A Hybrid Domain Decomposition Method and its Applications to Contact Problems
TR2009-924

by

Jungho Lee

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

New York University

September 2009

Professor Olof B. Widlund

Dedication

I dedicate this dissertation to my parents. Thank you for the sacrifices you have made for me; I owe you everything.

Acknowledgements

I thank my advisor Olof Widlund. Without his guidance, this dissertation would have been impossible. Thank you so much for your encouragement and invaluable advice, and I am very fortunate to have had an opportunity to work with you.

Abstract

Our goal is to solve nonlinear contact problems. We consider bodies in contact with each other divided into subdomains, which in turn are unions of elements. The contact surface between the bodies is unknown a priori, and we have a nonpenetration condition between the bodies, which is essentially an inequality constraint. We choose to use an active set method to solve such problems, which has both outer iterations in which the active set is updated, and inner iterations in which a (linear) minimization problem is solved on the current active face. In the first part of this dissertation, we review the basics of domain decomposition methods. In the second part, we consider how to solve the inner minimization problems. Using an approach based purely on FETI algorithms with only Lagrange multipliers as unknowns, as has been developed by the engineering community, does not lead to a scalable algorithm with respect to the number of subdomains in each body. We prove that such an algorithm has a condition number estimate which depends linearly on the number of subdomains across a body; numerical experiments suggest that this is the best possible bound. We also consider a new method based on the saddle point formulation of the FETI methods with both displacement vectors and Lagrange multipliers as unknowns. The resulting system is solved with a block-diagonal preconditioner which combines the one-level FETI and the BDDC methods. This approach allows the use of inexact solvers. We show that this new method is scalable with respect to the number of subdomains, and that its convergence rate depends only logarithmically on the number of degrees of freedom of the subdomains and bodies. In the last part of this dissertation, a model contact problem is solved by two approaches. The first one is a nonlinear algorithm which combines an active set method and the new method of Chapter 4. We also present a novel way of finding an initial active set. The second one uses the SMALBE algorithm, developed by Dostal et al. We show that the former approach has advantages over the latter.

Contents

List of Figures

List of Tables

Part I

The Basics

Chapter 1

Introduction

1.1 An Overview

Finite element discretizations of elliptic partial differential equations result in a very large, sparse algebraic system. Solving such a system directly can be very expensive. Even iterative methods, such as the conjugate gradient method when the system is symmetric and positive definite, can converge very slowly due to the large condition number of such systems. Therefore we precondition the system so that the preconditioned system has a much smaller condition number than the original system. Domain decomposition methods can be viewed as preconditioning techniques which can take advantage of modern parallel computers.

In domain decomposition methods, the original domain is split into potentially many subdomains, on each of which a small subproblem related to the original huge problem is solved directly. Data exchange occurs between neighboring subdomains. As the number of subdomains increases, we also need to solve a global problem, which involves a few unknowns for each subdomain, in order to prevent the convergence rate from deteriorating.

Domain decomposition methods can largely be divided into two categories, Schwarz methods (with overlapping domains) and substructuring methods (with nonoverlapping subdomains). In this dissertation, we will focus on the substructuring methods. We will concentrate on popular variants of the substructuring methods, namely one-level FETI (finite element tearing and interconnecting), FETI-DP (dual-primal FETI) and BDDC (balancing domain decomposition by constraints) methods.

We consider contact problems with multiple bodies. Contact problems are characterized by an active area of contact, which is unknown a priori, and inequality constraints, such as nonpenetration conditions; see [1, Section 3]. Therefore contact problems can be stated as energy minimization problems with inequality constraints. In this dissertation, we study two domain decomposition methods

under the assumption that an active set method is used for enforcing inequality constraints. In each step of an active set method the active set is updated, and a minimization problem on the current active set is solved approximately, until a desired accuracy is achieved. Thus, an active set method requires outer iterations in which the active set is updated and inner iterations in which a minimization problem is solved.

The two domain decomposition methods we present here deal with the inner minimization: the first is the FETI-FETI method, which is an obvious extension of the FETI methods to the context of the contact problems described above and has been used by the engineering community. The FETI-FETI method is shown not to be scalable with respect to the number of subdomains, both theoretically and numerically. A hybrid method is introduced next; it is a not-so-obvious, scalable alternative to the FETI-FETI method.

This dissertation is organized as follows.

- Part I: in Chapter 1, we provide the very basic functional analytic tools that are used throughout the theory of domain decomposition methods. In Chapter 2, we review the one-level FETI, FETI-DP and BDDC methods.
- Part II: in Chapter 3, we introduce a nonlinear model problem and the FETI-FETI method, and also provide an analysis of the convergence rate of the FETI-FETI method. In Chapter 4, we introduce the hybrid method and provide an eigenvalue analysis of its preconditioned operator.
- Part III: in Chapter 5, we solve the nonlinear model problem using a combination of an active set method and the hybrid method studied in Chapter 4. In Chapter 6, we transform the nonlinear model problem in its original primal form to its dual form in terms of Lagrange multipliers with bound and equality constraints, and solve it using the SMALBE (Semi-Monotonic Augmented Lagrangians for Bound and Equality constraints) algorithm introduced by Dostal. We compare these two methods.

1.2 Functional Analytic Tools

1.2.1 Sobolev Spaces

Assume Ω is a bounded domain in $\mathbb{R}^n, n = 2, 3$. $L^2(\Omega)$ is the space of all real-valued, measurable functions u which satisfy

$$\int_\Omega |u|^2 dx < \infty.$$

It is a Hilbert space with the scalar product

$$(u, v)_{L^2(\Omega)} = \int_\Omega uv dx$$

and an induced norm

$$||u||^2_{L^2(\Omega)} = (u, u)^2_{L^2(\Omega)} = \int_\Omega |u|^2 dx.$$

The space $H^1(\Omega) \subset L^2(\Omega)$ is a space of functions u such that $u, \nabla u \in L^2(\Omega)$, i.e.,

$$\int_\Omega |u|^2 dx < \infty \quad \text{and} \int_\Omega |\nabla u|^2 dx < \infty,$$

where ∇u is to be understood in terms of weak derivatives of u. The scaled H^1-norm of u is given by

$$||u||^2_{H^1(\Omega)} := \int_\Omega |\nabla u|^2 dx + \frac{1}{H^2_\Omega} \int_\Omega |u|^2 dx,$$

where H_Ω is the diameter of Ω; this scaling factor is obtained by dilation of a domain of unit diameter. The corresponding H^1- seminorm is defined by

$$|u|^2_{H^1(\Omega)} = \int_\Omega |\nabla u|^2 dx.$$

$H^1_0(\Omega)$ is a closure of $C^\infty_0(\Omega)$ in $H^1(\Omega)$, a subspace of functions in $H^1(\Omega)$ which vanish on the boundary in the $L^2(\partial\Omega)$ sense.

1.2.2 Trace and Extension Theorems

Assume Ω is a bounded Lipschitz domain in $\mathbb{R}^n, n = 2, 3$, and also $\Gamma \subset \partial\Omega$, a subset of positive measure. We define $H^{1/2}(\Gamma)$, the trace space of $H^1(\Omega)$, with the semi-norm and the full norm given by

$$|u|^2_{H^{1/2}(\Gamma)} = \int_\Gamma \int_\Gamma \frac{|u(x) - u(y)|^2}{|x - y|^d} dx dy,$$

and

$$||u||^2_{H^{1/2}(\Gamma)} = |u|^2_{H^{1/2}(\Omega)} + \frac{1}{H_\Gamma} ||u||^2_{L^2(\Gamma)},$$

where H_Γ is the diameter of Γ and d is the dimension of Ω.

1.2.3 Poincaré and Friedrichs' Inequalities

We first introduce the following theorem, which is [42, Lemme 2.7.1]:

Theorem 1.2.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and let $f_i, i = 1, \cdots, L, L \geq 1$ be functionals in $H^1(\Omega)$, such that, if u is constant in Ω,*

$$\sum_{i=1}^{L} |f_i(u)|^2 = 0 \Leftrightarrow u = 0.$$

Then, there exist constants C_1 and C_2, depending only on Ω and the functionals f_i, such that, for $u \in H^1(\Omega)$,

$$||u||^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)} + C_2 \sum_{i=1}^{L} |f_i(u)|^2.$$

The theorem follows from the compactness of $H^1(\Omega)$ in $L^2(\Omega)$, which in fact holds even for *John domains*; see [2]. *John domains* will be defined in Chapter 3. For the moment, we will only concentrate on Lipschitz domains. The following lemmas are special cases of Theorem 1.2.1, and are repeatedly used throughout the theory of domain decomposition methods.

Lemma 1.2.2 (Poincaré inequality)**.** *Let $\Omega \in \mathbb{R}^n$ be a bounded Lipschitz domain. Then, there exist constants C_1 and C_2, depending only on Ω, such that*

$$||u||^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)} + C_2 \left(\int_\Omega u dx \right)^2, \quad \forall u \in H^1(\Omega).$$

Lemma 1.2.3 (Friedrichs inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain, and $\Gamma \subset \partial\Omega$ have nonvanishing $(n-1)$- dimensional measure. Then, there exist constant C_1 and C_2, depending only on Ω and Γ, such that*

$$||u||^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)} + C_2 ||u||^2_{L^2(\Gamma)}, \quad \forall u \in H^1(\Omega).$$

1.3 Krylov Subspace Methods

In this dissertation, we will mainly use two Krylov subspace methods. One is the preconditioned conjugate gradient method for positive definite and symmetric problems. The other is the preconditioned conjugate residual method for symmetric, not necessarily positive definite, problems.

1.3.1 The Preconditioned Conjugate Gradient Method

We first introduce the conjugate gradient method in its original form and then its preconditioned version, following the presentation of [14, Section 3.1]. Suppose we want to solve the following problem:

$$\min_{x\in\mathbb{R}^n} f(x), \tag{1.1}$$

or, equivalently,

$$Ax = b, \tag{1.2}$$

where $f(x) = \frac{1}{2}x^TAx - b^Tx$, $b \in \mathbb{R}^n$ is a given vector and $A \in \mathbb{R}^{n\times n}$ is a symmetric and positive definite matrix. Suppose we have nonzero vectors $\{p^k\}_{k=1}^n$ which are *A-conjugate*, i.e.,

$$\langle p^i, Ap^j\rangle = \langle p^i, p^j\rangle_A = 0, \quad \forall i \neq j.$$

Such $p^1, \cdots, p^n$ are linearly independent and thus form a basis of $\mathbb{R}^n$. Any $x \in \mathbb{R}^n$ can be written in the form

$$x = \xi_1 p^1 + \cdots + \xi_n p^n. \tag{1.3}$$

Subsituting (1.3) into f, we obtain

$$\min_{x\in\mathbb{R}^n} f(x) = \min_{\xi_1\in\mathbb{R}} f(\xi_1 p^1) + \cdots + \min_{\xi_n\in\mathbb{R}} f(\xi_n p^n). \tag{1.4}$$

Thus the original problem (1.1), or (1.2), has been turned into n one-dimensional problems, and it is easy to see that the solution $\hat{x}$ of the original problem is given by

$$\hat{x} = \xi_1 p^1 + \cdots + \xi_n p^n,$$

where

$$\xi_i = b^Tp^i/(p_i)^TAp_i, \quad i = 1, \cdots, n.$$

Finding the exact solution $\hat{x}$ can be an enormous task when the dimension n is large. In such a case, it is natural to find an approximation $\widetilde{x}$ of $\hat{x}$ with an initial guess x^0 and just a few elements of the set $\{p^k\}_{k=1}^n$. Suppose we want to find a minimizer x^k of f in the set $S^k := x^0 + \text{Span}\{p^1, \cdots, p^k\}$. Any $x \in S^k$ can be written in the form

$$x = x^0 + \xi_1 p^1 + \cdots \xi_k p^k,$$

and substituting the expression above into f and using the A-conjugacy of $\{p^k\}_{k=1}^n$, we obtain

$$f(x) = f(x^0) \quad + \quad \left(\frac{1}{2}\xi_1^2(p^1)^TAp^1 + \xi_1(Ax^0 - b)^Tp^1\right) + \cdots$$

$$+ \quad \left(\frac{1}{2}\xi_k^2 (p^k)^T A p^k + \xi_k (Ax^0 - b)^T p^k\right).$$

With $g^0 := \nabla f(x^0) = Ax^0 - b$ and $f_0(x) := 1/2 x^T Ax - x^T g_0$, we have

$$f(x) = f(x^0) + f_0(\xi_1 p^1) + \cdots + f_0(\xi_k p^k)$$

and

$$f(x^k) = \min_{x \in S^k} f(x) = f(x^0) + \min_{\xi_1 \in \mathbb{R}} f(\xi_1 p^1) + \cdots + \min_{\xi_k \in \mathbb{R}} f(\xi_k p^k). \tag{1.5}$$

From (1.5) also follows

$$f(x^k) = \min_{x \in S^k} f(x) = f(x^{k-1}) + \min_{\xi \in \mathbb{R}} f(\xi p^k), \quad k \geq 1, \tag{1.6}$$

and that we can generate the approximations x^k iteratively. The *conjugate direction method* starts from an arbitrary initial guess x^0, and given x^{k-1}, we can find x^k with the formula

$$x^k = x^{k-1} - \alpha_k p^k, \quad \alpha_k = (g^0)^T p^i / (p^i)^T A p^i.$$

We also need an efficient way of generating $\{p^k\}_{k=1}^n$; in the conjugate gradient method, we apply the Gram-Schmidt process to the Krylov spaces

$$\mathcal{K}^k = \mathcal{K}^k(A, g^o) = \text{Span}\{g^0, Ag^0, \cdots, A^{k-1} g^0\}, \quad k = 1, \cdots, n,$$

to obtain a set of search directions. The complete conjugate gradient method for the solution of (1.1) is is described in Figure 1.1.

1. Initialize: choose $x^0 \in \mathbb{R}^n$, set $g^0 = Ax^0 - b, p^1 = g^0$

2. Iterate $k = 1, 2, 3, \cdots$, while $||g^{k-1}|| > 0$

$$\begin{aligned}
\alpha_k &= \langle g^{k-1}, g^{k-1} \rangle / \langle Ap^k, p^k \rangle \\
x^k &= x^{k-1} - \alpha_k p^k \\
g^k &= g^{k-1} - \alpha_k A p^k \\
\beta_k &= \langle g^k, g^k \rangle / \langle g^{k-1}, g^{k-1} \rangle = -\langle g^k, Ap^k \rangle / \langle Ap^k, p^k \rangle \\
p^{k+1} &= g^k + \beta_k p^k
\end{aligned}$$

Figure 1.1: CG (Conjugate Gradient) Algorithm

We have the following error bound for the conjugate gradient method; see [14, Theorem 3.2].

Theorem 1.3.1. *Let A be symmetric and positive definite and $\hat{x}$ the solution of (2.4). Let $x^k, k = 0, 1, 2, \cdots$ be the iterates of the conjugate gradient method. Then the error $e^k = x^k - \hat{x}$ satisfies*

$$||e^k||_A \leq 2 \left(\frac{\sqrt{\mathcal{K}(A)} - 1}{\sqrt{\mathcal{K}(A)} + 1} \right)^k ||e^0||_A, \tag{1.7}$$

where $\mathcal{K}(A)$ denotes the spectral condition number of A.

As Theorem 1.3.1 suggests, the conjugate gradient method in its original form can require many iterations for a desired accuracy when the condition number of the system matrix $\mathcal{K}(A)$ is large, which is often the case for system matrices which arise from PDE discretizations. We therefore consider the following transformed equation, which is equivalent to (1.2),

$$M^{-1/2} A M^{-1/2} y = M^{-1/2} b, \tag{1.8}$$

where M^{-1} is also symmetric and positive definite and $\mathcal{K}(M^{-1/2} A M^{-1/2}) \ll \mathcal{K}(A)$. Applying the conjugate gradient algorithm to (1.8), we obtain the preconditioned conjugate gradient method; see Figure 1.2.

1. Initialize: choose $x^0 \in \mathbb{R}^n$, set $g^0 = Ax^0 - b, z^0 = M^{-1} g^0, p^1 = z^0$

2. Iterate $k = 1, 2, 3, \cdots$, while $||g^{k-1}|| > 0$

$$\begin{aligned} \alpha_k &= \langle z^{k-1}, g^{k-1} \rangle / \langle Ap^k, p^k \rangle \\ x^k &= x^{k-1} - \alpha_k p^k \\ g^k &= g^{k-1} - \alpha_k A p^k \\ \beta_k &= \langle z^k, g^k \rangle / \langle z^{k-1}, g^{k-1} \rangle \\ p^{k+1} &= g^k + \beta_k p^k \end{aligned}$$

Figure 1.2: PCG (Preconditioned Conjugate Gradient) Algorithm

1. Initialize: choose $x^0 \in \mathbb{R}^n$, set $r^0 = b - Au^0, p^{-1} = 0, p^0 = M^{-1}r^0$

2. Iterate $k = 1, 2, 3, \cdots$, while $||g^{k-1}|| > 0$

$$\begin{aligned}
\beta &= \langle r^{k-1}, M^{-1}Ap^{k-1}\rangle / \langle Ap^{k-1}, M^{-1}Ap^{k-1}\rangle \\
u^k &= u^{k-1} + \beta p^{k-1} \\
r^k &= r^{k-1} - \beta Ap^{k-1} \\
\alpha_0 &= \langle AM^{-1}Ap^{k-1}, M^{-1}Ap^{k-1}\rangle / \langle Ap^{k-1}, M^{-1}Ap^{k-1}, M^{-1}Ap^{k-1}\rangle \\
\alpha_1 &= \langle AM^{-1}Ap^{k-1}, M^{-1}Ap^{k-2}\rangle / \langle Ap^{k-2}, M^{-1}Ap^{k-2}\rangle \\
p^k &= M^{-1}Ap^{k-1} - \alpha_0 p^{k-1} - \alpha_1 p^{k-2}
\end{aligned}$$

Figure 1.3: PCR (Preconditioned Conjugate Residual) Algorithm

1.3.2 The Preconditioned Conjugate Residual Method

The preconditioned conjugate residual method can be viewed as a generalization of the preconditioned conjugate gradient method, for symmetric, not necessarily positive definite, systems; it can essentially be derived by replacing the Euclidean scalar products of Figure 1.2 with certain bilinear forms. For a fine introduction, see [25, Section 9.5].

We consider the system (1.2), where $A \in \mathbb{R}^n$ is symmetric but not necessarily positive definite. Let $M \in \mathbb{R}^n$ be a symmetric, positive definite matrix. The preconditioned conjugate residual method is described in Figure 1.3.

We define

$$\mathcal{K}(M^{-1}A) = \frac{\mu_{max}}{\mu_{min}} = \frac{\max\{|\lambda| : \lambda \in \sigma(M^{-1}A)\}}{\min\{|\lambda| : \lambda \in \sigma(M^{-1}A)\}}, \tag{1.9}$$

where $\sigma(M^{-1}A)$ is the spectrum of $M^{-1}A$. We have the following result, which is taken from [43, C.6.2]; a proof can be found in [25, Section 9.5].

Lemma 1.3.2. *Let A be regular and symmetric and M symmetric and positive definite. Then, after k steps of the PCR algorithm, the norm of the residual is bounded by*

$$||M^{-1/2}r^k||_2 \leq \frac{2\rho^\mu}{1+\rho^{2\mu}}||M^{-1/2}r^0||_2,$$

where $\rho = \frac{\mathcal{K}-1}{\mathcal{K}+1}$ and $\mu \in \mathbb{Z}$, such that $k/2 - 1 < \mu \leq k/2$.

Chapter 2

Iterative Substructuring Methods with Nonoverlapping Subdomains

2.1 Introduction

The original FETI method, which later became to be known as the one-level FETI method, was first introduced by Farhat and Roux [24]. The discretization of an elliptic partial differential equation, with subdomain interface continuity conditions which are needed when a domain decomposition (with no coupling between subdomains) is introduced, can be formulated as a Karush-Kuhn-Tucker (KKT) system with the displacement vectors as primal unknowns and the Lagrange multipliers as dual unknowns. In this original method, the KKT system is reduced to an equation in terms of the Lagrange multipliers alone; this reduction process requires some care in case of the presence of subdomains lacking essential boundary conditions. The resulting equation is solved using the conjugate gradient method. The reduction process involves solving a Neumann problem exactly on each subdomain. The resulting algorithm is scalable, in the sense that the number of iterations needed to achieve a certain accuracy is independent of the number of subdomains, or, subproblems. Later Farhat, Mandel and Roux introduced a variant of the FETI method [23] with a *Dirichlet preconditioner*, in which an additional Dirichlet problem is solved exactly on each subdomain in the preconditioning step. The use of this preconditioner makes the resulting algorithm even less sensitive to the number of unknowns in each subproblem. Theoretical analysis of the convergence of the one-level FETI method with Dirichlet preconditioners was first carried out by Mandel and Tezaur in [40], and they showed that the conndition number $\mathcal{K}$ satisfies the following upper bound:

$$\mathcal{K} \leq C\left(1 + log\frac{H}{h}\right)^2, \tag{2.1}$$

where H, h are the diameters of a typical subdomain and a typical element, respectively. See also [6], [7] and [35].

The second generation of the FETI method, namely the dual-primal FETI (FETI-DP) method, was introduced for two-dimensional problems by Farhat, Lesoinne, Le Tallec, Pierson and Rixen in [22]. In this method, a certain degree of the continuity coupling between subdomains, also known as *primal constraints*, is introduced. The continuity of the displacement vectors at some select interface nodes is built into the problem formulation, as in *primal* methods, whereas the continuity at other interface nodes is imposed by the use of *dual* Lagrange multipliers as in the one-level FETI method; thus the name *dual-primal* FETI. In FETI-DP methods, sufficiently many primal constraints are introduced so that the resulting stiffness matrix for the entire system becomes invertible. In addition, these primal constraints also provide a coarse solver which is needed for the scalability of the algorithm. Mandel and Tezaur carried out a theoretical analysis for the FETI-DP method in the case of two-dimensional second and fourth order problems, and obtained the same condition number estimate (2.1).

The FETI-DP method is preferred to the original one-level FETI method for several reasons. Among them, a major advantage is that the introduction of the primal continuity constraints eliminates the need to solve singular problems. Solving singular subproblems is not a trivial matter; see [21]. Also, the FETI-DP method allows us great flexibility in choosing primal continuity constraints, which provide the coarse solver, and can make the algorithm robust with respect to the PDE coefficients.

The BDDC (Balancing Domain Decomposition by Constraints) method was first introduced by Dohrmann in [11]. It is a variant of the two-level Neumann-Neumann type preconditioner, with its coarse problem similar to that of the corresponding FETI-DP method; see [38]. Mandel and Dohrmann proved that the BDDC method has the same condition number estimate (2.1) in [39], using the abstract Schwarz framework.

The one-level FETI and the FETI-DP methods are *dual* substructuring methods, whereas the BDDC method is a *primal* substructuring method. This terminology follows from the fact that a minimization problem, in terms of the primal variables, is solved in the BDDC method, and an equivalent problem, in terms of the dual variables, is solved in the FETI methods.

In Chapters 3 and 4, we will consider the FETI-FETI and a hybrid methods, respectively. One-level FETI, FETI-DP and BDDC methods serve as building blocks of those methods and we therefore briefly describe them here.

2.2 Problem Setting

2.2.1 Domain Decomposition

We consider a second-order scalar elliptic problem on a bounded domain $\Omega \subset \mathbb{R}^n, n = 2, 3$. We denote the boundary of Ω by $\partial\Omega$, and assume that homogeneous Dirichlet boundary conditions are imposed on $\partial\Omega_D \subset \partial\Omega$, which is a subset of $\partial\Omega$ with a positive measure. Let $\partial\Omega_N := \partial\Omega \setminus \partial\Omega_D$ be its complement. The corresponding Sobolev space in which the solution will be found is $H_0^1(\Omega, \partial\Omega_D) := \{v \in H^1(\Omega) : u = 0 \quad \text{on} \quad \partial\Omega_D\}$. We find $u \in H_0^1(\Omega, \partial\Omega_D)$ such that

$$a(u, v) = f(v), \quad \forall v \in H_0^1(\Omega, \partial\Omega_D), \tag{2.2}$$

where

$$a(u, v) := \int_\Omega \rho(x)\nabla u \cdot \nabla v, \quad f(v) = \int_\Omega fv. \tag{2.3}$$

Note that (2.2) is equivalent to the following minimization problem:

$$\min_{u \in H_0^1(\Omega, \partial\Omega_D)} \frac{1}{2}a(u, u) - f(u). \tag{2.4}$$

We decompose Ω into N nonoverlapping subdomains $\Omega_i, i = 1, \cdots, N$, each of which is the union of shape-regular elements. The diameter of Ω_i is H_i, and we set $H = \max_i H_i$. The triangulation of the subdomain Ω_i is of diameter h_i, and we set $h = \max_i h_i$. We note that many of the estimates in this dissertation will be expressed in terms of the ratio H/h, which is to be interpreted as $\max_i H_i/h_i$. We also note that $(H_i/h_i)^n, \Omega_i \subset \mathbb{R}^n$ gives a measure of the number of degrees of freedom associated with Ω_i.

The finite element nodes on the boundaries of neighboring subdomains match across the interface $\Gamma := \cup_{i \neq j} \partial\Omega_i \cap \partial\Omega_j$. Γ is the union of

- faces, edges and vertices in three dimensions: faces, regarded as open subsets of Γ, are shared by two subdomains. Edges, regarded as open subsets of the boundaries of the faces, are shared by more than two subdomains. Vertices are endpoints of edges.

- edges and vertices in two dimensions: edges, regarded as open subsets of Γ, are shared by two subdomains. Vertices, as in three dimensions, are endpoints of edges.

We note that the nodal values on $\partial\Omega_D$ will always vanish and those on $\partial\Omega_N$ which belong to only one subdomain will effectively belong to the subdomain interior. They will be eliminated together with the interior degrees of freedom when the given linear system is reduced to a Schur complement system associated with the

interface Γ.
We assume that $\rho(x) = \rho_i \geq \rho_{min} > 0, \forall x \in \Omega_i, i = 1, \cdots, N$. We also introduce the corresponding set of interface nodes $\Gamma_h := \cup_{i \neq j} \partial\Omega_{i,h} \cap \partial\Omega_{j,h}$, where $\partial\Omega_{i,h}$ and $\partial\Omega_{j,h}$ are the sets of finite element nodes on $\partial\Omega_i$ and $\partial\Omega_j$, respectively. We also define local bilinear forms and linear functionals,

$$a^{(i)}(u,v) := \int_{\Omega_i} \rho(x) \nabla u \cdot \nabla v, \quad f^{(i)}(v) = \int_{\Omega_i} fv, \quad i = 1, \cdots, N. \tag{2.5}$$

We will consider linear elasticity problems in $\mathbb{R}^n, n = 2, 3$ as well. The equations of linear elasticity model the displacement of a linear elastic material under the action of external and internal forces. The elastic body occupies a bounded domain $\Omega \subset \mathbb{R}^n, n = 2, 3$. We denote its boundary by $\partial\Omega$ and assume that a part of it, $\partial\Omega_D$, is clamped, i.e., Dirichlet boundary conditions are imposed on $\partial\Omega_D$, and that the rest of the boundary, $\partial\Omega_N := \partial\Omega \setminus \partial\Omega_D$ is subject to a surface force $\mathbf{g}$, i.e., a natural boundary condition. We also introduce a body force $\mathbf{f}$, e.g., gravity. With $\mathbf{H^1}(\Omega) := (H^1(\Omega))^n, n = 2, 3$, the appropriate space for a variational formulation is the Sobolev space $\mathbf{H^1_0}(\Omega, \partial\Omega_D) := \{\mathbf{v} \in \mathbf{H^1}(\Omega) : \mathbf{v} = \mathbf{0} \quad \text{on} \quad \partial\Omega_D\}$. The linear elasticity problem consists in finding the displacement $\mathbf{u} \in \mathbf{H^1_0}(\Omega, \partial\Omega_D)$ of the elastic body Ω such that $\forall \mathbf{v} \in \mathbf{H^1_0}(\Omega, \partial\Omega_D)$,

$$\int_\Omega G(\mathbf{x}) \epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) d\mathbf{x} + \int_\Omega G(\mathbf{x}) \beta(\mathbf{x}) \text{div}\mathbf{u}\, \text{div}\mathbf{v} d\mathbf{x} = \langle \mathbf{F}, \mathbf{v} \rangle. \tag{2.6}$$

Here G and β are material parameters that depend on the Young's modulus $E > 0$ and the Poisson ratio $\nu \in (0, \frac{1}{2})$; we have $G = E/(1+\nu)$ and $\beta = \nu/(1-2\nu)$. The coefficients are also referred to as the Lamé parameters. In this dissertation, we only consider the case of compressible elasticity, which means that the Poisson ratio ν is bounded away from 1/2. Furthermore, $\epsilon_{ij}(\mathbf{u}) := \frac{1}{2}(\partial u_i/\partial x_j + \partial u_j/\partial x_i)$ are the elements of the linearized strain tensor, and

$$\epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) = \sum_{i,j=1}^{3} \epsilon_{ij}(\mathbf{u}) \epsilon_{ij}(\mathbf{v}), \quad \langle \mathbf{F}, \mathbf{v} \rangle := \int_\Omega \mathbf{f}^T \mathbf{v} d\mathbf{x} + \int_{\partial\Omega_N} \mathbf{g}^T \mathbf{v} d\sigma.$$

2.2.2 Finite Element Spaces

We discuss the choice of the space of finite element functions in one-level FETI, FETI-DP, and BDDC methods. We denote a standard finite element space of continuous, piecewise linear functions on Ω_i by $W^{(i)}$. We will always assume that these functions vanish on $\partial\Omega_D$. Each $W^{(i)}$ is decomposed into a subdomain interior

part $W_I^{(i)}$ and a subdomain interface part $W_\Gamma^{(i)}$:

$$W^{(i)} = W_I^{(i)} \oplus W_\Gamma^{(i)}.$$

We denote the associated product spaces by $W := \prod_{i=1}^N W^{(i)}, W_I := \prod_{i=1}^N W_I^{(i)}$, and $W_\Gamma := \prod_{i=1}^N W_\Gamma^{(i)}$.

The functions in W and W_Γ are in general discontinuous across the interface, whereas the finite element solutions are continuous across the interface Γ. Therefore we introduce $\widehat{W}$ and $\widehat{W}_\Gamma$, which are the continuous subspaces of W and W_Γ, respectively.

For the FETI-DP and BDDC methods, we will also need a subspace $\widetilde{W} \subset W$, intermediate between W and $\widehat{W}$, which consists of finite element functions which satisfy certain continuity constraints. The corresponding interface space is denoted by $\widetilde{W}_\Gamma$. In the two-dimensional case, we require the functions in $\widetilde{W}$ to be continuous at subdomain vertices. In the three-dimensional case, enforcing such vertex constraints alone makes the condition number of the resulting algorithm very sensitive to the number of degrees of freedom on each subdomain and we need different continuity constraints to obtain a better algorithm; we will give more details in Section 2.6.

We introduce the following decomposition of $\widetilde{W}_\Gamma$:

$$\widetilde{W}_\Gamma = W_\Delta \oplus \widehat{W}_\Pi = \left(\prod_{i=1}^N W_\Delta^{(i)} \right) \oplus \widehat{W}_\Pi,$$

where $\widehat{W}_\Pi$, a *primal subspace*, consists of continuous functions, and $W_\Delta^{(i)}$, a *dual subspace*, consists of functions which are allowed to be discontinuous across the interface. More precisely, $\widehat{W}_\Pi$ is spanned by subdomain vertex nodal basis functions, i.e., consists of functions which are nonzero on Γ only at subdomain vertices (*primal nodes*), in the two-dimensional case. Accordingly, $W_\Delta^{(i)} \in W_\Gamma^{(i)}$ consist of functions which are zero at the vertices of the subdomain Ω_i. The terminologies *primal* and *dual* indicate the fact that the continuity is imposed in the manner of *primal methods* at the *primal* nodes, and in the manner of *dual methods* (i.e., via Lagrange multipliers) at the *dual* nodes, respectively. In the three-dimensional case, we need to be more careful about the design of $\widehat{W}_\Pi$, i.e., the choice of *primal* constraints and $W_\Delta^{(i)}, i = 1, \cdots, N$, due to the reason mentioned above. W_Δ is the product space of $W_\Delta^{(i)}, i = 1, \cdots, N$ and we also define $W_\Pi = \prod_{i=1}^N W_\Pi^{(i)}$, where $W_\Pi^{(i)}$ is the local subspace of $\widehat{W}_\Pi$ for the subdomain $\Omega_i, i = 1, \cdots, N$. See Figure 2.1 for a depiction of $W, \widetilde{W}$, and $\widehat{W}$ in the two-dimensional case.

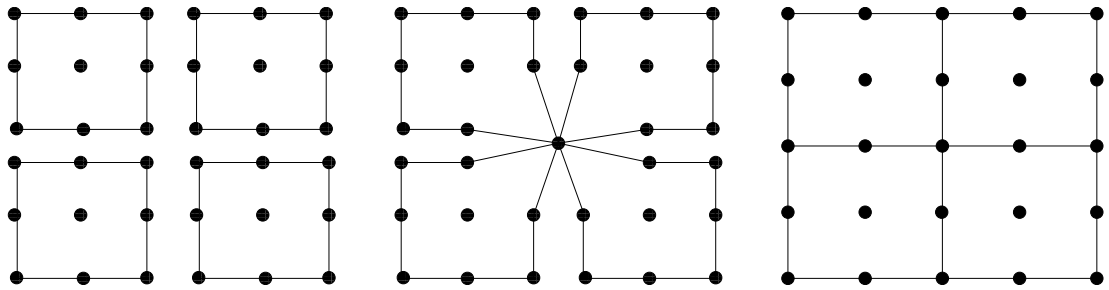
We note that we will not distinguish between a finite element function and its vector counterpart of nodal values.

For each subdomain $\Omega_i, i = 1, \cdots, N$, we assemble local stiffness matrices

$$A^{(i)} = \begin{bmatrix} A_{II}^{(i)} & A_{\Gamma I}^{(i)^T} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{bmatrix}$$

and local load vectors $f^{(i)}$ by integrating appropriate expressions over individual subdomains.

Figure 2.1: $W, \widetilde{W}$ and $\widehat{W}$



(a) W: One-Level (b) $\widetilde{W}$: FETI-DP (c) $\widehat{W}$: BDDC

2.3 Some Useful Operators

We routinely use several restriction, extension, and scaling operators in the theory of domain decomposition methods. These operators serve many purposes, and among them is the need to describe small subdomain problems in terms of the original huge problem on the entire domain.and to establish a connection between different finite element structures.

The restriction operator $R_\Gamma^{(i)}$ maps a vector of the product space W_Γ to its restriction in the subdomain space $W_\Gamma^{(i)}$. $\widetilde{R}_\Gamma^{(i)}$ and $\widehat{R}_\Gamma^{(i)}$ are similar, and represent restrictions from $\widetilde{W}_\Gamma$ and $\widehat{W}_\Gamma$, respectively, to $W_\Gamma^{(i)}$. $\widetilde{R}_\Gamma : \widetilde{W}_\Gamma \to W_\Gamma$ and $\widehat{R}_\Gamma : \widehat{W}_\Gamma \to W_\Gamma$ are the direct sums of $\widetilde{R}_\Gamma^{(i)}$ and $\widehat{R}_\Gamma^{(i)}$, respectively. $R_\Pi^{(i)}$ is the restriction operator from $\widehat{W}_\Pi$ to $W_\Pi^{(i)}$.

$R_{\Gamma\Delta}^{(i)}$ extracts the subdomain part corresponding to the subdomain space $W_\Delta^{(i)}$ from a vector in $\widehat{W}_\Gamma$. Similarly, $R_{\Gamma\Pi}$ extracts the part that corresponds to $\widehat{W}_\Pi$ from a vector in $\widehat{W}_\Gamma$. $\bar{R}_\Gamma : \widehat{W}_\Gamma \to \widetilde{W}_\Gamma$ is the direct sum of $R_{\Gamma\Delta}^{(i)}$ and $R_{\Gamma\Pi}$.

We also introduce scaling factors $\delta_i^\dagger(x)$ for each node $x \in \Gamma_h \cap \partial\Omega_{i,h}, i = 1, \cdots, N$: for $\gamma \in [1/2, \infty)$,

$$\delta_i^\dagger(x) = \frac{\rho_i^\gamma}{\sum_{j \in \mathcal{N}_x} \rho_j^\gamma}, \quad x \in \partial\Omega_{i,h} \cap \Gamma_h.$$

Here, $\mathcal{N}_x$ is the set of indices j of the subdomains such that $x \in \partial\Omega_{j,h}$.

We also introduce the scaled version of $R^{(i)}_{\Gamma\Delta}$, which we denote by $R^{(i)}_{D,\Gamma\Delta}$; each row of $R^{(i)}_{\Gamma\Delta}$ has exactly one nonzero entry corresponding to a node x on the subdomain interface. Multiplying each such entry with $\delta_i^\dagger(x)$ results in the scaled version $R^{(i)}_{D,\Gamma\Delta}$. $\bar{R}_{D,\Gamma}$ is the direct sum of $R^{(i)}_{D,\Gamma\Delta}$ and $R_{\Gamma\Pi}$. The scaled version of $\widehat{R}^{(i)}_\Gamma$, which we denote by $\widehat{R}^{(i)}_{D,\Gamma}$, is defined analogously, and so is their direct sum $\widehat{R}_{D,\Gamma}$.

It is easy to see that

$$\widehat{R}_\Gamma^T \widehat{R}_{D,\Gamma} = \widehat{R}_{D,\Gamma}^T \widehat{R}_\Gamma = I \quad \text{and} \quad \bar{R}_\Gamma^T \bar{R}_{D,\Gamma} = \bar{R}_{D,\Gamma}^T \bar{R}_\Gamma = I \quad \text{on} \quad \widehat{W}_\Gamma.$$

We define the averaging operators on $\widehat{W}_\Gamma$ and $\widetilde{W}_\Gamma$ by

$$\widehat{E}_D := \widehat{R}_\Gamma \widehat{R}_{D,\Gamma}^T \quad \text{and} \quad E_D := \bar{R}_\Gamma \bar{R}_{D,\Gamma}^T,$$

respectively.

We need to represent the difference of the values of the displacement unknowns at a node common to two or more subdomains, in the FETI methods. We usually use the symbol B, and add more symbols to indicate the space on which it acts. For instance, the matrix

$$B = [B^{(1)}, B^{(2)}, \cdots, B^{(N)}]$$

consists of elements $0, 1, -1$ such that $Bu = 0, u \in W$ if and only if the values of u associated with more than one subdomain boundary conincide. The columns of $B^{(i)}$, which correspond to the interior nodes of Ω_i, are zero. Thus, $B^{(i)} = [\ 0 \quad B^{(i)}_\Gamma\ ]$ when the interior degrees of freedom are ordered first. B_Γ is obtained by eliminating the zero columns of B which correspond to the interior nodes, or

$$B_\Gamma = [B^{(1)}_\Gamma, B^{(2)}_\Gamma, \cdots, B^{(N)}_\Gamma].$$

Similarly, $\widetilde{B}$ and $\widetilde{B}_\Gamma$ are jump operators which act on the spaces $\widetilde{W}$ and $\widetilde{W}_\Gamma$, respectively.

We also introduce the scaled versions of B_Γ and $\widetilde{B}_\Gamma$, which we denote by $B_{D,\Gamma}$ and $\widetilde{B}_{D,\Gamma}$, respectively. $B^{(i)}_{D,\Gamma}$ is obtained as follows: each nonzero entry of $B^{(i)}_\Gamma$ contributes to the Lagrange multiplier enforcing the continuity at a node $x \in \partial\Omega_i \cap \partial\Omega_j$ and is multiplied by $\delta_j^\dagger(x)$ to produce the corresponding $B^{(i)}_{D,\Gamma}$. $\widetilde{B}_{D,\Gamma}$ is

obtained in the same manner.

We define

$$P_D := \widetilde{B}_{D,\Gamma}^T \widetilde{B}_\Gamma.$$

Lemma 2.3.1.

$$E_D + P_D = I; \quad E_D^2 = E_D, P_D^2 = P_D; \quad E_D P_D = P_D E_D = 0$$

Proof. See [38, Lemma 1]. □

2.4 Schur Complement Systems and Discrete Harmonic Extensions

In the first step of many iterative substructuring algorithms, the unknowns in the interior of each subdomain are eliminated. In this step, the Schur complement with respect to the interface unknowns with their nodes on $\partial\Omega_i \cap \Gamma$ is introduced. For instance, consider the KKT system (2.9) for the one-level FETI method, where the local stiffness matrices are

$$A^{(i)} = \begin{bmatrix} A_{II}^{(i)} & A_{\Gamma I}^{(i)^T} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{bmatrix}, \quad i = 1, \cdots, N.$$

The corresponding local Schur complements are

$$S^{(i)} = A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)} A_{II}^{(i)^{-1}} A_{\Gamma I}^{(i)^T}, \quad i = 1, \cdots, N,$$

and the Schur complement for the entire system is $S := diag_{i=1}^N S^{(i)}$, a direct sum of $S^{(i)}, i = 1, \cdots, N$.

We introduce the space of *discrete harmonic functions*, which is an important subspace directly related to the Schur complements. A function $u^{(i)}$ is said to be discrete harmonic on Ω_i if

$$A_{II}^{(i)} u_I^{(i)} + A_{I\Gamma}^{(i)} u_\Gamma^{(i)} = 0.$$

From this definition, we can see that a discrete harmonic function in Ω_i is completely determined by its values on the subdomain boundary $\partial\Omega_i \cap \Gamma$. We use the notation $u^{(i)} := \mathcal{H}_i(u_\Gamma^{(i)})$ to indicate extending $u_\Gamma^{(i)}$ into the interior of Ω_i so that the resulting function is discrete harmonic in Ω_i, and call $\mathcal{H}_i$ the discrete harmonic extension operator on Ω_i. Also, we denote the piecewise discrete harmonic extension of u_Γ into the entire Ω by $\mathcal{H}(u_\Gamma)$.

Lemma 2.4.1. *Let $u_\Gamma^{(i)}$ be the restriction of a finite element function to $\partial\Omega_i \cap \Gamma$.*

Then, the discrete harmonic extension $u^{(i)} = \mathcal{H}_i(u_\Gamma^{(i)})$ *of* $u_\Gamma^{(i)}$ *into* Ω_i *satisfies*

$$u_\Gamma^{(i)^T} A^{(i)} u_\Gamma^{(i)} = \min_{v^{(i)}|_{\partial\Omega_i \cap \Gamma = u_\Gamma^{(i)}}} v^{(i)^T} A^{(i)} v^{(i)}$$

and

$$u_\Gamma^{(i)^T} S^{(i)} u_\Gamma^{(i)} = u^{(i)^T} A^{(i)} u^{(i)}$$

Analogously, if u_Γ *is the restriction of a finite element function to* Γ*, the piecewise discrete harmonic extension* $u = \mathcal{H}(u_\Gamma)$ *of* u_Γ *into the interior of the subdomains satisfies*

$$u^T A u = \min_{v|_\Gamma = u_\Gamma} v^T A v$$

and

$$u_\Gamma^T S u_\Gamma = u^T A u.$$

2.5 One-Level FETI methods

In this subsection, we review the one-level FETI method. We use the finite element functions in the space W to discretize the minimization problem (2.4) (or, equivalently, the variational problem (1.2)). Since the functions in W are in general discontinuous across the interface Γ, we need to enforce the continuity condition explicitly:

$$\min_{u \in W} \frac{1}{2} a(u,u) - f(u), \quad \text{subject to} \quad Bu = 0. \tag{2.7}$$

We can rewrite the minimization problem (2.7) using matrix notation:

$$\min_{u \in W} \frac{1}{2} u^T A u - f^T u, \quad \text{subject to} \quad Bu = 0, \tag{2.8}$$

where

$$A = \begin{bmatrix} A^{(1)} & & \\ & \ddots & \\ & & A^{(N)} \end{bmatrix}, \quad f = \begin{bmatrix} f^{(1)} \\ \vdots \\ f^{(N)} \end{bmatrix}.$$

Introducing a vector of Lagrange multipliers λ to enforce the continuity constraint $Bu = 0$, we obtain the following Karush-Kuhn-Tucker (KKT) system:

Find $(u, \lambda) \in W \times \text{range}(B)$, such that

$$\left.\begin{array}{rcrcl} Au & + & B^T\lambda & = & f \\ Bu & & & = & 0 \end{array}\right\}. \tag{2.9}$$

λ is unique only up to an additive element of $\ker(B^T)$. The space of Lagrange multipliers is therefore chosen as $\mathrm{range}(B)$.
Eliminating the interior unknowns in each subdomain, we obtain the following:

Find $(u_\Gamma, \lambda) \in W_\Gamma \times \mathrm{range}(B_\Gamma)$, such that

$$\left.\begin{array}{rcrcl} Su_\Gamma & + & B_\Gamma^T\lambda & = & g \\ B_\Gamma u_\Gamma & & & = & 0 \end{array}\right\}, \tag{2.10}$$

where

$$S = \begin{bmatrix} S^{(1)} & & \\ & \ddots & \\ & & S^{(N)} \end{bmatrix}, \quad S^{(i)} = A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)} A_{II}^{(i)^{-1}} A_{\Gamma I}^{(i)^T}, i = 1, \cdots, N,$$

$$g = \begin{bmatrix} g^{(1)} \\ \vdots \\ g^{(N)} \end{bmatrix}, \quad g^{(i)} = f_\Gamma^{(i)} - A_{\Gamma I}^{(i)} A_{II}^{(i)^{-1}} f_I^{(i)}, i = 1, \cdots, N.$$

In all FETI methods, we reduce the KKT system (2.10) to an equation of λ alone, by solving the first equation of (2.10) for u_Γ. The matrices A in (2.9) and S in (2.10), however, are singular in general, when there are subdomains with boundaries which do not intersect the Dirichlet boundary $\partial\Omega_D$. We call such subdomains floating. In such a case the solution of the first equation of (2.10) exists if and only if $g - B_\Gamma^T\lambda \in \mathrm{range}(S)$; this requirement leads to the introduction of a projection P, which will be introduced shortly. First, we introduce a matrix R such that $\mathrm{range}(R) = \ker(S)$:

$$R = \begin{bmatrix} R^{(1)} & & \\ & \ddots & \\ & & R^{(N)} \end{bmatrix},$$

where $R^{(i)}$ consists of the null vectors of $S^{(i)}, i = 1, \cdots, N$. Subdomains with nonsingular stiffness matrices do not contribute to the matrix R, i.e., $R^{(i)}$ is an empty matrix if $\partial\Omega_i \cap \partial\Omega_D \neq \varnothing$. We can now solve the first equation of (2.10) for u_Γ:

$$u_\Gamma = S^\dagger(g - B_\Gamma^T\lambda) + R\alpha \quad \text{if} \quad g - B_\Gamma^T\lambda \in \mathrm{range}(S) = \ker(S)^\perp = \mathrm{range}(R)^\perp, \tag{2.11}$$

where $S^\dagger$ is a pseudoinverse of S and α has to be determined. Substituting (2.11) into the second equation of (2.10), we obtain

$$B_\Gamma S^\dagger B_\Gamma^T\lambda = B_\Gamma S^\dagger g + B_\Gamma R\alpha. \tag{2.12}$$

We introduce the notation $F := B_\Gamma S^\dagger B_\Gamma^T, d := B_\Gamma S^\dagger g, G := B_\Gamma R, e := R^T g$ and $P := I - G(G^T G)^{-1} G^T$. Note that P is a projection operator with its range orthogonal to G. We apply this P to (2.12) to eliminate the term with α and rewrite the orthogonality condition in (2.11) to obtain the following:

$$\left.\begin{array}{rcl} PF\lambda & = & Pd \\ G^T\lambda & = & e. \end{array}\right\}. \tag{2.13}$$

We define the space

$$V := \{\mu \in \text{range}(B_\Gamma) : B_\Gamma^T \mu \in \text{range}(S)\} = \ker(G^T),$$

which we call the space of admissible increments, following Farhat, Chen and Mandel [20]. The one-level FETI method is a preconditioned conjugate gradient method applied to

$$PF\lambda = Pd, \quad \lambda \in \lambda_0 + V \tag{2.14}$$

where λ_0 is chosen such that $G^T\lambda_0 = e$. Here, we only consider the *Dirichlet preconditioner* $M_D^{-1} := B_{D,\Gamma} S B_{D,\Gamma}^T$. With this choice of preconditioner, the preconditioned operator of the one-level FETI method has the following condition number bound:

$$\mathcal{K} \leq C(1 + log(H/h))^2, \tag{2.15}$$

where $\mathcal{K}$ denotes the condition number of the preconditioned operator in the appropriate subspace and C is a constant which does not depend on H or h. For a proof of (2.15), see [40] or [43, Section 6.3]. Thus the convergence rate of the one-level FETI method depends only polylogarithmically on the number of elements across a subdomain.

2.6 FETI-DP methods

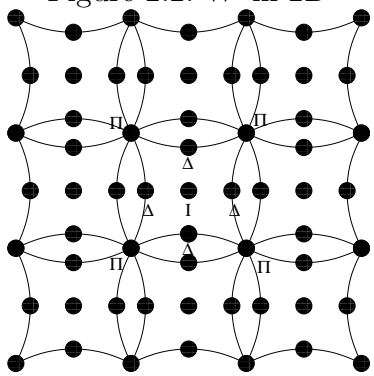
In this subsection, we closely follow the approach and notation of [38]. In the FETI-DP method, we use finite element functions in $\widetilde{W} = W_I \oplus \widetilde{W_\Gamma}$ to discretize (2.4), or equivalently, (2.3). As mentioned in Section 2.2.2, we can let $\widetilde{W_\Gamma}$ consist of functions which are continuous at subdomain vertices and obtain a scalable algorithm in the two-dimensional case, but enforcing such vertex constraints alone makes the resulting algorithm very sensitive to the number of degrees of freedom on each subdomain. In fact, the following upper bound for the condition number of the resulting preconditioned operator has been established:

$$\mathcal{K} \leq C\frac{H}{h}\left(1 + \frac{H}{h}\right)^2,$$

see [17], [37]. Numerical results also indicate that the linear factor H/h cannot be removed; see [22]. The remedy is to introduce the average values over certain edges and faces for scalar second-order elliptic problems, and even additional first order moments for linear elasticity problems with large jumps in PDE coefficients, as primal constraints in addition to or in replacement of the vertex constraints; see [36], [29], and [31].

We first provide a common framework for both two- and three-dimensional cases and then provide more details for the three-dimensional case, following [38] and [36, Section 4].

Figure 2.2: $\widetilde{W}$ in $2D$



We first note that the local stiffness matrices $A^{(i)}$ and the local load vectors $f^{(i)}$ can be written as follows:

$$A^{(i)} = \begin{bmatrix} A_{II}^{(i)} & A_{\Delta I}^{(i)^T} & A_{\Pi I}^{(i)^T} \\ A_{\Delta I}^{(i)} & A_{\Delta\Delta}^{(i)} & A_{\Pi I}^{(i)^T} \\ A_{\Pi I}^{(i)} & A_{\Pi\Delta}^{(i)} & A_{\Pi\Pi}^{(i)} \end{bmatrix}, \quad f^{(i)} = \begin{bmatrix} f_I^{(i)} \\ f_\Delta^{(i)} \\ f_\Pi^{(i)} \end{bmatrix}, \tag{2.16}$$

where I, Δ and Π indicate the index sets corresponding to the interior nodes, dual nodes, i.e., corresponding to $W_\Delta^{(i)}$, and primal nodes, i.e., corresponding to $W_\Pi^{(i)}$, respectively. We introduce the matrix $\widetilde{A}$, which can be thought of as the restriction of A, defined for the functions in W, to the subspace $\widetilde{W}$:

$$\widetilde{A} = \begin{bmatrix} A_{II}^{(1)} & A_{\Delta I}^{(1)^T} & & & & \widetilde{A}_{\Pi I}^{(1)^T} \\ A_{\Delta I}^{(1)} & A_{\Delta\Delta}^{(1)} & & & & \widetilde{A}_{\Pi\Delta}^{(1)^T} \\ & & \ddots & & & \vdots \\ & & & A_{II}^{(N)} & A_{\Delta I}^{(N)^T} & \widetilde{A}_{\Pi I}^{(N)^T} \\ & & & A_{\Delta I}^{(N)} & A_{\Delta\Delta}^{(i)} & \widetilde{A}_{\Pi\Delta}^{(N)^T} \\ \widetilde{A}_{\Pi I}^{(1)} & \widetilde{A}_{\Pi\Delta}^{(1)} & \cdots & \widetilde{A}_{\Pi I}^{(N)} & \widetilde{A}_{\Pi\Delta}^{(N)} & \widetilde{A}_{\Pi\Pi} \end{bmatrix}. \tag{2.17}$$

Here,

$$\widetilde{A}_{\Pi I}^{(i)} = R_{\Pi}^{(i)^T} A_{\Pi I}^{(i)}, \quad \widetilde{A}_{\Pi\Delta}^{(i)} = R_{\Pi}^{(i)^T} A_{\Pi\Delta}^{(i)}, \quad i = 1, \cdots, N,$$

and

$$\widetilde{A}_{\Pi\Pi} = \sum_{i=1}^{N} R_{\Pi}^{(i)^T} A_{\Pi\Pi}^{(i)} R_{\Pi}^{(i)}.$$

As in the one-level FETI method, we introduce a vector of Lagrange multipliers and obtain the following saddle point problem:

Find $(u, \lambda) \in \widetilde{W} \times \text{range}(\widetilde{B})$, such that

$$\left.\begin{array}{rcrcl} \widetilde{A}u & + & \widetilde{B}^T\lambda & = & f \\ \widetilde{B}u & & & = & 0 \end{array}\right\}. \tag{2.18}$$

Eliminating the interior unknowns of each subdomain from the system (2.18), we obtain:

Find $(u, \lambda) \in \widetilde{W_\Gamma} \times \text{range}(\widetilde{B}_\Gamma)$, such that

$$\left.\begin{array}{rcrcl} \widetilde{S}_\Gamma u_\Gamma & + & \widetilde{B}_\Gamma^T\lambda & = & g \\ \widetilde{B}_\Gamma u & & & = & 0 \end{array}\right\}. \tag{2.19}$$

We note that $\widetilde{S}_\Gamma$ for $\widetilde{W_\Gamma}$ can be defined as follows:

$$\widetilde{A}\begin{bmatrix} u_I^{(1)} \\ u_\Delta^{(1)} \\ \vdots \\ u_I^{(N)} \\ u_\Delta^{(N)} \\ u_\Pi \end{bmatrix} = \begin{bmatrix} 0 \\ (\widetilde{S}_\Gamma u_\Gamma)_\Delta^{(1)} \\ \vdots \\ 0 \\ (\widetilde{S}_\Gamma u_\Gamma)_\Delta^{(1)} \\ (\widetilde{S}_\Gamma u_\Gamma)_\Pi \end{bmatrix}, \tag{2.20}$$

where $u_\Gamma^T = [u_\Delta^{(1)^T} \cdots u_\Delta^{(N)^T} u_\Pi^T]$. $\widetilde{S}_\Gamma$ can also be regarded as the restriction of S, defined on W_Γ, to the subspace $\widetilde{W_\Gamma}$:

$$\widetilde{S}_\Gamma = \widetilde{R}_\Gamma^T S \widetilde{R}_\Gamma.$$

The matrix $\widetilde{A}$, and therefore also $\widetilde{S}_\Gamma$, are nonsingular, so we can solve the first equation of (2.19) for u_Γ without any difficulty and substitute the resulting equation into the second equation of (2.19):

$$\widetilde{B}_\Gamma \widetilde{S}_\Gamma^{-1} \widetilde{B}_\Gamma^T \lambda = -\widetilde{B}_\Gamma \widetilde{S}_\Gamma^{-1} g. \tag{2.21}$$

The Dirichlet preconditioner used in the FETI-DP algorithms to solve the equation (2.21) is of the form $\widetilde{B}_{D,\Gamma}\widetilde{S}_\Gamma\widetilde{B}_{D,\Gamma}^T$.

We now return to the issue of the choice of primal constraints in the three-dimensional case. We note that for scalar elliptic problems, enforcing the continuity of edge averages and vertex values leads to a scalable algorithm with its condition number bounded polylogarithmically in terms of the number of degrees of freedom on each subdomain, regardless of the coefficient jumps; see [37]. For linear elasticity problems, we are guaranteed to have such a condition number bound only when the material coefficient jumps are small, and we need a more elaborate choice of primal constraints to obtain robust condition number estimates independent of arbitrarily large jumps of the coefficients; see [36], [31].

There are two ways of enforcing additional continuity constraints besides vertex constraints, one using an extra set of Lagrange multipliers, and the other, with a change of basis. The second approach, in which appropriate average values or first order moments become explicit degrees of freedom, in general leads to a smaller and computationally more efficient coarse problem. Also, with the second approach, we can use the same implementation of the algorithm as described above; for details, see [36, Section 4]

We here illustrate how the change of basis is done when the average values over all the edges of a subdomain Ω_i are required to have common values across the interface. We closely follow the idea and notation of [38, Section 3.3] and [37, Section 4.2.1]. It is sufficient to consider the transformation for a single edge, say $\mathcal{E}$. Let u_E and $\hat{u}_E$ indicate the nodal displacement unknowns for the edge $\mathcal{E}$ in the original basis and the new basis, respectively. Denoting the change-of-basis matrix from the new basis to the original basis by T_E, we have

$$u_E = T_E\hat{u}_E.$$

T_E can be designed in many different ways, depending on which entry of $\hat{u}_E$ represents the average value of u_E: with $u_E^T = \begin{bmatrix} u_1 & \cdots & u_m & \cdots & u_l \end{bmatrix}$ and $\hat{u}_E^T = \begin{bmatrix} \hat{u}_1 & \cdots & \hat{u}_m & \cdots & \hat{u}_l \end{bmatrix}$, where node m can be any node on $\mathcal{E}$,

$$\begin{aligned} u_E &= T_E\hat{u}_E \\ \begin{bmatrix} u_1 \\ \vdots \\ u_m \\ \vdots \\ u_l \end{bmatrix} &= \begin{bmatrix} 1 & & & 1 & & & \\ & \ddots & & \vdots & & & \\ -1 & \cdots & & 1 & \cdots & -1 & \\ & & & \vdots & \ddots & & \\ & & & 1 & & 1 & \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \vdots \\ \hat{u}_m \\ \vdots \\ \hat{u}_l \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \hat{u}_m + \begin{bmatrix} \hat{u}_1 \\ \vdots \\ -\hat{u}_1 - \cdots - \hat{u}_{m-1} - \hat{u}_{m+1} - \cdots - \hat{u}_l \\ \vdots \\ \hat{u}_l^{(i)} \end{bmatrix}, \tag{2.22}$$

with $\hat{u}_m = 1/l \sum_{k=1}^{l} u_k$. Note that the edge unknowns have been separated into two parts, the first of which is a constant vector and the second of which has zero average. $\hat{u}_m$ will be added to the set of primal variables. The rest of the edge unknowns, i.e., $\begin{bmatrix} u_1 & \cdots & u_{m-1} & u_{m+1} & \cdots & u_l \end{bmatrix}$, will be the dual variables, along with the face nodal unknowns. Such a change of basis can be performed edge by edge, and we define $T_E^{(i)}$ as the change-of-basis matrix for all edges, i.e., a block-diagonal matrix consisting of such T_E described above for individual edges. Reordering the unknowns such that $u^{(i)^T} = \begin{bmatrix} u_I^{(i)^T} & u_{\overline{\Gamma}}^{(i)^T} & u_E^{(i)^T} \end{bmatrix}$, where $u_I^{(i)}$ are the interior unknowns, $u_{\overline{\Gamma}}^{(i)}$ are the interface unknowns excluding the unknowns corresponding to the edges, and $u_E^{(i)}$ are the unknowns corresponding to the collection of all edges, we now introduce the transformation for the entire subdomain Ω_i:

$$T^{(i)} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & T_E^{(i)} \end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} u_I^{(i)} \\ u_{\overline{\Gamma}}^{(i)} \\ u_E^{(i)} \end{bmatrix} = T^{(i)} \begin{bmatrix} u_I^{(i)} \\ u_{\overline{\Gamma}}^{(i)} \\ \hat{u}_E^{(i)} \end{bmatrix},$$

or, with $\hat{u}^{(i)^T} = \begin{bmatrix} u_I^{(i)^T} & u_{\overline{\Gamma}}^{(i)^T} & \hat{u}_E^{(i)^T} \end{bmatrix}$,

$$u^{(i)} = T^{(i)} \hat{u}^{(i)}.$$

In the following discussion, we will always assume that such a change of basis has already been performed in the three-dimensional case. The local stiffness matrices need to be modified accordingly, and we use $\overline{A}^{(i)} := T^{(i)^T} A^{(i)} T^{(i)}$ in lieu of $A^{(i)}$. We again rearrange $\hat{u}^{(i)}$ and $\overline{A}^{(i)}$ such that the interior unknowns come first, the primal unknowns last, and the dual unknowns in the middle; see (2.16). Note that the set of primal unknowns include not only the vertex nodal values but also the average values over the edges. We also need to transform the local load vector $f^{(i)}$, and reorder the resulting vector accordingly. We will use the original

notation $A^{(i)}, u^{(i)}$, and $f^{(i)}$ for the transformed and rearranged stiffness matrices, displacement vectors, and load vectors to keep the notation uncluttered. Now that we have obtained (2.16) for the three-dimensional case, we can derive the corresponding matrices, vectors, and operators for the FETI-DP algorithm for the three-dimensional case as well.

With $\widetilde{B}_{D,\Gamma}\widetilde{S}_\Gamma\widetilde{B}_{D,\Gamma}^T\widetilde{B}_\Gamma\widetilde{S}_\Gamma^{-1}\widetilde{B}_\Gamma^T$ as the preconditioner, we also have the following condition number bound,

$$\mathcal{K} \leq C(1 + log(H/h))^2,$$

for the preconditioned operator of the FETI-DP method. For a proof of this convergence bound for the two-dimensional case, see [41]. For three-dimensional scalar elliptic problems and linear elasticity problems, see e.g., [37] and [36], respectively.

2.7 BDDC methods

In this subsection, we review the BDDC method, following [45].
The discretized problem on the entire domain Ω is:

Find $(u_I, u_\Gamma) \in (W_I, \widehat{W}_\Gamma)$, such that

$$\begin{pmatrix} A_{II} & A_{\Gamma I}^T \\ A_{\Gamma I} & A_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} f_I \\ f_\Gamma \end{pmatrix}. \tag{2.23}$$

The equation (2.23) can be rewritten as

$$\begin{bmatrix} A_{II}^{(1)} & & & A_{\Gamma I}^{(1)^T}\widehat{R}_\Gamma^{(1)} \\ & \ddots & & \vdots \\ & & A_{II}^{(N)} & A_{\Gamma I}^{(N)^T}\widehat{R}_\Gamma^{(N)} \\ \widehat{R}_\Gamma^{(1)^T}A_{\Gamma I}^{(1)} & \cdots & \widehat{R}_\Gamma^{(N)^T}A_{\Gamma I}^{(N)} & \sum_{i=1}^N \widehat{R}_\Gamma^{(i)^T}A_{\Gamma\Gamma}^{(i)}\widehat{R}_\Gamma^{(i)} \end{bmatrix} \begin{bmatrix} u_I^{(1)} \\ \vdots \\ u_I^{(N)} \\ u_\Gamma \end{bmatrix} \begin{bmatrix} f_I^{(1)} \\ \vdots \\ f_I^{(N)} \\ \sum_{i=1}^N R_\Gamma^{(i)^T} f_\Gamma^{(i)} \end{bmatrix}. \tag{2.24}$$

Eliminating the interior unknowns of each subdomain, i.e., eliminating the upper left block of (2.24), we obtain

$$\widehat{S}_\Gamma u_\Gamma = g, \tag{2.25}$$

where

$$\begin{aligned} \widehat{S}_\Gamma &= \sum_{i=1}^N \widehat{R}_\Gamma^{(i)^T}(A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)}A_{II}^{(i)^{-1}}A_{\Gamma I}^{(i)^T})\widehat{R}_\Gamma^{(i)} \\ &= \sum_{i=1}^N \widehat{R}_\Gamma^{(i)^T} S^{(i)} \widehat{R}_\Gamma^{(i)} \end{aligned}$$

$$= \widehat{R}_\Gamma^T S \widehat{R}_\Gamma, \tag{2.26}$$

and

$$g = \sum_{i=1}^{N} \widehat{R}_\Gamma^{(i)^T} (f_\Gamma^{(i)} - A_{\Gamma I}^{(i)} A_{II}^{(i)^{-1}} f_I^{(i)}).$$

From (2.26), we can see that $\widehat{S}_\Gamma$ can be regarded as the restriction of S, defined on $W_\Gamma = \prod_{i=1}^{N} W_\Gamma^{(i)}$, to the continuous subspace $\widehat{W}_\Gamma$. We can also view $\widehat{S}_\Gamma$ as the restriction of $\widetilde{S}_\Gamma$ to $\widehat{W}_\Gamma$:

$$\widehat{S}_\Gamma = \bar{R}_\Gamma^T \widetilde{S}_\Gamma \bar{R}_\Gamma.$$

In the BDDC method, we use $M^{-1} = \bar{R}_{D,\Gamma}^T \widetilde{S}_\Gamma^{-1} \bar{R}_{D,\Gamma}$ as the preconditioner, and the preconditioned operator is

$$\bar{R}_{D,\Gamma}^T \widetilde{S}_\Gamma^{-1} \bar{R}_{D,\Gamma} \bar{R}_\Gamma^T \widetilde{S}_\Gamma \bar{R}_\Gamma. \tag{2.27}$$

Recall that the preconditioned operator for the FETI-DP method is

$$\widetilde{B}_{D,\Gamma} \widetilde{S}_\Gamma \widetilde{B}_{D,\Gamma}^T \widetilde{B}_\Gamma \widetilde{S}_\Gamma^{-1} \widetilde{B}_\Gamma^T. \tag{2.28}$$

Using Lemma 2.3.1, we can prove that (2.27) and (2.28) have essentially the same spectrum, and therefore the same condition number estimate, $\mathcal{K} \leq C(1 + log(H/h))^2$; see [38].

Part II

Auxiliary Linear Algorithms for Nonlinear Contact Problems

Chapter 3

FETI-FETI method

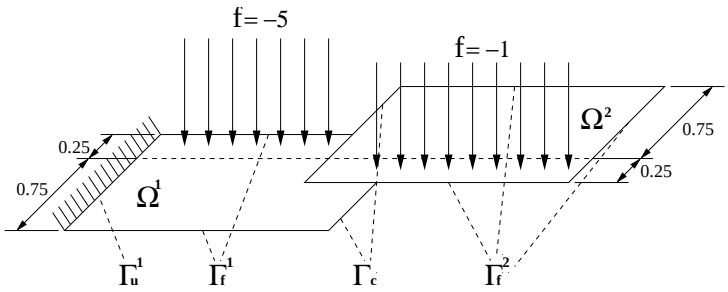
3.1 Introduction

Our ultimate goal is to solve contact problems with N bodies, $\Omega_1, \cdots, \Omega_N$. Contact problems are characterized by an active area of contact, which is unknown a priori, and inequality constraints such as non-penentration conditions; see [1]. We recall that the subdomain interface continuity constraints are of the form $Bu = 0$ in the FETI methods, which are due to the use of a domain decomposition algorithm and the fact that finite element functions that are used are not continuous across the interface. The introduction of the subdomains and the ensuing need for continuity constraints such as $Bu = 0$ are artificial in a sense. In contact problems, however, inequality constraints arise from the fact that we have multiple bodies and are inherent to the nature of the problem.

In this and the next chapters, we assume that we use an active set method to deal with the inequality constraints; for other ways of dealing with inequality constraints, see [1]. An active set method gives rise to a sequence of auxiliary equality constrained problems, in which some of the inequality constraints are replaced by corresponding equality constraints and the rest are ignored. An active set method has outer iterations in which the active set is updated, and a minimization problem on the current active face is solved in each inner iteration. The FETI-FETI method of this chapter and the hybrid method of the next chapter deal with the inner minimization problem.

3.2 The Model Problem

In this dissertation, we concentrate on scalar elliptic problems in two- and three-dimensions with inequality constraints. We present the following model problem as a motivation. It will actually be solved in Chapters 5 and 6:



$$
\begin{aligned}
\min \quad & \sum_{i=1}^{2} \left(\frac{1}{2} \int_{\Omega^i} |\nabla u^i|^2 dx - \int_{\Omega^i} f u^i dx \right) \\
\text{where} \quad & u^i \in H^1(\Omega^i), i = 1, 2, \quad \Omega^1 = (0,1) \times (0,1), \Omega^2 = (1,2) \times (0,1) \\
& u^1 = 0 \quad \text{on} \quad \Gamma_u^1 = \{0\} \times (0,1) \\
& u^2 - u^1 \geq 0 \quad \text{on} \quad \Gamma_c = \{1\} \times (0,1)
\end{aligned} \tag{3.1}
$$

The reason we consider only scalar elliptic problems is that the inequality constraints in scalar elliptic problems are much simpler than those in linear elasticity problems and their simplicity allows us to focus on the analysis of the preconditioned operator. In linear elasticity problems, the non-penentration conditions depend on the current configuration of the bodies and need to be updated in each iteration (see [1, Section 4]), whereas in this scalar problem the inequality condition is expresssed by a single equation such as $Bu \leq 0$.

We consider multiple bodies $\Omega_i, i = 1, \cdots, N$, each of which has many degrees of freedom and is decomposed into subdomains $\Omega_{i,j}, i = 1, \cdots, N, j = 1, \cdots, N_i$. The diameter of the body Ω_i is H_i^b, with $H_b = \max_i H_i^b$. The diameter of the subdomain $\Omega_{i,j}$ is $H_{i,j}^s$, with $H_s = \max_{i,j} H_{i,j}^s$. When there is no danger of confusion, we will use the notation H instead of H_s as in Chapter 2.

We assume that at least one body is clamped on part of its boundary, and denote the union of such fixed boundaries for the entire system by $\partial\Omega_D$. We assume that $\rho(x) = \rho_{i,j} \geq \rho_{min} > 0, \forall x \in \Omega_{i,j}, \forall i, j$. We also assume that the coefficient varies only moderately within the same body; in particular, we assume

$$
\rho_i := \max_j \rho_{i,j} \leq C \rho_{i,j}, \quad \forall i, j, \tag{3.2}
$$

where $C \geq 1$ is a constant independent of i.

We introduce two types of global interfaces: the first one is $\overline{\Gamma_{gl}} := \bigcup_{i \neq j} \partial\Omega_i \cap \partial\Omega_j$, and can be viewed as the potential contact area between the bodies: in the model problem, this is Γ_c. The second one, the "current" contact area, is denoted

by Γ_{gl}^k, where $\Gamma_{gl}^k \subset \overline{\Gamma_{gl}}$; the superscript k indicates the outer iteration of the active set method and reminds us that the "current" active set changes. In each outer iteration of the active set method some of the inequality constraints are adopted as the corresponding equality constraints and the rest are ignored, and $\Gamma_{gl,h}^k$ can be viewed as the collection of the nodes at which equality constraints are being imposed. We also introduce the local interfaces $\Gamma_{loc}^{(i)} := \bigcup_{j \neq k}(\partial\Omega_{i,j} \cap \partial\Omega_{i,k}), i = 1, \cdots, N$. We assume there are no traction forces and denote the union of the free boundaries by $\partial\Omega_F := (\cup_i \partial\Omega_i) \setminus (\partial\Omega_D \cup \overline{\Gamma_{gl}})$.

We denote a standard finite element space of continuous, piecewise linear functions on $\Omega_{i,j}$ by $W^{(i,j)}$. Each $W^{(i,j)}$ is decomposed into a subdomain interior part $W_I^{(i,j)}$ and a subdomain interface part $W_\Gamma^{(i,j)}$ for functions on $\partial\Omega_{i,j} \cap \overline{\Gamma_{gl}}$. $W_\Gamma^{(i,j)}$ is further decomposed into a primal subspace $W_\Pi^{(i,j)}$ and a dual subspace $W_\Delta^{(i,j)}$ in the style of the FETI-DP method. In the two-dimensional case, on which we will concentrate, we can impose the primal continuity at vertex nodes. In the three-dimensional case, we require the average values of all edges to be continuous in addition to the continuity at all vertex nodes. This choice of primal constraints leads to a scalable algorithm with the condition number estimate bounded above by $C(1 + log(H/h))^2$ in the scalar elliptic case [37]. We also note that all functions vanish on the Dirichlet boundary $\partial\Omega_D$. We define associated product spaces, $W_I^{(i)} := \prod_{j=1}^{N_i} W_I^{(i,j)}, W_\Gamma^{(i)} := \prod_{j=1}^{N_i} W_\Gamma^{(i,j)}, W_\Delta^{(i)} := \prod_{j=1}^{N_i} W_\Delta^{(i,j)}, W_\Pi^{(i)} := \prod_{j=1}^{N_i} W_\Pi^{(i,j)}$, and $\widehat{W}_\Pi^{(i)}$, which is the continuous subspace of $W_\Pi^{(i)}$.

We introduce spaces analogous to the $\widetilde{W}_\Gamma$ and $\widehat{W}_\Gamma$ of Section 2.2.2. Functions in $W_\Gamma^{(i)}$ are in general discontinuous across the local interface $\Gamma_{loc}^{(i)}$, and we define $\widehat{W}_\Gamma^{(i)}$ as the continuous subspace of $W_\Gamma^{(i)}$. $\widetilde{W}_\Gamma^{(i)} := W_\Delta^{(i)} \oplus \widehat{W}_\Pi^{(i)}$ is intermediate between $\widehat{W}_\Gamma^{(i)}$ and $W_\Gamma^{(i)}$.

3.3 Technical Tools

In the theory of domain decomposition methods, it had been previously assumed that each subdomain is a union of a small number of coarse triangles or tetrahedra. However, this assumption is often unrealistic, especially when the subdomains result from using a mesh partitioner. In such a case, subdomain boundaries may not even be uniformly Lipschitz.

However, there have been recent developments on the theory of domain decomposition methods with very weak assumptions about the regularity of the subdomains. In this section, we introduce some of the new results obtained by Dohrmann, Klawonn, Rheinbach, and Widlund. We also provide some technical lemmas that will be needed for the convergence analysis of the FETI-FETI method.

We first give the definition of a John domain; see Hajłasz [26] and the references

therein.

Definition 1 (John Domains)**.** *A domain $\Omega \subset \mathbb{R}^n$ - an open, bounded, and connected set - is a John domain if there exist a constant C_J and a distinguished central point $x_0 \in \Omega$ such that each $x \in \Omega$ can be joined to it by a rectifiable curve $\gamma : [0,1] \to \Omega$ with $\gamma(0) = x_0, \gamma(1) = x$ and $|x - \gamma(t)| \leq C_J \cdot distance(\gamma(t), \partial\Omega)$ for all $t \in [0,1]$.*

We note that the length of the boundary of a John domain can be arbitrarily much longer than its diameter, and also that a John domain can have cusps facing inwards, but not outwards. This means that if we have a union of several nonoverlapping subdomains, having any kind of cusp on the interface would lead to the existence of a subdomain which is not a John domain.

We define uniform domains, which are also known as Jones domains. It is known, and easy to see, that any uniform domain is a John domain. According to Jones [27, Theorem 4], uniform domains form the largest class of domains for which an extension theorem holds in two dimensions; see Lemma 3.3.1 below. We also note that the complement of a uniform domain is also a uniform domain; see [27, Theorem C].

Definition 2 (Uniform Domains)**.** *A domain $\Omega \subset \mathbb{R}^n$ is a uniform domain if there exists a constant C_U such that any pair of points $x_1 \in \Omega$ and $x_2 \in \Omega$ can be joined by a rectifiable curve $\gamma(t) : [0,1] \to \Omega$ with $\gamma(0) = x_1, \gamma(1) = x_2$ and the Euclidean arc length of $\gamma \leq C_U|x_1 - x_2|$ and $min_{i=1,2}|x_i - \gamma(t)| \leq C_U \cdot distance(\gamma(t), \partial\Omega)$ for all $t \in [0,1]$.*

3.3.1 Technical Tools - part I

In this subsection, we collect technical tools that are mostly from [43, Chapter 4], [12] and [33]. The first lemma that we introduce is [27, Theorem 1].

Lemma 3.3.1. *Let $\Omega \subset \mathbb{R}^n$ be a uniform domain. Then there exists a bounded linear operator $E_\Omega : H^1(\Omega) \to H^1(\mathbb{R}^n)$, which extends any element in $H^1(\Omega)$ to one defined for all of $\mathbb{R}^n$, i.e., $(E_\Omega u)|_\Omega = u, \quad \forall u \in H^1(\Omega)$. The norm of this operator depends only on $C_U(\Omega)$ and the dimension n.*

The following lemma is [33, Lemma 4.5], and we note that Lemma 3.3.1 is needed its the proof.

Lemma 3.3.2 (Extension Lemma)**.** *Let Ω_i and Ω_j be subsets of $\mathbb{R}^n$ and two subdomains with a common $(n-1)-$dimensional interface $\mathcal{E}^{ij}$. Furthermore, let Ω_i be a uniform domain, let $V_i^h = \{v_h \in W^h(\Omega_i) : v_h(x) = 0 \quad$ at all nodes of $\partial\Omega_i \setminus \mathcal{E}^{ij}\}$, and let $V_j^h = \{v_h \in W^h(\Omega_j) : v_h(x) = 0 \quad$ at all nodes of $\partial\Omega_j \setminus \mathcal{E}^{ij}\}$, where $W^h(\Omega_i)$*

is the standard finite element space of continuous, piecewise linear functions on Ω_i. Then, there exists an extension operator

$$E_{ji}^h : V_j^h \to V_i^h, \tag{3.3}$$

with the following properties:

1. $(E_{ji}^h u_h)|_{\Omega_j} = u_h, \quad \forall u_h \in V_j^h$
2. $||E_{ji}^h u_h||_{H^1(\Omega_i)} \leq C||u_h||_{H^1(\Omega_j)}, \quad \forall u_h \in V_j^h,$

where the constant $C \geq 0$ depends only on the uniformity parameter $C_U(\mathcal{C}\Omega_i)$ of the complement of Ω_i and the shape regularity of the finite elements and is otherwise independent of the finite element mesh sizes h_i and h_j and the diameters H_i and H_j.

We note that the inequalities of the following lemma are well known in the theory of iterative substructuring methods. Proofs for domains satisfying an interior cone condition are given in [3] and [8, Section 4.9] and a different proof is given in [43, Lemma 4.15]. For a proof that this inequality is sharp, see [5]. For a proof of the following lemma, which is for John domains, see [12].

Lemma 3.3.3 (Discrete Sobolev Inequality)**.** *Let $\Omega \subset \mathbb{R}^2$ be a John domain with diameter H. Then,*

$$\begin{aligned} ||u - \bar{u}_\Omega||^2_{L^\infty(\Omega)} &\leq C(1 + log(H/h))|u|^2_{H^1(\Omega)} \quad \textit{and} \\ ||u||^2_{L^\infty(\Omega)} &\leq C(1 + log(H/h))||u||^2_{H^1(\Omega)}, \quad \forall u \in W^h(\Omega). \end{aligned}$$

The constant C depends only on the John parameter $C_J(\Omega)$ of Ω and the shape regularity of elements.

Corollary 3.3.4. *Let $\Omega \subset \mathbb{R}^2$ be a John domain with diameter H. Then,*

$$||u - u(x_0)||^2_{L^2(\Omega)} \leq CH^2(1 + log(H/h))|u|^2_{H^1(\Omega)}, \quad \forall x_0 \in \Omega. \tag{3.4}$$

Proof. Note that

$$||u - u(x_0)||^2_{L^2(\Omega)} \leq 2||u - \bar{u}_\Omega||^2_{L^2(\Omega)} + 2||\bar{u}_\Omega - u(x_0)||^2_{L^2(\Omega)} \tag{3.5}$$

The first term on the right hand side of (3.5) can be estimated by an elementary Poincaré inequality, which holds for John domains:

$$||u - \bar{u}_\Omega||^2_{L^2(\Omega)} \leq CH^2|u|^2_{H^1(\Omega)}, \tag{3.6}$$

where C is a constant independent of the size of Ω. As for the second term on the

right hand side,

$$\begin{aligned} & ||\bar{u}_\Omega - u(x_0)||^2_{L^2(\Omega)} \\ \leq \; & |\Omega| ||u - \bar{u}_\Omega||^2_{L^\infty(\Omega)} \\ \leq \; & C|\Omega|(1 + log(H/h))|u|^2_{H^1(\Omega)}, \end{aligned} \tag{3.7}$$

where the second inequality follows from the discrete Sobolev inequality in Lemma 3.3.3. Combining (3.5),(3.6) and (3.7), we obtain (3.4). □

We need another tool to estimate energies coming from edge contributions in the two-dimensional case. For a proof of the following lemma for John domains, see [33, Lemma 4.4]. See [18] for the first proof of the same lemma for regular subdomains in two dimensions.

Lemma 3.3.5 (Edge Lemma)**.** *Let $\Omega_i \subset \mathbb{R}^2$ be a John domain, $\mathcal{E}^{ij} \subset \partial\Omega_i$ be an edge, and $\theta_{\mathcal{E}^{ij}} \in W^h(\Omega_i)$ be a finite element function which equals 1 at all nodes of $\mathcal{E}^{ij}$, vanishes at the other nodes on $\partial\Omega_i$, and is discrete harmonic in Ω_i. Then,*

$$|\mathcal{H}(\theta_{\mathcal{E}^{ij}} u)|^2_{H^1(\Omega_i)} \leq C(1 + log(H/h))^2 ||u||^2_{H^1(\Omega_i)}, \quad \forall u \in W^h(\Omega_i), \tag{3.8}$$

$$|\theta_{\mathcal{E}^{ij}}|^2_{H^1(\Omega_i)} \leq C(1 + log(H/h)) \tag{3.9}$$

and

$$||\theta_{\mathcal{E}^{ij}}||^2_{L^2(\Omega_i)} \leq CH^2(1 + log(H/h)). \tag{3.10}$$

Here, C depends only on the John parameter $C_J(\Omega_i)$ of Ω_i and the shape regularity of the finite elements. The logarithmic factor of (3.10) can be removed for P_1 elements if all angles of the triangulation are acute.

We also note that the bound of Lemma 3.3.5 is independent of the length of the edge $\mathcal{E}^{ij}$.

We would need similar face and edge lemmas to advance the theory for the three-dimensional case. Theory for irregular subdomains in three dimensions is not complete at this point. However, such bounds have been established for regular subdomains such as tetrahedra or cubes. Therefore we will assume that we have regular subdomains in the three-dimensional case in the rest of this chapter. The following lemma is taken from [43, Section 4.6.3] and [10], which deal with tetrahedral subdomains and cubic subdomains, respectively.

Lemma 3.3.6 (Face Lemma)**.** *Let $\Omega_i \subset \mathbb{R}^3$ be a tetrahedron or a cube, $\mathcal{F}^j \subset \partial\Omega_i$ be a face, and $\theta_{\mathcal{F}^j} \in W^h(\Omega_i)$ be a finite element function which equals 1 at all nodes of $\mathcal{F}^j$, vanishes at the other nodes on $\partial\Omega_i$, and is discrete harmonic in Ω_i. We then have*

$$|\mathcal{H}(\theta_{\mathcal{F}^j} u)|^2_{H^1(\Omega_i)} \leq C(1 + log(H/h))^2 ||u||^2_{H^1(\Omega_i)}, \quad \forall u \in W^h(\Omega_i), \tag{3.11}$$

and

$$|\theta_{\mathcal{F}^j}|^2_{H^1(\Omega_i)} \leq C(1 + log(H/h))H. \tag{3.12}$$

C is independent of H and h.

For the proof of the previous lemma, a function $\vartheta_{\mathcal{F}_j}$ is constructed, which coincides with $\theta_{\mathcal{F}_j}$ on $\partial\Omega_i$ and satisfies $|\nabla\vartheta_{\mathcal{F}_j}(x)| \leq C/r(x)$, where $r(x)$ denotes the distance between x and the edge of Ω_i closest to x. The lemma then follows from the fact that the harmonic extension is of minimal energy.

Suppose we have a square subset, $\mathcal{F}_0^j \subsetneq \mathcal{F}^j$. We will present an argument that we can replace $\mathcal{F}^j$ with $\mathcal{F}_0^j$ in Lemma 3.3.6 and obtain a bound which is uniformly bounded regardless of the ratio $|\mathcal{F}_0^j|/|\mathcal{F}^j|$, by constructing a function $\vartheta_{\mathcal{F}_0^j}$ with similar properties as $\vartheta_{\mathcal{F}^j}$. Here, we describe how such functions can be constructed for the case of a cubic subdomain, following [10]. We first consider functions which are not finite element functions, and then obtain their finite-element counterparts by linear interpolation. We will use the same notation for both functions.

We first sketch how $\vartheta_{\mathcal{F}^j}$ is constructed. We divide the cube into twenty-four tetrahedra by connecting its center C to all the vertices and the centers of all the faces, $C^k, k = 1, 2, \cdots, 6$, and drawing the diagonals on each face. The function $\vartheta_{\mathcal{F}^j}$ associated with the face $\mathcal{F}^j$ is defined to be $1/6$ at the center C, and we require $\vartheta_{\mathcal{F}^j}(C^k) = \delta_{jk}$, where δ_{jk} is the Kronecker symbol. $\vartheta_{\mathcal{F}^j}$ is linear on the segment CC^k. The values inside a tetrahedron defined by the segment CC^k and one edge of the face $\mathcal{F}^k$ are defined to be constant on the intersection of any plane throughout that edge and the tetrahedron, and the value is given by the value at the point of intersection between the plane and the segment CC^k.

We now consider $\theta_{\mathcal{F}_0^j}$. We can construct a cube, a proper subset of the cubic subdomain Ω_i, with $\mathcal{F}_0^j$ as its base; we denote it by T. In T, $\vartheta_{\mathcal{F}_0^j}$ is defined exactly the same way as $\vartheta_{\mathcal{F}^j}$ in Ω_i; we can complete the definition of $\vartheta_{\mathcal{F}_0^j}$ by extending its values by zero in $\Omega_i \setminus T$.

The following is [43, Lemma 4.19]. We note that the proof is quite elementary, and can be obtained using the energy-minimizing property of discrete harmonic extensions and inverse inequalities.

Lemma 3.3.7. *Let $\mathcal{E}$ be an edge of a subdomain $\Omega_i \subset \mathbb{R}^3$ and let $u \in W^h(\Omega_i)$. Then,*

$$\begin{aligned} |\mathcal{H}(\theta_{\mathcal{E}}u)|^2_{H^1(\Omega_i)} &\leq |I^h(\vartheta_{\mathcal{E}}u)|^2_{H^1(\Omega_i)} \leq C||I^h(\theta_{\mathcal{E}}u)||^2_{L^2(\mathcal{E})}, \\ |\mathcal{H}(\theta_{\mathcal{E}}u)|^2_{H^1(\Omega_i)} &\leq |I^h(\vartheta_{\mathcal{E}}u)|^2_{H^1(\Omega_i)} \leq C||u||^2_{L^2(\mathcal{E})}. \end{aligned}$$

The proof of the following lemma, which is [43, Lemma 4.16], is straightforward; we can take Lemma 3.3.3 given for two dimensions, and integrate along the third direction. It will be needed for the analysis of the three-dimensional case.

Lemma 3.3.8. *Let $\bar{u}_{\mathcal{E}}$ be the average value of u over $\mathcal{E}$, an edge of a regular domain $\Omega_i \subset \mathbb{R}^3$. Then,*

$$||u||^2_{L^2(\mathcal{E})} \leq C(1 + log(H/h))||u||^2_{H^1(\Omega_i)}$$

and

$$||u - \bar{u}_{\mathcal{E}}||^2_{L^2(\mathcal{E})} \leq C(1 + log(H/h))|u|^2_{H^1(\Omega_i)}.$$

We can obtain the following lemma by a use of the Cauchy-Schwarz inequality and Lemma 3.3.8; see [43, Lemma 4.21].

Lemma 3.3.9. *Let $\mathcal{E}$ be an edge of $\Omega_i \subset \mathbb{R}^3$. Then, for any $u \in W^h(\Omega_i)$,*

$$\begin{aligned} ||u - \bar{u}_{\mathcal{E}}||^2_{L^2(\partial\Omega_i)} &\leq CH(1 + log(H/h))|u|^2_{H^1(\Omega_i)}, \\ ||u - \bar{u}_{\mathcal{E}}||^2_{L^2(\Omega_i)} &\leq CH^2(1 + log(H/h))|u|^2_{H^1(\Omega_i)}. \end{aligned}$$

The following lemma is [4, Lemma 2.3], and can be proved by an inverse inequality and Sobolev embedding.

Lemma 3.3.10. *For all $u \in W^h(\Omega_i)$, where $\Omega_i \subset \mathbb{R}^3$,*

$$||u||^2_{L^\infty(\Omega_i)} \leq C(1/h)||u||^2_{H^1(\Omega_i)},$$

where C is independent of h and the diameter of Ω_i.

3.3.2 Technical Tools - part II

In this subsection, we present lemmas that are specific to the study of the FETI-FETI method.

We need a Poincaré-type inequality to treat the energy terms coming from the global interface between different bodies. We present such a lemma for the two-dimensional case. In the following lemma, which is an adaptation of [8, Lemma 10.6.6], we assume that we have Lipschitz subdomains.

Lemma 3.3.11. *Let $v \in \widetilde{W}^{(i)} = W_I^{(i)} \oplus \widetilde{W}_\Gamma^{(i)}$. Then*

$$||v||^2_{L^2(\Omega_i)} \leq C\left(H_b^2(1 + log(H_s/h))^2 \sum_j |v|^2_{H^1(\Omega_{i,j})} + \frac{1}{H_b^2}\left|\int_{\Omega_i} v dx\right|^2\right), \quad (3.13)$$

and

$$||v||^2_{L^2(\Omega_i)} \leq C\left(H_b^2(1 + log(H_s/h))^2 \sum_j |v|^2_{H^1(\Omega_{i,j})} + \left|\int_{\partial\Omega_i \cap \partial\Omega_D} v ds\right|^2\right). \quad (3.14)$$

Proof. We follow the idea of [8, Lemma 10.6.6]. Let c be piecewise constant in each subdomain of Ω_i, i.e., $c(x) = c_{i,j}, \forall x \in \Omega_{i,j}, j = 1, \cdots, N_i$. We define a function $Ec \in H^1(\Omega_i)$ as follows. On the local interface of Ω_i, Ec is defined as the average of c, i.e.,

$$Ec(x) = \frac{\sum_{j \in \mathcal{N}_x} c_{i,j}(x)}{|\mathcal{N}_x|}, \quad x \in \Gamma^{(i)}_{loc,h},$$

where $\mathcal{N}_x$ is the set of indices of the subdomains of Ω_i with x on their boundaries. Also, we set $Ec|_{\partial\Omega_i} = c|_{\partial\Omega_i}$

With $Ec|_{\partial\Omega_{i,j}}, j = 1, \cdots, N_i$ given as above, Ec is defined to be discrete harmonic in each subdomain of Ω_i. We have

$$\begin{aligned} & ||c||^2_{L^2(\Omega_i)} \\ \leq\ & 2||c - Ec||^2_{L^2(\Omega_i)} + 2||Ec||^2_{L^2(\Omega_i)} \\ \leq\ & 2\sum_j ||c - Ec||^2_{L^2(\Omega_{i,j})} + C\left(H_b^2 |Ec|^2_{H^1(\Omega_i)} + \frac{1}{H_b^2}\left|\int_{\Omega_i} Ec dx\right|^2\right), \end{aligned} \tag{3.15}$$

where the second inequality follows from a Poincaré inequality with scaling. We estimate the first term of (3.15):

$$\begin{aligned} & \sum_j ||c - Ec||^2_{L^2(\Omega_{i,j})} \\ \leq\ & C\sum_j \left(H_s^2 |c - Ec|^2_{H^1(\Omega_{i,j})} + H_s ||c - Ec||^2_{L^2(\partial\Omega_{i,j})}\right) \\ \leq\ & C\left(\sum_j H_s^2 |c - Ec|^2_{H^1(\Omega_{i,j})} + H_s \sum_{e \in \mathcal{E}^i(\Omega_i)} ||[[c]]_e||^2_{L^2(e)}\right), \end{aligned} \tag{3.16}$$

where the first inequality is a Friedrichs inequality with scaling, $\mathcal{E}^i(\Omega_i)$ denotes the set of interior edges of Ω_i and $[[c]]_e$ the jump of the function c across the edge e. Note that $c - Ec$ is constant on each edge of $\Omega_{i,j}, j = 1, \cdots, N_i$, and its values will often differ between different edges and vertices. Therefore we can write

$$(c - Ec)(x) = \sum_{\mathcal{E} \in \partial\Omega_{i,j}} (c - Ec)|_{\mathcal{E}} \theta_{\mathcal{E}}(x) + \sum_{\mathcal{V} \in \partial\Omega_{i,j}} (c - Ec)|_{\mathcal{V}} \theta_{\mathcal{V}}(x), \quad x \in \partial\Omega_{i,j,h}.$$

We note that the characteristic function $\theta_{\mathcal{E}}$ has already been defined in Lemma 3.3.5; $\theta_{\mathcal{V}}$ is defined analogously. It equals 1 at $\mathcal{V}$, vanishes at all other nodes on $\partial\Omega_{i,j}$ and is discrete harmonic in $\Omega_{i,j}$. Noting that $c - Ec$ is discrete harmonic in $\Omega_{i,j}$, we have

$$|c - Ec|^2_{H^1(\Omega_{i,j})}$$

$$\leq C\left(\sum_{\mathcal{E}}|(c-Ec)|_{\mathcal{E}}\theta_{\mathcal{E}}|^2_{H^1(\Omega_{i,j})}+\sum_{\mathcal{V}}|(c-Ec)|_{\mathcal{V}}\theta_{\mathcal{V}}|^2_{H^1(\Omega_{i,j})}\right). \quad (3.17)$$

We have

$$\begin{aligned} & |(c-Ec)|_{\mathcal{E}}\theta_{\mathcal{E}}|^2_{H^1(\Omega_{i,j})} \\ \leq\ & |[[c]]_{\mathcal{E}}||\theta_{\mathcal{E}}|^2_{H^1(\Omega_{i,j})} \\ \leq\ & \frac{1}{|\mathcal{E}|}||[[c]]_{\mathcal{E}}||^2_{L^2(\mathcal{E})}|\theta_{\mathcal{E}}|^2_{H^1(\Omega_{i,j})} \\ \leq\ & C(1+log(H_s/h))\frac{1}{|\mathcal{E}|}||[[c]]_{\mathcal{E}}||^2_{L^2(\mathcal{E})}, \end{aligned} \quad (3.18)$$

where the last inequality follows from Lemma 3.3.5. Using the fact that $|\theta_{\mathcal{V}}|^2_{H^1(\Omega_{i,j})} = O(1)$, we have

$$\begin{aligned} & |(c-Ec)|_{\mathcal{V}}\theta_{\mathcal{V}}|^2_{H^1(\Omega_{i,j})} \\ \leq\ & C|(c-Ec)|_{\mathcal{V}}|. \end{aligned} \quad (3.19)$$

We note that (3.19) can be absorbed into (3.18). Combining (3.17), (3.18), and (3.19), we obtain

$$\begin{aligned} |Ec|^2_{H^1(\Omega_{i,j})} &= |c-Ec|^2_{H^1(\Omega_{i,j})} \\ &\leq C(1+log(H_s/h))\frac{1}{|\mathcal{E}|}||[[c]]_{\mathcal{E}}||^2_{L^2(\mathcal{E})}. \end{aligned} \quad (3.20)$$

Combining (3.15), (3.16), and (3.20), we have

$$\begin{aligned} & ||c||^2_{L^2(\Omega_i)} \\ \leq\ & CH_b^2(1+log(H_s/h))\sum_{e\in\mathcal{E}^i(\Omega_i)}\frac{1}{|e|}||[[c]]_e||^2_{l^2(e)}+C\frac{1}{H_b^2}\left|\int_{\Omega_i}Ecdx\right|^2 \\ \leq\ & CH_b^2H_s^{-1}(1+log(H_s/h))\sum_{e\in\mathcal{E}^i(\Omega_i)}||[[c]]_e||^2_{L^2(e)}+C\frac{1}{H_b^2}\left|\int_{\Omega_i}Ecdx\right|^2. \end{aligned} \quad (3.21)$$

The rest of the proof of (3.13) is similar to the proof of [8, Lemma 10.6.7]; let v satisfy the assumption of the lemma and $\bar{v}$ be defined by

$$\bar{v}(x)=\frac{1}{|\Omega_{i,j}|}\int_{\Omega_{i,j}}vdx, \quad \forall x\in\Omega_{i,j}, \quad j=1,\cdots,N_i.$$

Then,

$$
\begin{aligned}
& ||v||^2_{L^2(\Omega_i)} \\
\leq\ & 2||v-\bar{v}||^2_{L^2(\Omega_i)} + 2||\bar{v}||^2_{L^2(\Omega_i)} \\
\leq\ & 2\sum_j ||v-\bar{v}||^2_{L^2(\Omega_{i,j})} + C\Big(H_b^2 H_s^{-1}(1+log(H_s/h)) \sum_{e\in\mathcal{E}^i(\Omega_i)} |||[[\bar{v}]]_e||^2_{L^2(e)} + \\
& \frac{1}{H_b^2}\left|\int_{\Omega_i} \bar{v}dx\right|^2\Big) \\
\leq\ & C\Big(\sum_j H_s^2 |v|^2_{H^1(\Omega_{i,j})} + H_b^2 H_s^{-1}(1+log(H_s/h)) \sum_{e\in\mathcal{E}^i(\Omega_i)} |||[[v]]_e||^2_{L^2(e)} + \\
& H_b^2 H_s^{-1}(1+log(H_s/h)) \sum_{e\in\mathcal{E}^i(\Omega_i)} |||[[v-\bar{v}]]_e||^2_{L^2(e)} + \frac{1}{H_b^2}\left|\int_{\Omega_i} v dx\right|^2\Big), \quad (3.22)
\end{aligned}
$$

where the second inequality follows from (3.21). We estimate the second to the last term:

$$
\begin{aligned}
& H_b^2 H_s^{-1}(1+log(H_s/h)) \sum_{e\in\mathcal{E}^i(\Omega_i)} |||[[v-\bar{v}]]_e||^2_{L^2(e)} \\
\leq\ & H_b^2 H_s^{-1}(1+log(H_s/h)) C \sum_j ||v-\bar{v}||^2_{L^2(\partial\Omega_{i,j})} \\
\leq\ & H_b^2 H_s^{-1}(1+log(H_s/h)) C \sum_j \left(H_s|v-\bar{v}|^2_{H^1(\Omega_{i,j})} + H_s^{-1}||v-\bar{v}||^2_{L^2(\Omega_{i,j})}\right) \\
\leq\ & C H_b^2 (1+log(H_s/h)) \sum_j |v-\bar{v}|^2_{H^1(\Omega_{i,j})}, \quad (3.23)
\end{aligned}
$$

where the second inequality follows from a trace theorem and the third a Poincaré inequality with scaling.

Assuming e is shared by $\Omega_{i,j}$ and $\Omega_{i,k}$ and $\mathcal{V}$ is a vertex of e, we have

$$
\begin{aligned}
& H_b^2 H_s^{-1}(1+log(H_s/h)) |||[[v]]_e||^2_{L^2(e)} \\
=\ & H_b^2 H_s^{-1}(1+log(H_s/h)) ||v_{i,j}-v_{i,k}||^2_{L^2(e)} \\
\leq\ & 2H_b^2 H_s^{-1}(1+log(H_s/h)) \left(||v_{i,j}-v_{i,j}(\mathcal{V})||^2_{L^2(e)} + ||v_{i,k}-v_{i,k}(\mathcal{V})||^2_{L^2(e)}\right) \\
\leq\ & C H_b^2 H_s^{-1}(1+log(H_s/h)) \left(H_s||v_{i,j}-v_{i,j}(\mathcal{V})||^2_{L^\infty(\Omega_{i,j})} + H_s||v_{i,k}-v_{i,k}(\mathcal{V})||^2_{L^\infty(\Omega_{i,k})}\right) \\
\leq\ & C H_b^2 (1+log(H_s/h))^2 \left(|v|^2_{H^1(\Omega_{i,j})} + |v|^2_{H^1(\Omega_{i,k})}\right), \quad (3.24)
\end{aligned}
$$

where the last inequality follows from [43, Lemma 4.15]. Combining (3.22), (3.23), and (3.24), we obtain (3.13).

Similarly, we have

$$
\begin{aligned}
& ||v||^2_{L^2(\Omega_i)} \\
\leq\ & 2||v-\bar{v}||^2_{L^2(\Omega_i)} + 2||\bar{v}||^2_{L^2(\Omega_i)} \\
\leq\ & 2\sum_j ||v-\bar{v}||^2_{L^2(\Omega_{i,j})} + C\Big(H_b^2 H_s^{-1}(1+log(H_s/h)) \sum_{e\in\mathcal{E}^i(\Omega_i)} |||[\bar{v}]]_e||^2_{L^2(\Omega_i)} \\
& + \left|\int_{\partial\Omega_i\cap\partial\Omega_D} \bar{v} ds\right|^2\Big) \\
\leq\ & 2\sum_j ||v-\bar{v}||^2_{L^2(\Omega_{i,j})} + C\Big(H_b^2 H_s^{-1}(1+log(H_s/h)) \sum_{e\in\mathcal{E}^i(\Omega_i)} |||[\bar{v}]]_e||^2_{L^2(\Omega_i)} + \\
& \left|\int_{\partial\Omega_i\cap\partial\Omega_D} v ds\right|^2 + \left|\int_{\partial\Omega_i\cap\partial\Omega_D} (v-\bar{v}) ds\right|^2\Big). \qquad (3.25)
\end{aligned}
$$

Letting $\mathcal{E}^b(\Omega_i)$ denote the set of exterior subdomain edges of Ω_i, we have

$$
\begin{aligned}
& \left|\int_{\partial\Omega_i\cap\partial\Omega_D} (v-\bar{v}) dx\right|^2 \\
\leq\ & |\partial\Omega_i\cap\partial\Omega_D| \int_{\partial\Omega_i\cap\partial\Omega_D} (v-\bar{v})^2 ds \\
=\ & |\partial\Omega_i\cap\partial\Omega_D| \sum_{e\in\mathcal{E}^b(\Omega_i)} ||v-\bar{v}||^2_{L^2(e)} \\
\leq\ & |\partial\Omega_i\cap\partial\Omega_D| \sum_j ||v-\bar{v}||^2_{L^2(\partial\Omega_{i,j})} \\
\leq\ & C|\partial\Omega_i\cap\partial\Omega_D| \sum_j H_s |v|^2_{H^1(\Omega_{i,j})} \\
\leq\ & C H_b H_s \sum_j |v|^2_{H^1(\Omega_{i,j})}. \qquad (3.26)
\end{aligned}
$$

Combining (3.25), (3.26), (3.23), and (3.24), we obtain (3.14). □

We note that $|\int_{\Omega_i} v dx|^2$ and $|\int_{\partial\Omega_i\cap\partial\Omega_D} v ds|^2$ of (3.13) and (3.14), respectively, can be replaced by $\sum_{i=1}^L |f_i(u)|^2$ (with a proper scaling factor), where $f_i, i = 1,\cdots,L, L\geq 1$ are functionals in $H^1(\Omega)$, such that, if u is constant in Ω,

$$\sum_{i=1}^L |f_i(u)|^2 = 0 \Leftrightarrow v = 0;$$

see Theorem 1.2.1.

We provide a trace theorem for functions in $\widetilde{W}^{(i)}$, for $\Omega_i \in \mathbb{R}^2$ with Lipschitz subdomains, $\Omega_{i,j}, j = 1,\cdots,N_i$.

Lemma 3.3.12. *Let $u \in \widetilde{W}^{(i)} = W_I^{(i)} \oplus \widetilde{W}_\Gamma^{(i)}$. Then we have*

$$||u||^2_{L^2(\partial\Omega_i)} \leq C \left(H_b(1 + log(H_s/h))^2 \sum_{j=1}^{N_i} |u|^2_{H^1(\Omega_{i,j})} + \frac{1}{H_b}||u||^2_{L^2(\Omega_i)} \right). \quad (3.27)$$

Proof. We follow the framework of the proof of Lemma 3.3.11. Let c be piecewise constant in each subdomain of Ω_i, i.e., $c(x) = c_{i,j}, \forall x \in \Omega_{i,j}, j = 1, \cdots, N_i$. We define the function Ec as before, i.e., Ec is defined to be the average of c on $\Gamma^{(i)}_{loc}$, $Ec|_{\partial\Omega_i} = c|_{\partial\Omega_i}$, and discrete harmonic in each subdomain of Ω_i. Then,

$$\begin{aligned}
& ||c||^2_{L^2(\partial\Omega_i)} \\
= \; & ||Ec||^2_{L^2(\partial\Omega_i)} \\
\leq \; & C \left(H_b |Ec|^2_{H^1(\Omega_i)} + \frac{1}{H_b}||Ec||^2_{L^2(\Omega_i)} \right) \\
\leq \; & C \left(H_b \sum_j |Ec|^2_{H^1(\Omega_{i,j})} + \frac{1}{H_b}||c||^2_{L^2(\Omega_i)} + \frac{1}{H_b}||c - Ec||^2_{L^2(\Omega_i)} \right) \\
\leq \; & C \Big(H_b \sum_j |Ec|^2_{H^1(\Omega_{i,j})} + \frac{1}{H_b}||c||^2_{L^2(\Omega_i)} + \frac{1}{H_b} H_s^2 \sum_j |c - Ec|^2_{H^1(\Omega_{i,j})} + \\
& \frac{1}{H_b} H_s \sum_{e \in \mathcal{E}^i(\Omega_i)} |||[[c]]_e||^2_{L^2(e)} \Big) \\
\leq \; & C \left(H_b H_s^{-1}(1 + log(H_s/h)) \sum_{e \in \mathcal{E}^i(\Omega_i)} |||[[c]]_e||^2_{L^2(e)} + \frac{1}{H_b}||c||^2_{L^2(\Omega_i)} \right), \quad (3.28)
\end{aligned}$$

where the second, fourth, and fifth inequalities follow from a trace theorem with scaling, (3.16), and (3.20), respectively. Now let v satisfy the assumption of the lemma and $\bar{v}$ be a piecewise constant function which is the average of v in each subdomain of Ω_i. Then,

$$\begin{aligned}
& ||v||^2_{L^2(\partial\Omega_i)} \\
\leq \; & 2||v - \bar{v}||^2_{L^2(\partial\Omega_i)} + 2||\bar{v}||^2_{L^2(\partial\Omega_i)} \\
\leq \; & C \Big(H_b H_s^{-1}(1 + log(H_s/h)) \sum_{e \in \mathcal{E}^i(\Omega_i)} |||[[\bar{v}]]_e||^2_{L^2(e)} + \frac{1}{H_b}||v||^2_{L^2(\Omega_i)} + \\
& ||v - \bar{v}||^2_{L^2(\partial\Omega_i)} \Big). \quad (3.29)
\end{aligned}$$

Using a trace theorem with scaling at the subdomain level, we obtain

$$||v - \bar{v}||^2_{L^2(\partial\Omega_i)}$$

$$
\begin{aligned}
&\leq \sum_j ||v-\bar{v}||^2_{L^2(\partial\Omega_{i,j})} \\
&\leq C\sum_j \left(H_s|v-\bar{v}|^2_{H^1(\Omega_{i,j})} + \frac{1}{H_s}||v-\bar{v}||^2_{L^2(\Omega_{i,j})} \right) \\
&\leq C\sum_j H_s|v|^2_{H^1(\Omega_{i,j})}. \qquad (3.30)
\end{aligned}
$$

Bounding $\sum_{e\in\mathcal{E}^i(\Omega_i)} |||[\bar{v}]]_e||^2_{L^2(e)}$ as in (3.23) and (3.24), we have

$$
\begin{aligned}
&||v||^2_{L^2(\partial\Omega_i)} \\
&\leq C\left(H_b(1+log(H_s/h))^2 \sum_j |v|^2_{H^1(\Omega_{i,j})} + \frac{1}{H_b}||v||^2_{L^2(\Omega_i)} \right). \qquad (3.31)
\end{aligned}
$$

□

Before we complete this section, we comment on the analysis of linear elasticity problems in two and three dimensions. Korn inequality is essential in any such analysis and here we present a version for Jones (i.e., uniform) domains. For a proof, see Durán and Muschietti [19].

Lemma 3.3.13 (Korn inequality for uniform domains)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded uniform domain. Then, there exists C, which depends only on the Jones parameter $C_U(\Omega)$ and the dimension n, such that*

$$
|\mathbf{u}|_{H^1(\Omega)} \leq C\sum_{i,j} ||\epsilon(\mathbf{u})_{ij}||^2_{L^2(\Omega)}
$$

for all $\mathbf{u} \in \{\mathbf{u} \in \mathbf{H^1}(\Omega) : \int_\Omega (\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i})dx = 0, i,j = 1,\cdots,n\}$.

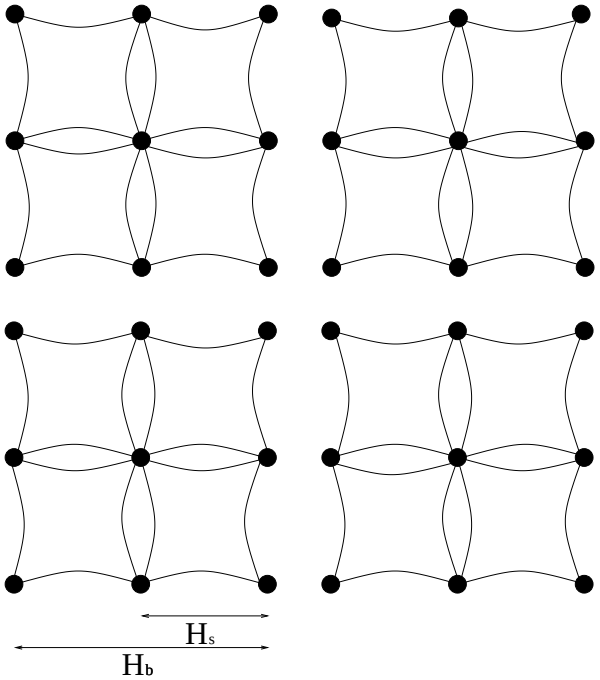
We would also need to consider different primal constraints on each body, especially for the case of three-dimensional problem; see [36] for details.

Using this Korn inequality, the analysis for a two-dimensional linear elasticity problem with Jones domains can be carried out without much difficulty. However, we do not have enough technical tools for the three-dimensional case at this point. For instance, face and edge lemmas similar to those of Lemma 3.3.5 would be essential; for such results for regular subdomains, see [43, Section 4.6]

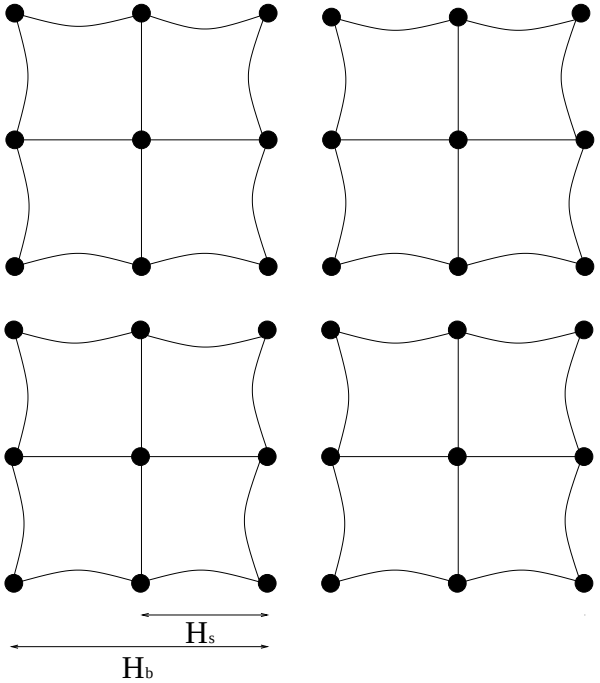
3.4 Algorithm

In the formulation of the FETI-FETI method in two dimensions, we will use the spaces $\widetilde{W}_c := \prod_{i=1}^N \widetilde{W}^{(i)} := \prod_{i=1}^N W_I^{(i)} \oplus \widetilde{W}_\Gamma^{(i)}$ and $\widetilde{W}_{\Gamma,c} := \prod_{i=1}^N \widetilde{W}_\Gamma^{(i)}$. See Figure 3.1(a).

Figure 3.1: $\widetilde{W}_c$ for FETI-FETI, $\widehat{W}_c$ for Hybrid



(a) FETI-FETI



(b) Hybrid

First, we introduce some notation which is very similar to those of the previous chapters. We introduce the matrices $A^{(i)}$, which are direct sums of the stiffness matrices $A^{(i,j)}, j = 1, \cdots, N_i$, for individual subdomains:

$$A^{(i)} = \begin{bmatrix} A^{(i,1)} & & \\ & \ddots & \\ & & A^{(i,N_i)} \end{bmatrix}, \quad i = 1, \cdots, N. \tag{3.32}$$

We also introduce the matrices $\widetilde{A}^{(i)}$, which are analogous to the matrix $\widetilde{A}$ introduced in (2.17). They are the restrictions of $A^{(i)}$ to $\widetilde{W}^{(i)} = W_I^{(i)} \oplus \widetilde{W}_\Gamma^{(i)}$:

$$\widetilde{A}^{(i)} = \begin{bmatrix} A_{II}^{(i,1)} & A_{\Delta I}^{(i,1)^T} & & & & \widetilde{A}_{\Pi I}^{(i,1)^T} \\ A_{\Delta I}^{(i,1)} & A_{\Delta\Delta}^{(i,1)} & & & & \widetilde{A}_{\Pi\Delta}^{(i,1)^T} \\ & & \ddots & & & \vdots \\ & & & A_{II}^{(i,N_i)} & A_{\Delta I}^{(i,N_i)^T} & \widetilde{A}_{\Pi I}^{(i,N_i)^T} \\ & & & A_{\Delta I}^{(i,N_i)} & A_{\Delta\Delta}^{(i,N_i)} & \widetilde{A}_{\Pi\Delta}^{(i,N_i)^T} \\ \widetilde{A}_{\Pi I}^{(i,1)} & \widetilde{A}_{\Pi\Delta}^{(i,1)} & \cdots & \widetilde{A}_{\Pi I}^{(i,N_i)} & \widetilde{A}_{\Pi\Delta}^{(i,N_i)} & \widetilde{A}_{\Pi\Pi}^{(i)} \end{bmatrix}.$$

Here,

$$\widetilde{A}_{\Pi I}^{(i,j)} = R_\Pi^{(i,j)^T} A_{\Pi I}^{(i,j)}, \quad \widetilde{A}_{\Pi\Delta}^{(i,j)} = R_\Pi^{(i,j)^T} A_{\Pi\Delta}^{(i,j)}, \quad i = 1, \cdots, N_i,$$

and

$$\widetilde{A}_{\Pi\Pi}^{(i)} = \sum_{j=1}^{N_i} R_\Pi^{(i,j)^T} A_{\Pi\Pi}^{(i,j)} R_\Pi^{(i,j)},$$

where $R_\Pi^{(i,j)} : \widehat{W}_\Pi^{(i)} \longrightarrow W_\Pi^{(i,j)}, j = 1, \cdots, N_i$.
A Schur complement $\widetilde{S}_\Gamma^{(i)}$ on $\widetilde{W}_\Gamma^{(i)}$ is obtained by eliminating the interior unknowns in each subdomain from $\widetilde{A}^{(i)}$; see (2.20). Note that $\widetilde{S}_\Gamma^{(i)}$ can also be regarded as the restriction of $S^{(i)}$ to $\widetilde{W}_\Gamma^{(i)}$, i.e.,

$$\widetilde{S}_\Gamma^{(i)} = \widetilde{R}_\Gamma^{(i)^T} S^{(i)} \widetilde{R}_\Gamma^{(i)},$$

where

$$S^{(i)} = \begin{bmatrix} S^{(i,1)} & & \\ & \ddots & \\ & & S^{(i,N_i)} \end{bmatrix}, \quad S^{(i,j)} = A_{\Gamma\Gamma}^{(i,j)} - A_{\Gamma I}^{(i,j)^T} A_{II}^{(i,j)^{-1}} A_{\Gamma I}^{(i,j)^T}, \quad j = 1, \cdots, N_i,$$

and $\widetilde{R}_\Gamma^{(i)} : \widetilde{W}_\Gamma^{(i)} \to W_\Gamma^{(i)}$, defined similarly as $\widetilde{R}_\Gamma$ of Section 2.3.

Recalling that we are using an active set method to deal with the inequality

conditions, we formulate the minimization problem on the current active set:

$$\min_{u\in\widetilde{W}_c}\frac{1}{2}u^T\widetilde{A}_c u-\widetilde{f}_c^T u,\quad \text{with}\quad Z^k\widetilde{B}_c u=0, \tag{3.33}$$

where

$$\widetilde{A}_c=\begin{bmatrix}\widetilde{A}^{(1)} & & \\ & \ddots & \\ & & \widetilde{A}^{(N)}\end{bmatrix},\quad u=\begin{bmatrix}u^{(1)}\\ \vdots\\ u^{(N)}\end{bmatrix},\quad \widetilde{f}_c=\begin{bmatrix}\widetilde{f}^{(1)}\\ \vdots\\ \widetilde{f}^{(N)}\end{bmatrix},$$

$$u^{(i)}\in\widetilde{W}^{(i)}=W_I^{(i)}\oplus\widetilde{W}_\Gamma^{(i)},\quad i=1,\cdots,N,$$

and

$$\widetilde{B}_c=\begin{bmatrix}B_{loc}\\ B_{gl}\end{bmatrix}=\begin{bmatrix}B_{loc}^{(1)} & \cdots & 0\\ 0 & \ddots & 0\\ 0 & \cdots & B_{loc}^{(N)}\\ B_{gl}^{(1)} & \cdots & B_{gl}^{(N)}\end{bmatrix},$$

$$B_{loc}^{(i)}=\begin{bmatrix}B_{loc}^{(i,1)} & \cdots B_{loc}^{(i,N_i)}\end{bmatrix},\quad i=1,\cdots,N,$$

$$Z^k=\begin{bmatrix}I & 0\\ 0 & Z_{gl}^k\end{bmatrix}.$$

$Z^k\widetilde{B}_c u=0$ in (3.33) indicates the continuity constraint across the local subdomain interface $\Gamma_{loc}^{(i)}, i=1,\cdots,N$, as well as the continuity constraint across the global area of contact Γ_{gl}^k. Z_{gl}^k is a square matrix obtained by replacing some of the diagonal entries of the identity matrix with zeros; only the entries corresponding to the nodes at which an equality is imposed are retained. We use the superscript k to remind us that Z_{gl}^k and Z^k change in each iteration of the active set method. We have $B_{loc}^{(i)}u^{(i)}=0, u^{(i)}\in\widetilde{W}^{(i)}=W_I^{(i)}\oplus\widetilde{W}_\Gamma^{(i)}$, exactly when the values associated with more than one subdomain on the body Ω_i coincide. Note that $B_{loc}^{(i)}$ has nonzero columns only for the components of $W_\Delta^{(i)}$.
We also introduce a scaled jump operator, $\widetilde{B}_{D,c}$:

$$\widetilde{B}_{D,c}=\begin{bmatrix}B_{loc,D}\\ B_{gl,D}\end{bmatrix}=\begin{bmatrix}B_{loc,D}^{(1)} & \cdots & 0\\ 0 & \ddots & 0\\ 0 & \cdots & B_{loc,D}^{(N)}\\ B_{gl,D}^{(1)} & \cdots & B_{gl,D}^{(N)}\end{bmatrix}$$

and

$$B_{loc,D}^{(i)} = \left[\begin{array}{ccc} B_{loc,D}^{(i,1)} & \cdots & B_{loc,D}^{(i,N_i)} \end{array} \right], \quad i = 1, \cdots, N.$$

$B_{loc,D}^{(i)}$ and $B_{gl,D}^{(i)}$ are obtained in the same manner as $B_{D,\Gamma}$ of the one-level FETI method (see Section 2.5). The nonzero entry of $B_{loc}^{(i,j)}$ associated with the Lagrange multipliers for the continuity at the node $x \in \partial\Omega_{i,j} \cap \Omega_{i,k}$ is multiplied by $\delta_{i,k}^{\dagger}(x) = \rho_{i,k}^{\gamma}(x) / \sum_{s \in \mathcal{N}_{x,loc}^{(i)}} \rho_{i,s}^{\gamma}(x)$, where $\mathcal{N}_{x,loc}^{(i)}$ is the set of indices of the subdomains of Ω_i with x on their boundary. The nonzero entry of $B_{gl}^{(i)}$ associated with the Lagrange multiplier for the continuity at the node $x \in \partial\Omega_i \cap \partial\Omega_j$ is multiplied by $\delta_j^{\dagger}(x) = \sum_{s \in \mathcal{N}_{x,loc}^{(j)}} \rho_{j,s}^{\gamma}(x) / \sum_{k \in \mathcal{N}_{x,gl}, t \in \mathcal{N}_{x,loc}^{(k)}} \rho_{k,t}^{\gamma}(x)$, where $\mathcal{N}_{x,gl}$ is the set of indices of the subdomains of any body which share the node x on their boundary.

Eliminating the interior unknowns in all subdomains of each body, we obtain the following reduced minimization problem,

$$\min_{u_\Gamma \in \widetilde{W}_{\Gamma,c}} \frac{1}{2} u_\Gamma^T \widetilde{S}_c u_\Gamma - \widetilde{g}_c^T u_\Gamma, \quad \text{with} \quad Z_\Gamma^k \widetilde{B}_{\Gamma,c} u_\Gamma = 0, \tag{3.34}$$

where

$$\widetilde{S}_c = \left[\begin{array}{ccc} \widetilde{S}_\Gamma^{(1)} & & \\ & \ddots & \\ & & \widetilde{S}_\Gamma^{(N)} \end{array} \right], \quad u_\Gamma = \left[\begin{array}{c} u_\Gamma^{(1)} \\ \vdots \\ u_\Gamma^{(N)} \end{array} \right], \quad u_\Gamma^{(i)} \in \widetilde{W}_\Gamma^{(i)}, i = 1, \cdots, N.$$

Z_Γ^k is obtained by removing some of the rows and columns of Z^k. This minimization problem is equivalent to the following KKT system:

$$\left[\begin{array}{cc} \widetilde{S}_c & (Z_\Gamma^k B_{\Gamma,c})^T \\ Z_\Gamma^k \widetilde{B}_{\Gamma,c} & 0 \end{array} \right] \left[\begin{array}{c} u_\Gamma \\ \lambda \end{array} \right] = \left[\begin{array}{c} \widetilde{g}_c \\ 0 \end{array} \right]. \tag{3.35}$$

It is natural to reduce this system to an equation for λ as in the one-level FETI method and solve it with the PCG method in a proper subspace, using the following preconditioner:

$$M_D^{-1} := Z_\Gamma^k \widetilde{B}_{D,\Gamma_c} \widetilde{S} \widetilde{B}_{D,\Gamma_c}^T Z_\Gamma^k.$$

The resulting method, which we name the FETI-FETI method, turns out not to be scalable with respect to the number of subdomains. We present a partial explanation for this phenomenon, following the framework of [43, Section 6.3].

Let $P_D := B_{D,\Gamma}^T B_\Gamma$, where B_Γ and $B_{D,\Gamma}$ are the jump operator and the scaled jump operator for the one-level FETI method, respectively, defined in Section 2.3. At the core of the eigenvalue analysis for the one-level FETI method is the following result:

Lemma 3.4.1. *For any $w \in \text{range}(S)$, we have*

$$|P_D w|_S^2 \leq C(1 + log(H/h))^2 |w|_S^2,$$

where C is independent of H, h.

For a proof, see [43, Section 6.2.3]. This lemma is used for bounding $\lambda_{max}(M_D^{-1}F)$ from above. It is easy to show that $\lambda_{min}(M_D^{-1}F) \geq 1$, and therefore that the condition number of the preconditioned operator for the one-level FETI method grows like $C(1 + log(H/h))^2$. In order to prove the existence of a convergence bound of the FETI-FETI method with a similar technique, we would need to bound $|\widetilde{P}_D^k w|_{\widetilde{S}_c}$ from above by $|w|_{\widetilde{S}_c}$, for $w \in range(\widetilde{S}_c)$. Here, $\widetilde{P}_D^k := \widetilde{B}_{D,\Gamma_c}^T Z_\Gamma^k \widetilde{B}_{\Gamma_c}$. We do this in the next section.

3.5 Condition Number Estimates

3.5.1 Convergence bound for the FETI-FETI method

In this section, we assume that our two-dimensional subdomains have Lipschitz boundaries. Recall that $\widetilde{P}_D^k := \widetilde{B}_{D,\Gamma_c}^T Z_\Gamma^k \widetilde{B}_{\Gamma,c}$, and thus the operator $\widetilde{P}_D^k$ changes in each step of the active set method; see Section 3.2 and Section 3.4. We prove

Lemma 3.5.1. *For any $w \in range(\widetilde{S}_c)$, we have*

$$|\widetilde{P}_D^k w|_{\widetilde{S}_c}^2 \leq C \frac{H_b}{H_s}(1 + log(H_s/h))^2 |w|_{\widetilde{S}_c}^2,$$

where $C > 0$ is a constant independent of H_b, H_s, h.

We first make some observations. We obtain the following formulae by modifying [43, (6.42)]:

$$(\widetilde{P}_D^k w(x))_{i,j} = \sum_{s \in \mathcal{N}_{x,loc}^{(i)}} \delta_{i,s}^\dagger(x)(w_{i,j}(x) - w_{i,s}(x)), \quad \text{if} \quad x \in \partial\Omega_{i,j} \cap \Gamma_{loc}^{(i)}, \tag{3.36}$$

$$(\widetilde{P}_D^k w(x))_i = \sum_{k \in \mathcal{N}_{x,gl}} \delta_k^\dagger(x)(w_i(x) - w_k(x)), \quad \text{if} \quad x \in \partial\Omega_i \cap \Gamma_{gl}^k. \tag{3.37}$$

Also, recalling that $\widetilde{P}_D^k := \widetilde{B}_{D,\Gamma_c}^T Z_\Gamma^k \widetilde{B}_{\Gamma_c}$,

$$(\widetilde{P}_D^k w(x))_i = 0, \quad \text{if} \quad x \in \partial\Omega_i \cap \overline{\Gamma_{gl}} \setminus \Gamma_{gl}^k. \tag{3.38}$$

The above equality is due to the fact that the nodes which do not belong to the current active set Γ_{gl}^k are *deactivated*. In (3.37) and (3.38), we do not specify

subdomain indices since we are considering nodes belonging to only one subdomain. However, in the following discussion, we sometimes do specify the relevant subdomain indices when necessary.

We follow the proof of [43, Lemma 6.3] very closely. However, one of the main differences between the proof there and the proof we present here is that we mainly work with H^1- seminorms instead of $H^{1/2}$- seminorms. Also, we do not have a Poincaré-type inequality such as [43, Lemma 6.2] for individual subdomains, since in our algorithm subdomains of the same body are connected by certain continuity constraints; instead, we use a Poincaré-type inequality (Lemma 3.3.11) for entire bodies.

Proof. Recalling that

$$|\widetilde{P}_D^k w|^2_{\widetilde{S}_c} = \sum_{i=1}^{N} |(\widetilde{P}_D^k w)_i|^2_{\widetilde{S}_\Gamma^{(i)}}, \quad |w|^2_{\widetilde{S}_c} = \sum_{i=1}^{N} |w_i|^2_{\widetilde{S}_\Gamma^{(i)}},$$

it suffices to show that

$$|(\widetilde{P}_D^k w)_i|^2_{\widetilde{S}_\Gamma^{(i)}} \leq C \frac{H_b}{H_s}(1 + log(H_s/h))^2 |w_i|^2_{\widetilde{S}_\Gamma^{(i)}}, \quad i = 1, 2, \cdots, N.$$

Furthermore,

$$|(\widetilde{P}_D^k w)_i|^2_{\widetilde{S}_\Gamma^{(i)}} = \sum_{j=1}^{N_i} |(\widetilde{P}_D^k w)_{i,j}|^2_{S^{i,j}},$$

where $(\widetilde{P}_D^k w)_{i,j}$ is the restriction of $(\widetilde{P}_D^k w)_i \in \widetilde{W}_\Gamma^{(i)}$ to $W_\Gamma^{(i,j)}$. Thus it suffices to examine each $|(\widetilde{P}_D^k w)_{i,j}|_{S^{(i,j)}}$ separately. For notational simplicity, let $v_{i,j}(x) := (\widetilde{P}_D^k w)_{i,j}$. We can see that the coefficients in (3.36) and (3.37) are constant on each individual edge while their values will differ between different edges. Also, $(\widetilde{P}_D^k w(x))_{i,j} = 0$ for a vertex node $x \in \Gamma_{loc}^{(i)}$, since we are imposing continuity at all vertices of the same body. Therefore it makes sense to write $v_{i,j}$ as a sum of functions each of which vanishes at all interface nodes outside a certain edge or a vertex. We can accomplish this by using characteristic finite element functions for individual edges and vertices; such characteristic functions for an edge and a vertex, $\theta_\mathcal{E}$ and $\theta_\mathcal{V}$, have already been introduced in Lemma 3.3.5 and Lemma 3.3.11, respectively. Construction of these finite element functions and the proof of their characteristics for three-dimensional problems can be found in [43, Chapter 4]. Construction of $\theta_\mathcal{E}$ and $\theta_\mathcal{V}$ for two-dimensional problems is analogous and here we just present their characteristics without any proofs.

Two-dimensional Case

Using the partition of unity,

$$v_{i,j} = \sum_{\mathcal{E}\subset\partial\Omega_{i,j}\cap\Gamma_{gl}^k} I^h(\theta_{\mathcal{E}} v_{i,j}) + \sum_{\mathcal{E}\subset\partial\Omega_{i,j}\cap\Gamma_{loc}^{(i)}} I^h(\theta_{\mathcal{E}} v_{i,j}) + \sum_{\mathcal{V}\subset\partial\Omega_{i,j}\cap\Gamma_{gl}^k} I^h(\theta_{\mathcal{V}} v_{i,j}).$$

We first consider the terms for the edges on the global boundary Γ_{gl}.
Edge Terms - Global Interface
Suppose $\mathcal{E}$ is shared by $\partial\Omega_{i,j}$ and $\partial\Omega_{k,l}$, where $i \neq k$. Then,

$$I^h(\theta_{\mathcal{E}} v_{i,j}) = I^h(\theta_{\mathcal{E}}\delta_k^\dagger(\mathcal{E})(w_{i,j} - w_{k,l})),$$

where $\delta_k^\dagger(\mathcal{E})$ is the constant value of $\delta_k^\dagger(x)$ on the edge $\mathcal{E}$.

$$|I^h(\theta_{\mathcal{E}} v_{i,j})|^2_{S^{(i,j)}} = \rho_{i,j}|\mathcal{H}(\theta_{\mathcal{E}} v_{i,j})|^2_{H^1(\Omega_{i,j})}. \tag{3.39}$$

We then have,

$$\begin{aligned}
& \rho_{i,j}|\mathcal{H}(\theta_{\mathcal{E}} v_{i,j})|^2_{H^1(\Omega_{i,j})} \\
= \;& \rho_{i,j}|\mathcal{H}(\theta_{\mathcal{E}}(\delta_k^\dagger(\mathcal{E})(w_{i,j} - w_{k,l})))|^2_{H^1(\Omega_{i,j})} \\
\leq \;& 2\rho_{i,j}\delta_k^\dagger(\mathcal{E})^2(|\mathcal{H}(\theta_{\mathcal{E}}(w_{i,j}))|^2_{H^1(\Omega_{i,j})} + |\mathcal{H}(\theta_{\mathcal{E}}(w_{k,l}))|^2_{H^1(\Omega_{i,j})}) \\
\leq \;& 2\min(\rho_{i,j},\rho_{k,l})(|\mathcal{H}(\theta_{\mathcal{E}}(w_{i,j}))|^2_{H^1(\Omega_{i,j})} + |\mathcal{H}(\theta_{\mathcal{E}}(w_{k,l}))|^2_{H^1(\Omega_{i,j})}).
\end{aligned}$$

Here, the second inequality follows from [36, Lemma 8.4]. We treat the first term using Lemma 3.3.5:

$$\begin{aligned}
& 2\min(\rho_{i,j},\rho_{k,l})|\mathcal{H}(\theta_{\mathcal{E}}(w_{i,j}))|^2_{H^1(\Omega_{i,j})} \\
\leq \;& 2\rho_{i,j}(1 + log(H_s/h))^2||\mathcal{H}(w_{i,j})||^2_{H^1(\Omega_{i,j})}
\end{aligned}$$

For the second term, we also need the extension lemma:

$$\begin{aligned}
& 2\min(\rho_{i,j},\rho_{k,l})|\mathcal{H}(\theta_{\mathcal{E}}(w_{k,l}))|^2_{H^1(\Omega_{i,j})} \\
\leq \;& 2C\rho_{k,l}|E^h_{kl,ij}(\mathcal{H}(\theta_{\mathcal{E}}(w_{k,l})))|^2_{H^1(\Omega_{i,j})} \\
\leq \;& 2C\rho_{k,l}||\mathcal{H}(\theta_{\mathcal{E}}(w_{k,l}))||^2_{H^1(\Omega_{k,l})} \\
\leq \;& 2C\rho_{k,l}(1 + log(H_s/h))^2||\mathcal{H}(w_{k,l})||^2_{H^1(\Omega_{k,l})}.
\end{aligned}$$

We sum (3.39) over j, k and l, which indicate the indices of all subdomains of Ω_i intersecting the global boundary $\overline{\Gamma_{gl}}$, the indices of all bodies sharing their boundaries with Ω_i and the collection of the indices of all their subdomains sitting

on their boundaries, respectively:

$$\begin{aligned}
& \sum_j |I^h(\theta_{\mathcal{E}} v)|^2_{S^{(i,j)}} \\
\leq\ & C(1+log(H_s/h))^2 \left(\sum_j \rho_{i,j} ||\mathcal{H}(w_{i,j})||^2_{H^1(\Omega_{i,j})} + \sum_{k,l} \rho_{k,l} ||\mathcal{H}(w_{k,l})||^2_{H^1(\Omega_{k,l})} \right) \\
=\ & C(1+log(H_s/h))^2 \left(\sum_j \rho_{i,j} |\mathcal{H}(w_{i,j})|^2_{H^1(\Omega_{i,j})} + \sum_{k,l} \rho_{k,l} |\mathcal{H}(w_{k,l})|^2_{H^1(\Omega_{k,l})} \right) \\
+\ & C(1+log(H_s/h))^2 H_s^{-2} \Big(\sum_j \rho_{i,j} ||\mathcal{H}(w_{i,j})||^2_{L^2(\Omega_{i,j})} + \\
& \sum_{k,l} \rho_{k,l} ||\mathcal{H}(w_{k,l})||^2_{L^2(\Omega_{k,l})} \Big). \qquad (3.40)
\end{aligned}$$

We control the L^2- terms of (3.40) using a similar argument as in [43, Lemma 3.10]. Using a Friedrichs inequality for each subdomain of Ω_i which intersects the global boundary, we get

$$\rho_{i,j} ||\mathcal{H}(w_{i,j})||^2_{L^2(\Omega_{i,j})} \leq C \left(H_s^2 \rho_{i,j} |\mathcal{H}(w_{i,j})|^2_{H^1(\Omega_{i,j})} + H_s \rho_{i,j} ||w_{i,j}||^2_{L^2(\partial\Omega_{i,j} \cap \partial\Omega_i)} \right). \qquad (3.41)$$

Summing over the boundary of Ω_i, we get

$$\begin{aligned}
& \sum_j \rho_{i,j} ||\mathcal{H}(w_{i,j})||^2_{L^2(\Omega_{i,j})} \\
\leq\ & C \left(H_s^2 \sum_j \rho_{i,j} |\mathcal{H}(w_{i,j})|^2_{H^1(\Omega_{i,j})} + H_s \rho_i ||\mathcal{H}(w_i)||^2_{L^2(\partial\Omega_i)} \right) \\
\leq\ & C \Big(H_s^2 \sum_j \rho_{i,j} |\mathcal{H}(w_{i,j})|^2_{H^1(\Omega_{i,j})} + H_b H_s (1+log(H_s/h))^2 \rho_i \sum_{s=1}^{N_i} |\mathcal{H}(w_{i,s})|^2_{H^1(\Omega_{i,s})} + \\
& \frac{H_s}{H_b} \rho_i \sum_{s=1}^{N_i} ||\mathcal{H}(w_{i,s})||^2_{L^2(\Omega_{i,s})} \Big) \\
\leq\ & C H_b H_s (1+log(H_s/h))^2 \rho_i \sum_j |\mathcal{H}(w_{i,j})|^2_{H^1(\Omega_{i,j})} + \\
& C \frac{H_s}{H_b} \rho_i \cdot H_b^2 (1+log(H_s/h))^2 \sum_j |\mathcal{H}(w_{i,j})|^2_{H^1(\Omega_{i,j})} \\
\leq\ & C H_b H_s (1+log(H_s/h))^2 \rho_i \sum_j |\mathcal{H}(w_{i,j})|^2_{H^1(\Omega_{i,j})}.
\end{aligned}$$

where the second inequality follows from Lemma 3.3.12 and the third from Lemma 3.3.11. We also note that we repeatedly use the fact that the PDE coefficients vary only moderately within the same body, i.e., (3.2).

Edge Terms - Local Interface

Suppose $\mathcal{E}$ is shared by $\partial\Omega_{i,j}$ and $\partial\Omega_{i,s}$. Then,

$$I^h(\theta_{\mathcal{E}} v_{i,j}) = I^h(\theta_{\mathcal{E}} \delta^\dagger_{i,s}(\mathcal{E})(w_{i,j} - w_{i,s})),$$

where $\delta^\dagger_{i,s}(\mathcal{E})$ is the constant value of $\delta^\dagger_{i,s}(x)$ on the edge $\mathcal{E}$.

With $\bar{w}_{i,j} := \int_{\Omega_{i,j}} w_{i,j}dx / \int_{\Omega_{i,j}} 1dx$ and $\bar{w}_{i,s} := \int_{\Omega_{i,s}} w_{i,s}dx / \int_{\Omega_{i,s}} 1dx$, we have

$$\begin{aligned}
& |I^h(\theta_{\mathcal{E}} v_{i,j})|^2_{S^{(i,j)}} \\
= \;& \rho_{i,j} |\mathcal{H}(\theta_{\mathcal{E}} v_{i,j})|^2_{H^1(\Omega_{i,j})} \\
= \;& \rho_{i,j} |\mathcal{H}(\theta_{\mathcal{E}} \delta^{\dagger}_{i,s}(\mathcal{E})(w_{i,j} - w_{i,s}))|^2_{H^1(\Omega_{i,j})} \\
= \;& \rho_{i,j} |\mathcal{H}(\theta_{\mathcal{E}} \delta^{\dagger}_{i,s}(\mathcal{E})((w_{i,j} - \bar{w}_{i,j}) - (w_{i,s} - \bar{w}_{i,s}) + (\bar{w}_{i,j} - \bar{w}_{i,s})))|^2_{H^1(\Omega_{i,j})} \\
\leq \;& 3\rho_{i,j} |\mathcal{H}(\theta_{\mathcal{E}} \delta^{\dagger}_{i,s}(\mathcal{E})(w_{i,j} - \bar{w}_{i,j})|^2_{H^1(\Omega_{i,j})} \\
+ \;& 3\rho_{i,j} |\mathcal{H}(\theta_{\mathcal{E}} \delta^{\dagger}_{i,s}(\mathcal{E})(w_{i,s} - \bar{w}_{i,s})|^2_{H^1(\Omega_{i,j})} \\
+ \;& 3\rho_{i,j} |\theta_{\mathcal{E}} \delta^{\dagger}_{i,s}(\mathcal{E})(\bar{w}_{i,j} - \bar{w}_{i,s})|^2_{H^1(\Omega_{i,j})}
\end{aligned} \tag{3.42}$$

We can estimate the first term using Lemma 3.3.5, a Poincaré inequality and [36, Lemma 8.4]:

$$\begin{aligned}
& \rho_{i,j} |\mathcal{H}(\theta_{\mathcal{E}} \delta^{\dagger}_{i,s}(\mathcal{E})(w_{i,j} - \bar{w}_{i,j})|^2_{H^1(\Omega_{i,j})} \\
\leq \;& C\rho_{i,j} \delta^{\dagger}_{i,s}(\mathcal{E})^2 (1 + log(H_s/h))^2 ||\mathcal{H}(w_{i,j}) - \bar{w}_{i,j}||^2_{H^1(\Omega_{i,j})} \\
\leq \;& C\rho_{i,j} (1 + log(H_s/h))^2 |\mathcal{H}(w_{i,j})|^2_{H^1(\Omega_{i,j})}.
\end{aligned} \tag{3.43}$$

For the second term, we need to use Lemma 3.3.2 in addition to the other lemmas:

$$\begin{aligned}
& \rho_{i,j} |\mathcal{H}(\theta_{\mathcal{E}} \delta^{\dagger}_{i,s}(\mathcal{E})(w_{i,s} - \bar{w}_{i,s})|^2_{H^1(\Omega_{i,j})} \\
\leq \;& \rho_{i,s} \delta^{\dagger}_{i,s}(\mathcal{E})^2 |E^h_{is,ij}(\mathcal{H}(\theta_{\mathcal{E}}(w_{i,s} - \bar{w}_{i,s}))|^2_{H^1(\Omega_{i,j})} \\
\leq \;& C\rho_{i,s} \delta^{\dagger}_{i,s}(\mathcal{E})^2 ||\mathcal{H}(\theta_{\mathcal{E}}(w_{i,s} - \bar{w}_{i,s})||^2_{H^1(\Omega_{i,s})} \\
\leq \;& C\rho_{i,s} (1 + log(H/h))^2 |\mathcal{H}(w_{i,s})|^2_{H^1(\Omega_{i,s})}.
\end{aligned} \tag{3.44}$$

For the last term,

$$|\theta_{\mathcal{E}}(\bar{w}_{i,j} - \bar{w}_{i,s})|^2_{H^1(\Omega_{i,j})} = |\theta_{\mathcal{E}}|^2_{H^1(\Omega_{i,j})} |(\bar{w}_{i,j} - \bar{w}_{i,s})|^2.$$

The energy of $\theta_{\mathcal{E}}$ can be estimated using Lemma 3.3.5. Adding and subtracting the common value $w_{i,j}(\mathcal{V}) = w_{i,s}(\mathcal{V})$, where $\mathcal{V}$ is an end point of the edge $\mathcal{E}$, we find that

$$|(\bar{w}_{i,j} - \bar{w}_{i,s})|^2 \leq 2|(\bar{w}_{i,j} - w_{i,j}(\mathcal{V}))|^2 + 2|(\bar{w}_{i,j} - w_{i,s}(\mathcal{V}))|^2.$$

We can estimate the first term on the right hand side using Lemma 3.3.3:

$$|(\bar{w}_{i,j} - w_{i,j}(\mathcal{V}))|^2 \leq ||w_{i,j} - \bar{w}_{i,j}||^2_{L^2(\Omega_{i,j})} + C(1 + log(H_s/h))|w_{i,j}|^2_{H^1(\Omega_{i,j})}.$$

The second term can be estimated in the same manner.

Vertex Terms

We note that

$$|\theta_{\mathcal{V}} v_{i,j}(\mathcal{V})|^2_{S^{(i,j)}} = \rho_{i,j}|\mathcal{H}(\theta_{\mathcal{V}} v_{i,j}(\mathcal{V}))|^2_{H^1(\Omega_{i,j})} = \rho_{i,j}|\theta_{\mathcal{V}} v_{i,j}(\mathcal{V})|^2_{H^1(\Omega_{i,j})}, \tag{3.45}$$

where the second equality follows from the fact that $v_{i,j}(\mathcal{V})$ is a constant and $\theta_{\mathcal{V}}$ is a discrete harmonic function. Denoting an auxiliary function which vanishes at every node in $\bar{\Omega}_{i,j,h}$ except at $\mathcal{V}$ where it assumes the value 1 by $\vartheta_{\mathcal{V}}$, we have

$$\begin{aligned} & \rho_{i,j}|\theta_{\mathcal{V}} v_{i,j}(\mathcal{V})|^2_{H^1(\Omega_{i,j})} \\ = \;& \rho_{i,j}|v_{i,j}(\mathcal{V})|^2|\theta_{\mathcal{V}}|^2_{H^1(\Omega_{i,j})} \\ \leq \;& \rho_{i,j}|v_{i,j}(\mathcal{V})|^2|\vartheta_{\mathcal{V}}|^2_{H^1(\Omega_{i,j})} \\ \leq \;& C\rho_{i,j}|v_{i,j}(\mathcal{V})|^2, \end{aligned} \tag{3.46}$$

where the first inequality follows from the minimality of the energy of the discrete harmonic functions and the second inequality from the fact that a nodal basis function in two dimensions has $O(1)$ energy. Using the formula (3.37) and Lemma 3.3.3,

$$\begin{aligned} & \rho_{i,j}|v_{i,j}(\mathcal{V})|^2 \\ = \;& \rho_{i,j}\left|w_i(\mathcal{V}) - \sum_{k\in\mathcal{N}_{\mathcal{V},gl}\setminus\{i\}} \delta^\dagger_k(\mathcal{V}) w_k(\mathcal{V})\right|^2 \\ \leq \;& \rho_{i,j}|\mathcal{N}_{\mathcal{V},gl}|\left(|w_i(\mathcal{V})|^2 + \sum_{k\in\mathcal{N}_{\mathcal{V},gl}\setminus\{i\}} \delta^\dagger_k(\mathcal{V})^2|w_k(\mathcal{V})|^2\right) \\ \leq \;& \rho_{i,j}|\mathcal{N}_{\mathcal{V},gl}|\left(||\mathcal{H}(w_{i,j})||^2_{L^\infty(\Omega_{i,j})} + \sum_{k\in\mathcal{N}_{\mathcal{V},gl}\setminus\{i\}} \delta^\dagger_k(\mathcal{V})^2||\mathcal{H}(w_{k,l})||^2_{L^\infty(\Omega_{k,l})}\right) \\ \leq \;& C\rho_{i,j}|\mathcal{N}_{\mathcal{V},gl}|(1+log(H_s/h))^2\Bigg(||\mathcal{H}(w_{i,j})||^2_{H^1(\Omega_{i,j})} + \\ & \sum_{k\in\mathcal{N}_{\mathcal{V},gl}\setminus\{i\}} \delta^\dagger_k(\mathcal{V})^2||\mathcal{H}(w_{k,l})||^2_{H^1(\Omega_{k,l})}\Bigg). \end{aligned} \tag{3.47}$$

And we proceed as usual. □

We now present a condition number estimate for the FETI-FETI method. We denote the subspace of Lagrange multipliers in which the preconditioned conjugate gradient method is performed by V^k:

$$V^k := \{\lambda \in range(Z^k_\Gamma \widetilde{B}_{\Gamma_c} : Z^k_\Gamma \widetilde{B}_{\Gamma_c}\lambda \in range(\widetilde{S}_c)\}.$$

Recall that $M^{-1}_D := Z^k_\Gamma \widetilde{B}_{D,\Gamma_c}\widetilde{S}_c\widetilde{B}^T_{D,\Gamma_c}Z^k_\Gamma$. Also, let $F := Z^k_\Gamma \widetilde{B}_{\Gamma_c}\widetilde{S}^\dagger_c\widetilde{B}^T_{\Gamma_c}Z^k_\Gamma$. Then we

have the following result:

Theorem 3.5.2. *For any $\lambda \in V^k$, we have*

$$\langle M_D \lambda, \lambda \rangle \leq \langle F\lambda, \lambda \rangle \leq C \frac{H_b}{H_s} (1 + log(H_s/h))^2 \langle M_D \lambda, \lambda \rangle,$$

where $C > 0$ is a constant independent of H_b, H_s, h.

In Section 4.4, we will present numerical results which imply that the the algebraic factor H_b/H_s in Theorem 3.5.2 cannot be removed.

Chapter 4

Hybrid method

In this chapter, we consider the hybrid method, which is a scalable alternative to the FETI-FETI method of Chapter 3. The hybrid method relies on the use of an *inexact solver*; before we start the description of the algorithm, we review other algorithms which use an inexact solver.

4.1 Some Algorithms Using Inexact Solvers

In FETI methods, we use a subdomain structure for which the continuity of the solution across the subdomain boundaries is achieved only at the convergence of the solution; see Figure 2.1. Thus the continuity condition needs to be enforced explicitly, which results in an energy minimization problem with an equality constraint, which is equivalent to a KKT system with displacement unknowns as primal variables and Lagrange multipliers as dual variables. This KKT system is reduced to an equation in Lagrange multipliers alone, and this reduction process amounts to solving a Neumann problem exactly on each subdomain. As this problem is solved by a CG method, an additional Dirichlet problem is solved exactly on each subdomain in the preconditioning step, which makes the convergence rate less sensitive to the number of unknowns in each subdomain.

The use of inexact Dirichlet solvers is possible without a radical change to the algorithm. However, the use of exact Neumann solvers is inherent to the structure of the KKT system and the use of inexact Neumann solvers would lead to a different problem to be solved. We can choose to keep the original KKT system the way it is and solve it with a suitable Krylov subspace method, for instance the preconditioned conjugate residual (PCR) method, with an efficient preconditioner.

The aforementioned approach was taken and successfully analyzed for the one-level FETI method by Klawonn and Widlund in [34]. In it, they solve a KKT system using the PCR method. They extended the convergence analysis of Mandel and Tezaur [40] for scalar, second-order elliptic equations to the system of

equations of linear elasticity, using the Korn's inequalities. Also, they analyzed the convergence rate of the PCR method using some of the results of Brezzi [9].

In FETI-DP methods, the continuity coupling between subdomains (primal constraints) results in a nonsingular system matrix and also provides a coarse solver which guarantees the scalability of the algorithm with respect to the number of subdomains. The coarse problem is solved exactly with a direct solver. However, the size of the coarse problem is usually proportional to the number of subdomains and can be very large when the number of subdomains is large, or the PDE coefficients are badly distributed and a large number of primal constraints is needed. In such a case solving the coarse problem can be quite costly. The BDDC method is very closely related to the FETI-DP method, and suffers from the same disadvantage when the size of the coarse problem is large.

This issue has been addressed by Xuemin Tu for the BDDC method in [45] for two dimensions and in [44] for three dimensions, and by Klawonn and Rheinbach in [30] for the FETI-DP method. In [45], Tu solves the coarse problem inexactly by grouping subdomains into *subregions* , i.e., introducing a higher level of hierarchy. The same strategy would not work for the FETI-DP method since the coarse problem is inherent to the formulation of the FETI-DP method, whereas in the BDDC method the coarse problem is solved in the preconditioning step. Therefore Klawonn and Rheinbach take a similar approach as in [34] and consider several different KKT systems, which allow the use of inexact solvers, using the subdomain structure of the FETI-DP method.

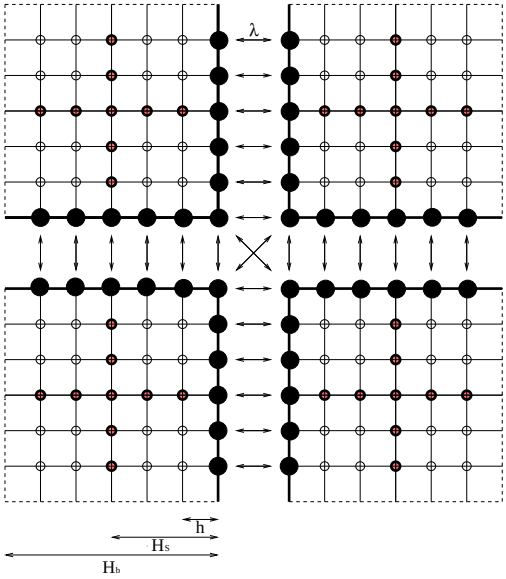
4.2 Algorithm and the Finite Element Space

We use the notation introduced in Chapter 3. We use finite element functions in the space $\widehat{W}_{\Gamma,c} := \prod_{i=1}^{N} \widehat{W}_{\Gamma}^{(i)}$; see Figure 3.1(b). $W_{\Gamma_{OL}}^{(i)}$ denotes a finite element space on $\partial\Omega_i \cap \overline{\Gamma_{gl}}$. Here, OL stands for One-Level; this is because we will use one-level FETI type preconditioners, which is obtained by regarding each body as a single subdomain. See Figure 4.1.

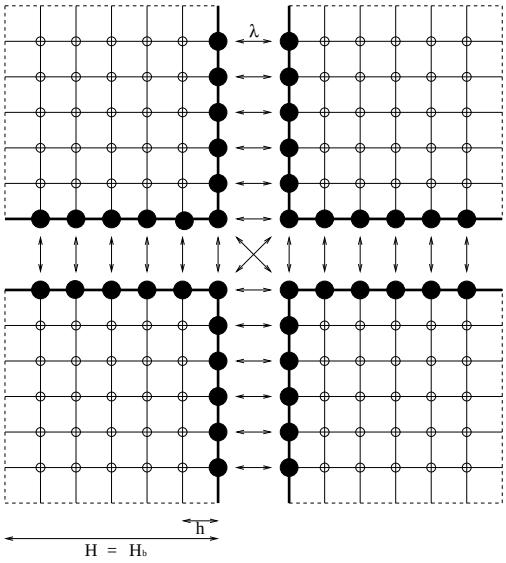
We introduce the Schur complement $\widehat{S}_{\Gamma}^{(i)}$ on $\widehat{W}_{\Gamma}^{(i)}$, which can be obtained by restricting $S^{(i)}$ to $\widehat{W}_{\Gamma}^{(i)}$:

$$\widehat{S}_{\Gamma}^{(i)} = \widehat{R}_{\Gamma}^{(i)^T} S^{(i)} \widehat{R}_{\Gamma}^{(i)}, i = 1, \cdots, N,$$

Figure 4.1: Subdomain structure of the hybrid and the one-Level FETI methods. In (a), the domain consists of 4 bodies ($\Omega_i, i = 1, 2, 3, 4$), each of which is divided into 4 subdomains ($\Omega_{i,j}, j = 1, 2, 3, 4$). In (b), the domain consists of 4 bodies, each of which is a single subdomain. Small and hollow dots in both (a) and (b) indicate interior nodes at which unknowns are eliminated; medium dots in (a) indicate nodes on the local interface between subdomains of the same body; big and solid dots in both (a) and (b) indicate the nodes on the global interface between the bodies. In both (a) and (b), arrows indicate Lagrange multipliers. Dotted lines indicate the Dirichlet boundary of the Poisson problem.



(a) Hybrid



(b) One-Level FETI

where $\widehat{R}_{\Gamma}^{(i)} : \widehat{W}_{\Gamma}^{(i)} \longrightarrow W_{\Gamma}^{(i)}$, the direct sum of the restrictions $\widehat{R}_{\Gamma}^{(i,j)} : \widehat{W}_{\Gamma}^{(i)} \longrightarrow W_{\Gamma}^{(i,j)}, j = 1, \cdots, N_i$. We also introduce restriction operators $\bar{R}_{\Gamma}^{(i)} : \widehat{W}_{\Gamma}^{(i)} \longrightarrow \widetilde{W}_{\Gamma}^{(i)}$,

$$\bar{R}_{\Gamma}^{(i)} = \begin{bmatrix} R_{\Gamma\Delta}^{(i,1)} \\ \vdots \\ R_{\Gamma\Delta}^{(i,N_i)} \\ R_{\Gamma\Pi}^{(i)} \end{bmatrix},$$

where $R_{\Gamma\Delta}^{(i,j)}$ extracts from a vector in $\widehat{W}_{\Gamma}^{(i)}$ the part that belongs to $W_{\Delta}^{(i,j)}$ and $R_{\Gamma\Pi}^{(i)} : \widehat{W}_{\Gamma}^{(i)} \longrightarrow \widehat{W}_{\Pi}^{(i)}$ is defined similarly. We also define the scaled versions $\bar{R}_{D,\Gamma}^{(i)}$:

$$\bar{R}_{D,\Gamma}^{(i)} = \begin{bmatrix} R_{D,\Gamma\Delta}^{(i,1)} \\ \vdots \\ R_{D,\Gamma\Delta}^{(i,N_i)} \\ R_{\Pi\Gamma}^{(i)} \end{bmatrix}.$$

Here, $R_{D,\Gamma\Delta}^{(i,j)}$ is obtained as follows: a nonzero entry of $R_{\Gamma\Delta}^{(i,j)}$, which corresponds to a node $x \in \partial\Omega_{i,j,h} \setminus \partial\Omega_{i,h}$, is multiplied by $\delta_{i,j}^{\dagger}(x)$, where

$$\delta_{i,j}^{\dagger}(x) := \frac{\rho_{i,j}^{\gamma}(x)}{\sum_{k \in \mathcal{N}_{x,loc}^{(i)}} \rho_{i,k}^{\gamma}(x)}.$$

The restriction of the minimization problem (3.34) in $\widetilde{W}_{\Gamma,c}$ to the subspace $\widehat{W}_{\Gamma,c}$ is as follows:

$$\min_{u_\Gamma \in \widehat{W}_{\Gamma,c}} \frac{1}{2} u_\Gamma^T \widehat{S}_c u_\Gamma - \widehat{g}_c^T u_\Gamma, \quad \text{with} \quad \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c} u_\Gamma = 0, \tag{4.1}$$

where

$$\widehat{S}_c = \begin{bmatrix} \widehat{S}_\Gamma^{(1)} & & \\ & \ddots & \\ & & \widehat{S}_\Gamma^{(N)} \end{bmatrix},$$

and $\widehat{Z}_\Gamma^k$ is obtained by removing irrelevant rows and columns from Z_Γ^k of (3.34). We remind the reader that we use the superscript k to indicate that Z_Γ^k and $\widehat{Z}_\Gamma^k$ change in each iteration of the active set method. Note that since the problem has been formulated in $\widehat{W}_{\Gamma,c}$, we do not need to impose continuity across the local subdomain interface, $\Gamma_{loc}^{(i)}$; continuity is already built into the structure. Therefore $\widehat{B}_{\Gamma,c}$ has nonzero entries only in the columns which correspond to a node on the global area of contact, $\overline{\Gamma_{gl}}$.

In addition to $\widehat{B}_{\Gamma,c}$ we define $B_{\Gamma_{OL}}$, which acts on vectors in the space $\prod_{i=1}^{N} W_{\Gamma_{OL}}^{(i)}$. This operator is needed in the preconditioner for the hybrid method. Recall that $\widehat{B}_{\Gamma,c}$ acts on vectors in the space $\prod_{i=1}^{N} \widehat{W}_{\Gamma}^{(i)}$ and only has rows corresponding to the Lagrange multipliers enforcing the continuity between different bodies, i.e., continuity across $\overline{\Gamma_{gl}}$. Thus $B_{\Gamma_{OL}}$ and $\widehat{B}_{\Gamma,c}$ differ only in that $\widehat{B}_{\Gamma,c}$ has a number of zero columns which correspond to the nodes on $\Gamma_{loc,h}$. We note that $B_{\Gamma_{OL}}$ can be regarded as the jump operator for the one-level FETI method viewing each body as a subdomain. We can define the scaled jump operator $B_{\Gamma_{OL},D}$ in the usual way.

Introducing a vector of Lagrange multipliers λ, we arrive at the following saddle point problem:

Find $(u_\Gamma, \lambda) \in \widehat{W}_{\Gamma,c} \times \text{range}(\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})$, such that

$$\begin{bmatrix} \widehat{S}_c & (\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \\ \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c} & 0 \end{bmatrix} \begin{bmatrix} u_\Gamma \\ \lambda \end{bmatrix} = \begin{bmatrix} \widehat{g}_c \\ 0 \end{bmatrix}. \tag{4.2}$$

We can solve (4.2) by reducing the system to an equation in λ alone in a proper subspace, but solving an equation of the form $\widehat{S}_c x = b$ is expensive. Instead, we keep the saddle point problem (4.2) the way it is and solve it by a Krylov subspace method which can deal with indefinite systems, such as the preconditioned conjugate residual (PCR) method. Due to the singularity of the matrix $\widehat{S}_c$, the solution of the upper part of the system (4.2) exists if and only if $\widehat{g}_c - (\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \lambda \in \text{range}(\widehat{S}_c)$. Most of the discussion here concerning this issue will be very similar to that of Section 2.5 on the one-level FETI method. As in the one-level FETI method, we introduce a matrix R_c such that $\text{range}(R_c) = ker(\widehat{S}_c)$:

$$R_c = \begin{bmatrix} \widehat{R}^{(1)} & & \\ & \ddots & \\ & & \widehat{R}^{(N)} \end{bmatrix},$$

where $\widehat{R}^{(i)}$ consists of the null vectors of $\widehat{S}_\Gamma^{(i)}, i = 1, \cdots, N$. In the PCR iterations, we will use an initial vector of Lagrange multipliers λ_0 which satisfies $\widehat{g}_c - (\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \lambda_0 \in \text{range}(\widehat{S}_c)$, and increments μ_i with $\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c}^T \mu_i \in \text{range}(\widehat{S}_c), i = 1, 2, \cdots$. Therefore the space of admissible increments is defined as follows:

$$V^k := \{\mu \in \text{range}(\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c}) : (\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \mu \in \text{range}(\widehat{S}_c)\} = \ker(G^{k^T}),$$

where $G^k := \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c} R_c$.

We introduce a projection operator P^k for the Lagrange mulipliers which is an

orthogonal projection from range($\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c}$) to $V^k = \ker(G^{k^T})$:

$$P^k := I - G^k (G^{k^T} G^k)^{-1} G^{k^T}.$$

We also introduce a subspace of $\widehat{W}_{\Gamma,c}$, $\widehat{W}_{\Gamma,R} := \text{range}(\widehat{S}_c)$. We rewrite (4.2) in terms of vectors in the subspace $\widehat{W}_{\Gamma,R} \times V^k$. First, noting that any admissible λ can be written as $\lambda = \lambda_0 + \mu, \mu \in V^k$, we rewrite the leading equation of (4.2) as

$$\widehat{S}_c u_\Gamma + (\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \mu = \widehat{g}_c - (\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \lambda_0. \tag{4.3}$$

Using (4.3) and $P^{k^T} \mu = P^k \mu = \mu$, we can rewrite (4.2):

$$\begin{bmatrix} \widehat{S}_c & (P^k \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \\ \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c} & 0 \end{bmatrix} \begin{bmatrix} u_\Gamma \\ \mu \end{bmatrix} = \begin{bmatrix} \widehat{g}_c - (\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \lambda_0 \\ 0 \end{bmatrix}. \tag{4.4}$$

The solution of (4.4) satisfies

$$\begin{bmatrix} \widehat{S}_c & (P^k \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \\ P^k \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c} & 0 \end{bmatrix} \begin{bmatrix} u_\Gamma \\ \mu \end{bmatrix} = \begin{bmatrix} \widehat{g}_c - (\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c})^T \lambda_0 \\ 0 \end{bmatrix}. \tag{4.5}$$

We use the system (4.5) in order to make sure that our iterates are in the subspace $\widehat{W}_{\Gamma,R} \times V^k$. However, the displacement variable is not uniquely defined by the system (4.5) since we are enforcing $P^k \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c} u_\Gamma = 0$, but not the original no-jump condition $\widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c} u_\Gamma = 0$; in this case, we can obtain a solution $u_\Gamma - R_c (G^{k^T} G^k)^{-1} G^{k^T} \widehat{Z}_\Gamma^k \widehat{B}_{\Gamma,c} u_\Gamma$ which satisfies all necessary requirements, with any given solution u_Γ of (4.5).

We now discuss our choice of the preconditioner. Let $A_{OL}^{(i)}$ denote the stiffness matrix for the entire body Ω_i: this needs to be distinguished from $A^{(i)}$, which is a direct sum of stiffness matrices for individual subdomains (see (3.32)). We have

$$A_{OL}^{(i)} = \begin{bmatrix} A_{II}^{(i)} & A_{\Gamma I}^{(i)^T} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{bmatrix}, \quad i = 1, \cdots, N,$$

where $A_{\Gamma\Gamma}^{(i)}$ is for the nodes on $\partial\Omega_i \cap \overline{\Gamma_{gl}}$ and $A_{II}^{(i)}$ is for all other nodes on Ω_i, etc. We define the corresponding Schur complement $S_{OL}^{(i)} := A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)} A_{II}^{(i)^{-1}} A_{\Gamma I}^{(i)^T}$ and also a block-diagonal matrix for the entire system consisting of the Schur complement

matrices for each body:

$$S_{OL} := \begin{bmatrix} S_{OL}^{(1)} & & \\ & \ddots & \\ & & S_{OL}^{(N)} \end{bmatrix}.$$

Noting that $A_{II}^{(i)}$ can be a large matrix and solving an equation of the form $A_{II}^{(i)}x = b$ can be quite expensive, we also introduce an *inexact* Schur complement

$$\widetilde{S}_{OL}^{(i)} := A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)} \widetilde{A}_{II}^{(i)^{-1}} A_{\Gamma I}^{(i)^T}, i = 1, \cdots, N,$$

and their direct sum, $\widetilde{S}_{OL} := \text{diag}_{i=1}^{N_i} \widetilde{S}_{OL}^{(i)}$, where $\widetilde{A}_{II}^{(i)^{-1}}$ is an inexact Dirichlet solver, which is defined as follows. Let $\widehat{W}_{\Gamma_0}^{(i)}$ and $W_{\Gamma_0}^{(i,j)}$ denote the space of continuous finite element functions on $\Gamma_{loc}^{(i)}$ and $\partial\Omega_{i,j} \cap \Gamma_{loc}^{(i)}$, which are similar to $\widehat{W}_{\Gamma}^{(i)}$ and $W_{\Gamma}^{(i,j)}$, respectively. Also, we define a restriction operator $R_{\Gamma_0}^{(i)} : \widehat{W}_{\Gamma_0}^{(i)} \to W_{\Gamma_0}^{(i,j)}$. After some symmmetric permutation, we can obtain

$$A_{II}^{(i)} = \begin{bmatrix} A_{II}^{(i,1)} & & & & A_{\Gamma_0 I}^{(i,1)^T} R_{\Gamma_0}^{(i,1)} \\ & A_{II}^{(i,2)} & & & A_{\Gamma_0 I}^{(i,2)^T} R_{\Gamma_0}^{(i,2)} \\ & & \ddots & & \vdots \\ & & & A_{II}^{(i,N_i)} & A_{\Gamma_0 I}^{(i,N_i)^T} R_{\Gamma_0}^{(i,N_i)} \\ R_{\Gamma_0}^{(i,1)^T} A_{\Gamma_0 I}^{(i,1)} & R_{\Gamma_0}^{(i,2)^T} A_{\Gamma_0 I}^{(i,2)} & \cdots & R_{\Gamma_0}^{(i,N_i)^T} A_{\Gamma_0 I}^{(i,N_i)} & \sum R_{\Gamma_0}^{(i,j)^T} A_{\Gamma_0\Gamma_0}^{(i,j)} R_{\Gamma_0}^{(i,j)} \end{bmatrix}. \tag{4.6}$$

The solution of $A_{II}^{(i)}x = b$ can be found by a block factorization. More precisely, with $\widehat{S}_{\Gamma_0}^{(i)} := \sum_{j=1}^{N_i} R_{\Gamma_0}^{(i,j)^T} (A_{\Gamma_0\Gamma_0}^{(i,j)} - A_{\Gamma_0 I}^{(i,j)} A_{II}^{(i,j)^{-1}} A_{\Gamma_0 I}^{(i,j)^T}) R_{\Gamma_0}^{(i,j)}$, we have

$$x_I^{(j)} = A_{II}^{(i,j)^{-1}} (b_I^{(j)} - A_{\Gamma_0 I}^{(i,j)^T} R_{\Gamma_0}^{(i,j)} x_{\Gamma_0}), \quad j = 1, \cdots, N_i, \tag{4.7}$$

where

$$\widehat{S}_{\Gamma_0}^{(i)} x_{\Gamma_0} = b_{\Gamma_0} - \sum_{j=1}^{N_i} R_{\Gamma_0}^{(i,j)^T} A_{\Gamma_0 I}^{(i,j)} A_{II}^{(i)^{-1}} b_I^{(j)}. \tag{4.8}$$

Solving (4.8) can be expensive; the solution of $\widetilde{A}_{II}^{(i)}\tilde{x} = b$ is defined as

$$\tilde{x}_I^{(j)} = A_{II}^{(i,j)^{-1}} (b_I^{(j)} - A_{\Gamma_0 I}^{(i,j)^T} R_{\Gamma_0}^{(i,j)} \tilde{x}_{\Gamma_0}), \quad j = 1, \cdots, N_i, \tag{4.9}$$

where

$$\tilde{x}_{\Gamma_0} = \widetilde{R}_{D,\Gamma_0}^{(i)^T} \widetilde{S}_{\Gamma_0}^{(i)^\dagger} \widetilde{R}_{D,\Gamma_0}^{(i)} \left(b_{\Gamma_0} - \sum_{j=1}^{N_i} R_{\Gamma_0}^{(i,j)^T} A_{\Gamma_0 I}^{(i,j)} A_{II}^{(i)^{-1}} b_I^{(j)} \right), \tag{4.10}$$

with $\widetilde{R}^{(i)}_{D,\Gamma_0}$ and $\widetilde{S}^{(i)}_{\Gamma_0}$ defined similarly as $\widetilde{R}^{(i)}_{D,\Gamma}$ and $\widetilde{S}^{(i)}_{\Gamma}$, respectively. We now introduce the following block-diagonal preconditioner for the system:

$$\mathcal{B}^{-1} = \begin{bmatrix} P_R M^{-1}_{BDDC} P_R & 0 \\ 0 & P^k M_D^{-1} P^k \end{bmatrix} \tag{4.11}$$

where $P_R := I - R_c(R_c^T R_c)^{-1} R_c^T$ is an orthogonal projection operator onto range$(\widehat{S}_c)$ and

$$M^{-1}_{BDDC} = \begin{bmatrix} \bar{R}^{(1)^T}_{D,\Gamma} \widetilde{S}^{(1)^\dagger}_{\Gamma} \bar{R}^{(1)}_{D,\Gamma} & & \\ & \ddots & \\ & & \bar{R}^{(N)^T}_{D,\Gamma} \widetilde{S}^{(N)^\dagger}_{\Gamma} \bar{R}^{(N)}_{D,\Gamma} \end{bmatrix},$$

$$M_D^{-1} = \widehat{Z}^k_\Gamma B_{\Gamma_{OL},D} \widetilde{S}_{OL} B^T_{\Gamma_{OL},D} \widehat{Z}^{k^T}_\Gamma.$$

We rewrite the KKT system (4.5):

$$\mathcal{A}x = \mathcal{F}, \tag{4.12}$$

where

$$\mathcal{A} := \begin{bmatrix} \widehat{S}_c & (P^k \widehat{Z}^k_\Gamma \widehat{B}_{\Gamma,c})^T \\ P^k \widehat{Z}^k_\Gamma \widehat{B}_{\Gamma,c} & 0 \end{bmatrix}, \quad x := \begin{bmatrix} u_\Gamma \\ \lambda \end{bmatrix} \quad \text{and} \quad \mathcal{F} := \begin{bmatrix} \widehat{g}_c \\ 0 \end{bmatrix}. \tag{4.13}$$

Our hybrid method is the preconditioned conjugate residual (PCR) method, to solve the preconditioned system

$$\mathcal{B}^{-1}\mathcal{A}x = \mathcal{B}^{-1}\mathcal{F}. \tag{4.14}$$

4.3 Convergence Number Estimates

The preconditioned system (4.14) is solved using the preconditioned conjugate residual (PCR) method. Suppose the following system is solved with the PCR method with the preconditioner $\mathcal{B}^{-1}$:

$$\mathcal{A}u = \mathcal{F}. \tag{4.15}$$

According to Lemma 1.3.2, we need to study the spectrum of the preconditioned operator $\mathcal{B}^{-1}\mathcal{A}$, which has the same spectrum as $\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2}$.

General Case We first study a general case, where

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}, \quad \mathcal{B}^{-1} = \begin{bmatrix} \widehat{A}^{-1} & 0 \\ 0 & \widehat{C}^{-1} \end{bmatrix},$$

assuming that

$$\alpha_0 u^T \widehat{A} u \leq u^T A u \leq \alpha_1 u^T \widehat{A} u, \quad \forall u. \tag{4.16}$$

We assume $A, \widehat{A}, \widehat{C}$ are real symmetric and positive definite, and B has full rank. Then,

$$\mathcal{B}^{-1/2} \mathcal{A} \mathcal{B}^{-1/2} = \begin{bmatrix} \widehat{A}^{-1/2} A \widehat{A}^{-1/2} & \widehat{A}^{-1/2} B^T \widehat{C} \\ \widehat{C}^{-1/2} B \widehat{A}^{-1/2} & 0 \end{bmatrix}.$$

In the following, we use the notation $\widetilde{A} := \widehat{A}^{-1/2} A \widehat{A}^{-1/2}$ and $\widetilde{B} := \widehat{C}^{-1/2} B \widehat{A}^{-1/2}$. Note that

$$\alpha_0 u^T u \leq u^T \widetilde{A} u \leq \alpha_1 u^T u, \quad \forall u. \tag{4.17}$$

We study the cases where $\widehat{A} = A$ and $\widehat{A} \neq A$ separately. When $\widehat{A} = A$, $\widetilde{A}$ is simply the identity matrix and

$$\mathcal{B}^{-1/2} \mathcal{A} \mathcal{B}^{-1/2} = \begin{bmatrix} I & \widetilde{B}^T \\ \widetilde{B} & 0 \end{bmatrix}. \tag{4.18}$$

Lemma 4.3.1. *Let $\mathcal{B}^{-1/2} \mathcal{A} \mathcal{B}^{-1/2}$ be defined as in (4.18). We then have*

$$\mathcal{K}(\mathcal{B}^{-1}\mathcal{A}) = \mathcal{K}(\mathcal{B}^{-1/2} \mathcal{A} \mathcal{B}^{-1/2}) = \frac{1/2 + \sqrt{1/4 + \lambda_{max}}}{-1/2 + \sqrt{1/4 + \lambda_{min}}},$$

where λ_{max} and λ_{min} are the largest and smallest eigenvalues of $\widetilde{B}^T \widetilde{B}$, respectively.

Proof. We consider the following eigenvalue problem:

$$\begin{bmatrix} I & \widetilde{B}^T \\ \widetilde{B} & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = t \begin{bmatrix} u \\ \lambda \end{bmatrix},$$

which is equivalent to

$$\begin{array}{rcl} u + \widetilde{B}^T \lambda & = & tu \\ \widetilde{B} u & = & t\lambda \end{array} \tag{4.19}$$

Substituting the second equation of (4.19) into the first, we obtain

$$u + t^{-1} \widetilde{B}^T \widetilde{B} u = tu.$$

Denoting the eigenvalues of $\widetilde{B}^T\widetilde{B}$ by $\lambda_i, i = 1, \cdots, n$, we obtain

$$(1 + \lambda_i/t - t)u = 0, \quad i = 1, \cdots, n.$$

Since $u = 0$ leads to $\lambda = 0$, we need to solve $1 + \lambda_i/t - t = 0, i = 1, \cdots, n$, which are equivalent to the quadratic equations $t^2 - t - \lambda_i = 0$. Their solutions are $1/2 \pm \sqrt{1/4 + \lambda_i}$ and thus

$$\sigma(\mathcal{B}^{-1}\mathcal{A}) = \{1/2 \pm \sqrt{1/4 + \lambda_i} : i = 1, \cdots, n\}.$$

Clearly,

$$\begin{aligned} \max\{|\lambda| : \lambda \in \sigma(\mathcal{B}^{-1}\mathcal{A})\} &= 1/2 + \sqrt{1/4 + \lambda_{max}} \text{ and} \\ \min\{|\lambda| : \lambda \in \sigma(\mathcal{B}^{-1}\mathcal{A})\} &= -1/2 + \sqrt{1/4 + \lambda_{min}}, \end{aligned}$$

where $\lambda_{max} := \max_{i=1}^n \lambda_i$ and $\lambda_{min} := \min_{i=1}^n \lambda_i$.

□

We now consider the case $\widehat{A} \neq A$. Then the eigenvalue analysis of $\mathcal{A}_1 := \mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2} = \begin{bmatrix} \widetilde{A} & \widetilde{B}^T \\ \widetilde{B} & 0 \end{bmatrix}$ is not as easy, and we left- and right- multiply this symmetrized preconditioned operator with $\mathcal{C}^{-1/2} = \begin{bmatrix} \widetilde{A}^{-1/2} & 0 \\ 0 & I \end{bmatrix}$ to obtain

$$\mathcal{A}_2 := \mathcal{C}^{-1/2}\mathcal{A}_1\mathcal{C}^{-1/2} = \begin{bmatrix} I & \widetilde{A}^{-1/2}\widetilde{B}^T \\ \widetilde{B}\widetilde{A}^{-1/2} & 0 \end{bmatrix}. \tag{4.20}$$

Eigenvalues of $\mathcal{A}_2$ can be analyzed in the same manner as in Lemma 4.3.1. To relate the spectrum of $\mathcal{A}_1$ to the spectrum of $\mathcal{A}_2$, we use the Courant-Fischer Minimax Theorem.

Theorem 4.3.2 (Courant-Fischer)**.** *Let $A \in R^{n\times n}$ be a symmetric matrix with real eigenvalues $\lambda_i, i = 1, \cdots, n$, which are ordered so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then*

$$\lambda_k = \max_{dim(V)=k} \min_{\substack{x\in V \\ x\neq 0}} \frac{x^T Ax}{x^T x} \tag{4.21}$$

$$\lambda_k = \min_{dim(V)=n-k+1} \max_{\substack{x\in V \\ x\neq 0}} \frac{x^T Ax}{x^T x} \tag{4.22}$$

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ denote the eigenvalues of $\mathcal{A}_2$, and $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_n$ the eigenvalues of $\mathcal{A}_1$. Suppose $\lambda_k > 0$ and $\lambda_{k+1} < 0$, where λ_k is the smallest positive eigenvalue of $\mathcal{A}_2$ and λ_{k+1} the largest negative eigenvalue of $\mathcal{A}_2$. Also, let

$q_i, i = 1, \cdots, n$ denote the eigenvectors of $\mathcal{A}_2$ such that $\mathcal{A}_2 q_i = \lambda_i q_i$ and $q_i^T q_j = \delta_{ij}, i, j = 1, \cdots, n$. Using (4.21) and the fact that $\mathcal{A}_1 = \mathcal{C}^{1/2}\mathcal{A}_2\mathcal{C}^{1/2}$ we have

$$\tilde{\lambda}_k = \max_{dim(V)=k} \min_{\substack{x\in V \\ x\neq 0}} \frac{x^T\mathcal{A}_1 x}{x^T x} = \max_{dim(V)=k} \min_{\substack{x\in V \\ x\neq 0}} \frac{(\mathcal{C}^{1/2}x)^T\mathcal{A}_2(\mathcal{C}^{1/2}x)}{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)} \frac{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)}{x^T x}$$

For $V := \mathcal{C}^{-1/2}\text{span}\{q^{(1)}, q^{(2)}, \cdots, q^{(k)}\}$, we have

$$\min_{\substack{x\in V \\ x\neq 0}} \frac{(\mathcal{C}^{1/2}x)^T\mathcal{A}_2(\mathcal{C}^{1/2}x)}{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)} \geq \lambda_k.$$

Noting that

$$\alpha_0 x^T x \leq x^T\mathcal{C}x \leq \alpha_1 x^T x, \quad \forall x \tag{4.23}$$

due to the definition of $\mathcal{C}$ and (4.17), we have

$$\min_{\substack{x\in V \\ x\neq 0}} \frac{(\mathcal{C}^{1/2}x)^T\mathcal{A}_1(\mathcal{C}^{1/2}x)}{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)} \frac{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)}{x^T x} \geq \lambda_k\alpha_0.$$

Taking the maximum over all k-dimensional subspaces on the left hand side of the previous equation, we obtain

$$\tilde{\lambda}_k \geq \lambda_k\alpha_0.$$

Similarly, using (4.22), we have

$$\tilde{\lambda}_{k+1} = \min_{dim(V)=n-k} \max_{\substack{x\in V \\ x\neq 0}} \frac{x^T\mathcal{A}_1 x}{x^T x} = \min_{dim(V)=n-k} \max_{\substack{x\in V \\ x\neq 0}} \frac{(\mathcal{C}^{1/2}x)^T\mathcal{A}_2(\mathcal{C}^{1/2}x)}{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)} \frac{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)}{x^T x}$$

For $V := \mathcal{C}^{-1/2}\{q^{(k+1)}, \cdots, q^{(n)}\}$, we have

$$\max_{\substack{x\in V \\ x\neq 0}} \frac{(\mathcal{C}^{1/2}x)^T\mathcal{A}_2(\mathcal{C}^{1/2}x)}{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)} \leq \lambda_{k+1}$$

and

$$\max_{\substack{x\in V \\ x\neq 0}} \frac{(\mathcal{C}^{1/2}x)^T\mathcal{A}_2(\mathcal{C}^{1/2}x)}{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)} \frac{(\mathcal{C}^{1/2}x)^T(\mathcal{C}^{1/2}x)}{x^T x} \leq \lambda_{k+1}\alpha_1.$$

Taking the minimum on the left hand side of the previous equation, we obtain

$$\tilde{\lambda}_{k+1} \leq \lambda_{k+1}\alpha_1.$$

By a similar argument,

$$\tilde{\lambda}_1 \leq \lambda_1\alpha_1 \quad \text{and} \quad \tilde{\lambda}_n \geq \lambda_n\alpha_0.$$

Letting λ'_{max} and λ'_{min} denote the maximum and the minimum eigenvalues of $\widetilde{A}^{-1/2}\widetilde{B}^T\widetilde{B}\widetilde{A}^{-1/2}$, respectively, we obtain

$$\begin{aligned}\mathcal{K}(\mathcal{A}_1) &= \frac{\max\{\tilde{\lambda}_1, |\tilde{\lambda}_n|\}}{\min\{\tilde{\lambda}_k, |\tilde{\lambda}_{k+1}|\}} \leq \frac{\alpha_1}{\alpha_0}\frac{\max\{\lambda_1, |\lambda_n|\}}{\min\{\lambda_k, |\lambda_{k+1}|\}} \\ = \quad \frac{\alpha_1}{\alpha_0}\mathcal{K}(\mathcal{A}_2) &\leq \frac{\alpha_1}{\alpha_0}\frac{1/2+\sqrt{1/4+\lambda'_{max}}}{-1/2+\sqrt{1/4+\lambda'_{min}}}, \end{aligned} \tag{4.24}$$

where the second inequality follows from the definition of $\mathcal{A}_2$ in (4.20) and Lemma 4.3.1. Noticing that

$$\begin{aligned}\lambda_{max}(\widetilde{A}^{-1/2}\widetilde{B}^T\widetilde{B}\widetilde{A}^{-1/2}) &\leq \lambda_{max}(\widetilde{B}^T\widetilde{B})\lambda_{max}(\widetilde{A}^{-1}) \\ \lambda_{min}(\widetilde{A}^{-1/2}\widetilde{B}^T\widetilde{B}\widetilde{A}^{-1/2}) &\geq \lambda_{min}(\widetilde{B}^T\widetilde{B})\lambda_{min}(\widetilde{A}^{-1}),\end{aligned}$$

and

$$\frac{1}{\alpha_1}u^Tu \leq u^T\widetilde{A}^{-1}u \leq \frac{1}{\alpha_0}u^Tu, \quad \forall u,$$

we rewrite (4.24) in terms of λ_{max} and λ_{min}, the maximum and the minimum eigenvalues of $\widetilde{B}^T\widetilde{B}$ and obtain:

$$\mathcal{K}(\mathcal{A}_1) \leq \frac{\alpha_1}{\alpha_0}\frac{1/2+\sqrt{1/4+\lambda_{max}/\alpha_0}}{-1/2+\sqrt{1/4+\lambda_{min}/\alpha_1}}. \tag{4.25}$$

Special Case We now use these results to study the convergence bound of our preconditioned system $\mathcal{B}^{-1}\mathcal{A}$, where $\mathcal{B}^{-1}$ and $\mathcal{A}$ are defined in (4.11) and (4.13), respectively. We have

$$A = \widehat{S}_\Gamma, \quad B = P^k\widehat{Z}_\Gamma^k\widehat{B}_{\Gamma,c}, \quad \widehat{A}^{-1} = P_R M_{BDDC}^{-1}P_R, \quad \widehat{C}^{-1} = P^k M_D^{-1}P^k.$$

Notice that $A, \widehat{A}$ and $\widehat{C}$ are now singular. However, this does not pose any problem, since in the application of the PCR method our iterates will be in a proper subspace in which those matrices will be nonsingular. From (4.25), we can see that the extreme eigenvalues of $\widetilde{B}^T\widetilde{B}$ and α_0, α_1 in (4.16) are important parameters, where $\widetilde{B}^T\widetilde{B} = \widehat{A}^{-1/2}B^T\widehat{C}^{-1}B\widehat{A}^{-1/2}$, which has the same spectrum as $B\widehat{A}^{-1}B^T\widehat{C}^{-1}$. In our case, $B\widehat{A}^{-1}B^T\widehat{C}^{-1}$ becomes $P^k\widehat{Z}_\Gamma^k\widehat{B}_\Gamma P_R M_{BDDC}^{-1}P_R\widehat{B}_\Gamma^T\widehat{Z}_\Gamma^{kT}P^k \cdot P^k M_D^{-1}P^k$. For a proof of the following lemma, see [38].

Lemma 4.3.3.

$$x^T\widehat{S}_\Gamma^{(i)\dagger}x \leq x^T\widetilde{R}_{D,\Gamma}^{(i)T}\widetilde{S}_\Gamma^{(i)^{-1}}\widetilde{R}_{D,\Gamma}^{(i)}x \leq C\left(1+log\left(\frac{H_s}{h}\right)\right)^2 x^T\widehat{S}_\Gamma^{(i)\dagger}x, \tag{4.26}$$

for all $x \in \text{range}(\widehat{S}_\Gamma^{(i)}), \quad i = 1, \cdots, N.$

Thus, we can study the spectrum of

$$\begin{aligned} & P^k \widehat{Z}_\Gamma^k \widehat{B}_\Gamma P_R \widehat{S}_\Gamma^\dagger P_R \widehat{B}_\Gamma^T \widehat{Z}_\Gamma^{kT} P^k \cdot P^k M_D^{-1} P^k \\ = \; & P^k \widehat{Z}_\Gamma^k \widehat{B}_\Gamma \widehat{S}_\Gamma^\dagger \widehat{B}_\Gamma^T \widehat{Z}_\Gamma^{kT} P^k \cdot P^k \widehat{Z}_\Gamma^k B_{\Gamma_{OL,D}} \widetilde{S}_{OL} B_{\Gamma_{OL,D}}^T \widehat{Z}_\Gamma^{kT} P^k \end{aligned}$$

instead of that of $P^k \widehat{Z}_\Gamma^k \widehat{B}_\Gamma P_R M_{BDDC}^{-1} P_R \widehat{B}_\Gamma^T \widehat{Z}_\Gamma^{kT} P^k \cdot P^k M_D^{-1} P^k$, where

$$\widehat{S}_\Gamma^\dagger = \begin{bmatrix} \widehat{S}_\Gamma^{(1)\dagger} & & \\ & \ddots & \\ & & \widehat{S}_\Gamma^{(N)\dagger} \end{bmatrix}.$$

Lemma 4.3.4.

$$\widehat{B}_\Gamma \widehat{S}_\Gamma^\dagger \widehat{B}_\Gamma^T \widehat{Z}_\Gamma^{kT} P^k = B_{\Gamma_{OL}} S_{OL}^\dagger B_{\Gamma_{OL}}^T \widehat{Z}_\Gamma^{kT} P^k.$$

Proof. Note that the solution of $\widehat{S}_\Gamma^{(i)} u_\Gamma^{(i)} = \widehat{B}_\Gamma^{(i)^T} v$, where $\widehat{B}_\Gamma^{(i)^T} v \in \text{range}(\widehat{S}_\Gamma^{(i)})$, can be obtained from the following equation:

$$\begin{bmatrix} A_{II}^{(i,1)} & & & & A_{\Gamma I}^{(i,1)^T} R_\Gamma^{(i,1)} \\ & A_{II}^{(i,2)} & & & A_{\Gamma I}^{(i,2)^T} R_\Gamma^{(i,2)} \\ & & \ddots & & \vdots \\ & & & A_{II}^{(i,N_i)} & A_{\Gamma I}^{(i,N_i)^T} R_\Gamma^{(i,N_i)} \\ R_\Gamma^{(i,1)^T} A_{\Gamma I}^{(i,1)} & R_\Gamma^{(i,2)^T} A_{\Gamma I}^{(i,2)} & \cdots & R_\Gamma^{(i,N_i)^T} A_{\Gamma I}^{(i,N_i)} & \sum_{j=1}^{N_i} R_\Gamma^{(i,j)^T} A_{\Gamma\Gamma}^{(i,j)} R_\Gamma^{(i,j)} \end{bmatrix} \begin{bmatrix} u_I^{(i,1)} \\ u_I^{(i,2)} \\ \vdots \\ u_I^{(i,N_i)} \\ \widehat{u}_\Gamma^{(i)} \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \widehat{B}_\Gamma^{(i)^T} v \end{bmatrix}, \tag{4.27}$$

where $R_\Gamma^{(i,j)} : \widehat{W}_\Gamma^{(i)} \longrightarrow W_\Gamma^{(i,j)}$ is a restriction operator. Noting that all entries of $\widehat{B}_\Gamma^{(i)^T} v$, corresponding to the nodes on $\Gamma_{loc}^{(i)}$, are zero and eliminating those entries results in $B_{\Gamma_{OL}}^{(i)^T} v$, we can rearrange the system (4.27):

$$A_{OL}^{(i)} u^{(i)} = \begin{bmatrix} A_{II}^{(i)} & A_{\Gamma I}^{(i)^T} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{bmatrix} \begin{bmatrix} u_I^{(i)} \\ u_\Gamma^{(i)} \end{bmatrix} = \begin{bmatrix} 0 \\ B_{\Gamma_{OL}}^{(i)^T} v \end{bmatrix} \tag{4.28}$$

where $u_\Gamma^{(i)}$ is the displacement on $\overline{\Gamma_{gl}} \cap \partial\Omega_i$. The equivalence of (4.27) and (4.28) shows that $\widehat{B}_\Gamma^{(i)} \widehat{S}_\Gamma^{(i)\dagger} \widehat{B}_\Gamma^{(i)^T} \widehat{Z}_\Gamma^{kT} P^k = B_{\Gamma_{OL}}^{(i)} S_{OL}^{(i)\dagger} B_{\Gamma_{OL}}^{(i)^T} \widehat{Z}_\Gamma^{kT} P^k$.

□

Due to Lemma 4.3.4, the operator that we need to study can be written as

$$P^k \widehat{Z}_\Gamma^k B_{\Gamma_{OL}} S_{OL}^\dagger B_{\Gamma_{OL}}^T \widehat{Z}_\Gamma^{k^T} P^k \cdot P^k \widehat{Z}_\Gamma^k B_{\Gamma_{OL,D}} \widetilde{S}_{OL} B_{\Gamma_{OL,D}}^T \widehat{Z}_\Gamma^{k^T} P^k.$$

The proof of the following lemma proceeds, line by line, as the proof of [43, Theorem 6.15].

Lemma 4.3.5. *For* $\forall \lambda \in \text{range}(P^k)$,

$$\begin{aligned} & \langle \lambda, \lambda \rangle \\ \leq\ & \langle P^k \widehat{Z}_\Gamma^k B_{\Gamma_{OL}} S_{OL}^\dagger B_{\Gamma_{OL}}^T \widehat{Z}_\Gamma^{k^T} P^k \cdot P^k \widehat{Z}_\Gamma^k B_{\Gamma_{OL,D}} \widetilde{S}_{OL} B_{\Gamma_{OL,D}}^T \widehat{Z}_\Gamma^{k^T} P^k \lambda, \lambda \rangle \\ \leq\ & C(1 + log(H_b/h))^2 (1 + log(H_s/h))^2 \langle \lambda, \lambda \rangle. \end{aligned}$$

We now can derive a concrete bound for (4.25), using Lemma 4.3.3 and 4.3.4. In our case, $\lambda_{max} = C(1 + log(H_b/h))^2(1 + log(H_s/h))^2, \lambda_{min} = 1, \alpha_1 = C(1 + log(H_s/h))^2$, and $\alpha_0 = 1$. Assuming H_b/h and H_s/h are large enough, we have

$$\frac{1}{2} + \sqrt{\frac{1}{2} + \frac{\lambda_{\max}}{\alpha_0}} \approx \sqrt{\frac{\lambda_{max}}{\alpha_0}}, \tag{4.29}$$

and

$$-\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\lambda_{min}}{\alpha_1}} = -\frac{1}{2} + \frac{1}{2}\sqrt{1 + 4\frac{\lambda_{min}}{\alpha_1}} = -\frac{1}{2} + \frac{1}{2}\left(1 + 2\frac{\lambda_{min}}{\alpha_1} + O\left(\left(4\frac{\lambda_{min}}{\alpha_1}\right)^2\right)\right). \tag{4.30}$$

Combining (4.25), (4.29), and (4.30), we have

$$\mathcal{K}(\mathcal{A}_1) \leq C(1 + log(H_b/h))(1 + log(H_s/h))^5. \tag{4.31}$$

We obtain

Theorem 4.3.6. *Let* $\mathcal{B}^{-1}, \mathcal{A}$, *and* $\mathcal{K}(\mathcal{B}^{-1}\mathcal{A})$ *be defined as in (4.11), (4.13), and (1.9), respectively. Then we have the following bound:*

$$\mathcal{K}(\mathcal{B}^{-1}\mathcal{A}) \leq C(1 + log(H_b/h))(1 + log(H_s/h))^5.$$

4.4 Numerical Experiments

Recall that an active set method consists of outer iterations, in which the active set is updated, and inner iterations, in which auxiliary equality constrained

Table 4.1: Results for the FETI-FETI method. cond, iter denote condition number estimates and the iteration counts, respectively. Area on which continuity is imposed between bodies: Γ, i.e., the entire interface for (I), and only a proper subset of Γ, Γ_0 for (II)

			(I)		(II)	
$1/H_b$	H_b/H_s	H_s/h	cond	iter	cond	iter
2	2	2	2.5536	7	1.9940	7
4			3.6188	12	2.8136	10
6			3.9929	13	2.8718	10
8			3.9004	13	2.7254	10
10			3.7063	13	2.6951	10
12			4.0142	13	2.7227	10
2	4	2	7.1076	10	4.9790	9
	6		12.07490	12	7.1625	10
	8		17.1343	13	7.7988	10
	10		22.2380	15	8.6543	11
	12		27.3672	14	12.2114	12
	14		32.5125	17	12.0330	12
	16		37.6688	19	15.4197	12
	18		42.8328	20	16.4836	12
2	2	4	4.5201	9	4.2654	9
		8	6.9059	10	6.3238	10
		16	9.8320	12	9.1107	12
		32	13.2897	13	12.1263	13
		64	17.2765	16	15.6644	14
		128	21.7917	18	23.3692	17

problems are solved on the current active set. In this section, we solve such auxiliary equality constrained problems using the FETI-FETI method and the hybrid method.

We solve the following minimization problem:

$$\min \sum_{i=1}^{N_b \times N_b} \left(\frac{1}{2} \int_{\Omega^i} |\nabla u^i|^2 dx - \int_{\Omega^i} f u^i dx \right), \tag{4.32}$$

where $\Omega_i \subset \mathbb{R}^2, i = 1, \cdots, N_b \times N_b$ are square bodies with side length $H_b := 1/N_b$ which form the system $\Omega = \bigcup_{i=1}^{N_b \times N_b} \Omega_i = [0,1] \times [0,1]$. We require $u^i \in$

Table 4.2: Results for the hybrid method. iter denotes the iteration counts. Area on which continuity is imposed between bodies: Γ, i.e., the entire interface for (I), and only a proper subset of Γ, Γ_0 for (II)

			(I)	(II)
$1/H_b$	H_b/H_s	H_s/h	iter	iter
2	2	2	10	10
4			12	11
6			12	11
8			11	11
10			11	11
12			11	11
2	4	2	10	10
	6		8	10
	8		8	10
	10		8	10
	12		8	9
	14		8	8
	16		7	8
	18		7	7
2	2	4	11	13
		8	13	15
		16	14	16
		32	15	17
		64	16	19
		128	17	20

$H^1(\Omega_i), u^i|_{\partial\Omega_i \cap \partial\Omega} = 0$. Each Ω_i is decomposed into $N_s \times N_s$ square subdomains, each of which is discretized by square bilinear elements of side length h. Also, $\Gamma := \cup_{i \neq j} \partial\Omega_i \cap \partial\Omega_j$ denotes the interface between the bodies.

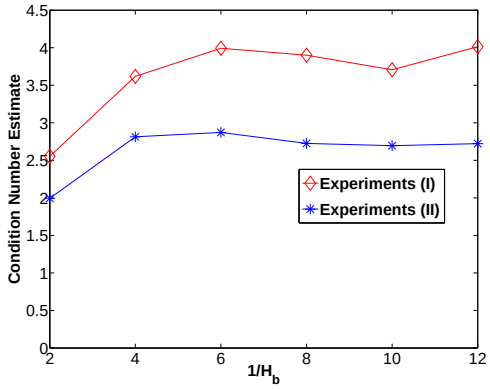
We consider two linearized problems, each with a different *contact area* between the bodies. In the first problem, the entire Γ is considered as the contact area, i.e., we require the continuity of the displacement vector across the entire Γ; see Figure 4.1. In the second problem, continuity is imposed only on the middle third of the faces between the bodies. The Krylov subspace methods of choice are the preconditioned conjugate gradient method for the FETI-FETI method and the preconditioned conjugate residual method for the hybrid method. All our experiments have been performed in MATLAB, and the stopping criterion is $||r_n||_2/||r_0||_2 < 10^{-6}$, where r_n and r_0 are the n th and initial residuals, respectively.

In Table 4.1, the results obtained with the FETI-FETI method are presented. We have three parameters; the number of bodies across Ω ($N_b = 1/H_b$), the number of subdomains across each body ($N_s = H_b/H_s$), and the number of elements across each subdomain (H_s/h). We vary one parameter while keeping the other two fixed. The results for the first set of experiments, with the entire Γ as the contact surface, are shown in Column (I); those for the second set of experiments with a reduced contact area shown in Column (II). We observe that the condition number estimates and the iteration counts are indepedent of $1/H_b$, linearly dependent on H_b/H_s, and logarithmically dependent on H_s/h. The condition numbers from Table 4.1 are also plotted in Figure 4.2.

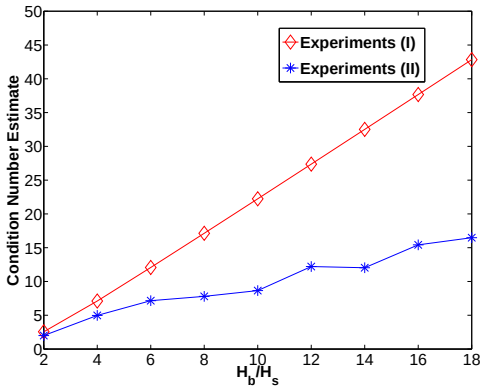
The same numerical results have been obtained independently by Klawonn and Rheinbach; see [32] and [28].

In Table 4.2, the results for the hybrid method are shown. We vary the parameters the same way we did for the FETI-FETI method. We observe that the iteration counts are independent of $1/H_b$ and logarithmically dependent on H_s/h. The iteration counts from Table 4.2 are also plotted in Figure 4.3.

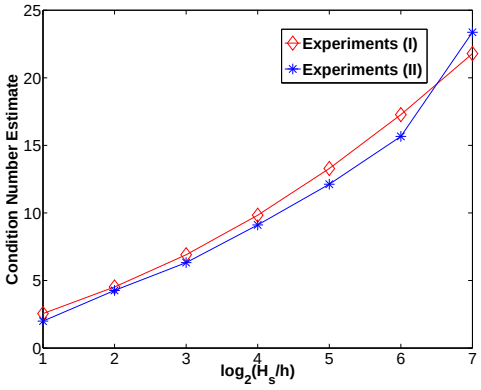
Figure 4.2: Condition number estimates for the FETI-FETI method. Area on which continuity is imposed between bodies: Γ, i.e., the entire interface for (I), and only a proper subset of Γ, Γ_0 for (II)



(a) $1/H_b$ varies

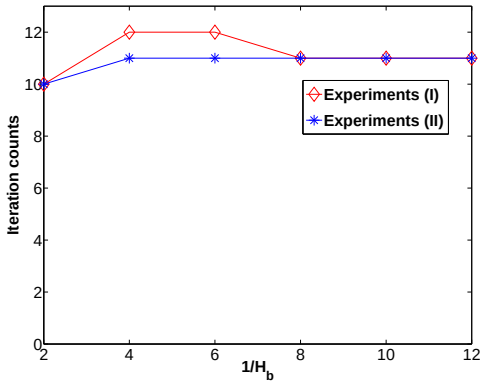


(b) H_b/H_s varies

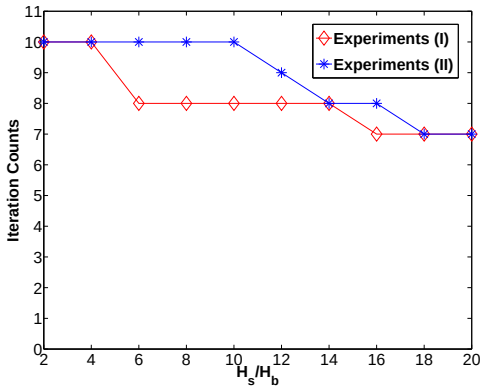


(c) H_s/h varies

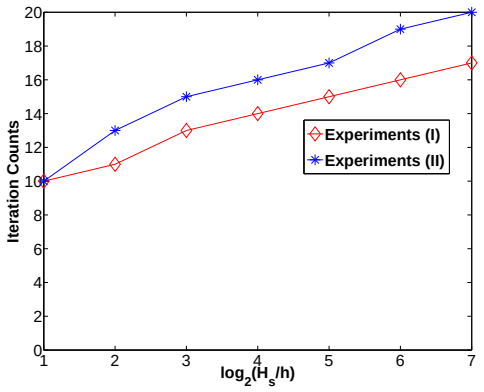
Figure 4.3: Iteration counts for the hybrid method. Area on which continuity is imposed between bodies: Γ, i.e., the entire interface for (I), and only a proper subset of Γ, Γ_0 for (II)



(a) $1/H_b$ varies



(b) H_b/H_s varies



(c) H_s/h varies

Part III

Solving a Nonlinear Contact Problem

Chapter 5

Active set method combined with the hybrid method

An active set method can often be slow due to a poor guess of the optimal active set. We discuss how to find a reasonably good initial active set. We first recall that our minimization problem (3.1) can be written as

$$\min_{u_\Gamma \in \widehat{W}_{\Gamma,c}} \frac{1}{2} u_\Gamma^T \widehat{S}_c u_\Gamma - \widehat{g}_c^T u_\Gamma, \quad \text{with} \quad \widehat{B}_{\Gamma,c} u_\Gamma \leq 0, \tag{5.1}$$

in terms of the primal variables. We can also reformulate (5.1) in terms of the dual variables:

$$\min \frac{1}{2} \lambda^T \widehat{B}_{\Gamma,c} \widehat{S}_c^\dagger \widehat{B}_{\Gamma,c}^T \lambda - \widehat{d}_c^T \lambda, \quad \text{with} \quad \lambda \geq 0, \tag{5.2}$$

where $\widehat{d}_c = \widehat{B}_{\Gamma,c}^T \widehat{S}_c^\dagger \widehat{g}_c$.

In our experiments, we solve the dual problem (5.2) approximately and use the resulting solution, denoted by $\hat{\lambda}$, to obtain the initial active set.

More precisely, we inexactly solve the unconstrained version of (5.2)

$$\min \frac{1}{2} \lambda^T \widehat{B}_{\Gamma,c} \widehat{S}_c^\dagger \widehat{B}_{\Gamma,c}^T \lambda - \widehat{d}_c^T \lambda \tag{5.3}$$

with the preconditioned conjugate gradient method, with $\widehat{B}_{D,\Gamma,c} \widehat{S}_c \widehat{B}_{D,\Gamma,c}^T$ as the preconditioner, where $\widehat{B}_{D,\Gamma,c}$ is the scaled version of $\widehat{B}_{\Gamma,c}$. We denote the resulting solution by λ^*. Subsequently, we orthogonally project λ^* onto the feasible region in the *transformed coordinate system*, which is associated with the preconditioner $\widehat{B}_{D,\Gamma,c} \widehat{S}_c \widehat{B}_{D,\Gamma,c}^T$, and denote such a projection by $\hat{\lambda}$.

In Figure 5.1, we illustrate the projection of λ^* in the original and the transformed coordinate systems. The concentric ellipses on the left of Figure 5.1 indicate the level sets of $f(\lambda) := \frac{1}{2} \lambda^T \widehat{B}_{\Gamma,c} \widehat{S}_c^\dagger \widehat{B}_{\Gamma,c}^T \lambda - \widehat{d}_c^T \lambda$, whereas the concentric circles on

the right of Figure 5.1 indicate the level sets of the transformed function

$$\bar{f}(\bar{\lambda}) := \frac{1}{2}\bar{\lambda}^T M^{-1/2}\widehat{B}_{\Gamma,c}\widehat{S}_c^{\dagger}\widehat{B}_{\Gamma,c}^T M^{-1/2}\bar{\lambda} - \widehat{d}_c^T M^{-1/2}\bar{\lambda} = \frac{1}{2}\bar{\lambda}^T\bar{\lambda} - \widehat{d}_c^T M^{-1/2}\bar{\lambda},$$

where M^{-1} is the pseudoinverse of the system matrix, i.e., $M^{-1} := (\widehat{B}_{\Gamma,c}\widehat{S}_c\widehat{B}_{\Gamma,c}^T)^{\dagger}$. The feasible region $\Omega_B := \{\lambda : \lambda \geq 0\}$ has been transformed into $\widetilde{\Omega}_B := \{\bar{\lambda} : M^{-1/2}\bar{\lambda} \geq 0\}$.

Whereas the projection of λ^* onto Ω_B in the original coordinate system does not necessarily coincide with $\tilde{\lambda}$, the minimizer of the inequality constrained problem (5.2), the projection of λ^* onto $\tilde{\Omega}_B$ in the transformed coordinate system coincides with $\tilde{\lambda}$. When $M^{-1} \neq (\widehat{B}_{\Gamma,c}\widehat{S}_c\widehat{B}_{\Gamma,c}^T)^{\dagger}$ we cannot expect this to happen, but we can expect the projection of λ^* in the transformed coordinate system to be a better approximation of $\tilde{\lambda}$ than the projection in the original coordinate system, and we will use the projection in the transformed coordinate system as our initial vector.

However, we note that we do not solve (5.3) exactly in practice, due to the presence of $\widehat{S}_c^{\dagger}$; solving an equation of the form $\widehat{S}_c x = b$ with a given right hand side b can be expensive. Instead, we replace $\widehat{S}_c^{\dagger}$ with M_{BDDC}^{-1}, and solve the resulting problem:

$$\min \frac{1}{2}\lambda^T\widehat{B}_{\Gamma,c}M_{BDDC}^{-1}\widehat{B}_{\Gamma,c}^T\lambda - \widehat{d}_c^T\lambda. \tag{5.4}$$

We still denote the solution of (5.4) by λ^* and the projection of λ^* onto the feasible region $\tilde{\Omega}_B := \{\bar{\lambda} : M^{-1/2}\bar{\lambda} \geq 0\}$ by $\hat{\lambda}$, where $M^{-1} := \widehat{B}_{D,\Gamma,c}\widehat{S}_c\widehat{B}_{D,\Gamma,c}^T$.

We recall the KKT conditions for (5.1), which are satisfied by an optimal pair (u_Γ, λ):

$$\begin{aligned} \widehat{B}_{\Gamma,c}u_\Gamma &\leq 0 \\ \lambda &\geq 0 \\ \lambda^T(\widehat{B}_{\Gamma,c}u_\Gamma) &= 0 \\ \widehat{S}_c u_\Gamma - \widehat{g}_c + \widehat{B}_{\Gamma,c}^T\lambda &= 0 \end{aligned} \tag{5.5}$$

The second and the third equations of (5.5) indicate that $\lambda_i > 0$ implies $(\widehat{B}_{\Gamma,c}u_\Gamma)_i = 0$. This motivates us to set

$$\text{initial active set} \quad = \{i : \hat{\lambda}_i > 0\}. \tag{5.6}$$

This turns out to be a good estimate of the optimal active set. Once the initial active set is chosen, we solve the corresponding saddle point problem of the form (4.2) with the preconditioned conjugate residual method. If the resulting solution is infeasible or does not satisfy the KKT conditions, we modify the active set and repeat the process. In other words, we

1. solve (5.4) with the PCG method, iterating until the norm of the residual has been reduced by a factor of 10^{-5}; denote the resulting solution by λ^*.

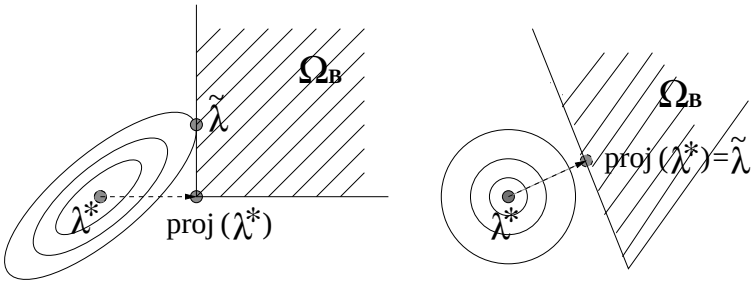
2. Project λ^* onto $\tilde{\Omega}_B$ in the transformed coordinate system as described above, to obtain the projection $\hat{\lambda}$, and the initial active set $\{i : \hat{\lambda}_i > 0\}$.

3. With $x_0^T = [u_0^T \quad \lambda_0^T] = [0 \quad \hat{\lambda}^T]$ as initial vectors, solve a saddle point problem of the form (4.2), where $\widehat{Z}_\Gamma^k$ of (4.2) is determined by the current active set. Iterate until the norm of the residual has been reduced by a factor of 10^{-5}.

4. If the u part of the resulting solution is not feasible or not optimal, modify the active set accordingly, following Figure 6.3, and repeat the process.

We have solved the nonlinear model problem (3.1) with the method described above; the results are reported in Table 5.1.

Figure 5.1: Projection of λ^* onto the feasible region in original and transformed coordinates, respectively. When preconditioner = inverse of the system matrix (as shown in right), projection of the solution of unconstrained problem, λ^*, onto the feasible region is the solution of the constrained problem, $\tilde{\lambda}$. Therefore we can expect $\text{proj}(\lambda^*) \approx \tilde{\lambda}$ with a good preconditioner.

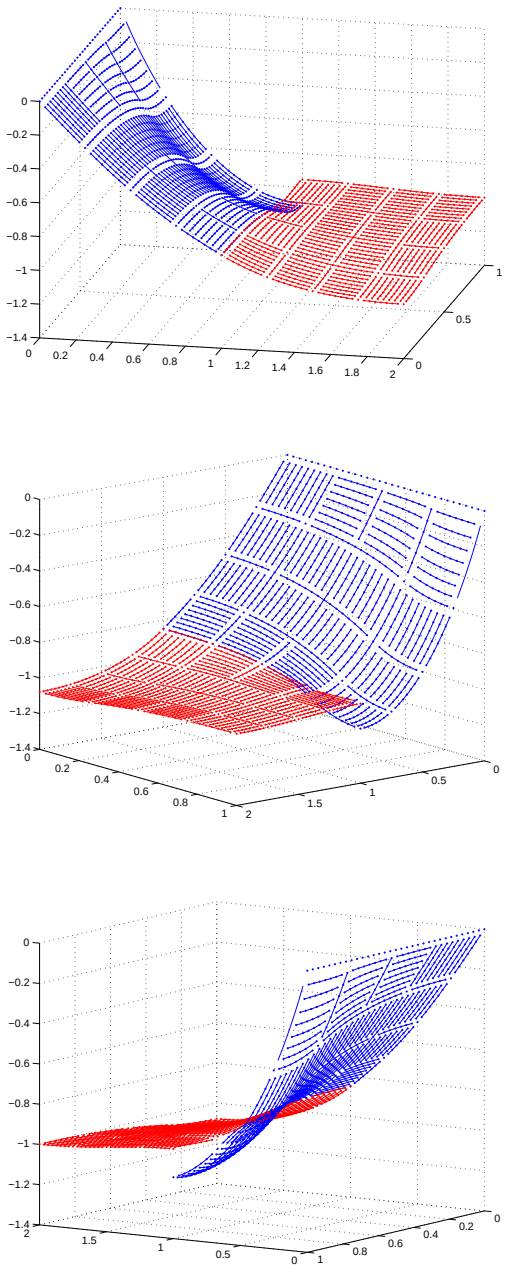


In Table 5.1, notice that the number of inner minimizations does not increase rapidly as we increase the number of elements per subdomain or the number of subdomains per body (or membrane, in the model problem), which is an indication of the scalability of the hybrid algorithm.

Table 5.1: Results: active set method + hybrid method. *PCG it.* denotes the number of PCG iterations needed to solve (5.4) until the norm of the residual has been reduced by 10^{-5}. *outer it.* denotes the number of outer iterations of the active set method; *inner it.* denotes the number of iterations needed to solve the inner minimization problems via PCR method on the active faces identified in the outer iterations. *total it.* denotes the sum of the number of inner iterations.

$N_{sub}(1/H)$	H/h	$N_{dof}(\lambda)$	$N_{dof}(total)$	PCG it.	outer it.	inner it.	total it.
16(4)	4	17	561	5	2	16 16	32
16(4)	8	33	2145	6	2	21 19	40
16(4)	12	49	4753	6	1	25 21 21	67
16(4)	16	65	8385	7	4	29 24 22 22	97
16(4)	20	81	13041	7	4	27 25 23 22	97
64(8)	4	33	2145	7	2	19 19	38
64(8)	8	65	8385	8	1	25	25
64(8)	12	97	18721	8	1	28	28
64(8)	16	129	33153	9	1	31	31
64(8)	20	161	51681	9	2	32 28	60
144(12)	4	49	4753	8	1	21	21
144(12)	8	97	18721	10	2	26 24	50
144(12)	12	145	41905	11	3	30 26 27	83
144(12)	16	193	74305	11	4	32 30 28 28	118
144(12)	20	241	115921	11	4	35 31 29 29	124
256(16)	4	65	8385	9	1	21	21
256(16)	8	129	33153	11	1	29	29
256(16)	12	193	74305	13	2	33 27	60
256(16)	16	257	131841	13	2	35 30	65
256(16)	20	321	205761	14	3	39 32 32	103
400(20)	4	81	13041	10	2	23 22	45
400(20)	8	161	51681	12	3	29 27 25	81
400(20)	12	241	115921	14	3	33 32 30	95
400(20)	16	321	205761	14	4	38 34 31 31	134
400(20)	20	401	321201	15	5	39 38 32 32 32	173

Figure 5.2: Solution of the model problem, from different angles. $N_{sub} = 16, H/h = 8$.



Chapter 6

SMALBE Algorithm for Bound and Equality Constraints

By the duality theorem of [14, Chapter 2], the nonlinear model problem (3.1) can be reformulated in terms of Lagrange multipliers as a bound and equality constrained problem; we will give this reformulation in Section 6.4. We first describe the SMALBE (Semi-Monotonic Augmented Lagrangians for Bound and Equality constraints) algorithm for convex quadratic problems with bound and equality constraints, developed by Dostal [13],[14, Chapter 6]. In Section 6.4, we will use this algorithm to solve the problem (3.1).

We wish to solve a bound and equality constrained problem of the following form:

$$\min_{u \in \Omega_{BE}} f(x), \tag{6.1}$$

where $f(x) = \frac{1}{2}x^T Ax - b^T x, A \in \mathbb{R}^{n\times n}$ is a positive definite, symmetric matrix, $b, l \in \mathbb{R}^n, \Omega_{BE} := \{x \in \mathbb{R}^n : Bx = c \quad \text{and} \quad x \geq l\}, B \in \mathbb{R}^{m\times n}$, and $c \in \text{Im}B$. To allow the possibility that not all components of x are bound constrained, we admit $l_i = -\infty$. In the SMALBE algorithm, the equality constraints and the bound constraints are treated separately. In particular, the SMALBE algorithm has features of the SMALE (Semi-Monotonic Augmented Lagrangians for Equality constraints) algorithm and the MPRGP (Modified Proportioning and Reduced Gradient Projection) algorithm, which are methods for convex quadratic problems with equality constraints and bound constraints, respectively. We first review these two methods.

6.1 SMALE Algorithm

The SMALE algorithm is for convex quadratic problems with equality constraints. Suppose we want to solve

$$\min_{u \in \Omega_E} f(x),$$

where $\frac{1}{2}x^T Ax - b^T x, A \in \mathbb{R}^{n \times n}$ is a positive definite, symmetric matrix, $b \in \mathbb{R}^n, \Omega_E := \{x \in \mathbb{R}^n : Bx = c\}, B \in \mathbb{R}^{m \times n}$, and $c \in \mathrm{Im}B$.

We introduce the *augmented Lagrangian* penalty function $L : \mathbb{R}^{n+m+1} \to \mathbb{R}$ which is defined by

$$L(x, \lambda, \rho) = f(x) + (Bx - c)^T \lambda + \frac{\rho}{2}||Bx - c||^2. \tag{6.2}$$

and g, its gradient with respect to x, by

$$g(x, \lambda, \rho) := \nabla_x L(x, \lambda, \rho) = Ax - b + B^T(\lambda + \rho(Bx - c)). \tag{6.3}$$

The SMALE algorithm, described in Figure 6.1, can be viewed as an inexact augmented Lagrangian method with adaptive precision control. In Step 2, we can use any convergent algorithm for minimizing strictly convex quadratic functions, such as the conjugate gradient method. The SMALE algorithm has both outer (update of λ and ρ) and inner (finding x, given λ and ρ) loops, and the number of iterations required for the convergence is bounded for both in terms of a few parameters, e.g., $\lambda_{min}(A)$, the smallest eigenvalue of the Hessian of the cost function.

Let $\mathcal{T}$ denote any set of indices and assume that for any $t \in \mathcal{T}$ we have a minimization problem

$$\text{minimize} \quad f_t(x) \quad \text{s.t.} \quad x \in \Omega_t \tag{6.4}$$

where $\Omega_t = \{x \in \mathbb{R}^{n_t} : B_t x = 0\}, f_t(x) = \frac{1}{2}x^T A_t x - b_t^T x$, with $A_t \in \mathbb{R}^{n_t \times n_t}$ positive definite and symmetric, $B_t \in \mathbb{R}^{m_t \times n_t}$, and $b_t, x \in \mathbb{R}^{n_t}$. Then we have the following

1. Initialize: choose $\eta > 0, \beta > 1, M > 0, \rho_0 > 0, \lambda^0 \in \mathbb{R}^m, k = 0$

2. Iterate $k = 0, 1, 2, \cdots$, while $||g(x^k, \lambda^k, \rho_k)|| > M\epsilon||b||$ or $||Bx^k|| > \epsilon||b||$, where $\epsilon > 0$ is a given tolerance (inner iteration with adaptive precision control):
   find x^k such that

$$||g(x^k, \lambda^k, \rho_k)|| \leq \min\{M||Bx^k - c||, \eta\}.$$

3. Update the Lagrange multipliers:

$$\lambda^{k+1} = \lambda^k + \rho_k(Bx^k - c).$$

4. Update ρ provided the increase of the Lagrangian is not sufficient:
   if $k > 0$ and
   $$L(x^k, \lambda^k, \rho_k) < L(x^{k-1}, \lambda^{k-1}, \rho_{k-1}) + \frac{\rho_k}{2}||Bx^k - c||^2$$
   $\rho_{k+1} = \beta\rho_k$
   else
   $\rho_{k+1} = \rho_k$

Figure 6.1: SMALE algorithm

result, [14, Theorem 4.21], which shows that the number of outer iterations required for the convergence of the algorithm to a certain accuracy is bounded in terms of $\lambda_{min}(A_t)$.

Theorem 6.1.1. *Let $\{x_t^k\}, \{\lambda_t^k\}$ and $\{\rho_{t,k}\}$ be generated by the SMALE algorithm of Figure 6.1 for (6.4) with $||b_t|| \geq \eta_t > 0, \beta > 1, M > 0, \rho_{t,0} = \rho_0 > 0, \lambda_t^0 = 0$. Let $0 < a_{min}$ be a given constant. Finally, let the class of problems (6.4) satisfy*

$$a_{min} \leq \lambda_{min}(A_t),$$

where $\lambda_{min}(A_t)$ denotes the smallest eigenvalue of A_t, and denote

$$a = (2 + s)/(a_{min}\rho_0)$$

where $s \geq 0$ is the smallest integer such that $\beta^s \rho_0 \geq M^2/a_{min}$. Then for each $\epsilon > 0$ there are indices $k_t, t \in \mathcal{T}$, such that

$$k_t \leq a/\epsilon^2 + 1$$

and $x_t^{k_t}$ is an approximate solution of (6.4) satisfying

$$||g_t(x_t^{k_t}, \lambda_t^{k_t}, \rho_{t,k_t})|| \leq M\epsilon||b_t|| \quad \textit{and} \quad ||B_t x_t^{k_t}|| \leq \epsilon||b_t||. \tag{6.5}$$

We have another result, which shows that the number of inner iterations needed to find x^k which satisfies our requirements is bounded:

Theorem 6.1.2. *Let $\{x_t^k\}, \{\lambda_t^k\}$ and $\{\rho_{t,k}\}$ be generated by the SMALE algorithm of Figure 6.1 for (6.4) with $||b_t|| \geq \eta_t > 0, \beta > 1, M > 0, \rho_{t,0} = \rho_0 > 0, \lambda_t^0 = 0$. Let*

$0 < a_{min} < a_{max}$ *and* $0 < B_{max}$ *be given constants. Let Step 2 be implemented by the conjugate gradient method which generates the iterates* $x_t^{k,0}, x_t^{k,1}, \cdots, x_t^{k,l} = x_t^k$ *starting from* $x_t^{k,0} = x_t^{k-1}$ *with* $x_t^{-1} = 0$*, where* $l = l(k,t)$ *is the first index satisfying either*

$$||g(x_t^{k,l}, \lambda_t^k, \rho_k)|| \leq M||B_t x_t^{k,l}||$$

or

$$||g(x_t^{k,l}, \lambda_t^k, \rho_k)|| \leq \epsilon M||b_t||.$$

Finally, let the class of problems (6.4) satisfy

$$a_{min} \leq \lambda_{min}(A_t) \leq \lambda_{max}(A_t) = ||A_t|| \leq a_{max} \quad \textit{and} \quad ||B_t|| \leq B_{max}.$$

Then the Algorithm generates an approximate solution $x_t^{k_t}$ *of any problem (6.4) which satisfies (6.5) at* $O(1)$ *matrix-vector multiplications by the Hessian of the augmented Lagrangian* L_t *for (6.4).*

6.2 MPRGP Algorithm

The MPRGP algorithm is used for convex quadratic problems with bound constraints. Suppose we want to solve

$$\min_{u \in \Omega_B} f(x) \tag{6.6}$$

where $\frac{1}{2}x^T Ax - b^T x, A \in \mathbb{R}^{n \times n}$ is a positive definite, symmetric matrix, $b, l \in \mathbb{R}^n, \Omega_B := \{x \in \mathbb{R}^n : x \geq l\}$. Again, we allow $l_i = -\infty$. By the duality theorem of

[14, Chapter 2], we know that an optimal KKT pair $(\hat{x}, \hat{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfies

$$\begin{aligned} -\hat{x} + l &\leq 0 \\ \hat{\lambda} &\geq 0 \\ \hat{\lambda}^T(-\hat{x} + l) &= 0 \\ L_x(\hat{x}, \hat{\lambda}) &= 0, \end{aligned}$$

where

$$L(x, \lambda) := \frac{1}{2} x^T A x - b^T x + \lambda^T(-x + l).$$

These conditions are equivalent to

$$A\hat{x} - b \geq 0 \quad \text{and} \quad (A\hat{x} - b)^T(\hat{x} - l) = 0, \tag{6.7}$$

or, componentwise,

$$\hat{x}_i = l_i \Rightarrow \hat{g}_i \geq 0 \quad \text{and} \quad \hat{x}_i > l_i \Rightarrow \hat{g}_i = 0, \quad i = 1, \cdots, n, \tag{6.8}$$

where $\hat{g}_i = (A\hat{x} - b)_i$. The KKT conditions (6.8) determine some important subsets of $\mathcal{N} = \{1, 2, \cdots, n\}$, the set of all indices. We define an *active set* of x as the set of all indices for which $x_i = l_i$;

$$\mathcal{A}(x) := \{i \in \mathcal{N} : x_i = l_i\},$$

and a *free set* of x, as the complement of the active set:

$$\mathcal{F}(x) := \{i \in \mathcal{N} : x_i \neq l_i\}.$$

We also introduce two subsets of $\mathcal{A}(x)$,

$$\mathcal{B}(x) = \{i \in \mathcal{N} : x_i = l_i, g_i > 0\}, \quad \mathcal{B}_0(x) = \{i \in \mathcal{N} : x_i = l_i, g_i \geq 0\},$$

which are called a *binding set* and a *weakly binding set*, respectively. We decompose the part of the gradient $g(x) = Ax - b$ which violates the KKT conditions into the *free gradient* ϕ and the *chopped gradient* β, which are defined by

$$\begin{aligned} \phi_i(x) = g_i(x) \quad \text{for} \quad i \in \mathcal{F}(x), \qquad \phi_i(x) = 0 \quad \text{for} \quad i \in \mathcal{A}(x), \\ \beta_i(x) = 0 \quad \text{for} \quad i \in \mathcal{F}(x), \qquad \beta_i(x) = g_i^-(x) \quad \text{for} \quad i \in \mathcal{A}(x), \end{aligned} \tag{6.9}$$

where $g_i^- := \min\{g_i, 0\}$. Introducing the *projected gradient* $g^P(x) := \phi(x) + \beta(x)$, we can rewrite the KKT condition as

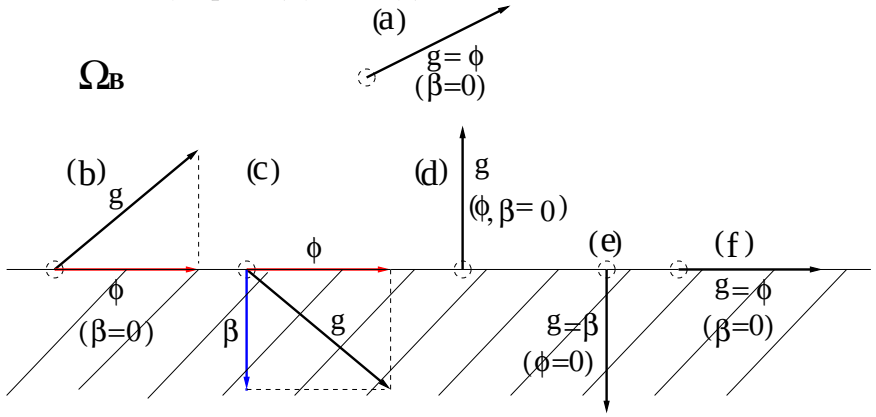
$$g^P(x) = 0.$$

Note that ϕ and β are orthogonal to each other and $-\phi$ and $-\beta$ are feasible descent directions for f; see Figure 6.2.

Among the methods for convex quadratic problems with bound constraints such as (6.6) are the active set method, Polyak's algorithm, and the gradient projection method. We will briefly describe these methods, since the MPRGP algorithm has features borrowed from these methods.

In an active set method, also known as a working set method, we solve a sequence of auxiliary equality constrained problems defined by a subset of the set $\mathcal{N}$. This task would have been very simple if we knew $\mathcal{A}(\hat{x})$ a priori, but since this is usually not the case, we start out by guessing which inequalities would be active

Figure 6.2: Gradient splitting: the projected gradient g^P is decomposed into ϕ and β. (a): the iterate is strictly inside the feasible region, so $\phi = g, \beta = 0$. (b),(c),(d),(e): the iterate lies on a face, and ϕ is always defined as the tangential component of g to the face. The normal component of g to the face, if its negative is a feasible direction, equals β; if not, $\beta = 0$.



in the solution $\hat{x}$.

Let $I \subset \mathcal{N}$ denote the set of indices of the bounds l_i which are predicted to be active in the solution, and let

$$\mathcal{W}_I = \{y \in \mathbb{R}^n : y_i = l_i, i \in I\}.$$

In this dissertation, we will call the predicted set of active bounds I and $\mathcal{W}_I$ an *active set* and an *active face*, respectively. The complete active set method is described in Figure 6.3.

Polyak's algorithm, described in Figure 6.4, differs from the active set method mainly in that we do not wait until an auxiliary, equality-constrained problem is solved to test the feasibility of the intermediate solution. We perform conjugate gradient iterations on $\mathcal{A}(x^k)$, the active set of $x^k, k = 0, 1, 2, \cdots$, if $||\phi(x^k)|| > 0$, i.e., if it is *worthwhile* to stay on the current active set; $||\phi(x^k)|| = 0$ indicates that

1. Initialize: choose $x^0 \in \Omega_B$, set $I^0 = \mathcal{B}_0(x^0), k = 0$

2. Iterate $k = 0, 1, 2, \cdots$ while $||g^P(x^k)|| > 0$
   Minimize in face $\mathcal{W}_{I^k}$:

$$\hat{y} = \arg \min_{y \in \mathcal{W}_{I^k}} f(y)$$

   if $\hat{y} \in \Omega_B$,
   set $x^{k+1} = \hat{y}, I^{k+1} = \mathcal{B}_0(x^{k+1})$.
   else
   set x^{k+1} as the cross point of the line segment $\overline{x^k \hat{y}}$ and Ω_B, and $I^{k+1} = \mathcal{A}(x^{k+1})$.

Figure 6.3: Active set method

we have reached the optimal point on the active set and we either have reached the solution or should leave the face. If the conjugate gradient iterate, say y, is feasible, we accept it as the next iterate, i.e., set $x^{k+1} = y$, and if not, we take the cross point of the line segment $\overline{x^k \hat{y}}$ and Ω_B as x^{k+1}. If $||\phi(x^k)|| = 0$, we take $-\beta(x^k)$ as the search direction and take $x^{k+1} := x^k - \alpha_{cg}\beta(x^k)$, where α_{cg} is the minimizer of $f(x^k - \alpha\beta(x^k))$.

In the gradient projection method, described in Figure 6.5, the next iterate is always defined by $x^{k+1} = P_{\Omega_B}(x^k - \bar{\alpha}g(x^k))$, where P_{Ω_B} denotes the projection operator onto Ω_B, $\bar{\alpha}$ is a fixed steplength, determined by the spectrum of the matrix A, the Hessian of the cost function.

Polyak's algorithm and the gradient projection method have different benefits. A nice feature of the gradient projection method is that we have a convergence rate bounded in terms of the spectrum of the matrix A, whereas Polyak's algorithm does

1. Initialize: choose $x^0 \in \Omega_B$, set $g = Ax^0 - b, p = g^P(x^0), k = 0$

2. Iterate $k = 0, 1, 2, \cdots$, while $||g^P(x^k)|| > 0$
   if $||\phi(x^k)|| > 0$
   $\alpha_{cg} = g^T p / p^T A p, y = x^k - \alpha_{cg} p$
   $\alpha_f = \max\{\alpha : x^k - \alpha p \in \Omega_B\} = \min\{(x_i^k - l_i)/p_i : p_i > 0\}$
   if $\alpha_{cg} \leq \alpha_f$ (conjugate gradient step)
   $x^{k+1} = y, g = g - \alpha_{cg} A p,$
   $\beta = \phi(y)^T A p / p^T A p, p = \phi(y) - \beta p$
   else (expansion step)
   $x^{k+1} = x^k - \alpha_f p, g = g - \alpha_f A p, p = \phi(x^{k+1})$
   end
   else
   $d = \beta(x^k), \alpha_{cg} = g^T d / d^T A d,$
   $x^{k+1} = x^k - \alpha_{cg} d, g = g - \alpha_{cg} A d, p = \phi(x^{k+1})$
   end

Figure 6.4: Polyak's algorithm

Given a positive definite, symmetric matrix $A \in \mathbb{R}^{n \times n}$ and $b, l \in \mathbb{R}^n$

1. Initialize: choose $x^0 \in \Omega_B, \bar{\alpha} \in (0, 2||A||^{-1}), k = 0$

2. Iterate $k = 0, 1, 2, \cdots$, while $||g^P(x^k)||$ is not small
   Gradient projection step:

$$x^{k+1} = P_{\Omega_B}(x^k - \bar{\alpha} g(x^k))$$

Figure 6.5: Gradient projection algorithm

not come with such an upper bound. On the other hand, the gradient projection method does not use the concept of active sets so it is prone to switch faces many times before it gets close to the solution.

The MPRGP algorithm combines the good features of Polyak's algorithm and the gradient projection method. Some notable differences between the MPRGP algorithm and Polyak's algorithm are that the magnitude of the free gradient $\phi(x^k)$ and that of the chopped gradient $\beta(x^k)$ are compared in each iteration, and if $||\phi(x^k)||$ is considerably larger than $\beta(x^k)$, our search for the next iterate stays on the current active face determined by $\mathcal{A}(x^k)$. If not, we leave the face and take $-\beta(x^k)$ as the search direction. The MPRGP algorithm is described in Figure 6.6. The complete description of the MPRGP algorithm and the criterion for deciding whether to leave the face or not requires a few more definitions, and we leave them out to keep the exposition simple and refer the reader to [14, Chapter 5]. The following theorem gives us a convergence rate for the MPRGP algorithm; see [13, Theorem 5.9].

Theorem 6.2.1. *Let $\Gamma > 0$ be a given constant, let λ_{min} denote the smallest eigenvalue of A, and let $\{x^k\}$ denote the sequence generated by the MPRGP algorithm of Figure 6.6 with $\bar{\alpha} \in (0, ||A||^{-1}]$. Then*

$$f(x^{k+1}) - f(\hat{x}) \leq \eta_\Gamma (f(x^k) - f(\hat{x})),$$

where $\hat{x}$ denotes the unique solution of (6.6) and

$$\eta_\Gamma = 1 - \frac{\bar{\alpha}\lambda_{min}}{2 + 2\hat{\Gamma}^2} \tag{6.10}$$

Choose $\bar{\alpha} \in (0, 2||A||^{-1})$

1. Initialize: choose $x^0 \in \Omega_B$, set $g = Ax^0 - b, p = \phi(x^0), k = 0$
2. Iterate $k = 0, 1, 2, \cdots$, while $g^P(x^k) > 0$
   if $\phi(x^k)$ *dominates*

$$\begin{aligned} \alpha_{cg} &= g^T p / p^T A p, y = x^k - \alpha_{cg} p \\ \alpha_f &= \max\{\alpha : x^k - \alpha p \in \Omega_B\} = \min\{(x_i^k - l_i)/p_i : p_i > 0\} \end{aligned}$$

   if $\alpha_{cg} \leq \alpha_f$ (conjugate gradient step)
   $x^{k+1} = y, g = g - \alpha_{cg} A p,$
   $\beta = \phi(y)^T A p / p^T A p, p = \phi(y) - \beta p$
   else (expansion step)
   $x^{k+1} = P_{\Omega_B}(x^k - \bar{\alpha}\phi(x^k))$
   $g = Ax^{k+1} - b, p = \phi(x^{k+1})$
   end
else (proportioning step)
   $d = \beta(x^k), \alpha_{cg} = g^T d / d^T A d$
   $x^{k+1} = x^k - \alpha_{cg} d, g = g - \alpha_{cg} A d, p = \phi(x^{k+1})$
end

Figure 6.6: MPRGP algorithm

with $\hat{\Gamma} = max\{\Gamma, \Gamma^{-1}\}$. *The error in the* $A-$ *norm is bounded by*

$$||x^k - \hat{x}||_A^2 \leq 2\eta_\Gamma^k ||x^0 - \hat{x}||_A^2 \leq 2\left(1 - \frac{1}{4\mathcal{K}(A)}\right)^k ||x^0 - \hat{x}||_A^2. \tag{6.11}$$

We note that for a typical *unconstrained* convex quadratic problem, the conjugate gradient method has the following relative error bound:

$$||x^k - \hat{x}||_A^2 \leq 2\left(\frac{\sqrt{\mathcal{K}(A)} - 1}{\sqrt{\mathcal{K}(A)} + 1}\right)^k ||x^0 - \hat{x}||_A^2, \tag{6.12}$$

see Section 1.3.1. We also note that

$$1 - \frac{1}{4\mathcal{K}(A)} > \frac{\sqrt{\mathcal{K}(A)} - 1}{\sqrt{\mathcal{K}(A)} + 1}$$

and thus the MPRGP algorithm in general converges more slowly than the CG method for the same cost function; this is because the MPRGP algorithm has the expansion and the proportioning steps in addition to the CG steps (see Figure 6.6), which are not as effective as the CG steps at decreasing the cost function.

6.3 SMALBE Algorithm

The SMALBE algorithm is for convex quadratic problems with bound and equality constraints, such as (6.1). It is a modification of the SMALE algorithm; the only difference is in Step 2. Not surprisingly, we use the same functions L and $g = \nabla L_x$ as defined in (6.2) and (6.3) to describe the algorithm. From the KKT conditions, a feasible vector $x \in \Omega_{BE}$ is a solution of (6.1) if and only if

$$g \geq 0 \quad \text{and} \quad g^T(x - l) = 0,$$

or equivalently

$$g^P = 0,$$

where g^P is the projected gradient of g, as defined in Section 6.2. The SMALBE algorithm for (6.1) is described in Figure 6.7. We have results concerning the upper bound for the number of iterations required for the convergence of the outer and the inner iterations, which are analogous to those of Section 6.1; see [14, Chapter 6].

Given a symmetric positive definite matrix $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{m\times n}$, n-vectors b, l.

1. Initialize: choose $\eta > 0, \beta > 1, M > 0, \rho_0 > 0, \lambda^0 \in \mathbb{R}^m, k = 0$

2. Iterate $k = 0, 1, 2, \cdots$,
   Find $x^k \geq l$ such that
   $$||g^P(x^k, \lambda^k, \rho_k)|| \leq \min\{M||Bx^k||, \eta\}$$

3. Update the Lagrange multipliers:
   $$\lambda^{k+1} = \lambda^k + \rho_k B x^k$$

4. Update ρ provided the increase of the Lagrangian is not sufficient:
   if $k > 0$ and
   $L(x^k, \lambda^k, \rho_k) < L(x^{k-1}, \lambda^{k-1}, \rho_{k-1}) + \frac{\rho_k}{2}||Bx^k - c||^2$ then
   $\rho_{k+1} = \beta\rho_k$
   else
   $\rho_{k+1} = \rho_k$
   end

Figure 6.7: SMALBE algorithm

6.4 Numerical Experiments with SMALBE

In this section, we reformulate the problem (3.1) in terms of the Lagrange multipliers as a bound and equality constrained problem and solve it with the SMALBE algorithm described in the previous section. Following [16], we decompose Ω_1 and Ω_2 into subdomains in the style of FETI-DP methods, which is the approach taken in the FETI-FETI method. Eliminating the interior unknowns of each subdomain

as usual and using the notation of Chapter 3, we obtain

$$\min_{u_\Gamma \in \widetilde{W}_{\Gamma,c}} \frac{1}{2} u_\Gamma^T \widetilde{S}_c u_\Gamma - \widetilde{g}_c^T u_\Gamma, \quad \text{with} \quad B_I u_\Gamma \leq 0 \quad \text{and} \quad B_E u_\Gamma = 0, \tag{6.13}$$

where

$$\widetilde{S}_c = \begin{bmatrix} \widetilde{S}_\Gamma^{(1)} & \\ & \widetilde{S}_\Gamma^{(2)} \end{bmatrix}, u_\Gamma = \begin{bmatrix} u_\Gamma^{(1)} \\ u_\Gamma^{(2)} \end{bmatrix}, \quad u_\Gamma^{(i)} \in \widetilde{W}_\Gamma^{(i)}, i = 1, 2,$$

and

$$\widetilde{B}_{\Gamma,c} = \begin{bmatrix} B_E \\ B_I \end{bmatrix},$$

such that $B_E u_\Gamma = 0$ enforces the continuity condition between the subdomains of the same body and $B_I u_\Gamma \leq 0$ enforces the *nonpenetration* condition between Ω_1 and Ω_2.

The primal problem (6.13) is equivalent to the following dual problem:

$$\min \ \theta(\lambda) \quad \text{subject to} \quad \lambda_I \geq 0 \quad \text{and} \quad R^T(\tilde{g}_c - \widetilde{B}_{\Gamma,c}^T \lambda) = 0, \tag{6.14}$$

where

$$range(R) = ker(\widetilde{S}_c),$$

$$\theta(\lambda) = \frac{1}{2} \lambda^T F \lambda - \lambda^T \tilde{d}$$

with

$$F := \widetilde{B}_{\Gamma,c} \widetilde{S}_c^\dagger \widetilde{B}_{\Gamma,c}^T, \quad \widetilde{d} = \widetilde{B}_{\Gamma,c} \widetilde{S}_c^\dagger \widetilde{g}_c.$$

We modify the dual problem (6.14) further to obtain a version more suitable for

computations. We adopt the notation

$$\widetilde{G} = R^T \widetilde{B}_{\Gamma,c}^T, \quad \widetilde{e} = R^T \widetilde{g}_c$$

and let T denote a matrix which defines the orthonormalization of the rows of $\widetilde{G}$, so that the matrix

$$G = T\widetilde{G}$$

has orthonormal rows. With $e = T\widetilde{e}$, we can rewrite (6.14) as

$$\min \frac{1}{2}\lambda^T F\lambda - \lambda^T \widetilde{d} \quad \text{s.t.} \quad \lambda_I \geq 0 \quad \text{and} \quad G\lambda = e. \tag{6.15}$$

The following lemma, taken from [15], allows us to reformulate the problem (6.15) in a vector space, not an affine space:

Lemma 6.4.1. *Let $\widetilde{B}_{\Gamma,c}$ be such that the negative entries of B_I are in the columns that correspond to the nodes in the floating body Ω_2. Then there is $\widetilde{\lambda}$ such that $\widetilde{\lambda}_I \geq 0$ and $G\widetilde{\lambda} = e$.*

Using Lemma 6.4.1, (6.15) can be rewritten as

$$\min \frac{1}{2}\lambda^T F\lambda - \lambda^T d \quad \text{s.t.} \quad \lambda_I \geq -\widetilde{\lambda}_I \quad \text{and} \quad G\lambda = 0, \tag{6.16}$$

where $d = \widetilde{d} - F\widetilde{\lambda}$. Finally, we observe that (6.16) is equivalent to the following problem:

$$\min \frac{1}{2}\lambda^T (PFP + \bar{\rho}Q)\lambda - \lambda^T P\widetilde{d} \quad \text{s.t.} \quad \lambda_I \geq 0 \quad \text{and} \quad G\lambda = e, \tag{6.17}$$

where $\bar{\rho}$ is an arbitrary positive constant and

$$Q = G^T G \quad \text{and} \quad P = I - Q.$$

We now use the SMALBE algorithm to solve the problem (6.17), with the MPRGP algorithm to solve auxiliary bound constrained problems; the results are presented in Table 6.1.

Ω_1, Ω_2 are decomposed into $N \times N$ square subdomains, each of which is discretized by conforming piecewise quadratic elements with square elements of side-length h. The computations were performed with the parameters

$$M = 1, \quad \rho_0 = 30, \quad \Gamma = 1, \quad \text{and} \quad \epsilon = 10^{-5},$$

and the stopping criterion for the outer iterations was

$$||g^P(\lambda^k)|| \leq \epsilon ||b|| \quad \text{and} \quad ||B\lambda^k|| \leq \epsilon ||b||.$$

We record the number of outer iterations, the number of inner iterations for each outer iteration, and the total number of inner iterations. We note that in each SMALBE iteration, the first outer iteration requires the largest number of inner iterations. This is because in the beginning the iterate is normally on a wrong active face and we end up taking many expansion steps and proportioning steps, which slow down the convergence of the MPRGP algorithm; see the discussion at the end of Section 6.2. After a while, however, the iterate settles down on the right face and we mainly take CG steps, which explains smaller number of inner iterations.

Table 6.1: Results: SMALBE

$N_{sub}(1/H)$	H/h	$N_{dof}(\lambda)$	$N_{dof}(total)$	outer it.	max(inner it.)	total it.
16(4)	4	161	561	4	24 6 5 6	41
16(4)	8	369	2145	5	59 6 8 7 8	88
16(4)	12	577	4753	5	38 9 8 11 8	74
16(4)	16	785	8385	4	50 13 12 13	88
16(4)	20	993	13041	6	47 13 12 9 12 8	101
64(8)	4	705	2145	5	32 6 6 6 6	56
64(8)	8	1633	8385	6	49 9 8 9 7 7	89
64(8)	12	2561	18721	7	55 9 13 7 9 9 9	111
64(8)	16	3489	33153	8	84 11 13 11 11 9 9 9	157
64(8)	20	4417	51681	8	39 13 12 13 10 12 12 10	121
144(12)	4	1633	4753	6	50 10 7 7 7 5	86
144(12)	8	3793	18721	7	35 9 11 9 8 8 7	87
144(12)	12	5953	41905	8	39 13 12 7 8 8 8 8	103
144(12)	16	8113	74305	10	40 11 21 12 10 9 8 8 9 8	136
144(12)	20	10273	115921	11	44 8 21 14 12 12 10 8 13 8 8	158
256(16)	4	2945	8385	6	49 12 6 10 7 7	91
256(16)	8	6849	33153	8	32 13 10 11 12 8 8 8	102
256(16)	12	10753	74305	10	34 19 17 10 12 10 10 8 9 8	137
256(16)	16	14657	131841	11	12 17 12 13 8 9 11 8 8	160
256(16)	20	18561	205761	12	50 15 21 13 20 10 11 10 12 12 8 10	192
400(20)	4	4641	13041	7	40 11 10 10 6 8 6	91
400(20)	8	10801	51681	9	41 16 11 11 10 8 9 6 9	121
400(20)	12	16961	115921	11	34 8 15 15 11 11 10 9 6 9 9	137
400(20)	16	23121	205761	13	51 13 15 10 15 13 13 10 11 9 6 9 9	184
400(20)	20	29281	321201	14	230 10 16 18 15 22 11 11 10 13 9 11 9 11	396

We also compare the total iteration counts of Table 5.1 and Table 6.1; note that for certain combinations of $N_{sub}(1/H)$ and H/h the SMALBE algorithm requires twice as many iterations as the combination of an active set method and the hybrid method.

Bibliography

[1] Philip Avery, Gert Rebel, Michel Lesoinne, and Charbel Farhat. A numerically scalable dual-primal substructuring method for the solution of contact problems–part i: the frictionless case. *Comput. Methods Appl. Mech. Engrg*, 193(23-26):2403–2426, 2004.

[2] O. V. Besov. On the compactness of embeddings of weighted Sobolev spaces on a domain with an irregular boundary. *Dokl. Akad. Nauk*, 376(6):727–732, 2001.

[3] J. H. Bramble, J. E. Pasciak, and A. H. Schatz. The construction of preconditioners for elliptic problems by substructuring. I. *Math. Comp.*, 47(175):103–134, 1986.

[4] James H. Bramble and Jinchao Xu. Some estimates for a weighted L^2 projection. *Math. Comp.*, 56(194):463–476, 1991.

[5] S. C. Brenner and L.-Y. Sung. Discrete Sobolev and Poincaré inequalities via Fourier series. *East-West J. Numer. Math.*, 8(2):83–92, 2000.

[6] Susanne C. Brenner. A new look at FETI. In *Domain decomposition methods in science and engineering (Lyon, 2000)*, Theory Eng. Appl. Comput.

Methods, pages 41–51. Internat. Center Numer. Methods Eng. (CIMNE), Barcelona, 2002.

[7] Susanne C. Brenner. An additive Schwarz preconditioner for the FETI method. *Numer. Math.*, 94(1):1–31, 2003.

[8] Susanne C. Brenner and L. Ridgway Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[9] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 8(R-2):129–151, 1974.

[10] Mario A. Casarin. Schwarz preconditioners for spectral and mortar finite element methods with applications to incompressible fluids. *PhD thesis, Courant Institute of Mathematical Sciences*, (TR-717), March 1996.

[11] Clark R. Dohrmann. A preconditioner for substructuring based on constrained energy minimization. *SIAM J. Sci. Comput.*, 25(1):246–258, 2003.

[12] Clark R. Dohrmann, Axel Klawonn, and Olof B. Widlund. Domain decomposition for less regular subdomains: overlapping Schwarz in two dimensions. *SIAM J. Numer. Anal.*, 46(4):2153–2168, 2008.

[13] Z. Dostál. An optimal algorithm for bound and equality constrained quadratic programming problems with bounded spectrum. *Computing*, 78(4):311–328, 2006.

[14] Zdeněk Dostál. *Optimal quadratic programming algorithms. With applications to variational inequalities.*, volume 23 of *Springer Optimization and Its Applications*. Springer, New York, 2009.

[15] Zdeněk Dostál and David Horák. Scalable FETI with optimal dual penalty for semicoercive variational inequalities. In *Current trends in scientific computing (Xi'an, 2002)*, volume 329 of *Contemp. Math.*, pages 79–88. Amer. Math. Soc., Providence, RI, 2003.

[16] Zdeněk Dostál, David Horák, and Dan Stefanica. A scalable FETI-DP algorithm for a semi-coercive variational inequality. *Comput. Methods Appl. Mech. Engrg.*, 196(8):1369–1379, 2007.

[17] Maksymilian Dryja, Barry F. Smith, and Olof B. Widlund. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM J. Numer. Anal.*, 31(6):1662–1694, 1994.

[18] Maksymilian Dryja and Olof B. Widlund. Some domain decomposition algorithms for elliptic problems. In *Iterative methods for large linear systems (Austin, TX, 1988)*, pages 273–291. Academic Press, Boston, MA, 1990.

[19] Ricardo G. Durán and Maria Amelia Muschietti. The Korn inequality for Jones domains. *Electron. J. Differential Equations*, pages No. 127, 10 pp., 2004.

[20] Charbel Farhat, Po-Shu Chen, and Jan Mandel. A scalable lagrange multiplier based domain decomposition method for time-dependent problems. *Internat. J. Numer. Methods Engrg.*, 38:3831–3853, 1995.

[21] Charbel Farhat and Michel Géradin. On the general solution by a direct method of a large-scale singular system of linear equations: application to the analysis of floating structures. *Internat. J. Numer. Methods Engrg.*, 41(4):675–696, 1998.

[22] Charbel Farhat, Michel Lesoinne, Patrick LeTallec, Kendall Pierson, and Daniel Rixen. FETI-DP: a dual-primal unified FETI method. I. A faster alternative to the two-level FETI method. *Internat. J. Numer. Methods Engrg.*, 50(7):1523–1544, 2001.

[23] Charbel Farhat, Jan Mandel, and François-Xavier Roux. Optimal convergence properties of the FETI domain decomposition method. *Comput. Methods Appl. Mech. Engrg.*, 115(3-4):365–385, 1994.

[24] Charbel Farhat and François-Xavier Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *Internat. J. Numer. Methods Engrg.*, 32:1205–1227, 1991.

[25] Wolfgang Hackbusch. *Iterative solution of large sparse systems of equations*, volume 95 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1994. Translated and revised from the 1991 German original.

[26] Piotr Hajłasz. Sobolev inequalities, truncation method, and John domains. In *Papers on analysis*, volume 83 of *Rep. Univ. Jyväskylä Dep. Math. Stat.*, pages 109–126. Univ. Jyväskylä, Jyväskylä, 2001.

[27] Peter W. Jones. Quasiconformal mappings and extendability of functions in Sobolev spaces. *Acta Math.*, 147(1-2):71–88, 1981.

[28] Axel Klawonn and Oliver Rheinbach. Highly scalable parallel domain decomposition methods with an application to biomechanics. to appear.

[29] Axel Klawonn and Oliver Rheinbach. A parallel implementation of dual-primal FETI methods for three-dimensional linear elasticity using a transformation of basis. *SIAM J. Sci. Comput.*, 28(5):1886–1906, 2006.

[30] Axel Klawonn and Oliver Rheinbach. Inexact FETI-DP methods. *Internat. J. Numer. Methods Engrg.*, 69(2):284–307, 2007.

[31] Axel Klawonn and Oliver Rheinbach. Robust FETI-DP methods for heterogeneous three dimensional elasticity problems. *Comput. Methods Appl. Mech. Engrg.*, 196(8):1400–1414, 2007.

[32] Axel Klawonn and Oliver Rheinbach. A hybrid approach to 3-level feti. *Proc. Appl. Math. Mech.*, 8(1):10841–10843, 2008.

[33] Axel Klawonn, Oliver Rheinbach, and Olof B. Widlund. An analysis of a FETI-DP algorithm on irregular subdomains in the plane. *SIAM J. Numer. Anal.*, 46(5):2484–2504, 2008.

[34] Axel Klawonn and Olof B. Widlund. A domain decomposition method with Lagrange multipliers and inexact solvers for linear elasticity. *SIAM J. Sci. Comput.*, 22(4):1199–1219, 2000.

[35] Axel Klawonn and Olof B. Widlund. FETI and Neumann-Neumann iterative substructuring methods: connections and new results. *Comm. Pure Appl. Math.*, 54(1):57–90, 2001.

[36] Axel Klawonn and Olof B. Widlund. Dual-primal FETI methods for linear elasticity. *Comm. Pure Appl. Math.*, 59(11):1523–1572, 2006.

[37] Axel Klawonn, Olof B. Widlund, and Maksymilian Dryja. Dual-primal FETI methods for three-dimensional elliptic problems with heterogeneous coefficients. *SIAM J. Numer. Anal.*, 40(1):159–179, 2002.

[38] Jing Li and Olof B. Widlund. FETI-DP, BDDC, and block Cholesky methods. *Internat. J. Numer. Methods Engrg.*, 66(2):250–271, 2006.

[39] Jan Mandel and Clark R. Dohrmann. Convergence of a balancing domain decomposition by constraints and energy minimization. *Numer. Linear Algebra Appl.*, 10(7):639–659, 2003. Dedicated to the 70th birthday of Ivo Marek.

[40] Jan Mandel and Radek Tezaur. Convergence of a substructuring method with Lagrange multipliers. *Numer. Math.*, 73(4):473–487, 1996.

[41] Jan Mandel and Radek Tezaur. On the convergence of a dual-primal substructuring method. *Numer. Math.*, 88(3):543–558, 2001.

[42] Jindřich Nečas. *Les méthodes directes en théorie des équations elliptiques*. Masson et Cie, Éditeurs, Paris, 1967.

[43] Andrea Toselli and Olof Widlund. *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2005.

[44] Xuemin Tu. Three-level BDDC in three dimensions. *SIAM J. Sci. Comput.*, 29(4):1759–1780, 2007.

[45] Xuemin Tu. Three-level BDDC in two dimensions. *Internat. J. Numer. Methods Engrg.*, 69(1):33–59, 2007.