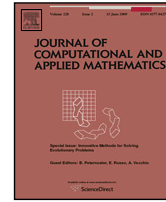




Contents lists available at ScienceDirect

# Journal of Computational and Applied Mathematics

journal homepage: [www.elsevier.com/locate/cam](http://www.elsevier.com/locate/cam)

## A scalable FETI–DP algorithm with non-penetration mortar conditions on contact interface

Zdeněk Dostál<sup>a</sup>, David Horák<sup>a</sup>, Dan Stefanica<sup>b,\*</sup><sup>a</sup> FEI VŠB–Technical University Ostrava, CZ-70833 Ostrava, Czech Republic<sup>b</sup> Baruch College, City University of New York, New York, NY 10010, United States

### ARTICLE INFO

#### Article history:

Received 30 November 2007

Received in revised form 13 November

2008

#### MSC:

65N30

65N55

#### Keywords:

Coercive variational inequalities

FETI–DP algorithms

Mortar finite elements

Modified proportioning algorithms

### 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.

© 2009 Elsevier B.V. All rights reserved.

### 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 in [1] as a parallel solver for problems described by elliptic partial differential equations. Later, Farhat, Mandel, and Roux [2] 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 [3,4] with improved scaling properties. By projecting the Lagrange multipliers in each iteration onto an auxiliary space to enforce continuity of the primal solutions at the crosspoints, Farhat, Mandel and Tezaur [5] obtained a faster converging FETI method for plate and shell problems. FETI algorithms were also implemented for, e.g., Helmholtz problems [6,7], linear elasticity with inexact solvers [8], Maxwell's equations [9,10], and Stokes problems [11–13].

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 [14] or by specialized quadratic programming methods [15–17]. Numerical experiments with these algorithms indicated their numerical scalability. Recently, using new results

\* Corresponding author.

E-mail addresses: [zdenek.dostal@vsb.cz](mailto:zdenek.dostal@vsb.cz) (Z. Dostál), [david.horak@vsb.cz](mailto:david.horak@vsb.cz) (D. Horák), [Dan\\_Stefanica@baruch.cuny.edu](mailto:Dan_Stefanica@baruch.cuny.edu) (D. Stefanica).

on the rate of convergence of improved versions [18,19] of the active set based proportioning algorithm [20], numerical scalability was proved for two variants of this algorithm in [21–23].

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 in [24]; see also [25]. 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 [26]. Farhat et al. [27,14,28] 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 [29,30] 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 in [31]. Biorthogonal mortar elements with slightly better computational properties were later developed in [32,33]. 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. [34,35,9,36,37].

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 [18,19] of the active set based proportioning algorithm [20]. 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 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 positive effect of the normalization of the constraints is enhanced in the preconditioner proposed in [38,39]. See also Stefanica and Klawonn [40]. However, the normalization is not required when the non-penetration is implemented by nodal constraints [29]. 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 in [41]. Kornhuber, Krause and Wohlmuth [42–44] introduced an algorithm based on monotone multigrid with scalable solution of auxiliary linear problems. Combining multigrid ideas and approximate projections, Schöberl [45,46] 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  $\Gamma^1$  and  $\Gamma^2$ , respectively. We denote by  $\Gamma_u^i$ ,  $\Gamma_f^i$ , and  $\Gamma_c^i$  the fixed, free, and potential contact parts of  $\Gamma^i$ ,  $i = 1, 2$ ; see Fig. 1a. Let  $\Gamma_c = \Gamma_c^1 \cup \Gamma_c^2$ .

The Sobolev space of the first order on  $\Omega^i$  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

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

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

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

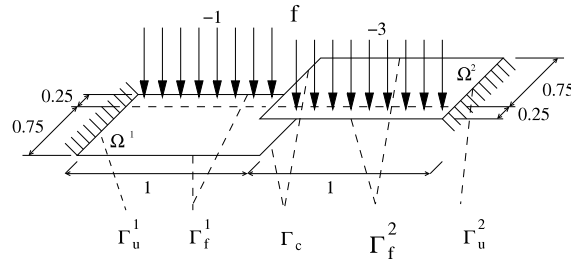


Fig. 1a. Coercive model problem.

We define the symmetric bilinear form  $a(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$  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  $\Omega^i$ ,  $i = 1, 2$ . We define the linear form  $\ell(\cdot) : \mathcal{H} \rightarrow \mathbb{R}$  by

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

and consider the following problem:

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

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 Fig. 1a and the left edge of the membrane  $\Omega^2$  is not allowed to penetrate below the right edge of the membrane  $\Omega^1$ . Note that  $\Gamma_u^i$ , the parts of  $\partial\Omega^i$ ,  $i = 1, 2$ , where Dirichlet conditions are prescribed, have positive Lebesgue measures. Thus, the quadratic form  $a(\cdot, \cdot)$  is coercive and the solution of problem (1) exists and is unique; cf. [47].

### 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  $\Omega^1$  and  $\Omega^2$ , 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 [48,34,49,50]. Weak continuity would be enforced by way of orthogonality of the jump of the mortar functions across the interfaces of the mortar partitions within  $\Omega^1$  and  $\Omega^2$ , respectively. This would be done using so-called mortar conditions. Across the potential contact boundary  $\Gamma_c$ , 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 [51] 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  $\Omega^i$ ,  $i = 1, 2$ , is partitioned on a rectangular grid into subdomains of diameter on the order of  $H$ . The restrictions of  $W$  to  $\Omega^1$  and  $\Omega^2$  are  $Q_1$  finite element spaces of comparable mesh sizes of order  $h$ , corresponding to the subdomain grids in  $\Omega^1$  and  $\Omega^2$ . Note that the subdomain grids do not necessarily match across the potential contact interface  $\Gamma_c$ . 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  $\Gamma_c$  are not crosspoints; see Fig. 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  $\Omega^1$  and  $\Omega^2$  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

$$B_E v = 0,$$

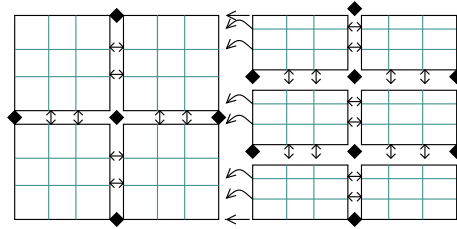


Fig. 1b. Decomposition of domains.

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  $\Gamma_c$ , the meshes on  $\Gamma_c^1$  and  $\Gamma_c^2$  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  $\Gamma_c^1$ , while the sides on  $\Gamma_c^2$  are called mortar sides. For each nonmortar side  $\gamma$ , we require the non-penetration conditions

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

where  $\Psi(\gamma)$  is a space of test functions having the same dimension as the number of interior nodes on  $\gamma$ . 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 [33]. From a theoretical point of view, the nonmortar sides could have been chosen on  $\Gamma_c^2$  as well, which would have rendered the sides on  $\Gamma_c^1$  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  $v_{\gamma}^1$  be the vector of the nodal values of  $v$  on  $\gamma \subset \Gamma_c^1$  and let  $v_{\gamma}^2$  be the vector of those nodal values of  $v$  on  $\Gamma_c^2$  that are opposite  $\gamma$ . The matrix formulation of (2) is

$$M_{\gamma} v_{\gamma}^1 - N_{\gamma} v_{\gamma}^2 \leq 0. \quad (3)$$

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  $\Gamma_c$ . All these inequality constraints can be written in matrix formulation as

$$B_I^{\#} v \leq 0, \quad (4)$$

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

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

$$B_I v \leq 0. \quad (5)$$

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

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

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 [29]. The main difference is the use of mortar non-penetration conditions. Also, global degrees of freedom are considered for the corner nodes on  $\Gamma_c$ .

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:

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

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  $\lambda_E$  and  $\lambda_l$  to enforce the inequality and redundancy constraints as follows:

$$L(v_r, v_c^g, \lambda_E, \lambda_l) = \frac{1}{2} [v_r^T \quad (L_c v_c^g)^T] K \begin{bmatrix} v_r \\ L_c v_c^g \end{bmatrix} - [v_r^T \quad (L_c v_c^g)^T] \begin{bmatrix} f_r \\ f_c \end{bmatrix} + v^T B_E^T \lambda_E + v^T B_l^T \lambda_l. \quad (8)$$

Let  $B_{l,r}$  and  $B_{l,c}$  be the matrices made of the columns of  $B_l$  corresponding to  $v_r$  and  $v_c$ , respectively; define  $B_{E,r}$  and  $B_{E,c}$  similarly. Then  $B_l = [B_{l,r} \quad B_{l,c}]$  and  $B_E = [B_{E,r} \quad B_{E,c}]$ . We can also group the parts of the Lagrange multiplier matrices  $B_l$  and  $B_E$  together with respect to the corner and remainder nodes as follows:

$$B_l = \begin{bmatrix} B_{l,r} \\ B_{E,r} \end{bmatrix} \quad \text{and} \quad B_c = \begin{bmatrix} B_{l,c} \\ B_{E,c} \end{bmatrix}.$$

The corner nodes on  $\Gamma_c^{-1}$  (with the exception for the end points of  $\Gamma_c$ ) 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  $\Gamma_c^2$ . This represents a natural departure from the algorithm suggested in [29]. As a result,  $B_{l,c}$  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

$$\lambda = \begin{bmatrix} \lambda_l \\ \lambda_E \end{bmatrix}.$$

Note that

$$\begin{bmatrix} B_l \\ B_E \end{bmatrix} = \begin{bmatrix} B_{l,r} & B_{l,c} \\ B_{E,r} & B_{E,c} \end{bmatrix} = [B_r \quad B_c].$$

Then

$$v^T B_E^T \lambda_E + v^T B_l^T \lambda_l = v_r^T B_r^T \lambda + (v_c^g)^T L_c^T B_c^T \lambda.$$

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  $L(v_r, v_c^g, \lambda_E, \lambda_l)$  over  $v_r$ , we rewrite (8) as

$$L(v_r, v_c^g, \lambda_E, \lambda_l) = \frac{1}{2} (v_r^T K_{rr} v_r + 2v_r^T K_{rc} L_c v_c^g + (v_c^g)^T L_c^T K_{cc} L_c v_c^g) - v_r^T f_r - (v_c^g)^T L_c^T f_c + v_r^T B_r^T \lambda + (v_c^g)^T L_c^T B_c^T \lambda$$

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  $v_c^g$ :

$$\begin{aligned} L_c(v_c^g, \lambda_E, \lambda_l) &= \frac{1}{2} (v_c^g)^T L_c^T K_{cc} L_c v_c^g - (v_c^g)^T L_c^T f_c + (v_c^g)^T L_c^T B_c^T \lambda - \frac{1}{2} (f_r - K_{rc} L_c v_c^g - B_r^T \lambda)^T K_{rr}^{-1} (f_r - K_{rc} L_c v_c^g - B_r^T \lambda) \\ &= \frac{1}{2} (v_c^g)^T K_{cc}^* v_c^g - (v_c^g)^T (\tilde{F}_{rc}^T \lambda + f_c^*) - \frac{1}{2} (f_r - B_r^T \lambda)^T K_{rr}^{-1} (f_r - B_r^T \lambda), \end{aligned}$$

where we used the following notations related to those from [24]:

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

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

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

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

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

Thus, maximizing  $L_\lambda$  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

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

### 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  $\bar{B} = [B_r \ 0]$  and let  $\bar{K}$  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

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

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  $\Omega^1$  and  $\Omega^2$ .

Note that  $F = \bar{B} \bar{K}^{-1} \bar{B}^T$  and therefore we find, as in Lemma 4.3 of [26], that

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

Let  $\|w\|_p$  and  $\|\lambda\|_2$  be the Euclidean norms of the primal and dual variables  $w$  and  $\lambda$ , respectively. Since  $w$  is a finite element function,

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

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

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

From (12)–(14), we find that

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

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

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

Using the last two chains of relations, we conclude that

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

Recall that

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

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  $\bar{B}^\#$  the matrix similar to  $\bar{B}$  without normalizing the rows of  $B_I$ , i.e., using  $B_I^\#$  instead of  $B_I$ :

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

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

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

$$\|\bar{B}^\top \lambda\|_2^2 \approx \|\lambda\|_2^2. \tag{17}$$

Or else, for non-normalized inequality conditions,

$$Ch^2 \|\lambda\|_2^2 \leq \|(\bar{B}^\#)^\top \lambda\|_2^2 \leq C \|\lambda\|_2^2. \tag{18}$$

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  $F^\# = \bar{B}^\# K^{-1} (\bar{B}^\#)^\top$ , 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

$$C \frac{\|\bar{B}^\top \lambda\|_2^2}{\|\lambda\|_2^2} \leq \lambda_{\min}(F) \leq \lambda_{\max}(F) \leq C \left(\frac{H}{h}\right)^2 \frac{\|\bar{B}^\top \lambda\|_2^2}{\|\lambda\|_2^2}. \tag{23}$$

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).  $\square$

**Proof of Lemma 1.** It is easy to see that

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

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

$$\|\bar{B}^\top \lambda\|_2^2 = \|B_{I,r}^\top \lambda_I + B_{E,r}^\top \lambda_E\|_2^2 = \|B_{I,r}^\top \lambda_I\|_2^2 + \|B_{E,r}^\top \lambda_E\|_2^2. \tag{24}$$

The entries in each row of  $B_{E,r}^\top$  are 0, except for two entries which are either 1 or  $-1$ . Therefore,  $\|B_{E,r}^\top \lambda_E\|_2^2 = 2\|\lambda_E\|_2^2$  and

$$\|\bar{B}^\top \lambda\|_2^2 \approx \|B_{I,r}^\top \lambda_I\|_2^2 + \|\lambda_E\|_2^2. \tag{25}$$

Let  $\gamma(j), j = 1 : N_{nm}$ , be the nonmortar sides on  $\Gamma_c$ , 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

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

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

$$\|B_{I,r}^T \lambda_I\|_2^2 \leq 2 \sum_{i=1}^{N_{nm}} \|B_{\gamma(i),r}^T \lambda_{I,j}\|_2^2. \tag{27}$$

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

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

where  $\text{Id}_{I,j}$  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., [32]. 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

$$\|B_{\gamma(j),r}^T \lambda_{I,j}\|_2^2 = \|\lambda_{I,j}\|_2^2 + \|P_{\gamma(j)}^T \lambda_{I,j}\|_2^2. \tag{29}$$

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

$$\|B_{I,r}^T \lambda_I\|_2^2 \geq \sum_{i=1}^{N_{nm}} \|\lambda_{I,j}\|_2^2 = \|\lambda_I\|_2^2. \tag{30}$$

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

$$\begin{aligned} \|P_{\gamma(j)}^T \lambda_{I,j}\|_2^2 &= \sup_{\psi \neq 0} \frac{\langle P_{\gamma(j)}^T \lambda_{I,j}, \psi \rangle^2}{\|\psi\|_2^2} = \sup_{\psi \neq 0} \frac{\langle \lambda_{I,j}, P_{\gamma(j)} \psi \rangle^2}{\|\psi\|_2^2} \\ &\leq \|\lambda_{I,j}\|_2^2 \sup_{\psi \neq 0} \frac{\|P_{\gamma(j)} \psi\|_2^2}{\|\psi\|_2^2}. \end{aligned} \tag{31}$$

Let  $\zeta(j)$  be the union of mortar sides from  $\Gamma_c^2$  opposite  $\gamma(j)$ . Then  $\psi$  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  $\psi$  on  $\gamma(j)$ . From the stability of the mortar projection we find that

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

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  $\psi$  and  $P_{\gamma(j)} \psi$  are vectors of nodal values of first order finite element functions, we find from (32) that

$$\|P_{\gamma(j)} \psi\|_2^2 \approx \frac{C}{h_{\gamma(j)}} \|P_{\gamma(j)} \psi\|_{L^2(\gamma(j))}^2 \leq \frac{C}{h_{\gamma(j)}} \|\psi\|_{L^2(\zeta(j))}^2 \leq C \frac{h_{\zeta(j)}}{h_{\gamma(j)}} \|\psi\|_2^2 \leq C \|\psi\|_2^2.$$

Therefore, from (31), it follows that

$$\|P_{\gamma(j)}^T \lambda_{I,j}\|_2^2 \leq C \|\lambda_{I,j}\|_2^2.$$

Using (29), we find that

$$\|\lambda_{I,j}\|_2^2 \leq \|B_{\gamma(j),r}^T \lambda_{I,j}\|_2^2 \leq C \|\lambda_{I,j}\|_2^2.$$



A bound for the norm of  $B_{I,r}^T \lambda_I$  can now be established using (27) and (30), and the fact that  $\|\lambda_I\|_2^2 = \sum_{i=1}^{N_{mm}} \|\lambda_{I,j}\|_2^2$  satisfies the following inequality:

$$\|\lambda_I\|_2^2 \leq \|B_{I,r}^T \lambda_I\|_2^2 \leq C \|\lambda_I\|_2^2. \tag{33}$$

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

$$\|\bar{B}^T \lambda\|_2^2 \approx \|B_{I,r}^T \lambda_I\|_2^2 + \|\lambda_E\|_2^2 \approx \|\lambda_I\|_2^2 + \|\lambda_E\|_2^2 = \|\lambda\|_2^2.$$

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

$$Ch^2 \|\lambda_I\|_2^2 \leq \|(B_{I,r}^\#)^T \lambda_I\|_2^2 \leq Ch^2 \|\lambda_I\|_2^2.$$

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

$$\|(\bar{B}^\#)^T \lambda\|_2^2 \approx \|(B_{I,r}^\#)^T \lambda_I\|_2^2 + \|\lambda_E\|_2^2,$$

we conclude that (18) is established.  $\square$

### 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., [52]. 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  $\lambda$  of  $\Theta(\lambda)$  by  $n$ , and let  $g$  be the gradient of  $\Theta(\lambda)$  at  $\lambda$ , i.e.,

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

The unique solution  $\bar{\lambda}$  of (34) is fully determined by the Karush-Kuhn-Tucker (KKT) optimality conditions; cf. [52]. To describe the KKT conditions in more detail, let  $N = \{1, 2, \dots, n\}$ , let  $I$  denote the set of indices of the constrained variables from problem (34), and let  $E = N \setminus I$  denote the set of indices of the unconstrained variables. Thus,

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

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

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

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

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

We call  $\lambda$  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  $\lambda$  by

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

Let us briefly describe the algorithm [19] for the solution of problem (34) that combines the proportioning algorithm [20] with gradient projections [45]. 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

$$\lambda^{k+1} = P_+(\lambda^k - \bar{\alpha} \varphi(\lambda^k)),$$

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

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

Note that

$$P_+(\lambda - \bar{\alpha}\varphi(\lambda)) = \lambda - \bar{\alpha}\tilde{\varphi}(\lambda).$$

We call the iterate  $\lambda^k$  *strictly proportional* if the following inequality holds:

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

The test (37) is used to decide which component of the projected gradient  $v(\lambda^k)$  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  $\alpha_{cg}$  is chosen to minimize  $f(\lambda^k - \alpha\beta(\lambda^k))$  with respect to  $\alpha$ , i.e.,

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

The purpose of the proportioning step is to remove indices 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 [53] which is constructed recurrently. The recurrence starts (or restarts) with  $p^s = \varphi(\lambda^s)$  whenever  $\lambda^s$  is generated by the expansion step or the proportioning step. If  $p^k$  is known, then  $p^{k+1}$  is given by the formulae [53]

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

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  $\lambda^0$  be an  $n$ -vector such that  $\lambda_i^0 \geq 0$ , for  $i \in I$ , let  $\bar{\alpha} \in (0, \|F\|^{-1}]$ , and let  $\Gamma > 0$  be given. For  $k \geq 0$  and  $\lambda^k$  known, choose  $\lambda^{k+1}$  as follows:

- (i) If  $v(\lambda^k) = 0$ , set  $\lambda^{k+1} = \lambda^k$ .
- (ii) If  $\lambda^k$  is strictly proportional and  $v(\lambda^k) \neq 0$ , try to generate  $\lambda^{k+1}$  by the conjugate gradient step. If  $\lambda_i^{k+1} \geq 0$  for  $i \in I$ , then accept it, else generate  $\lambda^{k+1}$  by the expansion step.
- (iii) If  $\lambda^k$  is not strictly proportional, define  $\lambda^{k+1}$  by proportioning.

For details about the implementation of the algorithm, we refer the reader to [19]. The basic properties of the algorithm are summed up in the following theorem:

**Theorem 2.** Let  $\Gamma > 0$  be a given constant and let  $\hat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}$ . Let  $\alpha_1 = \lambda_{\min}(F)$ , let  $\bar{\lambda}$  be the unique solution of (34), and denote by  $\{\lambda^k\}$  the sequence generated by Algorithm 1 with  $\bar{\alpha} \in (0, \|F\|^{-1}]$ . The following statements hold:

- (i) The rate of convergence in the energy norm defined by  $\|\lambda\|_F^2 = \lambda^\top F\lambda$  is given by

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

where

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

- (ii) If the solution  $\bar{\lambda}$  satisfies the strict complementarity conditions, i.e., if  $\bar{\lambda}_i = 0$  implies  $g_i(\bar{\lambda}) \neq 0$ , then there exists  $k \geq 0$  such that  $\lambda^k = \bar{\lambda}$ .
- (iii) If  $\Gamma$  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  $\lambda^k = \bar{\lambda}$ .

**Proof.** See [19].  $\square$

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

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

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

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

**Proof.** Under the assumptions of the theorem,

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

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

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

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

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

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

$$C \|\lambda_{H,h}^k - \bar{\lambda}_{H,h}\|^2 \leq \|\lambda_{H,h}^k - \bar{\lambda}_{H,h}\|_F^2. \tag{43}$$

Then (40) follows from (41)–(43).  $\square$

## 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  $\Omega^1$  and  $\Omega^2$  match, or do not match, across  $\Gamma_c$ , the potential contact interface;
- to investigate how the convergence of the algorithm depends on the choice of nonmortar sides either on  $\Gamma_c^1$  or on  $\Gamma_c^2$ .

We used the traction function  $f$  specified by

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

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

For matching subdomain partitions across  $\Gamma_c$ , we partitioned  $\Omega^1$  and  $\Omega^2$  into  $1 \times 1$ ,  $2 \times 2$ , and  $4 \times 4$  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  $\Omega^1$ , corresponding to  $H_2/h_2 \in \{7, 13, 25\}$ , respectively, in  $\Omega^2$ . 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.

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  $\Gamma_c^1$  or on  $\Gamma_c^2$ . The difference between the two methods is given by the number of mortar conditions, and therefore by the number of Lagrange multipliers  $\lambda_l$  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 size, i.e., on  $\Gamma_c^2$ . The number of iterations in this case was larger by about fifty percent than

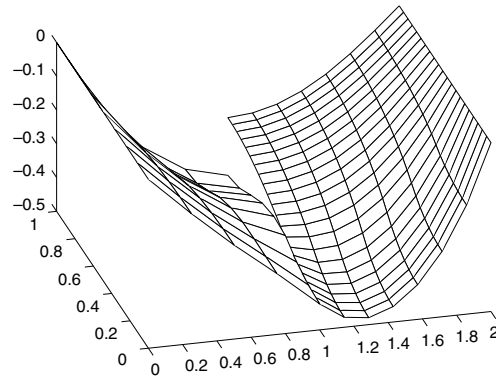


Fig. 2a. Solution corresponding to a 4-subdomain partition on  $\Omega^1$  and a 7-subdomain partition on  $\Omega^2$ .

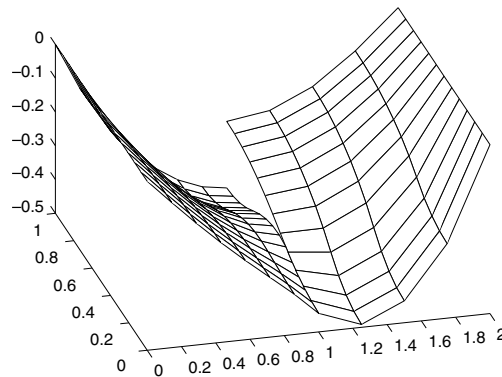


Fig. 2b. Solution corresponding to a 7-subdomain partition on  $\Omega^1$  and a 4-subdomain partition on  $\Omega^2$ .

Table 1

Convergence results: Matching subdomain partitions across  $\Gamma_c$ .

$N_1$	$N_2$	$\frac{H_1}{h_1}$	$\frac{H_2}{h_2}$	Nonmortars on $\Gamma_c^1$				Nonmortars on $\Gamma_c^2$			
				Iter	Primal	Dual	Corners	Iter	Primal	Dual	Corners
$1 \times 1$	$1 \times 1$	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
$2 \times 2$	$2 \times 2$	4	7	20	356	44	9	30	256	50	9
		8	13	26	1 108	92	9	36	1 108	102	9
		16	25	31	3860	188	9	51	3860	206	9
$4 \times 4$	$4 \times 4$	4	7	28	1 424	230	39	42	1 424	242	39
		8	13	46	4 432	486	39	63	4 432	506	39
		16	25	60	15 440	998	39	74	15 440	1034	39

in the case when the nonmortars were chosen on the coarser local mesh, i.e., on  $\Gamma_c^1$ ; 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 too coarse a 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  $\Gamma_c$  do not match,  $\Omega^1$  was partitioned into  $1 \times 2$ ,  $2 \times 4$ , and  $4 \times 8$  rectangles, corresponding to partitions of  $\Omega^2$  into  $1 \times 3$ ,  $2 \times 5$ , and  $4 \times 11$  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.

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  $\Gamma_c^1$ , and on  $\Gamma_c^2$ . The number of iterations was larger when more non-penetration conditions were required, i.e., when the number of nodes on

**Table 2**  
Convergence results: Non-matching subdomain partitions across  $\Gamma_c$ .

$N_1$	$N_2$	$\frac{h_1}{h_1}$	$\frac{h_2}{h_2}$	Nonmortars on $\Gamma_c^1$				Nonmortars on $\Gamma_c^2$			
				Iter	Primal	Dual	Corners	Iter	Primal	Dual	Corners
$1 \times 2$	$1 \times 3$	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	2 606	95	5	60	2 606	137	5
		25	16	41	2 219	104	5	41	2 219	101	5
$2 \times 4$	$2 \times 5$	4	7	34	840	122	22	52	840	140	22
		7	4	43	762	125	22	38	762	116	22
		8	13	49	2 608	256	22	68	2 608	288	22
		13	8	58	2 378	261	22	51	2 378	248	22
		16	25	57	9 072	524	22	87	9 072	584	22
		25	16	67	8 298	533	22	63	8 298	512	22
$4 \times 8$	$4 \times 11$	4	7	49	3 616	620	90	65	3 616	662	90
		7	4	56	3 148	581	90	50	3 148	566	90
		8	13	59	11 216	1298	90	88	11 216	1374	90
		13	8	71	9 836	1233	90	65	9 836	1214	90
		16	25	78	38 992	2 654	90	125	38 992	2 798	90
		25	16	94	34 348	2 537	90	90	34 348	2 510	90

**Table 3**  
Convergence results: Normalized  $B_I$  vs. non-normalized  $B_I^\#$ .

$N_1$	$N_2$	$\frac{h_1}{h_1}$	$\frac{h_2}{h_2}$	Normalized $B_I$				Non-normalized $B_I^\#$			
				Iter	Primal	Dual	Corners	Iter	Primal	Dual	Corners
$1 \times 1$	$1 \times 1$	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
$2 \times 2$	$2 \times 2$	4	7	20	356	44	9	48	256	44	9
		8	13	26	1 108	92	9	118	1 108	92	9
		16	25	31	3 860	188	9	268	3 860	188	9
$4 \times 4$	$4 \times 4$	4	7	28	1 424	230	39	106	1 424	230	39
		8	13	46	4 432	486	39	263	4 432	486	39
		16	25	60	15 440	998	39	743	15 440	998	39

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  $\Gamma_c$  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  $\Gamma_c^1$ . To make the comparison easy, we also included the convergence results from Table 1 for the algorithm using the normalized matrix  $B_I$ .

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.

**Acknowledgments**

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.

**References**

[1] Charbel Farhat, François-Xavier Roux, An unconventional domain decomposition method for an efficient parallel solution of large-scale finite element systems, *SIAM J. Sci. Comput.* 13 (1992) 379–396.  
 [2] Charbel Farhat, Jan Mandel, François-Xavier Roux, Optimal convergence properties of the FETI domain decomposition method, *Comput. Methods Appl. Mech. Eng.* 115 (1994) 367–388.  
 [3] Axel Klawonn, Olof B. Widlund, FETI and Neumann–Neumann iterative substructuring methods: Connections and new results, *Comm. Pure Appl. Math.* 54 (1) (2001) 57–90.

- [4] Daniel Rixen, Extended preconditioners for the FETI method applied to constrained problems, *Int. J. Num. Meth. Eng.* 54 (2002) 1–26.
- [5] Jan Mandel, Radek Tezaur, Charbel Farhat, A scalable substructuring method by Lagrange multipliers for plate bending problems, *SIAM J. Numer. Anal.* 36 (1999) 1370–1391.
- [6] Charbel Farhat, Antonini P. Macedo, Michel Lesoinne, A two-level domain decomposition method for the iterative solution of high frequency exterior Helmholtz problems, *Numer. Math.* 85 (2000) 283–308.
- [7] Leopoldo Franca, Antonini P. Macedo, A two-level finite element method and its application to the Helmholtz equation, *Int. J. Numer. Meth. Eng.* 43 (1998) 23–32.
- [8] Axel Klawonn, Olof B. Widlund, A domain decomposition method with Lagrange multipliers for linear elasticity, *SIAM J. Sci. Comput.* 22 (2000) 1199–1219.
- [9] Francesca Rapetti, Andrea Toselli, A FETI preconditioner for two dimensional edge element approximations of Maxwell's equations on non-matching grids, *SIAM J. Sci. Comput.* 23 (1) (2001) 92–108.
- [10] Andrea Toselli, Axel Klawonn, A FETI domain decomposition method for Maxwell's equations with discontinuous coefficients in two dimensions, *SIAM J. Numer. Anal.* 39 (3) (2001) 932–956.
- [11] Hyea-Hyun Kim, 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, J. Xu (Eds.), *Proceedings of the 15th International Conference on Domain Decomposition Methods*, in: *Lectures Notes in Computational Science and Engineering*, Springer-Verlag, Berlin, 2004, pp. 353–360.
- [12] Jing Li, A Dual–Primal FETI method for incompressible Stokes equations, Technical Report 816, Courant Institute of Mathematical Sciences, Department of Computer Sciences, 2001.
- [13] B. Verecke, H. Bavestrello, D. Dureisseix, An extension of the FETI domain decomposition method for incompressible and nearly incompressible problems, *Comput. Methods Appl. Mech. Engrg.* 192 (2003) 3409–3429.
- [14] David Dureisseix, Charbel Farhat, A numerically scalable domain decomposition method for the solution of frictionless contact problems, *Int. J. Numer. Meth. Eng.* 50 (2001) 2643–2666.
- [15] Zdeněk Dostál, Francisco A.M. Gomes, Sandra A. Santos, Duality based domain decomposition with natural coarse space for variational inequalities, *J. Comput. Appl. Math.* 126 (2000) 397–415.
- [16] Zdeněk Dostál, Francisco A.M. Gomes, Sandra A. Santos, Solution of contact problems by FETI domain decomposition with natural coarse space projection, *Comput. Meth. Appl. Mech. Eng.* 190 (2000) 1611–1627.
- [17] Zdeněk Dostál, David Horák, Scalability and FETI based algorithm for large discretized variational inequalities, *Math. Comput. Simul.* 3–6 (2003) 347–357.
- [18] Zdeněk Dostál, A proportioning based algorithm for bound constrained quadratic programming with the rate of convergence, *Numer. Algorithms* 34 (2–4) (2003) 293–302.
- [19] Zdeněk Dostál, Joachim Schöberl, Minimizing quadratic functions over non-negative cone with the rate of convergence and finite termination, *Comput. Optim. Appl.* 30 (2005).
- [20] Zdeněk Dostál, Box constrained quadratic programming with proportioning and projections, *SIAM J. Optim.* 7 (1997) 871–887.
- [21] Zdeněk Dostál, 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, J. Xu (Eds.), *Proceedings of the 15th International Conference on Domain Decomposition Methods*, in: *Lectures Notes in Computational Science and Engineering*, Springer-Verlag, Berlin, 2004, pp. 487–494.
- [22] Zdeněk Dostál, David Horák, Scalable FETI with optimal dual penalty for a variational inequality, *Numer. Linear Algebra Appl.* 111 (5–6) (2004) 455–472.
- [23] Zdeněk Dostál, David Horák, Theoretically supported scalable FETI for numerical solution of variational inequalities, *SIAM J. Numer. Anal.* 45 (2007) 500–513.
- [24] Charbel Farhat, Michel Lesoinne, Patrick Le Tallec, Kendall Pierson, 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 (2001) 1523–1544.
- [25] Maksymilian Dryja, Axel Klawonn, Olof B. Widlund, Dual–Primal FETI methods for three-dimensional elliptic problems with heterogeneous coefficients, *SIAM J. Numer. Anal.* 40 (2002) 159–179.
- [26] Jan Mandel, Radek Tezaur, On the convergence of a Dual–Primal substructuring method, *Numer. Math.* 88 (2001) 543–558.
- [27] Philip Avery, Gert Rebel, Michel Lesoinne, 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 (2004) 1367–1384.
- [28] K.F. Traore, C. Farhat, M. Lesoinne, 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.
- [29] Zdeněk Dostál, David Horák, Dan Stefanica, A scalable FETI–DP algorithm for a coercive variational inequality, *IMA J. Appl. Numer. Math.* 54 (2005).
- [30] Zdeněk Dostál, David Horák, Dan Stefanica, A scalable FETI–DP algorithm for semi-coercive variational inequalities, *Comput. Methods Appl. Mech. Engrg.* 196 (2007).
- [31] Christine Bernardi, Yvon Maday, Anthony T. Patera, A new non conforming approach to domain decomposition: The mortar element method, in: H. Brezzi et al. (Ed.), *Nonlinear Partial Differential Equations and their Applications*, Paris, 1994, pp. 13–51.
- [32] Barbara Wohlmuth, A mortar finite element method using dual spaces for the Lagrange multiplier, *SIAM J. Numer. Anal.* 38 (2000) 989–1012.
- [33] Barbara I. Wohlmuth, Discretization methods and iterative solvers based on domain decomposition, in: *Lectures Notes in Computational Science and Engineering*, vol. 17, Springer Verlag, 2001.
- [34] Maksymilian Dryja, Olof B. Widlund, A generalized FETI–DP method for the mortar discretization of elliptic problems, in: Ismael Herrera, David E. Keyes, Olof B. Widlund, Robert Yates (Eds.), *Fourteenth International Conference on Domain Decomposition Methods*, 2002, pp. 27–38. <http://www.ddm.org>.
- [35] Catherine Lacour, Yvon Maday, Two different approaches for matching nonconforming grids: The mortar element method and the FETI method, *BIT* 37 (1997) 720–738.
- [36] Dan Stefanica, A numerical study of FETI algorithms for mortar finite element methods, *SIAM J. Sci. Comput.* 23 (4) (2001) 1135–1160.
- [37] Dan Stefanica, FETI and FETI–DP methods for spectral and mortar spectral elements: A performance comparison, *J. Sci. Comput.* 17 (2002) 629–637.
- [38] Catherine Lacour, *Analyse et Résolution Numérique de Méthodes de Sous-Domains Non Conformés pour des Problèmes de Plaques*, Ph.D. Thesis, Université Pierre et Marie Curie, Paris, 1997.
- [39] Catherine Lacour, Iterative substructuring preconditioners for the mortar finite element method, in: Petter E. Bjørstad, Magne Espedal, David Keyes (Eds.), *Domain Decomposition Methods in Sciences and Engineering* (Proceedings from the Ninth International Conference, June 1996, Bergen, Norway), John Wiley & Sons, 1997, pp. 406–412.
- [40] Dan Stefanica, Axel Klawonn, The FETI method for mortar finite elements, in: C.-H. Lai, P. Bjørstad, M. Cross, O. Widlund (Eds.), *Eleventh International Conference on Domain Decomposition Methods*, 1999, pp. 121–129. <http://www.ddm.org>.
- [41] Jan Mandel, Étude algébrique d'une méthode multigrille pour quelques problèmes de frontière libre, *C. R. Acad. Sci. Paris* 298 (1984) 469–472.
- [42] Ralf Kornhuber, Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems, Teubner, Stuttgart, 1997.
- [43] Ralf Kornhuber, Rolf Krause, Adaptive multilevel methods for Signorini's problem in linear elasticity, *Comput. Vis. Sci.* 4 (2001) 9–20.
- [44] Barbara Wohlmuth, Rolf Krause, Monotone methods on non-matching grids for non-linear contact problems, *SIAM J. Sci. Comput.* 25 (2003) 324–347.
- [45] Joachim Schöberl, Solving the Signorini problem on the basis of domain decomposition techniques, *Computing* 60 (1998) 323–344.
- [46] Joachim Schöberl, Efficient contact solvers based on domain decomposition techniques, *Comput. Math. Appl.* 42 (2001) 1217–1228.
- [47] I. Hlaváček, J. Haslinger, J. Nečas, J. Lovíšek, *Solution of Variational Inequalities in Mechanics*, Springer Verlag, Berlin, 1988.
- [48] Maksymilian Dryja, Olof B. Widlund, A FETI–DP method for a mortar discretization of elliptic problems, in: Luca Pavarino, Andrea Toselli (Eds.), *Proceedings of the ETH Workshop on Domain Decomposition Methods*, Zurich, 2001, Springer-Verlag, 2002.

- [49] 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, J. Xu (Eds.), *Proceedings of the 15th International Conference on Domain Decomposition Methods*, in: *Lectures Notes in Computational Science and Engineering*, Springer-Verlag, Berlin, 2004, pp. 377–384.
- [50] Hyea Hyun Kim, A FETI–DP formulation of three dimensional elasticity problems with mortar discretization, *SIAM J. Sci. Comput.* 46 (5) (2008).
- [51] Dan Stefanica, Parallel FETI algorithms for mortars, *J. Appl. Numer. Math.* 54 (2005) 266–279.
- [52] Dimitri P. Bertsekas, *Nonlinear Programming*, second ed., Athena Scientific, Belmont, 1999.
- [53] Owe Axelsson, *Iterative Solution Methods*, Cambridge University Press, New York, 1994.