tilde{U} := U_{/\ker(B^\top)}$  and  $\tilde{U}^* = \text{range}(B)$ , then we can seek (formally)  $\lambda \in \tilde{U}$ . Due to the lack of the projection, there is no coarse problem, and so using the (scaled) Dirichlet preconditioner only results in a method that is not scalable, i.e., it deteriorates when the number  $s$  of subdomains grows.

We now reinstall a coarse problem using an outer projection. This technique is also called deflation (cf. [VSMW01, KR12]) and was first used for a FETI method by Farhat, Chen, and Mandel [FCM95] (see also [FM98]) in the context of dynamic elasticity. Let  $R_i$  span the kernel of  $S_i^K$  and define  $R = [R_i]_{i=1}^s$  and  $G := QBR$ , where  $Q : U^* \rightarrow U$  is a suitable SPD operator. The outer projector  $P : \tilde{U} \rightarrow \tilde{U}$  is then given by

$$P := I - G(G^\top F G)^{-1} G^\top F,$$

**Algorithm 8:** FETI/BETI algorithm for a dynamic problem based on PCG

---


$$\lambda^{(0)} = G(G^\top F G)^{-1} G^\top d$$

$$r^{(0)} = d - F \lambda^{(0)}$$

$$k = 0$$

**repeat**

$$z^{(k)} = M_{\text{sD}}^{-1} r^{(k)}$$

$$s^{(k)} = P z^{(k)}$$

$$p^{(k)} = s^{(k)} + \beta_{k-1} p^{k-1} \quad \text{where} \quad \beta_{-1} = 0, \quad \beta_{k-1} = \frac{\langle r^{(k)}, s^{(k)} \rangle}{\langle r^{(k-1)}, s^{(k-1)} \rangle}$$

for  $k > 0$

$$\lambda^{(k)} = \lambda^{(k-1)} + \alpha_k p^{(k)} \quad \text{where} \quad \alpha_k = \frac{\langle r^{(k)}, s^{(k)} \rangle}{\langle \widehat{F} p^{(k)}, p^{(k)} \rangle}$$

$$r^{(k+1)} = r^{(k)} - \alpha_k \widehat{F} p^{(k)}$$

$$k = k + 1$$

**until** *stopping criterion fulfilled for  $r^{(k)}$*

---

cf. [FM98, Sect. 5] (where  $\underline{Q} = I$ ) and [Tos01, Sect. 3]. To assemble the (sparse) matrix corresponding to  $(G^\top F G)$ , one has to apply the operator to the kernel vectors. To do this efficiently in a parallel regime, one needs similar techniques as described in Remark 2.36. Since  $\text{range}(P)$  and  $\text{range}(\underline{I} - P)$  are  $F$ -orthogonal, we can split  $\lambda = \lambda_0 + \widetilde{\lambda}$  where  $\lambda_0 \in \text{range}(I - P)$  and  $\widetilde{\lambda} \in \text{range}(P)$ . Due to the  $F$ -orthogonality,

$$\lambda_0 = G(G^\top F G)^{-1} G^\top d.$$

To solve for  $\widetilde{\lambda}$ , we apply a PCG to the (SPD) equation

$$P^\top F \widetilde{\lambda} = P^\top d$$

with initial value 0 and with the preconditioner  $P M_{\text{sD}}^{-1}$ , where

$$M_{\text{sD}}^{-1} = B_D S B_D^\top, \quad S = \text{diag}(S_i)_{i=1}^s.$$

As for the algorithm in Sect. 2.3, one can rewrite this PCG algorithm as an iteration for the original variable  $\lambda$ , see Algorithm 8. Thanks to the choice of the initial value, one needs to perform only one projection step (cf. [FM98] and [Tos01, Lemma 3.1]).

**Including Advection.** A generalization of the above algorithm to the non-symmetric problem (2.97) with  $\beta \neq 0$  (but further assumptions on  $\beta$  and  $\gamma$ ) can be found in [Tos01], where the PCG is replaced by a preconditioned GMRes method. For a BDDC method we refer to [TL08].

**The Case  $\beta = 0, \gamma \geq 0$ .** Assume that on some subdomains  $\gamma = 0$  and on the remaining ones  $\gamma > 0$ . This means that some of the Steklov-Poincaré operators already have kernels, and some do not. A method called *generalized FETI* has been

proposed by Farhat and Mandel [FM97], see also [Mat08, Sect. 4.3]. It introduces a projection that includes both the existing kernels and the artificial ones.

**Unification by Dual-Primal Methods.** We note that using a dual-primal approach (see Chap. 5), no case distinction is necessary, at least for the symmetric problems, cf. Sect. 5.3.5.1.

**Acceleration for Dynamic Problems.** In [FCR94] it has been proposed to accelerate the iterative solves by recycling the Krylov space of previous time steps. This technique can be used for FETI, Neumann-Neumann, FETI-DP, and BDDC.

### 2.8.2.2 Continuum Mechanics

**Linear Elasticity.** We consider the primal formulation of linear elasticity in variational form, for simplicity with homogeneous Dirichlet boundary conditions: find the displacement  $\mathbf{u} \in H^1(\Omega)^d$ ,  $\mathbf{u}|_{\Gamma_D} = 0$  such that

$$\int_{\Omega} \mathbf{C} \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}) \, dx = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, dx + \int_{\Gamma_N} \mathbf{t}_N \cdot \mathbf{v} \, ds \quad \forall \mathbf{v} \in H^1(\Omega)^d, \mathbf{v}|_{\Gamma_D} = 0,$$

where  $\boldsymbol{\varepsilon}(\mathbf{v}) = \frac{1}{2}(\nabla \mathbf{v} + (\nabla \mathbf{v})^\top)$  is the linearized strain tensor and  $\mathbf{C}$  the (linear) material tensor due to Hooke's law, such that  $\boldsymbol{\sigma}(\mathbf{u}) = \mathbf{C} \boldsymbol{\varepsilon}(\mathbf{u})$  is the stress tensor. It depends on the Young modulus  $E$  and the Poisson ratio  $\nu$  in the usual way, see e.g. [Bra01, SDH04] or [TW05, Sect. A.6.2]. Here, we assume that the Poisson ratio  $\nu$  is bounded away from the incompressible limit  $1/2$ .

The classical FETI method was originally introduced for the above problem, cf. [FR91, FR94]. When using the continuous piecewise linear finite elements for the components of the displacement  $\mathbf{u}$ , the derivation of FETI is in large parts analogous to the presentation in Sect. 2.2. The crucial difference lies in the local kernels which can have dimension of 0 up to 6 in three dimensions. Here lies an advantage of all-floating (total) FETI (cf. [DHK06]), where the local kernel is always the space of rigid body modes,

$$\mathcal{RB} := \begin{cases} \left\{ \mathbf{a} + b \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix} : \mathbf{a} \in \mathbb{R}^2, b \in \mathbb{R} \right\} & \text{for } d = 2, \\ \left\{ \mathbf{a} + \mathbf{b} \times \mathbf{x} : \mathbf{a}, \mathbf{b} \in \mathbb{R}^3 \right\} & \text{for } d = 3. \end{cases}$$

For an analysis of FETI for linear elasticity see [KW00] and [TW05, Sect. 8.5]. Here, the key ingredient is a spectral equivalence between the stiffness matrix of linear elasticity and the stiffness matrix of the vector Laplacian, which holds at least for vectors orthogonal to the kernel of the elasticity matrix, cf. [KW00, Lemma 5]. Using that equivalence, the FETI analysis can in large parts be reduced to the scalar elliptic case of Sect. 2.6.

For the FETI-DP method (where the local kernels do not have to be known explicitly) we refer to [FLL<sup>+</sup>01, KR06, KW06, KR07b, KR10]. For the BETI method see [Of06, Of08, OS09]. For FETI(-DP) and BDDC for mortar discretizations of linear elasticity see [Kim08a, Kim08c, Kim08b]. Finally, for a novel approach using the Smith factorization, see [CDNQ12].

**Plate and Shell Problems.** The FETI method was generalized for plate and shell problems by Farhat, Mandel, and Tezaur, see [MTF99, FM98]. Without a special treatment, however, the convergence is not satisfying, and a further outer projection of the Lagrange multipliers is necessary, see [FM98, FCMR98]. Due to the introduction of this second level, the resulting method is now called FETI-2. Note also that this method was important for the development of the dual-primal methods (cf. Chap. 5).

FETI-DP and BDDC methods for plate problems can be found in [FLL<sup>+</sup>01, BCLP10]. For an analysis see also [Bre03b].

**Almost Incompressible Elasticity.** In the almost incompressible case of elasticity, the Poisson ratio is close to  $1/2$ , which needs special treatment. Here we refer to [KRW07, PWZ10, GKR12], see also the related papers [DW09, DW10].

**Nonlinear Elasticity and Contact Problems.** There are three major sources of nonlinearities in solid mechanics:

- (i) *Geometric nonlinearities* (due to large deformations),
- (ii) *Nonlinear material laws* (other than Hooke's linear law),
- (iii) *Nonlinearities due to contact* (introducing active or inactive constraints).

FETI type methods for contact problems can, e.g., be found in [BDS08, DHK<sup>+</sup>05, DKV<sup>+</sup>10, HKD04, JKR12]. For FETI type methods for nonlinear material laws in biomechanics, we refer to [BKRS08, BBK<sup>+</sup>09, KNRV11].

**The Stokes Problem.** For FETI-DP and BDDC for Stokes we refer to [PW02, Li05, LW06a, KLP10a, KLP10b, KL10, ŠSB<sup>+</sup>11], see also [TW05, Sect. 9.4.2]. For a novel approach using the Smith factorization, see [DNR08, DNR09].

**Porous Media Flow.** For BDDC methods on mixed and hybrid discretizations of porous media flow problems, also called Darcy's problem (cf. [TW05, Sect. A.7.2]), we refer to [Tu05, Tu07a, Tu11, Sou12].

### 2.8.2.3 Acoustic Scattering

The acoustic scattering problem is governed by the (scalar) Helmholtz equation

$$-\Delta u - k^2 u = f,$$

where  $k$  is the wave number. Besides Dirichlet and Neumann conditions, one often considers non-reflecting boundary conditions modeling waves that are only

outgoing. Besides the “standard” difficulties of this indefinite problem (including a proper discretization), there are more difficulties when considering non-overlapping domain decomposition: local Dirichlet and Neumann problems can be unsolvable, when  $k^2$  hits an eigenvalue of the corresponding local Helmholtz problem. A remedy was proposed by Farhat, Macedo, and Tezaur [FMT99], see also [FML00a]. The local problems are consistently supplemented with Robin boundary conditions, which guarantee the solvability of the local problems. This method is now called FETI-H. The dual-primal generalization, FETI-DPH, was introduced later in [FLLA05, FATL05]. The generalization to BEM discretizations (“BETI-H”) was investigated by Steinbach and Windisch, see [Win10, SW11a, SW11b]. For a related method called FETI-2LM method see [FML<sup>+</sup>00b].

#### 2.8.2.4 Electromagnetic Problems

Eddy current problems are governed by the equation

$$\mathbf{curl}(\alpha \mathbf{curl} \mathbf{u}) + \beta \mathbf{u} = \mathbf{f},$$

with  $\alpha, \beta > 0$ , where  $\mathbf{u}$  is a vector field in  $H(\mathbf{curl})$ , usually discretized by Nédélec edge elements. Neumann-Neumann and FETI type methods have been first investigated by Toselli and coworkers [Tos99, Tos00, RT01, TK01, Tos06]. The results were further refined by Dohrmann and Widlund [DW12a, DW12b].

In electromagnetic scattering problems, the coefficient  $\beta$  above is negative. For a BETI approach on such kind of problems we refer to [Win10, SW12].

Finite and Boundary Element Tearing and  
Interconnecting Solvers for Multiscale Problems

Pechstein, C.

2013, XIV, 322 p., Hardcover

ISBN: 978-3-642-23587-0