Elsevier Editorial System(tm) for Journal of Computational and Applied Mathematics
Manuscript Draft

Manuscript Number:

Title: A Scalable FETI--DP Algorithm with Non--penetration Mortar Conditions on Contact Interface

Article Type: Research Paper

Section/Category: 65N Except 65N06, 65N12, 65N15, 65N40

Keywords: coercive variational inequalities;
FETI--DP algorithms;
mortar finite elements;
modified proportioning algorithms

Corresponding Author: Dr. Z. Dostal,

Corresponding Author's Institution: Technical University

First Author: Z. Dostal

Order of Authors: Z. Dostal; David Horak; Dan Stefanica

Manuscript Region of Origin:

Abstract:

A Scalable FETI–DP Algorithm with Non–penetration Mortar Conditions on Contact Interface

Zdeněk Dostál*, David Horák† and Dan Stefanica‡§

November 30, 2007

November 30, 2007

Abstract

By combining FETI algorithms of dual-primal type with recent results for bound constrained quadratic programming problems, we develop an optimal algorithm for the numerical solution of coercive variational inequalities. The model problem is discretized using non-penetration conditions of mortar type across the potential contact interface, and a FETI–DP algorithm is formulated. The resulting quadratic programming problem with bound constraints is solved by a scalable algorithm with a known rate of convergence given in terms of the spectral condition number of the quadratic problem. Numerical experiments for non–matching meshes across the contact interface confirm the theoretical scalability of the algorithm.

Key words. coercive variational inequalities, FETI–DP algorithms, mortar finite elements, modified proportioning algorithms

AMS(MOS) subject classifications. 65N30, 65N55

1 Introduction

Finite Element Tearing and Interconnecting (FETI)–based domain decomposition methods are efficient tools for the numerical solution of complex engineering problems. The FETI method was originally proposed by Farhat and Roux [23] as a parallel solver for problems described by elliptic partial differential equations. Later, Farhat, Mandel, and Roux [22] modified the basic FETI method by introducing so-called natural coarse grid projections to obtain a numerically scalable algorithm. The performance of the algorithm was further enhanced by using preconditioners [28, 37] with improved scaling properties. By projecting the Lagrange multipliers in each iteration onto an auxiliary space to enforce continuity of the

*FEI VŠB-Technical University Ostrava, CZ-70833 Ostrava, Czech Republic. E-mail address: zdenek.dostal@vsb.cz.

†FEI VŠB-Technical University Ostrava, CZ-70833 Ostrava, Czech Republic. E-mail address: david.horak@vsb.cz.

‡Baruch College, City University of New York, New York NY 10010. E-mail address: Dan_Stefanica@baruch.cuny.edu. Corresponding author.

§D. Horák and Z. Dostál were supported by the grants GA CR 201/07/0294, AS CR 1ET400300415, and the Ministry of Education of the Czech Republic No. MSM6198910027. D. Stefanica was supported by the National Science Foundation under Grant NSF–DMS–0103588, by the Research Foundation of the City University of New York under Awards PSC-CUNY 65463–00 34 and 66529–00 35, and by the Eugene M. Lang Foundation.

primal solutions at the crosspoints, Farhat, Mandel and Tezaur [35] obtained a faster converging FETI method for plate and shell problems. FETI algorithms were also implemented for, e.g., Helmholtz problems [21, 24], linear elasticity with inexact solvers [27], Maxwell's equations [36, 44], and Stokes problems [26, 32, 46].

The key ingredient of the FETI method is the decomposition of the computational domain into non-overlapping subdomains that are "glued" by Lagrange multipliers. The primal variables are eliminated by solving possibly singular local problems. The original problem reduces to a small, relatively well conditioned, typically equality constrained quadratic programming problem that is solved iteratively. If the procedure is applied to the discretized variational inequality describing the equilibrium of a system of elastic bodies in contact, not only the dimension of the problem is reduced, but also the original inequality constraints describing the non-penetration of the bodies reduce to the bound constraints. The resulting problem can be solved efficiently either by direct iterations [19] or by specialized quadratic programming methods [7, 8, 9]. Numerical experiments with these algorithms indicated their numerical scalability. Recently, using new results on the rate of convergence of improved versions [6, 13] of the active set based proportioning algorithm [5], numerical scalability was proved for two variants of this algorithm by Dostál and Horák [10, 11, 12].

The Dual-Primal FETI method (FETI–DP) is a variant of the FETI method which does not require solving singular problems to eliminate the primal variables. The FETI–DP method was introduced by Farhat et al. [20]; see also [16]. For two dimensional scalar problems, the continuity of the primal solution at crosspoints is implemented directly into the formulation of the primal problem so that one degree of freedom is considered at each crosspoint shared by more than two adjacent subdomains. The continuity of the primal variables across the rest of the subdomain interfaces is once again enforced by Lagrange multipliers. After eliminating the primal variables, the problem reduces to a small, unconstrained, strictly convex quadratic programming problem that is solved iteratively. An attractive feature of FETI–DP is that the resulting quadratic programming problem is unconstrained and its conditioning may be further improved by preconditioning [34]. Farhat et al. [1, 19, 45] introduced a FETI–DP based algorithm, FETI–C, for solving contact problems arising in structural mechanics. This method is based on Newton type iterations, and its scalability was established experimentally. The FETI–DP method was recently combined by the present authors [14, 15] with the aforementioned results on the solution of the bound constrained quadratic programming problems to develop a scalable algorithm for the solution of both coercive and semicoercive contact problems.

In many practical applications, such as the contact shape optimization and the transient problems, it is difficult and computationally expensive to generate matching meshes on the contact interface. A natural way to implement the non–penetration conditions on potentially non–matching contact interfaces is by using constraints of mortar type. The mortar finite element methods are non-conforming finite elements first introduced by Bernardi, Maday, and Patera in [3]. Biorthogonal mortar elements with slightly better computational properties were later developed by Wohlmuth [47, 49]. Mortars are well suited for parallel computing and have several advantages over conforming finite elements. For example, mesh generation is more flexible and can be made quite simple on individual subregions and local refinement of finite element models using mortar methods is straightforward. A large number of domain decomposition methods have been extended to mortar discretizations in order to take advantage of the inherent flexibility of the mortars. For example, it was shown both numerically and theoretically that the FETI and FETI–DP methods for mortar finite elements perform similarly to the case of conforming finite elements; cf. [18, 31, 36, 40, 41].

In this paper, we develop a scalable algorithm of FETI–DP type for solving a coercive

variational inequality obtained by discretizing a model contact problem by using an efficient implementation of mortar methods. The FETI–DP methodology is first applied to the discretized elliptic variational inequality to obtain a strictly convex quadratic programming problem with non-negativity constraints. This problem is then solved efficiently by recently proposed improvements [6, 13] of the active set based proportioning algorithm [5]. The rate of convergence of these algorithms can be bounded in terms of the spectral condition number of the Hessian of the quadratic problem. The scalability of the resulting algorithm can therefore be established provided that suitable bounds on the condition number of the Hessian exist. We present such estimates in terms of the decomposition parameter H and the discretization parameter h. We also obtain a bound on the number of conjugate gradient iterations required for finding the solution of the discretized variational inequality to a given precision. If the rows of the discretized constraint matrix is are normalized and if we keep the ratio H/h fixed, it is proved that this bound is independent of both the decomposition of the computational domain and the discretization. Let us recall that the normalization is not required when the non-penetration is implemented by nodal constraints [14]. We report numerical results that are in agreement with the theory and confirm the numerical scalability of our algorithm.

We note that the effort to develop scalable solvers for variational inequalities is not limited to FETI–type methods. For example, multigrid ideas were used early on by Mandel [33]. Kornhuber, Krause and Wohlmuth [29, 30, 48] introduced an algorithm based on monotone multigrid with scalable solution of auxiliary linear problems. Combining multigrid ideas and approximate projections, Schöberl [38, 39] introduced an algorithm for which linear complexity was established.

The rest of the paper is structured as follows. The model problem introduced in Section 2 is discretized using mortars in the subsequent section. A FETI–DP type algorithm is presented in Section 4 and spectral bounds on the resulting operator are established in Section 5. The solution to our FETI–DP method is obtained by a modified proportioning algorithm with reduced gradient projections, as presented in Section 6. Numerical results confirming the scalability of our algorithm are reported in Section 7.

2 Model problem

The computational domain for our model problem is $\Omega = \Omega^1 \cup \Omega^2$, where $\Omega^1 = (0,1)\times(0,1)$ and $\Omega^2 = (1,2)\times(0,1)$, with boundaries Γ^1 and Γ^2, respectively. We denote by Γ^i_u, Γ^i_f, and Γ^i_c the fixed, free, and potential contact parts of Γ^i, $i = 1, 2$; see Figure 1a. Let $\Gamma_c = \Gamma^1_c \cup \Gamma^2_c$.

The Sobolev space of the first order on Ω^i is denoted by $H^1(\Omega^i)$ and the space of Lebesgue square integrable functions is denoted by $L^2(\Omega^i)$. Let $V = V^1 \times V^2$, with

$$V^i = \left\{ v^i \in H^1(\Omega^i) : v^i = 0 \quad \text{on} \quad \Gamma^i_u \right\}, \quad i = 1, 2.$$

Let $\mathcal{K} \subset V$ be a closed convex subset of $\mathcal{H} = H^1(\Omega^1) \times H^1(\Omega^2)$ defined by

$$\mathcal{K} \;=\; \left\{ (v^1, v^2) \in V : v^2 - v^1 \geq 0 \quad \text{on} \quad \Gamma_c \right\}.$$

We define the symmetric bilinear form $a(\cdot,\cdot) : \mathcal{H} \times \mathcal{H} \to R$ by

$$a(u,v) = \sum_{i=1}^{2} \int_{\Omega^i} \left(\frac{\partial u^i}{\partial x_1} \frac{\partial v^i}{\partial x_1} + \frac{\partial u^i}{\partial x_2} \frac{\partial v^i}{\partial x_2} \right) d\Omega.$$

Let $f \in L^2(\Omega)$ be a given function and $f^i \in L^2(\Omega^i)$, $i = 1, 2$, be the restrictions of f to Ω^i, $i = 1, 2$. We define the linear form $l(\cdot) : \mathcal{H} \to R$ by

$$\ell(v) = \sum_{i=1}^{2} \int_{\Omega^i} f^i v^i d\Omega$$

and consider the following problem:

$$\text{Find} \quad \min \ \frac{1}{2} a(u,u) - \ell(u) \quad \text{subject to} \quad u \in \mathcal{K}. \tag{1}$$

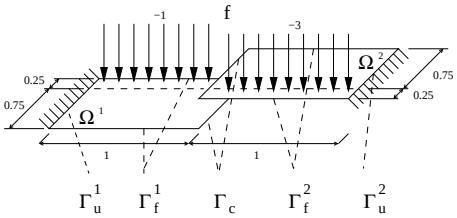

Figure 1a: Coercive model problem

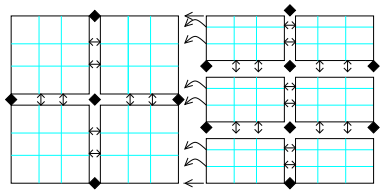
Figure 1b: Decomposition of domains

The solution of the model problem may be interpreted as the displacement of two membranes under the traction f. The membranes are fixed on the outer edges as in Figure 1a and the left edge of the membrane Ω^2 is not allowed to penetrate below the right edge of the membrane Ω^1. Note that Γ_u^i, the parts of $\partial\Omega^i$, $i = 1, 2$, where Dirichlet conditions are prescribed, have positive Lebesgue measure. Thus, the quadratic form $a(\cdot, \cdot)$ is coercive and the solution of problem (1) exists and is unique; cf. [25].

3 A mortar finite element discretization

Mortar finite elements are nonconforming finite elements that allow for a nonconforming decomposition of the computational domain into subdomains with nonmatching grids across the partition interface, and, at the same time, for the optimal coupling of different variational approximations in different subregions. Here, by optimality we mean that the global error is bounded by the sum of the local approximation errors on each subregion. The most general such partition for our model problem would consist of introducing mortar spaces on both Ω^1 and Ω^2, by partitioning these domains into nonoverlapping rectangular subdomains. For the FETI–DP type algorithms considered here, the partitions would need to be geometrically conforming, i.e, the intersection between the closures of any two rectangular subdomains should be either empty, or consist of a vertex or of an entire edge, and the mortars would need to be of the first type, i.e., continuous at the corner nodes; see [17, 18, 42]. Weak continuity would be enforced by way of orthogonality of the jump of the mortar functions across the interfaces of the mortar partitions within Ω^1 and Ω^2, respectively. This would be done using so–called mortar conditions. Across the potential contact boundary Γ_c, the non–penetration condition would also be replaced by inequalities related to mortar conditions. More details on mortar conditions follow shortly in this section.

However, for a contact problem like our model problem (1), the most efficient algorithms only require mortar conditions across the contact interface, while the bodies in contact may be discretized using continuous finite elements; see [43] for a detailed study of the

computational complexity and numerical performance of FETI–type algorithms for mortar methods. We are going to use this type of discretization throughout the paper.

The discrete space W is therefore constructed as follows: Each domain Ω^i, $i = 1, 2$, is partitioned on a rectangular grid into subdomains of diameter on the order of H. The restrictions of W to Ω^1 and Ω^2 are Q_1 finite element spaces of comparable mesh sizes of order h, corresponding to the subdomain grids in Ω^1 and Ω^2. Note that the subdomain grids do not necessarily match across the potential contact interface Γ_c. We call a crosspoint either a corner that belongs to four subdomains, or a corner that belongs to two subdomains and is located on $\partial\Omega^1 \setminus \Gamma_u^1$ or on $\partial\Omega^2 \setminus \Gamma_u^2$. The nodes corresponding to the end points of Γ_c are not crosspoints; see Figure 1b. An important feature for developing FETI–DP type algorithms is that a single global degree of freedom is considered at each crosspoint, while two degrees of freedom are introduced at all the other matching nodes across subdomain edges.

Let $v \in W$. The continuity of v in Ω^1 and Ω^2 is enforced at every interface node that is not a crosspoint. For simplicity, we also denote by v the nodal values vector of $v \in W$. In matrix notation, the continuity conditions can be written as

$$B_E v \;=\; 0,$$

where each row of the matrix B_E enforces continuity at one node on the subdomain interface where multiple degrees of freedom were considered. Thus, each row of B_E has only two nonzero entries, equal to 1 and -1, respectively.

The continuity of v at crosspoints is enforced by using a global vector of degrees of freedom v_c^g and a global-to-local map L_c with one nonzero entry equal to 1 in each row. Thus, if v_c is the vector of crosspoint nodal values of v, we require that $v_c = L_c v_c^g$.

Across the potential contact interface Γ_c, the meshes on Γ_c^1 and Γ_c^2 need not match. The non–penetration condition $[v] = v^1 - v^2 \leq 0$ is replaced by mortar conditions as follows: We call nonmortar sides all the sides on Γ_c^1, while the sides on Γ_c^2 are called mortar sides. For each nonmortar side γ, we require the non–penetration conditions

$$\int_\gamma [v]\ \psi\ ds \;\leq\; 0, \quad \forall\ \psi \in \Psi(\gamma), \tag{2}$$

where $\Psi(\gamma)$ is a space of test functions having the same dimension as the number of interior nodes on γ. For details on the special choice of discontinuous test functions corresponding to the biorthogonal mortars used in the numerical experiments from Section 7, we refer the reader to [49]. From a theoretical point of view, the nonmortar sides could have been chosen on Γ_c^2 as well, which would have rendered the sides on Γ_c^1 mortar sides. The effect of different choices of nonmortar sides on the numerical performance of our method is investigated in Section 7.

For algorithmic purposes, we derive a matrix formulation for the non–penetration mortar conditions (2). Let v_γ^1 be the vector of the nodal values of v on $\gamma \subset \Gamma_c^1$ and let v_γ^2 be the vector of those nodal values of v on Γ_c^2 that are opposite γ. The matrix formulation of (2) is

$$M_\gamma v_\gamma^1 \;-\; N_\gamma v_\gamma^2 \;\leq\; 0. \tag{3}$$

Let $B_\gamma^{\#} = [0\ M_\gamma\ 0\ -N_\gamma]$. Since the mortar conditions (2) are only related to interior nodes on the nonmortar edges, two extra conditions are necessary to enforce non–penetration at nodes corresponding to the endpoints of Γ_c. All these inequality constraints can be written in matrix formulation as

$$B_I^{\#} v \leq 0, \tag{4}$$

where $B_I^{\#}$ has one horizontal block $B_\gamma^{\#}$ corresponding to (2) for each nonmortar side γ, and two more rows for the non–penetration conditions at the endpoints of Γ_c. Note that the norm of each row in $B_\gamma^{\#}$ is on the order of the mesh size on γ, $h_\gamma \approx h$, while every row of B_E has norm of order 1.

Let B_γ be obtained from $B_\gamma^{\#}$ by normalizing every row of $B_\gamma^{\#}$. Also, let B_I be the matrix having one block B_γ for each nonmortar γ, together with the two extra rows as before. By construction, the matrix formulation (4) of the non–penetration conditions is equivalent to

$$B_I v \ \leq \ 0. \tag{5}$$

We show theoretically in Section 5 and experimentally in Section 7 that the FETI–DP algorithm proposed here is scalable, if normalized non-penetration conditions of the form (5) are used, and is not scalable, if non-normalized conditions of the form (4) are enforced.

The discretized version of problem (1) with the auxiliary domain decomposition has the form

$$\min \frac{1}{2} v^T K v - v^T f \quad \text{subject to} \quad B_I v \leq 0 \quad \text{and} \quad B_E v = 0, \tag{6}$$

where K is the positive definite stiffness matrix corresponding to the model problem and f represents the discrete analog of the linear form $\ell(\cdot)$.

4 A FETI–DP method with mortar non–penetration conditions

To solve (6), we propose a variant of the algorithm introduced in [14]. The main difference is the use of mortar non–penetration conditions. Also, global degrees of freedom are considered for the corner nodes on Γ_c.

We partition the nodal values of $v \in W$ into crosspoint nodal values, denoted by v_c, and remainder nodal values, denoted by v_r. Recall that the continuity conditions at crosspoints, i.e., at subdomain corners, are enforced by using a global vector of degrees of freedom v_c^g such that $v_c = L_c v_c^g$. Therefore,

$$v \ = \ \begin{bmatrix} v_r \\ v_c \end{bmatrix} \ = \ \begin{bmatrix} v_r \\ L_c v_c^g \end{bmatrix}.$$

Problem (6) can be written as a constrained minimization problem as follows:

$$\min \frac{1}{2} v^T K v \ - \ v^T f \quad \text{subject to} \quad B_I v \leq 0, \ \ B_E v = 0, \quad \text{and} \quad v_c = L_c v_c^g. \tag{7}$$

Let f_c and f_r be the parts of the right hand side f corresponding to crosspoints and remainder nodes, respectively. The Lagrangian associated with problem (7) can be expressed using Lagrange multipliers λ_E and λ_I to enforce the inequality and redundancy constraints as follows:

$$\begin{aligned} L(v_r, v_c^g, \lambda_E, \lambda_I) \ \ = \ \ & \frac{1}{2} \begin{bmatrix} v_r^T & (L_c v_c^g)^T \end{bmatrix} K \begin{bmatrix} v_r \\ L_c v_c^g \end{bmatrix} \ - \ \begin{bmatrix} v_r^T & (L_c v_c^g)^T \end{bmatrix} \begin{bmatrix} f_r \\ f_c \end{bmatrix} \\ & + \ v^T B_E^T \lambda_E \ + \ v^T B_I^T \lambda_I. \end{aligned} \tag{8}$$

Let $B_{I,r}$ and $B_{I,c}$ be the matrices made of the columns of B_I corresponding to v_r and v_c, respectively; define $B_{E,r}$ and $B_{E,c}$ similarly. Then $B_I = [B_{I,r} \ B_{I,c}]$ and $B_E = [B_{E,r} \ B_{E,c}]$.

We can also group the parts of the Lagrange multiplier matrices B_I and B_E together with respect to the corner and remainder nodes as follows:

$$B_r \;=\; \left[\begin{array}{c} B_{I,r} \\ B_{E,r} \end{array}\right] \quad \text{and} \quad B_c \;=\; \left[\begin{array}{c} B_{I,c} \\ B_{E,c} \end{array}\right].$$

The corner nodes on Γ_c^1 (with the exception for the end points of Γ_c) belong to two subdomains. We associate one global degree of freedom to each such corner. A similar procedure is applied to the corner nodes on Γ_c^2. This represents a natural departure from the algorithm suggested in [14]. As a result, $B_{I,c}$ becomes a nonzero matrix. Since the equality constraints are not related to crosspoints, $B_{E,c} = 0$, as in the conforming finite element case. Let

$$\lambda \;=\; \left[\begin{array}{c} \lambda_I \\ \lambda_E \end{array}\right].$$

Note that

$$\left[\begin{array}{c} B_I \\ B_E \end{array}\right] \;=\; \left[\begin{array}{cc} B_{I,r} & B_{I,c} \\ B_{E,r} & B_{E,c} \end{array}\right] \;=\; [B_r \; B_c].$$

Then

$$v^T B_E^T \lambda_E + v^T B_I^T \lambda_I \;=\; v_r^T B_r^T \lambda \;+\; (v_c^g)^T L_c^T B_c^T \lambda.$$

Let K_{rr}, K_{rc}, and K_{cc} be the blocks of K corresponding to the decomposition of v into v_r and v_c. To minimize $L(v_r, v_c^g, \lambda_E, \lambda_I)$ over v_r, we rewrite (8) as

$$\begin{aligned} L(v_r, v_c^g, \lambda_E, \lambda_I) \;&=\; \frac{1}{2}\left(v_r^T K_{rr} v_r + 2 v_r^T K_{rc} L_c v_c^g + (v_c^g)^T L_c^T K_{cc} L_c v_c^g\right) \\ &\quad - v_r^T f_r - (v_c^g)^T L_c^T f_c \;+\; v_r^T B_r^T \lambda \;+\; (v_c^g)^T L_c^T B_c^T \lambda \end{aligned}$$

and obtain that v_r is a solution of

$$K_{rr} v_r \;+\; K_{rc} L_c v_c^g \;-\; f_r + B_r^T \lambda \;=\; 0.$$

Note that K_{rr} is a positive definite submatrix of K since each subdomain has at least one corner node. We end up with the following Lagrangian to minimize over v_c^g:

$$\begin{aligned} L_c(v_c^g, \lambda_E, \lambda_I) \;&=\; \frac{1}{2}(v_c^g)^T L_c^T K_{cc} L_c v_c^g \;-\; (v_c^g)^T L_c^T f_c \;+\; (v_c^g)^T L_c^T B_c^T \lambda \\ &\quad - \frac{1}{2}\left(f_r - K_{rc} L_c v_c^g - B_r^T \lambda\right)^T K_{rr}^{-1}\left(f_r - K_{rc} L_c v_c^g - B_r^T \lambda\right) \\ &=\; \frac{1}{2}(v_c^g)^T K_{cc}^* v_c^g \;-\; (v_c^g)^T \left(\widetilde{F}_{I_{rc}}^T \lambda + f_c^*\right) \\ &\quad - \frac{1}{2}\left(f_r - B_r^T \lambda\right)^T K_{rr}^{-1}\left(f_r - B_r^T \lambda\right), \end{aligned}$$

where we used the following notations related to those from [20]:

$$\begin{aligned} F_{I_{rr}} \;&=\; B_r K_{rr}^{-1} B_r^T; \\ \widetilde{F}_{I_{rc}} \;&=\; B_r K_{rr}^{-1} K_{rc} L_c \;-\; B_c L_c; \\ K_{cc}^* \;&=\; L_c^T (K_{cc} - K_{rc}^T K_{rr}^{-1} K_{rc}) L_c; \\ f_c^* \;&=\; L_c^T (f_c - K_{rc}^T K_{rr}^{-1} f_r). \end{aligned}$$

The solution to the minimization of $L_c(v_c^g, \lambda_E, \lambda_I)$ over v_c^g must satisfy

$$K_{cc}^* v_c^g \;-\; \widetilde{F}_{I_{rc}}^T \lambda \;-\; f_c^* \;=\; 0.$$

This problem is solvable since, for a coercive problem, K_{cc}^* is a positive definite matrix. The corresponding minimal value of $L_c(v_c^g, \lambda_E, \lambda_I)$ is

$$L_\lambda(\lambda_E, \lambda_I) \;=\; \frac{1}{2}\left(f_c^* + \widetilde{F}_{I_{rc}}^T\lambda\right)^T (K_{cc}^*)^{-1}\left(f_c^* + \widetilde{F}_{I_{rc}}^T\lambda\right) \;-\; \frac{1}{2}\left(f_r - B_r^T\lambda\right)^T K_{rr}^{-1}\left(f_r - B_r^T\lambda\right).$$

Thus, maximizing L_λ over $\lambda_I \geq 0$ is equivalent to finding

$$\min_{\lambda_I \geq 0} \Theta(\lambda), \tag{9}$$

where

$$\Theta(\lambda) = \frac{1}{2}\lambda^T F\lambda - \lambda^T b \tag{10}$$

with

$$\begin{aligned} F &= F_{I_{rr}} + \widetilde{F}_{I_{rc}}(K_{cc}^*)^{-1}\widetilde{F}_{I_{rc}}^T; \\ b &= \widetilde{F}_{I_{rc}}(K_{cc}^*)^{-1}f_c^* - B_r K_{rr}^{-1} f_r. \end{aligned} \tag{11}$$

5 Bounds on the spectrum of F

In this section, we derive bounds on the spectrum of F that will be used in the following section for the convergence analysis of the modified proportioning algorithm required to solve the bound constrained quadratic problem (9).

Let $\overline{B} = [B_r\ 0]$ and let $\overline{K}$ be the stiffness matrix corresponding to the model problem on a finite element discretization where continuity is required at the corners but no other continuity is required across the subdomain edges. From inverse inequalities and Poincaré's inequality, it follows that

$$\frac{C}{H^2}||w||^2_{L^2(\Omega)} \;\leq\; \langle \overline{K}w, w\rangle \;\leq\; \frac{C}{h^2}||w||^2_{L^2(\Omega)}, \quad \forall w \in W, \tag{12}$$

where $\langle\cdot,\cdot\rangle$ is the notation for the Euclidean inner product. Here and throughout the paper, C is a generic constant independent of h, H, and the number of subdomains in the partitions of Ω^1 and Ω^2.

Note that $F = \overline{B}\ \overline{K}^{-1}\ \overline{B}^T$ and therefore we find as in Lemma 4.3 of [34] that

$$\langle F\lambda, \lambda\rangle \;=\; \sup_{w\in W} \frac{\langle \overline{B}^T\lambda, w\rangle^2}{\langle \overline{K}w, w\rangle}. \tag{13}$$

Let $||w||_{l^2}$ and $||\lambda||_{l^2}$ be the Euclidean norms of the primal and dual variables w and λ, respectively. Since w is a finite element function,

$$||w||^2_{L^2(\Omega)} \;\approx\; h^2||w||^2_{l^2}. \tag{14}$$

For $w = \overline{B}^T\lambda$ in (14),

$$||\overline{B}^T\lambda||^2_{L^2(\Omega)} \approx h^2||\overline{B}^T\lambda||^2_{l^2}. \tag{15}$$

From (12), (13), and (14), we find that

$$\langle F\lambda, \lambda\rangle \;\leq CH^2 \sup_{w\in W} \frac{\langle \overline{B}^T\lambda, w\rangle^2}{||w||^2_{L^2(\Omega)}} \;\leq CH^2||\overline{B}^T\lambda||^2_{l^2} \sup_{w\in W} \frac{||w||^2_{l^2}}{||w||^2_{L^2(\Omega)}} \;\leq C\left(\frac{H}{h}\right)^2 ||\overline{B}^T\lambda||^2_{l^2}.$$

Let $w_0 = \overline{B}^T \lambda$. From (12) and (13), and using (15), it follows that

$$\langle F\lambda, \lambda \rangle \ \geq Ch^2 \sup_{w \in W} \frac{\langle \overline{B}^T \lambda, w \rangle^2}{||w||^2_{L^2(\Omega)}} \ \geq \ Ch^2 \ \frac{\langle \overline{B}^T \lambda, w_0 \rangle^2}{||w_0||^2_{L^2(\Omega)}} \ = Ch^2 \ \frac{||\overline{B}^T \lambda||^4_{l^2}}{||\overline{B}^T \lambda||^2_{L^2(\Omega)}} \ \approx C||\overline{B}^T \lambda||^2_{l^2}.$$

From (16) and (16), we conclude that

$$C||\overline{B}^T \lambda||^2_{l^2} \ \leq \ \langle F\lambda, \lambda \rangle \ \leq \ C \left(\frac{H}{h} \right)^2 \ ||\overline{B}^T \lambda||^2_{l^2}. \tag{16}$$

Recall that

$$\overline{B} \ = \ [B_r \ 0] \ = \ \begin{bmatrix} B_{I,r} & 0 \\ B_{E,r} & 0 \end{bmatrix}.$$

We want to analyze the importance of the normalization of the mortar non-penetration conditions for the performance of our algorithm. To this end, let us denote by $\overline{B}^{\#}$ the matrix similar to $\overline{B}$ without normalizing the rows of B_I, i.e., using $B_I^{\#}$ instead of B_I:

$$\overline{B}^{\#} \ = \ \begin{bmatrix} B_{I,r}^{\#} & 0 \\ B_{E,r} & 0 \end{bmatrix}.$$

The major difference between $\overline{B}$ and $\overline{B}^{\#}$ is contained in the following lemma:

Lemma 1. If the mortar non-penetration conditions are normalized, then

$$||\overline{B}^T \lambda||^2_{l^2} \ \approx \ ||\lambda||^2_{l^2}. \tag{17}$$

Else, for non-normalized inequality conditions,

$$Ch^2 ||\lambda||^2_{l^2} \ \leq \ ||(\overline{B}^{\#})^T \lambda||^2_{l^2} \ \leq \ C||\lambda||^2_{l^2}. \tag{18}$$

As soon as Lemma 1 is established (see proof below), we can show the following result which will be used in the convergence estimates for the modified proportioning algorithm suggested in the next section:

Theorem 1. The following bounds on the spectrum, norm, and condition number of the operator F given by (11) hold:

$$C \ \leq \ \lambda_{min}(F) \ \leq \ \lambda_{max}(F) \ \leq \ C \left(\frac{H}{h} \right)^2; \tag{19}$$

$$\kappa(F) \ \leq \ C \left(\frac{H}{h} \right)^2; \quad ||F|| \ \leq \ C \left(\frac{H}{h} \right)^2. \tag{20}$$

Let $F^{\#} = \overline{B}^{\#} \ \overline{K}^{-1} \ (\overline{B}^{\#})^T$, the operator corresponding to F if the non–normalized mortar inequality matrix $B_I^{\#}$ is used in the algorithm instead of B_I. The condition number estimate for $F^{\#}$ deteriorates as follows:

$$Ch^2 \ \leq \ \lambda_{min}(F^{\#}) \ \leq \ \lambda_{max}(F^{\#}) \ \leq \ C \left(\frac{H}{h} \right)^2; \tag{21}$$

$$\kappa(F^{\#}) \ \leq \ \frac{C}{h^2} \ \left(\frac{H}{h} \right)^2; \quad ||F^{\#}|| \ \leq \ C \left(\frac{H}{h} \right)^2. \tag{22}$$

Proof. From (16) we find that

$$C\frac{||\overline{B}^T\lambda||_{l^2}^2}{||\lambda||_{l^2}^2} \;\leq\; \lambda_{min}(F) \;\leq\; \lambda_{max}(F) \;\leq\; C\left(\frac{H}{h}\right)^2\frac{||\overline{B}^T\lambda||_{l^2}^2}{||\lambda||_{l^2}^2}. \tag{23}$$

Then (19) follows from the estimate (17) of Lemma 1. A similar inequality to (23) also holds for $F^{\#}$, and (21) follows as before from the inequality (18) of Lemma 1.

Since F and $F^{\#}$ are symmetric, (20) and (22) follow immediately from (19) and (21). □

Proof of Lemma 1 It is easy to see that

$$\overline{B}^T\lambda \;=\; \begin{bmatrix} B_{I,r}^T\lambda_I \\ 0 \end{bmatrix} \;+\; \begin{bmatrix} B_{E,r}^T\lambda_E \\ 0 \end{bmatrix}.$$

All the nonzero primal variables corresponding to $B_{I,r}^T\lambda_I$ are located on Γ_c, while the primal variables corresponding to $B_{E,r}^T\lambda_E$ are all on the interface of the partitions of Ω^1 and Ω^2. Then,

$$||\overline{B}^T\lambda||_{l^2}^2 \;=\; ||B_{I,r}^T\lambda_I + B_{E,r}^T\lambda_E||_{l^2}^2 \;=\; ||B_{I,r}^T\lambda_I||_{l^2}^2 \;+\; ||B_{E,r}^T\lambda_E||_{l^2}^2. \tag{24}$$

The entries in each row of $B_{E,r}^T$ are 0, except for two entries which are either 1 or -1. Therefore, $||B_{E,r}^T\lambda_E||_{l^2}^2 = 2||\lambda_E||_{l^2}^2$ and

$$||\overline{B}^T\lambda||_{l^2}^2 \;\approx\; ||B_{I,r}^T\lambda_I||_{l^2}^2 \;+\; ||\lambda_E||_{l^2}^2. \tag{25}$$

Let $\gamma(j)$, $j = 1 : N_{nm}$, be the nonmortar sides on Γ_c, where N_{nm} denotes the number of nonmortar sides. Let $B_{\gamma(j)}$ be the matrix of the normalized mortar conditions corresponding to $\gamma(j)$, and let $\lambda_{I,j}$ be the Lagrange multipliers corresponding to $\gamma(j)$. We denote by $B_{\gamma(j),r}$ the part of $B_{\gamma(j)}$ corresponding to the remainder nodes. Then

$$B_{I,r} \;=\; \begin{bmatrix} B_{\gamma(1),r} \\ \vdots \\ B_{\gamma(N_{nm}),r} \end{bmatrix} \quad\text{and}\quad B_{I,r}^T\,\lambda_I \;=\; \sum_{i=1}^{N_{nm}} B_{\gamma(j),r}^T\,\lambda_{I,j}. \tag{26}$$

For every node on Γ_c, there are at most two vectors with nonzero entries at that node from among the vectors $\{B_{\gamma(j),r}^T\,\lambda_{I,j}\}_{i=1:N_{nm}}$. Therefore,

$$||B_{I,r}^T\,\lambda_I||_{l^2}^2 \;\leq\; 2\sum_{i=1}^{N_{nm}} ||B_{\gamma(j),r}^T\,\lambda_{I,j}||_{l^2}^2. \tag{27}$$

Recall, from Section 3, that $B_{\gamma(j)}$ is obtained from $B_{\gamma(j)}^{\#} = [0\;\; M_{\gamma(j)}\;\; 0\;\; -N_{\gamma(j)}]$ by normalizing the rows of $B_{\gamma(j)}^{\#}$. Thus, for biorthogonal mortars, the part of $B_{\gamma(j)}$ corresponding to the remainder nodes can be expressed as

$$B_{\gamma(j),r} \;=\; [0\;\; \mathrm{Id}_{I,j}\;\; 0\;\; -P_{\gamma(j)}],$$

where $\mathrm{Id}_{I,j}$ is the identity matrix of size equal to the number of Lagrange multipliers corresponding to $\gamma(j)$, i.e., the length of $\lambda_{I,j}$; and $P_{\gamma(j)}$ is the mortar projection matrix corresponding to $\gamma(j)$; see, e.g., [47]. Thus,

$$B_{\gamma(j),r}^T\,\lambda_{I,j} \;=\; \begin{bmatrix} 0 \\ \lambda_{I,j} \\ 0 \\ -P_{\gamma(j)}^T\,\lambda_{I,j} \end{bmatrix} \tag{28}$$

and therefore

$$||B^T_{\gamma(j),r}\ \lambda_{I,j}||^2_{l^2}\ =\ ||\lambda_{I,j}||^2_{l^2}\ +\ ||P^T_{\gamma(j)}\ \lambda_{I,j}||^2_{l^2}. \tag{29}$$

From (26), (28), and (29) we find that

$$||B^T_{I,r}\ \lambda_I||^2_{l^2}\ \geq\ \sum_{i=1}^{N_{nm}} ||\lambda_{I,j}||^2_{l^2}\ =\ ||\lambda_I||^2_{l^2}. \tag{30}$$

To estimate the norm of $P^T_{\gamma(j)}\ \lambda_{I,j}$, we use the L^2–stability of the mortar projection. It is easy to see that

$$\begin{aligned} ||P^T_{\gamma(j)}\ \lambda_{I,j}||^2_{l^2} &= \sup_{\psi\neq 0} \frac{\langle P^T_{\gamma(j)}\ \lambda_{I,j},\psi\rangle^2}{||\psi||^2_{l^2}} = \sup_{\psi\neq 0} \frac{\langle \lambda_{I,j}, P_{\gamma(j)}\ \psi\rangle^2}{||\psi||^2_{l^2}} \\ &\leq ||\lambda_{I,j}||^2_{l^2}\ \sup_{\psi\neq 0} \frac{||P_{\gamma(j)}\ \psi||^2_{l^2}}{||\psi||^2_{l^2}}. \end{aligned} \tag{31}$$

Let $\zeta(j)$ be the union of mortar sides from Γ^2_c opposite $\gamma(j)$. Then ψ corresponds to a vector of nodal values on $\zeta(j)$ and $P_{\gamma(j)}\psi$ is the vector of nodal values of the mortar projection of ψ on $\gamma(j)$. From the stability of the mortar projection we find that

$$||P_{\gamma(j)}\ \psi||^2_{L^2(\gamma(j))}\ \leq\ C||\psi||^2_{L^2(\zeta(j))}. \tag{32}$$

Let $h_{\gamma(j)}$ and $h_{\zeta(j)}$ be the mesh sizes on $\gamma(j)$ and on $\zeta(j)$, respectively. We recall that the meshes across any nonmortar side were assumed to be of order h. Therefore, $h_{\zeta(j)}/h_{\gamma(j)}$ is uniformly bounded. Using the fact that ψ and $P_{\gamma(j)}\ \psi$ are vectors of nodal values of first order finite element functions, we find from (32) that

$$||P_{\gamma(j)}\ \psi||^2_{l^2}\ \approx\ \frac{C}{h_{\gamma(j)}}\ ||P_{\gamma(j)}\ \psi||^2_{L^2(\gamma(j))}\ \leq\ \frac{C}{h_{\gamma(j)}}\ ||\psi||^2_{L^2(\zeta(j))}\ \leq\ C\frac{h_{\zeta(j)}}{h_{\gamma(j)}}\ ||\psi||^2_{l^2}\ \leq\ C||\psi||^2_{l^2}.$$

Therefore, from (31), it follows that

$$||P^T_{\gamma(j)}\ \lambda_{I,j}||^2_{l^2}\ \leq\ C||\lambda_{I,j}||^2_{l^2}.$$

Using (29), we find that

$$||\lambda_{I,j}||^2_{l^2}\ \leq\ ||B^T_{\gamma(j),r}\ \lambda_{I,j}||^2_{l^2}\ \leq\ C||\lambda_{I,j}||^2_{l^2}.$$

A bound for the norm of $B^T_{I,r}\lambda_I$ can now be established using (27), (30), and the fact that $||\lambda_I||^2_{l^2} = \sum_{i=1}^{N_{nm}} ||\lambda_{I,j}||^2_{l^2}$ satisfies the following inequality:

$$||\lambda_I||^2_{l^2}\ \leq\ ||B^T_{I,r}\ \lambda_I||^2_{l^2}\ \leq\ C||\lambda_I||^2_{l^2}. \tag{33}$$

Using (25) and (33), we can establish (17):

$$||\overline{B}^T\lambda||^2_{l^2}\ \approx\ ||B^T_{I,r}\lambda_I||^2_{l^2}\ +\ ||\lambda_E||^2_{l^2}\ \approx\ ||\lambda_I||^2_{l^2} + ||\lambda_E||^2_{l^2}\ =\ ||\lambda||^2_{l^2}.$$

For the case when non–normalized mortar conditions are used across the contact interface, i.e., when $B^{\#}_I$ and $\overline{B}^{\#}$ are used instead of B^I and $\overline{B}$, the only difference is in the scaling of the rows of $B^{\#}_I$ by h. In other words, we obtain, instead of (33), that

$$Ch^2||\lambda_I||^2_{l^2}\ \leq\ ||(B^{\#}_{I,r})^T\ \lambda_I||^2_{l^2}\ \leq\ Ch^2||\lambda_I||^2_{l^2}.$$

Since (25) also holds for this case, i.e.,

$$||(\overline{B}^{\#})^T\lambda||^2_{l^2}\ \approx\ ||(B^{\#}_{I,r})^T\lambda_I||^2_{l^2}\ +\ ||\lambda_E||^2_{l^2},$$

we conclude that (18) is established. □

6 Solution of bound constrained quadratic programming problems and numerical scalability

In this section, we introduce and analyze an algorithm for solving the bound constrained quadratic programming problem (9), i.e., find

$$\min_{\lambda_I \geq 0} \Theta(\lambda), \tag{34}$$

where $\Theta(\lambda) = \frac{1}{2}\lambda^T F\lambda - \lambda^T b$. It is well known that a solution of the problem (34) always exists, and is necessarily unique; see, e.g., [4]. Let us briefly review some results on applying an active set strategy to solving bound constrained quadratic programming problems.

To simplify our notations, let us denote the dimension of the argument λ of $\Theta(\lambda)$ by n, and let g be the gradient of $\Theta(\lambda)$ at λ, i.e.,

$$g \;=\; g(\lambda) \;=\; F\lambda - b. \tag{35}$$

The unique solution $\overline{\lambda}$ of (34) is fully determined by the Karush-Kuhn-Tucker (KKT) optimality conditions; cf. [4]. To describe the KKT conditions in more detail, let $N = \{1, 2, \ldots, n\}$, let I denote the set of indexes of the constrained variables from problem (34), and let $E = N \setminus I$ denote the set of indexes of the unconstrained variables. Thus,

$$\overline{\lambda}_i = 0 \;\text{ and }\; i \in I \;\text{ implies }\; \overline{g}_i \geq 0, \quad \text{ and } \overline{\lambda}_i > 0 \;\text{ or }\; i \in E \;\text{ implies } \overline{g}_i = 0. \tag{36}$$

The set $A(\lambda)$ of all indexes $i \in I$ for which $\lambda_i = 0$ is called the *active set* of λ, i.e.,

$$A(\lambda) = \{i \in I : \lambda_i = 0\}.$$

The complement $F(\lambda) = N \setminus A(\lambda)$ of $A(\lambda)$ is called the *free set* of λ.

To enable an alternative reference to the KKT conditions (36), we introduce the *free gradient* $\varphi(\lambda)$ and the *chopped gradient* $\beta(\lambda)$ of λ, defined by

$$\varphi_i(\lambda) \;=\; \begin{cases} g_i(\lambda), & \text{for} \quad i \in F(\lambda) \\ 0, & \text{for} \quad i \in A(\lambda) \end{cases} \quad \text{and} \quad \beta_i(\lambda) \;=\; \begin{cases} 0, & \text{for} \quad i \in F(\lambda) \\ g_i^-(\lambda), & \text{for} \quad i \in A(\lambda) \end{cases}$$

where $g_i^- = \min\{g_i, 0\}$. The KKT conditions (36) are satisfied if and only if the *projected gradient* $\nu(\lambda) = \varphi(\lambda) + \beta(\lambda)$ is equal to zero, i.e., $\nu(\lambda) = 0$.

We call λ a *feasible* vector if $\lambda_i \geq 0$ for $i \in I$. The projection P_+ to the set of feasible vectors is defined for any n-vector λ by

$$P_+(\lambda)_i \;=\; \begin{cases} \max\{\lambda_i, 0\}, & \text{for} \quad i \in I \\ \lambda_i, & \text{for} \quad i \in E \end{cases}$$

Let us briefly describe the algorithm [13] for the solution of problem (34) that combines the proportioning algorithm [5] with gradient projections [38]. We use a given constant $\Gamma > 0$, a test to decide about leaving the face, and three types of steps to generate a sequence of iterates $\{\lambda^k\}$ that approximate the solution of (34).

The *expansion step* is defined by

$$\lambda^{k+1} \;=\; P_+\left(\lambda^k - \overline{\alpha}\varphi(\lambda^k)\right),$$

with a fixed steplength of size $\overline{\alpha} \in (0, ||F||^{-1}]$. This step may expand the current active set. To describe it without any reference to P_+, we introduce, for any feasible λ, the *reduced free gradient* $\widetilde{\varphi}(\lambda)$ defined by

$$\widetilde{\varphi}_i \;=\; \widetilde{\varphi}_i(\lambda) = \;=\; \begin{cases} \min\{\lambda_i/\overline{\alpha}, \varphi_i\}, & \text{for} \quad i \in I \\ \varphi_i, & \text{for} \quad i \in E \end{cases}$$

Note that

$$P_+ \left(\lambda - \overline{\alpha}\varphi(\lambda)\right) = \lambda - \overline{\alpha}\widetilde{\varphi}(\lambda).$$

We call the iterate λ^k *strictly proportional* if the following inequality holds:

$$||\beta(\lambda^k)||^2 \leq \Gamma^2 \widetilde{\varphi}(\lambda^k)^\top \varphi(\lambda^k) \tag{37}$$

The test (37) is used to decide which component of the projected gradient $\nu(\lambda^k)$ will be reduced in the next step.

The *proportioning step* is defined by

$$\lambda^{k+1} = \lambda^k - \alpha_{cg}\beta(\lambda^k).$$

The steplength α_{cg} is chosen to minimize $f(\lambda^k - \alpha\beta(\lambda^k))$ with respect to α, i.e.,

$$\alpha_{cg} \;=\; \frac{\beta(\lambda^k)^T g(\lambda^k)}{\beta(\lambda^k)^T F \beta(\lambda^k)}.$$

The purpose of the proportioning step is to remove indexes from the active set.

The *conjugate gradient step* is defined by

$$\lambda^{k+1} = \lambda^k - \alpha_{cg} p^k$$

where p^k is the conjugate gradient direction [2] which is constructed recurrently. The recurrence starts (or restarts) with $p^s = \varphi(\lambda^s)$ whenever λ^s is generated by the expansion step or the proportioning step. If p^k is known, then p^{k+1} is given by the formulae [2]

$$p^{k+1} = \varphi(\lambda^k) - \gamma p^k, \quad \text{with} \quad \gamma = \frac{\varphi(\lambda^k)^\top F p^k}{(p^k)^\top F p^k}.$$

The conjugate gradient steps are used to carry out efficiently the minimization in the face $W_J = \{\lambda : \lambda_i = 0 \text{ for } i \in J \}$ given by $J = A(\lambda^s)$.

The algorithm that we use may now be described as follows:

Algorithm 1 Modified proportioning with reduced gradient projections (MPRGP).
Let λ^0 be an n-vector such that $\lambda^0_i \geq 0$, for $i \in I$, let $\overline{\alpha} \in (0, \|F\|^{-1}]$, and let $\Gamma > 0$ be given. For $k \geq 0$ and λ^k known, choose λ^{k+1} as follows:
(i) If $\nu(\lambda^k) = 0$, set $\lambda^{k+1} = \lambda^k$.
(ii) If λ^k is strictly proportional and $\nu(\lambda^k) \neq 0$, try to generate λ^{k+1} by the conjugate gradient step. If $\lambda^{k+1}_i \geq 0$ for $i \in I$, then accept it, else generate λ^{k+1} by the expansion step.
(iii) If λ^k is not strictly proportional, define λ^{k+1} by proportioning.

For details about the implementation of the algorithm, we refer the reader to [13]. The basic properties of the algorithm are summed up in the following theorem:

Theorem 2. Let $\Gamma > 0$ be a given constant and let $\widehat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}$. Let $\alpha_1 = \lambda_{min}(F)$, let $\overline{\lambda}$ be the unique solution of (34), and denote by $\{\lambda^k\}$ the sequence generated by Algorithm 1 with $\overline{\alpha} \in (0, \|F\|^{-1}]$. The following statements hold:
(i) The rate of convergence in the energy norm defined by $||\lambda||^2_F = \lambda^T F \lambda$ is given by

$$||\lambda^k - \overline{\lambda}||^2_F \;\leq\; 2\eta^k \;\left(\Theta(\lambda^0) - \Theta(\overline{\lambda})\right), \tag{38}$$

where

$$\eta \;=\; 1 - \frac{\overline{\alpha}\alpha_1}{2 + 2\widehat{\Gamma}^2} \;\geq\; 1 - \frac{1}{\kappa(F)(2 + 2\widehat{\Gamma}^2)} \;\geq\; 1 - \frac{1}{4\kappa(F)}. \tag{39}$$

(ii) If the solution $\overline{\lambda}$ satisfies the strict complementarity conditions, i.e., if $\overline{\lambda}_i = 0$ implies $g_i(\overline{\lambda}) \neq 0$, then there exists $k \geq 0$ such that $\lambda^k = \overline{\lambda}$.
(iii) If Γ and the spectral condition number $\kappa(F)$ of F satisfy

$$\Gamma \ \geq \ 2\left(\sqrt{\kappa(F)}+1\right),$$

then there exists $k \geq 0$ such that $\lambda^k = \overline{\lambda}$.

Proof. See [13]. □

Theorem 3. Let C_1 and Γ denote given positive numbers, let C be a constant that satisfies (19), i.e., $C \leq \lambda_{min}(F)$, and let $\overline{\alpha} \in (0,\ C^{-1}C_1^{-2}]$.

We denote by $\{\lambda^i_{H,h}\}$ the iterates generated by Algorithm 1 with the initial approximation $\lambda^0 = \lambda^0_{H,h} = 0$ for the solution $\overline{\lambda}_{H,h}$ of problem (34)

Then there exists $\overline{\eta} < 1$ independent of h and H such that $H/h \leq C_1$ implies

$$\|\lambda^k_{H,h} - \overline{\lambda}_{H,h}\| \ \leq \ \frac{\overline{\eta}^k}{C^2}\|b\|^2. \tag{40}$$

Proof. Under the assumptions of the theorem,

$$\Theta(\overline{\lambda}_{H,h}) \ = \ \min\{\Theta(\lambda) : \lambda_I \geq 0\} \ \geq \ \Theta(F^{-1}b) \ = \ -\frac{1}{2}b^TF^{-1}b \ \geq \ -\frac{1}{2C}\|b\|^2,$$

since $C \leq \lambda_{min}(F)$. Recall that $\{\lambda^i_{H,h}\}$ denotes the iterates generated by Algorithm 1 with initial approximation $\lambda^0 = \lambda^0_{H,h} = 0$. From Theorem 2, we obtain that

$$\|\lambda^k_{H,h} - \overline{\lambda}_{H,h}\|^2_F \ \leq \ 2\eta^k\left(\Theta(\lambda^0_{H,h}) - \Theta(\overline{\lambda}_{H,h})\right) \ \leq \ \frac{\eta^k}{C}\|b\|^2, \tag{41}$$

where η is defined by (39). From (19), $\alpha_1 = \lambda_{min}(F) \geq C$, and we find that

$$\eta \ = \ 1 - \frac{\overline{\alpha}\alpha_1}{2+2\widehat{\Gamma}^2} \ \leq \ 1 - \frac{\overline{\alpha}C}{2+2\widehat{\Gamma}^2} \ = \ \overline{\eta} \ < \ 1. \tag{42}$$

Since $C \leq \lambda_{min}(F)$,

$$C\|\lambda^k_{H,h} - \overline{\lambda}_{H,h}\|^2 \ \leq \ \|\lambda^k_{H,h} - \overline{\lambda}_{H,h}\|^2_F. \tag{43}$$

Then (40) follows from (41–43). □

7 Numerical experiments

In this section, we report results for the numerical solution of the model coercive contact problem to illustrate the performance of our FETI–DP algorithm implemented in MATLAB. The goals of our experiments were as follows:

- to establish numerical evidence for the scalability of the algorithm;
- to compare the performance of the method for the cases when the subdomain partitions of Ω^1 and Ω^2 match, or do not match, across Γ_c, the potential contact interface;
- to investigate how does the convergence of the algorithm depend on the choice of nonmortar sides either on Γ^1_c or on Γ^2_c.

We used the traction function f specified by

$$f(x_1,x_2)=\left\{\begin{array}{ccl} -3 & \text{for} & (x_1,x_2)\in(0,1)\times[0.75,1) \\ 0 & \text{for} & (x_1,x_2)\in(0,1)\times[0,0.75) \text{ and } (x_1,x_2)\in(1,2)\times[0.25,1) \\ -1 & \text{for} & (x_1,x_2)\in(1,2)\times[0,0.25) \end{array}\right\}.$$

The solutions of our benchmarks with different decomposition and discretizations are in Figure 2.

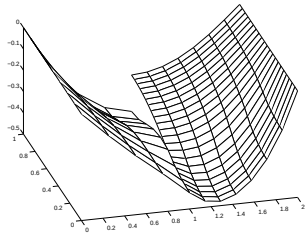

Figure 2a: Solution corresponding to a 4–subdomain partition on Ω^1 and a 7–subdomain partition on Ω^2

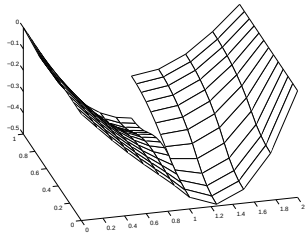

Figure 2b: Solution corresponding to a 7–subdomain partition on Ω^1 and a 4–subdomain partition on Ω^2

For matching subdomain partitions across Γ_c, we partitioned Ω^1 and Ω^2 into 1×1, 2×2, and 4×4 squares each, corresponding to $H_1=H_2\in\{1,1/2,1/4\}$. To avoid perfectly matching meshes, the number of nodes on each side of the square subdomains was chosen to be $H_1/h_1\in\{4,8,16\}$, in Ω^1, corresponding to $H_2/h_2\in\{7,13,25\}$, respectively, in Ω^2. In Table 1, we report the iteration count, i.e., the number of the conjugate gradient iterations required for the convergence of the solution of the problem to the given precision, as well as the size of the primal problem, of the dual problem, and the number of global corner degrees of freedom, i.e., the size of the coarse problem corresponding to solving a linear system for K_{cc}^*, for each partition described above.

Table 1: Convergence results: Matching subdomain partitions across Γ_c

				Nonmortars on Γ_c^1				Nonmortars on Γ_c^2			
N_1	N_2	$\frac{H_1}{h_1}$	$\frac{H_2}{h_2}$	Iter	primal	dual	corners	Iter	primal	dual	corners
1×1	1×1	4	7	6	89	5	0	13	89	8	0
		8	13	11	277	9	0	19	277	14	0
		16	25	16	965	17	0	25	965	26	0
2×2	2×2	4	7	20	356	44	9	30	256	50	9
		8	13	26	1108	92	9	36	1108	102	9
		16	25	31	3860	188	9	51	3860	206	9
4×4	4×4	4	7	28	1424	230	39	42	1424	242	39
		8	13	46	4432	486	39	63	4432	506	39
		16	25	60	15440	998	39	74	15440	1034	39

The algorithm converged after a small number of iterations for all partitions considered. For a fixed number of nodes per subdomain edge, i.e., for H_1/h_1 and H_2/h_2 simultaneously fixed, the number of iterations increased moderately when the number of subdomains quadrupled. Thus, numerical scalability of our method was observed for practical applications, and we inferred that the unspecified constants in Theorem 3 were not large.

The scalability of the method was observed regardless of whether the nonmortars were chosen on Γ_c^1 or on Γ_c^2. The difference between the two methods is given by the number of mortar conditions, and therefore by the number of Lagrange multipliers λ_I and by the size of the dual problem. There are more mortar conditions when the nonmortars are chosen on the edges with finer local mesh, i.e., on Γ_c^2. The number of iterations in this case was larger by about fifty percent than in the case when the nonmortars were chosen on the coarser local mesh, i.e., on Γ_c^1; see Table 1. This was due in part to the fact that the mortar non–penetration conditions had more of a local nature in the case of a finer local mesh. Therefore, little was gained by having more such conditions. This holds true with, possibly, the exception of a too coarse mesh on the nonmortar sides, i.e., $H_1 = H_2 = 1$ with $H_1/h_1 = 2$ or $H_1/h_1 = 3$. As a matter of fact, in this case, penetration may even occur at points on the contact interface due to the relative lack of non-penetration conditions.

For the case when the subdomain partitions across Γ_c do not match, Ω^1 was partitioned into 1×2, 2×4, and 4×8 rectangles, corresponding to partitions of Ω^2 into 1×3, 2×5, and 4×11 rectangles, respectively. The number of nodes on each side of the square subdomains was chosen to be, alternatively, in the set $\{(4,7),(8,13),(16,25)\}$; see Table 2 for more details. The iteration counts, the sizes of the primal and dual problems, and the sizes of the coarse problem are reported in Table 2.

Table 2: Convergence results: Non-matching subdomain partitions across Γ_c

				Nonmortars on Γ_c^1				Nonmortars on Γ_c^2			
N_1	N_2	$\frac{H_1}{h_1}$	$\frac{H_2}{h_2}$	Iter	primal	dual	corners	Iter	primal	dual	corners
1×2	1×3	4	7	15	242	23	5	41	242	35	5
		7	4	28	203	26	5	23	203	23	5
		8	13	29	750	47	5	48	750	69	5
		13	8	37	635	52	5	36	635	49	5
		16	25	33	2606	95	5	60	2606	137	5
		25	16	41	2219	104	5	41	2219	101	5
2×4	2×5	4	7	34	840	122	22	52	840	140	22
		7	4	43	762	125	22	38	762	116	22
		8	13	49	2608	256	22	68	2608	288	22
		13	8	58	2378	261	22	51	2378	248	22
		16	25	57	9072	524	22	87	9072	584	22
		25	16	67	8298	533	22	63	8298	512	22
4×8	4×11	4	7	49	3616	620	90	65	3616	662	90
		7	4	56	3148	581	90	50	3148	566	90
		8	13	59	11216	1298	90	88	11216	1374	90
		13	8	71	9836	1233	90	65	9836	1214	90
		16	25	78	38992	2654	90	125	38992	2798	90
		25	16	94	34348	2537	90	90	34348	2510	90

As before, for fixed number of nodes per subdomain edge, i.e., for H_1/h_1 and H_2/h_2 simultaneously fixed, the number of iterations increased moderately when the number of subdomains roughly quadrupled. Thus, numerical scalability of our method was once again observed, independent of whether the nonmortar sides were chosen on Γ_c^1, and on Γ_c^2. The number of iterations was larger when more non–penetration conditions were required, i.e., when the number of nodes on the nonmortars was larger. This was due to the fact that, for a mesh that is fine enough, some of the mortar non-penetration conditions become less

relevant.

In the experiments presented above, the rows of the matrix B_I were normalized as discussed in Section 5. We conclude this section by presenting numerical evidence that the performance of our FETI–DP method deteriorates unless the rows have norms of similar order; see Theorems 1 and 2 for an explanation of this phenomenon. In Table 3, we present the convergence results for matching subdomain partitions across Γ_c for the case when B_I is not normalized, i.e., when $B_I^{\#}$ is used instead of B_I in our algorithm. The nonmortar sides were chosen to be on Γ_c^1. To make the comparison easy, we also included the convergence results from Table 1 for the algorithm using the normalized matrix B_I.

Table 3: Convergence results: Normalized B_I vs. Non–normalized $B_I^{\#}$

				Normalized B_I				Non–normalized $B_I^{\#}$			
N_1	N_2	$\frac{H_1}{h_1}$	$\frac{H_2}{h_2}$	Iter	primal	dual	corners	Iter	primal	dual	corners
1×1	1×1	4	7	6	89	5	0	8	89	5	0
		8	13	11	277	9	0	14	277	9	0
		16	25	16	965	17	0	25	965	17	0
2×2	2×2	4	7	20	356	44	9	48	256	44	9
		8	13	26	1108	92	9	118	1108	92	9
		16	25	31	3860	188	9	268	3860	188	9
4×4	4×4	4	7	28	1424	230	39	106	1424	230	39
		8	13	46	4432	486	39	263	4432	486	39
		16	25	60	15440	998	39	743	15440	998	39

It is easy to see that the performance of the algorithm with non–normalized inequality constraints was much poorer and that this algorithm did not seem to be scalable. These numerical results are consistent with the theoretical condition number estimate from Theorem 1.

References

[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. Eng.*, 193:1367–1384, 2004.

[2] Owe Axelsson. *Iterative Solution Methods.* Cambridge University Press, New York, 1994.

[3] Christine Bernardi, Yvon Maday, and Anthony T. Patera. A new non conforming approach to domain decomposition: The mortar element method. In H. Brezzi et al., editor, *Nonlinear partial differential equations and their applications*, pages 13–51. Paris, 1994.

[4] Dimitri P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, 1999. Second Edition.

[5] Zdeněk Dostál. Box constrained quadratic programming with proportioning and projections. *SIAM J. Optimiz.*, 7:871–887, 1997.

[6] Zdeněk Dostál. A proportioning based algorithm for bound constrained quadratic programming with the rate of convergence. *Numerical Algorithms*, 34(2–4):293–302, 2003.

[7] Zdeněk Dostál, Francisco A. M. Gomes, and Sandra A. Santos. Duality based domain decomposition with natural coarse space for variational inequalities. *J. Comput. Appl. Math.*, 126:397–415, 2000.

[8] Zdeněk Dostál, Francisco A. M. Gomes, and Sandra A. Santos. Solution of contact problems by FETI domain decomposition with natural coarse space projection. *Comput. Meth. Appl. Mech. Eng.*, 190:1611–1627, 2000.

[9] Zdeněk Dostál and David Horák. Scalability and FETI based algorithm for large discretized variational inequalities. *Math. Comput. Simul.*, 3–6:347–357, 2003.

[10] Zdeněk Dostál and David Horák. On scalable algorithms for numerical solution of variational inequalities based on FETI and semi-monotonic augmented Lagrangians. In R. Kornhuber, R.H.W. Hoppe, D.E. Keyes, J. Périaux, O. Pironneau, and J. Xu, editors, *Proceedings of the 15th International Conference on Domain Decomposition Methods*, pages 487–494, Berlin, 2004. Springer-Verlag. Lectures Notes in Computational Science and Engineering.

[11] Zdeněk Dostál and David Horák. Scalable FETI with optimal dual penalty for a variational inequality. *Numer. Lin. Alg. Appl.*, 11(5–6):455–472, 2004.

[12] Zdeněk Dostál and David Horák. Theoretically supported scalable FETI for numerical solution of variational inequalities. *SIAM Journal on Numerical Analysis*, 45:500–513, 2007.

[13] Zdeněk Dostál and Joachim Schöberl. Minimizing quadratic functions over non-negative cone with the rate of convergence and finite termination. *Computational Optimization and Applications*, 30, 2005.

[14] Zdeněk Dostál, David Horák, and Dan Stefanica. A scalable FETI–DP algorithm for a coercive variational inequality. *IMA J. Appl. Numer. Math.*, 54, 2005.

[15] Zdeněk Dostál, David Horák, and Dan Stefanica. A scalable feti–dp algorithm for semi-coercive variational inequalities. *Computer Methods in Applied Mechanics and Engineering*, 196, 2007.

[16] Maksymilian Dryja, Axel Klawonn, and Olof B. Widlund. Dual-Primal FETI methods for three-dimensional elliptic problems with heterogeneous coefficients. *SIAM J. Numer. Anal.*, 40:159–179, 2002.

[17] Maksymilian Dryja and Olof B. Widlund. A FETI–DP method for a mortar discretization of elliptic problems. In Luca Pavarino and Andrea Toselli, editors, *Proceedings of the ETH Workshop on Domain Decomposition Methods, Zurich, 2001*. Springer-Verlag, 2002.

[18] Maksymilian Dryja and Olof B. Widlund. A generalized FETI–DP method for the mortar discretization of elliptic problems. In Ismael Herrera, David E. Keyes, Olof B. Widlund, and Robert Yates, editors, *Fourteenth International Conference on Domain Decomposition Methods*, pages 27–38. ddm.org, 2002.

[19] David Dureisseix and Charbel Farhat. A numerically scalable domain decomposition method for the solution of frictionless contact problems. *Int. J. Numer. Meth. Eng.*, 50:2643–2666, 2001.

[20] Charbel Farhat, Michel Lesoinne, Patrick Le Tallec, Kendall Pierson, and Daniel Rixen. FETI-DP: A Dual-Primal unified FETI method – part I: A faster alternative to the two-level FETI method. *Int. J. Numer. Meth. Eng.*, 50:1523–1544, 2001.

[21] Charbel Farhat, Antonini P. Macedo, and Michel Lesoinne. A two-level domain decomposition method for the iterative solution of high frequency exterior Helmholtz problems. *Numer. Math.*, 85:283–308, 2000.

[22] Charbel Farhat, Jan Mandel, and François-Xavier Roux. Optimal convergence properties of the FETI domain decomposition method. *Comput. Methods Appl. Mech. Eng.*, 115:367–388, 1994.

[23] Charbel Farhat and François-Xavier Roux. An unconventional domain decomposition method for an efficient parallel solution of large-scale finite element systems. *SIAM J. Sci. Comp.*, 13:379–396, 1992.

[24] Leopoldo Franca and Antonini P. Macedo. A two-level finite element method and its application to the Helmholtz equation. *Int. J. Numer. Meth. Eng.*, 43:23–32, 1998.

[25] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek. *Solution of Variational Inequalities in Mechanics*. Springer Verlag, Berlin, 1988.

[26] Hyea-Hyun Kim and Chang-Ock Lee. A FETI-DP formulation for two-dimensional stokes problem on nonmatching grids. In R. Kornhuber, R.H.W. Hoppe, D.E. Keyes, J. Périaux, O. Pironneau, and J. Xu, editors, *Proceedings of the 15th International Conference on Domain Decomposition Methods*, pages 353–360, Berlin, 2004. Springer-Verlag. Lectures Notes in Computational Science and Engineering.

[27] Axel Klawonn and Olof B. Widlund. A domain decomposition method with Lagrange multipliers for linear elasticity. *SIAM J. Sci. Comput.*, 22:1199–1219, 2000.

[28] 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.

[29] Ralf Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. Teubner, Stuttgart, 1997.

[30] Ralf Kornhuber and Rolf Krause. Adaptive multilevel methods for Signorini's problem in linear elasticity. *Comp. Visual. Sci.*, 4:9–20, 2001.

[31] Catherine Lacour and Yvon Maday. Two different approaches for matching nonconforming grids: the mortar element method and the FETI method. *BIT*, 37:720–738, 1997.

[32] Jing Li. A Dual-Primal FETI method for incompressible Stokes equations. Technical Report 816, Courant Institute of Mathematical Sciences, Department of Computer Sciences, 2001.

[33] Jan Mandel. Étude algébrique d'une méthode multigrille pour quelques problèmes de frontière libre. *C. R. Acad. Sci. Paris*, 298:469–472, 1984.

[34] Jan Mandel and Radek Tezaur. On the convergence of a Dual-Primal substructuring method. *Numer. Math.*, 88:543–558, 2001.

[35] Jan Mandel, Radek Tezaur, and Charbel Farhat. A scalable substructuring method by Lagrange multipliers for plate bending problems. *SIAM J. Numer. Anal.*, 36:1370–1391, 1999.

[36] Francesca Rapetti and Andrea Toselli. A FETI preconditioner for two dimensional edge element approximations of Maxwell's equations on non-matching grids. *SIAM J. Sci. Comp.*, 23(1):92–108, 2001.

[37] Daniel Rixen. Extended preconditioners for the FETI method applied to constrained problems. *Int. J. Num. Meth. Eng.*, 54:1–26, 2002.

[38] Joachim Schöberl. Solving the Signorini problem on the basis of domain decomposition techniques. *Computing*, 60:323–344, 1998.

[39] Joachim Schöberl. Efficient contact solvers based on domain decomposition techniques. *Comp. Math. Appl.*, 42:1217–1228, 2001.

[40] Dan Stefanica. A numerical study of FETI algorithms for mortar finite element methods. *SIAM J. Sci. Comp.*, 23(4):1135–1160, 2001.

[41] Dan Stefanica. FETI and FETI–DP methods for spectral and mortar spectral elements: A performance comparison. *J. Sci. Comp.*, 17:629–637, 2002.

[42] Dan Stefanica. Choosing nonmortars: Does it influence the performance of FETI–DP algorithms? In R. Kornhuber, R.H.W. Hoppe, D.E. Keyes, J. Périaux, O. Pironneau, and J. Xu, editors, *Proceedings of the 15th International Conference on Domain Decomposition Methods*, pages 377–384, Berlin, 2004. Springer-Verlag. Lectures Notes in Computational Science and Engineering.

[43] Dan Stefanica. Parallel FETI algorithms for mortars. *J. Appl. Numer. Math.*, 54:266–279, 2005.

[44] Andrea Toselli and Axel Klawonn. A FETI domain decomposition method for Maxwell's equations with discontinuous coefficients in two dimensions. *SIAM J. Numer. Anal.*, 39(3):932–956, 2001.

[45] K. F. Traore, C. Farhat, M. Lesoinne, and D. Dureisseix. A domain decomposition method with Lagrange multipliers for the massively parallel solution of large-scale contact problems. In *Proceedings of the Fifth World Congress on Computational Mechanics (WCCM V)*, 2002.

[46] B. Vereecke, H. Bavestrello, and D. Dureisseix. An extension of the FETI domain decomposition method for incompressible and nearly incompressible problems. *Comput. Methods Appl. Mech. Engrg.*, 192:3409–3429, 2003.

[47] Barbara Wohlmuth. A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM J. Numer. Anal.*, 38:989–1012, 2000.

[48] Barbara Wohlmuth and Rolf Krause. Monotone methods on non-matching grids for non-linear contact problems. *SIAM J. Sci. Comp.*, 25:324–347, 2003.

[49] Barbara I. Wohlmuth. *Discretization methods and iterative solvers based on domain decomposition*. Springer Verlag, 2001. Lectures Notes in Computational Science and Engineering, Vol. 17.