Duality-based domain decomposition with natural coarse-space for variational inequalities *

Zdeněk Dostál † Francisco A. M. Gomes Neto ‡
Sandra A. Santos §

October 5, 1998

*This work was supported by grant FAPESP No. 97/12676-4, CNPq, FINEP, FAEP–UNICAMP, and by grants GA ČR No. 201/97/0421 and 101/98/0535.

†Technical University of Ostrava, Ostrava, on visit to Department of Applied Mathematics, IMECC – UNICAMP, University of Campinas, (Zdenek.Dostal@vsb.cz)

‡Department of Applied Mathematics, IMECC – UNICAMP, University of Campinas, CP 6065, 13081–970 Campinas SP, Brazil (chico@ime.unicamp.br).

§Department of Mathematics, IMECC – UNICAMP, University of Campinas, CP 6065, 13081–970 Campinas SP, Brazil (sandra@ime.unicamp.br).

Abstract

An efficient non-overlapping domain decomposition algorithm of Neumann-Neumann type for solving variational inequalities arising from the elliptic boundary value problems with inequality boundary conditions has been presented. The discretized problem is first turned by the duality theory of convex programming into a quadratic programming problem with bound and equality constraints and the latter is further modified by means of orthogonal projectors to the natural coarse space introduced recently by Farhat and Roux. The resulting problem is then solved by an augmented Lagrangian type algorithm with an outer loop for the Lagrange multipliers for the equality constraints and an inner loop for the solution of the bound constrained quadratic programming problems. The projectors are shown to guarantee an optimal rate of convergence of iterative solution of auxiliary linear problems. Reported theoretical results and numerical experiments indicate high numerical and parallel scalability of the algorithm.

Key words and phrases:

Domain decomposition, natural coarse subspace, variational inequalities, quadratic programming

1 Introduction

Duality based domain decomposition methods proved to be practical and efficient tools for parallel solution of elliptic boundary value problems [18, 19, 30]. A given spatial domain is partitioned into non-overlapping subdomains, for each subdomain is defined an elliptic problem with Neumann boundary conditions on the subdomain interfaces, and intersubdomain field continuity is enforced via Lagrange multipliers. The Lagrange multipliers are evaluated by solving a relatively well conditioned dual problem of small size that may be efficiently solved by a suitable variant of the conjugate gradient algorithm. The first practical implementations by Farhat and Roux [18, 19] exploited only the favorable distribution of the spectrum of the matrix of the smaller problem [29] known also as the dual Schur complement matrix, but such algorithm was efficient only with a small number of subdomains. Later, they introduced a "natural coarse problem" whose solution was implemented by auxiliary projectors so that the resulting algorithm became optimal [20, 30].

It has been soon observed that the duality based domain decomposition methods may be at least as successful for the solution of variational inequalities as they are for linear problems. The first observation was that the duality not only reduces the dimension and improves conditioning of the original problem, but also reduces all the inequalities to simple bounds on variables [8, 9, 13] so that the dual problem may be solved much more efficiently than the primal problem [22, 10, 12].

Our goal here is to exploit the projector to the solution of the "natural coarse problem" to improve numerical scalability of our previous algorithms for numerical solution of variational inequalities. In particular, it turns out that application of the projectors decomposes the Hessian of the augmented Lagrangian so that it has at most one point of the spectrum outside the span of the spectrum of the dual Schur complement, and that iterative solution of auxiliary linear problems with such Hessian has an optimal rate of convergence. Since this feature of the algorithm is not exploited in the solution of linear problems, we believe that the algorithm will be at least as useful for the solution of variational inequalities as the related algorithm by Farhat and Roux is for linear problems.

The applications include the problem of finding the stresses and displacements of a system of linear elastic bodies without friction [28, 27] or the contact problem with a "given" Coulomb friction [27] that may be used to compute the solution of contact problems with Coulomb friction [27, 15]. Some other problems of this type may be found in Duvant and Lions [16] or Glowinski et al. [25].

To simplify our exposition, we restrict our considerations to a simple semicoercive model problem, that is described both in continuous and discrete versions in Sections 2 and 3. Then, in section 4, we use duality theory to reduce the discretized problem to a quadratic programming problem with simple bounds and equality constraints. The modification of the problem to enhance projectors to the natural coarse space is described in Section 5 together with results about distribution of the spectrum of the augmented Lagrangian of the modified problem. Details on the quadratic programming algorithms that we use are given in Section 6. First we describe the algorithm for quadratic programming problems with equality constraints and simple bounds. The approximations of the Lagrange multipliers for the equalities are generated in the outer loop of the augmented Lagrangian algorithm while the bound constrained subproblems are solved in the inner loop. We have adapted the basic scheme proposed by Conn, Gould and Toint [6] for the solution of more general problems. However, we have used the special structure of our problem to improve the performance of the algorithm. The precision of the solution of the auxiliary problems in the inner loop is controlled by the norm of feasibility of the current iterate and an estimate of the rate of convergence is given that has no term that accounts for the inexact solution of the auxiliary problems. Then we describe the algorithm for solving the bound constrained quadratic programming problems [22, 10, 21, 23, 4] in the inner loop. Our active set type algorithm generates search directions by the conjugate gradient mehod with optional preconditioning [3], exploits the projections on the feasible set, and uses the adaptive precision control for computing the solution of auxiliary problems. Theoretical results on convergence, robustness and optimality of the algorithm are reported. Results of numerical experiments that demonstrate the power of our algorithms are given in Section 7. Finally, in Section 8, some comments and conclusions are presented.

2 A Model Problem

To simplify our exposition, we shall restrict our attention to a variational inequality problem arising from the variational formulation of a model problem with inequality boundary conditions. For completeness, we shall also briefly sketch the derivation of the inequality and results on the existence and uniqueness of the solution.

In particular, we shall consider the problem of finding a sufficiently smooth u so that

$$-\Delta u = f \quad \text{in} \quad \Omega = \Omega^1 \cup \Omega^2 \tag{2.1}$$

$$u^1 = 0 \quad \text{on} \quad \Gamma_u^1 \tag{2.2}$$

$$\frac{\partial u^i}{\partial n_i} = 0 \quad \text{on} \quad \Gamma_f^i,\ i = 1, 2 \tag{2.3}$$

$$u^2 - u^1 \geq 0 \quad \text{on} \quad \Gamma_c = \Gamma_c^1 = \Gamma_c^2 \tag{2.4}$$

$$\frac{\partial u^2}{\partial n_2} \geq 0 \quad \text{on} \quad \Gamma_c \tag{2.5}$$

$$\frac{\partial u^2}{\partial n_2}\left(u^2 - u^1\right) = 0 \quad \text{on} \quad \Gamma_c \tag{2.6}$$

$$\frac{\partial u^1}{\partial n_1} + \frac{\partial u^2}{\partial n_2} = 0 \quad \text{on} \quad \Gamma_c. \tag{2.7}$$

Here

$$\Omega^1 = (0,1) \times (0,1) \quad \text{and} \quad \Omega^2 = (1,2) \times (0,1)$$

denote open domains with boundaries Γ^1, Γ^2 and their parts Γ_u^1, Γ_f^i, Γ_c^i, formed by the sides of $\Omega^i, i = 1, 2$ as in Figure 1, and $n_k(x)$ denote the components of the outer unit normal at $x \in \Gamma^i$.

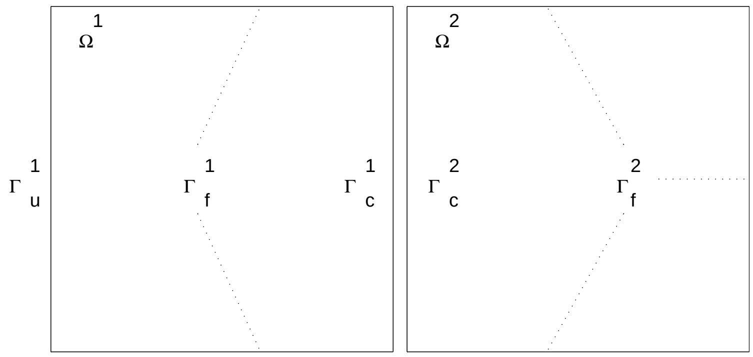

Figure 1: Domain of the model problem

The solution u of (2.1)-(2.7) may be interpreted as a vertical displacement of two membranes stretched by normalized horizontal forces and pressed together by vertical forces with density f. The inequality (2.4) describes the non-interpenetration of the adjacent edges of the membranes, with the edge of the right membrane above the edge of the left membrane. The right membrane can press the left membrane down (2.5) at the points that are in contact (2.6)-(2.7). If there is no contact at $x \in \Gamma_c$, i.e. $u^2(x) > u^1(x)$, then the membranes are stretched by the horizontal force in the same way as at $x \in \Gamma_f^i$. Other interpretations may be found in [16, 25].

To derive the variational inequality whose smooth solutions satisfy (2.1)-(2.7), let $H^1(\Omega^i)$ denote the Sobolev space of first order on the space $L^2(\Omega^i)$ of the functions on Ω^i whose squares are integrable in the sense of Lebesgue. Thus, $f \in H^1(\Omega^i)$ iff both f and its generalized first derivatives belong to $L^2(\Omega^i)$. Let

$$V^1 = \left\{ v \in H^1(\Omega^1) : v^1 = 0 \quad \text{on} \quad \Gamma_u^1 \right\}$$

denote the closed subspace of $H^1(\Omega^1)$, $V^2 = H^1(\Omega^2)$, and let

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

denote the closed subspace and the closed convex subset of $\mathcal{H} = H^1(\Omega^1) \times H^1(\Omega^2)$, respectively. The relations on the boundaries are in terms of traces [16, 27]. On $\mathcal{H}$ we shall define a symmetric bilinear form

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

and a linear form

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

Let u denote a smooth solution of (2.1)-(2.7). After multiplication of (2.1) by $v \in V$ and application of the Green theorem with simplifications based on (2.2)-(2.3), we get

$$a(u, v) - \ell(v) = \int_{\Gamma_c} \left\{ \frac{\partial u^1}{\partial n_1} v^1 + \frac{\partial u^2}{\partial n_2} v^2 \right\} d\Gamma.$$

In particular, for $v = w - u$ and $w \in \mathcal{K}$,

$$a(u, w - u) - \ell(w - u) = \int_{\Gamma_c} \left\{ \frac{\partial u^1}{\partial n_1} (w^1 - u^1) + \frac{\partial u^2}{\partial n_2} (w^2 - u^2) \right\} d\Gamma. \tag{2.8}$$

At the points of Γ_c with $u^1 > u^2$ we have, due to (2.6)-(2.7),

$$\frac{\partial u^1}{\partial n_1} = \frac{\partial u^2}{\partial n_2} = 0, \tag{2.9}$$

so that the integrand in (2.8) vanishes at such points. At the points of Γ_c with $u^1 = u^2$ we have, by (2.5) and (2.7),

$$\frac{\partial u^1}{\partial n_1}(w^1 - u^1) + \frac{\partial u^2}{\partial n_2}(w^2 - u^2) = \frac{\partial u^1}{\partial n_1}w^1 + \frac{\partial u^2}{\partial n_2}w^2 = \frac{\partial u^2}{\partial n_2}(w^2 - w^1) \geq 0. \quad (2.10)$$

Thus the integral in (2.8) is nonnegative for any $w \in \mathcal{K}$ and the solution u of (2.1)-(2.7) solves also the problem of finding $u \in \mathcal{K}$ such that

$$a(u, w - u) - \ell(w - u) \geq 0 \quad \text{for all} \quad w \in \mathcal{K}. \quad (2.11)$$

Using the well known technique described e.g. in [16, 25, 27], it is also possible to prove that any smooth solution $u \in \mathcal{K}$ of (2.11) is a solution of (2.1)-(2.7). Since the expression on the left of inequality (2.11) is the gradient of the energy functional

$$J(v) = \frac{1}{2}a(v, v) - \ell(v)$$

at u, it follows that problem (2.11) is equivalent to the problem

$$\min J(v) \quad \text{s.t.} \quad v \in \mathcal{K}. \quad (2.12)$$

Let us briefly examine the existence of solution of (2.12). First, it may be checked that $J(v)$ is convex but not coercive, i.e. $\|v\| \to \infty$ for $v \in \mathcal{K}$ does not necessarily imply $J(v) \to \infty$. To this purpose, let us define $e \in \mathcal{K}$ by

$$e(x^1, x^2) = (0, 1) \quad \text{for} \quad x^i \in \overline{\Omega}^i \quad (2.13)$$

with $\overline{\Omega}^i$ denoting the closure of Ω^i. Since $\lambda e \in \mathcal{K}$ for $\lambda \geq 0$ and

$$J(\lambda e) = -\lambda \int_\Omega f\, e\, d\Omega = -\lambda \int_{\Omega^2} f\, d\Omega,$$

it follows that $J(\lambda e)$ does not increase with λ if the last integral is non-negative. However, using the technique of [27, section 1.1.6] , it can be proved that $J(v)$ is coercive on $\mathcal{K}$ provided

$$\int_{\Omega^2} f\, d\Omega < 0. \quad (2.14)$$

The well known result on the existence and uniqueness of the minimum of convex coercive functionals (e.g. [24]) then guarantees that problems (2.11) and (2.12) have unique solution if f satisfies (2.14). Thus, in what follows, we shall assume that f satisfies condition (2.14).

3 Discretization and Domain Decomposition

Let (η_h, τ_h) define a partitioning of Ω into triangles $T_j \in \tau_h$ with vertices at $N_k \in \eta_h$ that matches the decomposition of Ω into Ω^1 and Ω^2 so that Γ_c is covered by the sides of adjacent triangles.

For $i = 1, 2$, let P_h^i denote the piecewise linear finite element subspaces of $H^1(\Omega^i)$, let $V_h^i = P_h^i \cap V^i$, and define

$$V_h = V_h^1 \times V_h^2 \quad \text{and} \quad \mathcal{K}_h = \mathcal{K} \cap V_h \tag{3.1}$$

so that the solution of the problem (2.12) is approximated by the finite element problem of finding

$$\min J(v_h) \quad \text{s.t.} \quad v_h \in \mathcal{K}_h. \tag{3.2}$$

The functions $p_h^i \in P_h^i$ are fully determined by the values $x_k^i = p_h^i(N_k^i)$ at the nodes $N_k^i \in \overline{\Omega}^i$. In particular, assuming that the nodes of $\Omega^i \setminus \Gamma_u$ are indexed independently by indices $1, 2, \ldots, s_i$ and denoting by e_k^i the functions of the standard basis od V_k^i so that $e_k^i(N_j^i) = \delta_{kj}$ (the Kronecker symbol), we can write any $v_h^i \in V_h^i$ in the form

$$v_h^i = \sum_{k=1}^{s_i} x_k^i e_k^i. \tag{3.3}$$

Substituting (3.3) into the expressions for $J(u)$ gives

$$J(v_h) = \frac{1}{2} x^T A x - f^T x$$

with $A = \operatorname{diag}[A^1, A^2]$ symmetric positive semidefinite matrix, $A^i = [a_{jk}^i], a_{jk}^i = a(e_j^i, e_k^i)$,

$$f = \begin{bmatrix} f^1 \\ f^2 \end{bmatrix}, \qquad x = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix},$$

$f^i = [f_j^i], f_j^i = \ell(e_j^i)$ and $x^i = [x_j^i]$. The vector x is the vector of the nodal unknowns.

To complete the discretization of (2.2), we have to describe conditions on x_k^i that correspond to

$$\sum_{i=1}^{2} \sum_{k=1}^{s_i} x_k^i e_k^i \in \mathcal{K}_h.$$

To this end, notice that the nodes on the interface Γ_c are doubled as in Figure 2. Thus, condition $x_j^1 \leq x_k^2$ may be written in the form

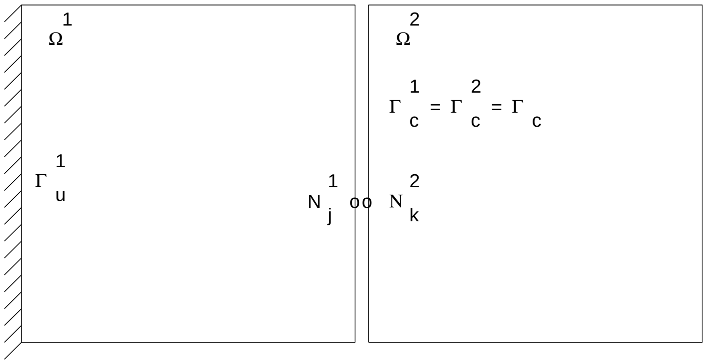

Figure 2: Matching nodes

$$bx \leq 0$$

with $b = [b^1, b^2]$ a row vector with zero entries except $b^1_j = 1$ and $b^2_k = -1$. Forming rows b_ℓ for all m couples of nodes on Γ_c and denoting

$$B_I = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix},$$

we get the discretized version of problem (2.12) as follows

$$\min \frac{1}{2} x^T A x - f^T x \quad \text{s.t.} \quad B_I x \leq 0. \tag{3.4}$$

The matrix A is positive semidefinite. It may be easily verified that the kernel of A is spanned by the discrete analogue of e given by (2.13).

So far, we have used only the natural decomposition of the spatial domain Ω into Ω^1 and Ω^2. However, we can optionally decompose each Ω^i into subdomains $\Omega^{i,1}, \ldots, \Omega^{i,p_i}$ with interfaces $\Gamma^{i,jk}$ as in Figure 3.

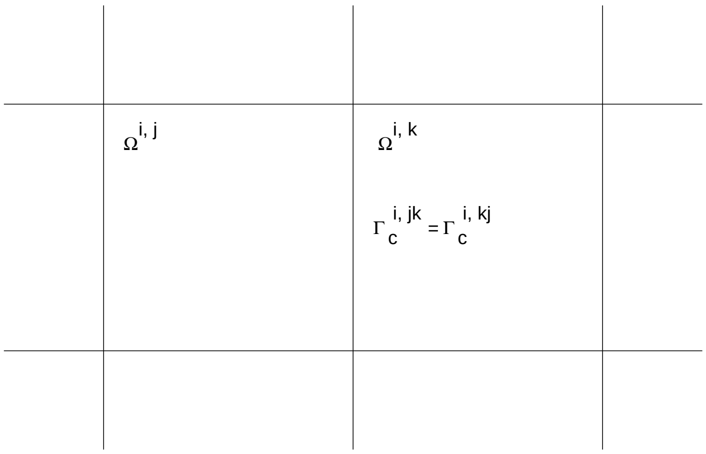

Figure 3: Auxiliary decomposition of Ω^i

Let us assume that we are given the auxiliary decomposition of each subdomain Ω^i that is compatible with the partitioning (η_h, τ_h) so that each subdomain $\Omega^{i,j}$ is partitioned by a subset of (η_h, τ_h). Indexing contiguously the nodes and entries of corresponding vectors in subdomains $\Omega^{i,j}$ and using the finite element discretization of the problem (2.12) with the basis functions that are zero extensions of $P_h(\Omega^{i,j})$ for $i = 1, 2$ and $j = 1, \ldots, p_i$, we get as above a vector f and a matrix A such that

$$A = \operatorname{diag}(A^1, A^2) = \operatorname{diag}(A_1, \ldots, A_{p_1+p_2}) \tag{3.5}$$

and

$$J(v_h) = \frac{1}{2}x^T A x - f^T x \tag{3.6}$$

for all piecewise linear functions v_h that are continuous in the subdomains $\Omega^{i,j}$.

To enforce continuity across $\Gamma^{i,jk}$, let us denote by B_E a matrix that has, for each node $N \in \eta_h \cap \Gamma^{i,jk}$, a row of zeros except 1 and -1 at the positions that correspond to the indices of N in $\overline{\Omega}^{i,j}$ and $\overline{\Omega}^{i,k}$, respectively. Some care should be taken with the corner nodes that belong to four subdomains to keep the rows of B independent. In our algorithms, we join such four nodes with global indices i, j, k, l by the rows of the matrix

$$\begin{bmatrix} \ldots & \sqrt{2}/2 & \ldots & \sqrt{2}/2 & \ldots & -\sqrt{2}/2 & \ldots & -\sqrt{2}/2 & \ldots \\ \ldots & 1 & \ldots & -1 & \ldots & 0 & \ldots & 0 & \ldots \\ \ldots & 0 & \ldots & 0 & \ldots & 1 & \ldots & -1 & \ldots \end{bmatrix}$$

with zero columns except i, j, k, l. If the nodes do not belong to the contact interface, the rows are used to form B_E, otherwise the last two rows go to B_E to join the couples of nodes on each side and the first row goes to B_I possibly replacing all other rows with nonzero columns i, j, k, l.

After the modifications described above, we get the full rank matrices B_I and B_E and the discretized version of problem (2.12) with auxiliary decomposition

$$\min \frac{1}{2}x^T A x - f^T x \quad \text{s.t.} \quad B_I x \le 0 \quad \text{and} \quad B_E x = 0. \tag{3.7}$$

4 Dual Formulation

The Lagrangian associated with problem (3.7) is

$$L(x, \lambda_I, \lambda_E) = \frac{1}{2}x^T A x - f^T x + \lambda_I^T B_I x + \lambda_E^T B_E x, \tag{4.1}$$

where λ_I and λ_E are the Lagrange multipliers associated with inequalities and equalities, respectively. Introducing notation

$$\lambda = \begin{bmatrix} \lambda_I \\ \lambda_E \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_I \\ B_E \end{bmatrix},$$

we can observe that B is a full rank matrix and write the Lagrangian briefly as

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

It is well known [5] that (3.7) is equivalent to the saddle point problem

$$\text{Find} \quad (\overline{x}, \overline{\lambda}) \quad \text{s.t.} \quad L(\overline{x}, \overline{\lambda}) = \sup_{\lambda_I \geq 0} \inf_x L(x, \lambda). \tag{4.2}$$

For fixed λ, the Lagrange function $L(\cdot, \lambda)$ is convex in the first variable and the minimizer x of $L(\cdot, \lambda)$ satisfies

$$Ax - f + B^T \lambda = 0. \tag{4.3}$$

Equation (4.3) has a solution iff

$$f - B^T \lambda \in \mathrm{Im} A, \tag{4.4}$$

which can be expressed more conveniently by means of a matrix R whose columns span the null space of A as

$$R^T (f - B^T \lambda) = 0. \tag{4.5}$$

Matrix R may be formed directly so that each floating subdomain is assigned to a row of R with ones in positions of the nodal variables that belong to the subdomain and zeros elsewhere. It may be checked that $R^T B^T$ is a full rank matrix. The matrix R may be also extracted from A [17].

Now assume that λ satisfies (4.4) and denote by $A^\dagger$ any matrix that satisfies

$$A A^\dagger A = A. \tag{4.6}$$

It may be verified directly that if x solves (4.3), then there is a vector α such that

$$x = A^\dagger (f - B^T \lambda) + R\alpha. \tag{4.7}$$

After substituting expression (4.7) into problem (4.2) and changing signs, we shall get the minimization problem

$$\min \Theta(\lambda) \quad \text{s.t.} \quad \lambda_I \geq 0 \quad \text{and} \quad R^T (f - B^T \lambda) = 0, \tag{4.8}$$

where

$$\Theta(\lambda) = \frac{1}{2}\lambda^T B A^\dagger B^T \lambda - \lambda^T B A^\dagger f. \tag{4.9}$$

Though any matrix $A^\dagger$ that satisfies (4.6), such as the Moore-Penrose pseudoinverse, may be used to get (4.7), we must be more cautious when we consider effective solving our problem. Farhat and Roux [18, 19] proposed to use as $A^\dagger$ the left generalized inverse that satisfies (4.6),

$$A^\dagger = \mathrm{diag}(A_1^\dagger, \ldots, A_{p_1+p_2}^\dagger),$$

where $A_i^\dagger = A_i^{-1}$ whenever A_i is non-singular. If A_i is singular than it is easy to check that there is a permutation matrix P_i and a non-singular matrix T_i such that

$$P_i^T A_i P_i = \begin{pmatrix} T_i & S_i \\ S_i^T & S_i^T T_i^{-1} S_i \end{pmatrix}$$

and that

$$A_i^\dagger = P_i^T \begin{pmatrix} T_i^{-1} & 0 \\ 0 & 0 \end{pmatrix} P_i$$

satisfies (4.6). In this case, $A^\dagger$ does not necessarily satisfy the other identities that define the Moore-Penrose pseudoinverse.

Once the solution $\overline{\lambda}, \overline{\mu}$ of (4.8) is known, the vector $\overline{x}$ that solves (4.2) can be evaluated, provided that the vector α of (4.7) is known. To find the formula that computes α, notice that the solution of (3.7) satisfies

$$\widetilde{B}_I \overline{x} = 0 \qquad \text{and} \qquad B_E \overline{x} = 0, \tag{4.10}$$

where $\widetilde{B}_I$ is the matrix formed by the rows b_i of B_I that correspond to active constraints, the latter being characterized by $\overline{\lambda}_i = 0$. After substituting $\overline{\lambda}$ into (4.7) and multiplying on the left by $\widetilde{B} = [\widetilde{B}_I^T, B_E^T]^T$, we get the equation

$$\widetilde{B} A^\dagger (f - B^T \overline{\lambda}) + \widetilde{B} R \alpha = 0 \tag{4.11}$$

that determines α. Solving the normal equation [26] to the equation (4.11) for α then yields

$$\alpha = -(R^T \widetilde{B}^T \widetilde{B} R)^{-1} R^T \widetilde{B}^T \widetilde{B} A^\dagger (f - B^T \overline{\lambda}) \tag{4.12}$$

which can be substituted into (4.7) to get $\overline{x}$.

Using the fact that $R^T B^T$ is a full rank matrix, it may be verified that the Hessian of Θ is positive definite. Moreover, the Hessian is the same as that of the basic FETI method by Farhat and Roux [18, 19], so that its spectrum is relatively favorably distributed for application of the conjugate gradient method [29].

If there is no optional domain decomposition, then there is neither B_E nor λ_E in (4.8) and (4.9), but the structure of the dual problem remains the same in the sense that it is still minimization of a strictly convex quadratic function with respect to simple bounds and equality constraints. Thus, what follows in the next section applies to the solution of both problems (3.4) and (3.7).

5 Modifications

Even though problem (4.8) is much more suitable for computations than (3.7) and was used for efficient solution of the discretized variational inequalities [13], further improvement may be achieved by adapting some simple observations and the results of Farhat, Mandel and Roux [20]. We shall reformulate problem (4.8) in such a way that if we apply the conjugate gradient method to minimize the aumented Lagrangian, the spectral distribution of the Hessian guarantees that the rate of convegence does not depend on either penalization or discretization parameters.

Let us denote

$$\begin{aligned} F &= BA^{\dagger}B^T, & \widetilde{d} &= BA^{\dagger}f, \\ \widetilde{G} &= R^T B^T, & \widetilde{e} &= R^T f \end{aligned}$$

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

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

has orthogonal rows. After denoting

$$e = T\widetilde{e},$$

problem (4.8) reads

$$\min \quad \frac{1}{2}\lambda^T F\lambda - \lambda^T \widetilde{d} \quad \text{s.t} \quad \lambda_I \geq 0 \quad \text{and} \quad G\lambda = e. \tag{5.1}$$

Next we shall transform the problem of minimization on the subset of the affine space to that on the subset of the vector space by means of arbitrary $\widetilde{\lambda}$ that satisfies

$$G\widetilde{\lambda} = e.$$

To this purpose, we shall look for the solution of (5.1) in the form $\lambda = \mu + \widetilde{\lambda}$. Since

$$\frac{1}{2}\lambda^T F\lambda - \lambda^T\widetilde{d} = \frac{1}{2}\mu^T F\mu - \mu^T(\widetilde{d} - F\widetilde{\lambda}) + \frac{1}{2}\widetilde{\lambda}^T F\widetilde{\lambda} - \widetilde{\lambda}^T\widetilde{d},$$

problem (5.1) is, after returning to the old notation, equivalent to

$$\min \quad \frac{1}{2}\lambda^T F\lambda - d^T\lambda \quad \text{s.t} \quad G\lambda = 0 \quad \text{and} \quad \lambda_I \geq -\widetilde{\lambda_I}. \tag{5.2}$$

with $d = \widetilde{d} - F\widetilde{\lambda}$.

To assess our progress, let us compare the distribution of the spectrum of the Hessians $H_1 = F + \rho\widetilde{G}^T\widetilde{G}$ and $H_2 = F + \rho G^T G$ of the augmented Lagrangians for problems (4.8) and (5.2), respectively. Let us assume that the eigenvalues of F are in the interval $[a, b]$ and that the nonzero eigenvalues of $\widetilde{G}^T\widetilde{G}$ are in $[\gamma, \delta]$. For each square matrix A, let $\sigma(A)$ denote its spectrum. Using the analysis of [12, 11], it follows that

$$\sigma(H_1) \subseteq [a, b] \cup [a + \rho\gamma, b + \rho\delta] \quad \text{and} \quad \sigma(H_2) \subseteq [a, b] \cup [a + \rho, b + \rho].$$

If ρ is sufficiently large and $\gamma < \delta$, than the spectrum of H_1 is distributed in two intervals with the larger one on the right. In this case, the analysis of Axelsson [1] shows that the rate of convergence of conjugate gradients for minimization of the quadratic function with the Hessian H_1 depends on the penalization parameter ρ. However, the situation is much more favorable for minimization of the quadratic function with the Hessian H_2 since, in this case, the spectrum is always distibuted in two intervals of the same length. It follows by analysis of Axelsson [1] that the rate of convergence is governed by the effective condition number $\overline{\kappa}(H_2) = 4b/a$ so that the number k of the conjugate gradient iterations that are necessary to reduce the gradient of the augmented Lagrangian for (5.2) by ϵ satisfies

$$k \leq \frac{1}{2}\text{int}\left(\sqrt{\frac{4b}{a}}\ln\left(\frac{2}{\epsilon}\right) + 1\right). \tag{5.3}$$

This bound is only two times greater than the bound obtained when minimizing with F and does not depend on the penalization parameter ρ.

Our final step is based on the observation that the augmented Lagrangian for problem (5.2) may be decomposed by the orthogonal projectors

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

on the image space of G^T and on the kernel of G, respectively. Problem (5.2) is then equivalent to

$$\min \quad \frac{1}{2}\lambda^T PFP\lambda - \lambda^T Pd \quad \text{s.t} \quad G\lambda = 0 \quad \text{and} \quad \lambda_I \geq -\widetilde{\lambda_I}, \tag{5.4}$$

and the Hessian $H_3 = PFP + \rho Q$ of the augmented Lagrangian

$$L(\lambda, \mu, \rho) = \frac{1}{2}\lambda^T(PFP + \rho Q)\lambda - \lambda^T Pd + \mu^T G\lambda \tag{5.5}$$

is decomposed by projectors P and Q whose image spaces are invariant subspaces of H_3. If $[a_P, b_P]$ denotes the interval that contains the non-zero eigenvalues of PFP, it follows that the eigenvalues of H_3 satisfy

$$\sigma(H_3) \subseteq [a_P, b_P] \cup \{\rho\} \quad \text{and} \quad [a_P, b_P] \subseteq [a, b] \tag{5.6}$$

so that, by the analysis of Axelsson [2], the number k of conjugate gradient iterations that are necessary to reduce the gradient of the augmented Lagrangian (5.5) by ϵ satisfies

$$k \leq \frac{1}{2}\text{int}\left(\sqrt{\frac{b_P}{a_P}}\ln\left(\frac{2}{\epsilon}\right) + 3\right). \tag{5.7}$$

The bound (5.7) is qualitatively better than (5.3). Analysis of the FETI method by Farhat, Mandel and Roux [20] implies that, for the regular decomposition,

$$\frac{b_P}{a_P} \leq \text{const}\ \frac{H}{h}, \tag{5.8}$$

where h and H are the mesh and subdomain diameters, respectively. Examining (5.7) and (5.8), we conclude that the rate of convergence for unconstrained minimization of the augmented Lagrangian (5.5) does not depend on either the penalization parameter ρ or the discretization parameter h provided the aspect ratio of the discretization and decomposition is close to one and the ratio H/h is kept bounded by a constant. More discussion of the conjugate gradient method applied to systems with augmented Lagrangian including numerical experiments may be found in [11]. More detailed derivaton of (5.8) may be found in [20].

The idea of using projectors to preconditioning is close to the older idea of preconditioning by projector [7], but only the discovery of the role of the natural coarse space [20] revealed its full power.

6 Solution of Bound and Equality Constrained Quadratic Programming Problems

Our development of an efficient algorithm for the solution of (5.4) is based on the observation that the solution of such problems may be reduced by the augmented Lagrangian technique [6, 12] to the solution of a sequence of quadratic programming (QP) problems with simple bounds, and that the latter can be solved much more efficiently than more general QP problems due to the possibility to use projections and results on adaptive precision control in the active set strategy [22, 21, 23, 4, 10]. Here we shall briefly review these results.

To simplify our notation, let us denote $F_P = PFP$ so that the augmented Lagrangian for problem (5.4) and its gradient are given by

$$L(\lambda, \mu, \rho) = \frac{1}{2}\lambda^T F_P \lambda - \lambda^T P d + \mu^T G \lambda + \frac{1}{2}\rho ||Q\lambda||^2$$

and

$$g(\lambda, \mu, \rho) = F_P \lambda - Pd + G^T(\mu + \rho G \lambda),$$

respectively. Also, let I be the set that contains the indices of the constrained entries of λ. Then, the *projected gradient* $g^P = g^P(\lambda, \mu, \rho)$ of L at λ is given componentwise by

$$g_i^P = g_i \quad \text{for} \quad \lambda_i > -\widetilde{\lambda_i} \quad \text{or} \quad i \notin I \quad \text{and} \quad g_i^P = g_i^- \quad \text{for} \quad \lambda_i = -\widetilde{\lambda_i} \quad \text{and} \quad i \in I$$

with $g_i^- = \min(g_i, 0)$.

The algorithm that we propose here may be considered a variant of the algorithm proposed by Conn, Gould and Toint [6] for identification of stationary points of more general problems. However, our algorithm is modified in order to exploit the specific structure of our problem to get improved performance. The most important of such modifications consists in including the adaptive precision control of auxiliary problems in Step 1.

All of the parameters that must be defined prior to the application of the algorithm are listed in step 0. Typical values of these parameters for our model problem are given in brackets.

Algorithm 6.1. (Simple bound and equality constraints)

Step 0. { Initialization of parameters} Set $0 < \alpha < 1$ $[\alpha = 0.1]$ for equality precision update, $1 < \beta$ $[\beta = 10]$ for penalty update, $\rho_0 > 0$ $[\rho_0 = 10^4]$ for initial penalty parameter, $\eta_0 > 0$ $[\eta_0 = 0.1]$ for initial equality precision, $M > 0$ $[M = 10^4]$ for balancing ratio, μ^0 $[\mu^0 = 0]$ and $k = 0$.

Step 1. Find λ^k so that
$||g^P(\lambda^k, \mu^k, \rho_k)|| \le M||G\lambda^k||$.

Step 2. If $||g^P(\lambda^k, \mu^k, \rho_k)||$ and $||G\lambda^k||$ are sufficiently small
then λ^k is the solution.

Step 3. If $||G\lambda^k|| \le \eta_k$
Step 3a. then $\mu^{k+1} = \mu^k + \rho_k G\lambda^k, \ \ \rho_{k+1} = \rho_k, \ \ \eta_{k+1} = \alpha\eta_k$
Step 3b. else $\rho_{k+1} = \beta\rho_k, \ \ \eta_{k+1} = \eta_k$
end if.

Step 4. Increase k and return to Step 1.

An implementation of Step 1 is carried out by the minimization of the augmented Lagrangian L subject to $\lambda_I \ge -\widetilde{\lambda_I}$ by means of the algorithm that we shall describe later. The unique solution $\overline{\lambda} = \overline{\lambda}(\mu, \rho)$ of this auxiliary problem satisfies the Karush-Kuhn-Tucker conditions

$$g^P(\overline{\lambda}, \mu, \rho) = 0. \tag{6.1}$$

Even though there are many parameters in Algorithm 6.1, only ρ_0 and M seem to be essentially problem dependent. Theoretical results of [14, Corollary 2.7] suggest that the rate of convergence of the Lagrange multipliers increases very little with M considerable less than $||(GF^{-1}G^T)^{-1}|| \leq ||F||$ and that ρ_0 should be larger than M. The algorithm is designed so that ρ adjusts to the other parameters, including M. Let us recall here that the large penalty parameter does not necessarily slow down the convergence of the conjugate gradient iterations [11].

The salient feature of this algorithm is that it deals with each type of constraint completely separately and that it accepts inexact solutions for the auxiliary box constrained problems in Step 1. For parallel implementation, it is necessary to keep the factors that form F_P since the latter is just used in the matrix-vector products. The action of $A^\dagger$ may be evaluated by means of a Cholesky decomposition. Besides, the matrix G of problem (5.4) that is used to compute the projections P and Q may be generated by means of the QR decomposition of $\widetilde{G}$. We shall present more details on implementation elsewhere.

The algorithm has been proved [12] to converge for any set of parameters that satisfy the prescribed relations. Moreover, it has been proved that the asymptotic rate of convergence is the same as for the algorithm with exact solution of auxiliary quadratic programming problems (i.e. $M = 0$) and that the penalty parameter is uniformly bounded. These results with the above discussion on elimination of the negative effect of penalization give theoretical support to algorithm 6.1.

In the rest of this section we shall describe in more detail implementation of Step 1 of Algorithm 6.1. In particular, we shall assume that μ and ρ are fixed and we shall denote

$$\theta(\lambda) = L(\lambda, \mu, \rho).$$

Let us recall that we assume that the set of indices of our dual variables λ_i is decomposed into two disjoint sets I and E with I denoting the indices of the constrained entries of λ, and let us denote by $\mathcal{A}(\lambda)$ and $\mathcal{F}(\lambda)$ the *active set* and *free set* of indices of λ, respectively, i.e.

$$\mathcal{A}(\lambda) = \{i \in I : \lambda_i = -\widetilde{\lambda_i}\} \quad \text{and} \quad \mathcal{F}(\lambda) = \{i : \lambda_i > -\widetilde{\lambda_i} \ \text{ or } \ i \in E\}. \tag{6.2}$$

The *chopped gradient* g^C and the *inner gradient* g^I of $\theta(\lambda)$ are defined by

$$\begin{aligned} g_i^I &= g_i \text{ for } i \in \mathcal{F}(\lambda) \text{ and } g_i^I = 0 \text{ for } i \in \mathcal{A}(\lambda) && (6.3)\\ g_i^C &= 0 \text{ for } i \in \mathcal{F}(\lambda) \text{ and } g_i^C = g_i^- \text{ for } i \in \mathcal{A}(\lambda). && (6.4) \end{aligned}$$

Hence the Karush-Kuhn-Tucker conditions for the solution of the problem of finding

$$\min \ \theta(\lambda) \quad \text{s.t.} \quad \lambda_I \geq -\widetilde{\lambda_I} \tag{6.5}$$

are satisfied iff the projected gradient $g^P = g^I + g^C$ vanishes.

An efficient algorithm for the solution of convex QP problems with simple bounds has been proposed independently by Friedlander and Martínez [22, 21, 23, 4] and Dostál [10]. The algorithm may be considered a modification of the Polyak algorithm that controls the precision of the solution of auxiliary problems by the norm of g^C in each inner iterate y^i.

If for $\Gamma > 0$ the inequality

$$||g^C(y^i)|| \leq \Gamma ||g^I(y^i)||$$

holds then we call y^i *proportional* [10]. The algorithm explores the face

$$W_J = \{y : y_i = -\widetilde{\lambda_i} \text{ for } i \in J\}$$

with a given active set $J \subseteq I$ as long as the iterates are proportional. If y^i is not proportional, we generate y^{i+1} by means of the descent direction $d^i = -g^C(y^i)$ in a step that we call proportioning, and then we continue exploring the new face defined by $J = \mathcal{A}(y^{i+1})$. The class of algorithms driven by proportioning may be defined as follows.

Algorithm 6.2 (General Proportioning Scheme - GPS)
Let a feasible λ^0 and $\Gamma > 0$ $[\Gamma = 1]$ be given. For $i > 0$, choose λ^{i+1} by the following rules:
(i) If λ^i is not proportional, define λ^{i+1} by proportioning.
(ii) If λ^i is proportional, choose a feasible λ^{i+1} so that

$$\theta(\lambda^{i+1}) \leq \theta(\lambda^i)$$

and λ^{i+1} satisfies at least one of the conditions: $\mathcal{A}(\lambda^i) \subset \mathcal{A}(\lambda^{i+1}), \lambda^{i+1}$ is not proportional, or λ^{i+1} minimizes θ subject to $\lambda \in W_J$, $J = \mathcal{A}(\lambda^i)$.

The set relation $\subset$ is used in the strict sense so that it is satisfied if the set on the left is a proper subset of the set on the right. Basic theoretical results have been proved in [22, 10, 21, 23, 4].

Theorem 6.3. *Let λ^k denote an infinite sequence generated by Algorithm GPS with given λ^0 and $\Gamma > 0$. Let $\theta(\lambda)$ be a strictly convex quadratic function. Then the following statements are true:*

(i) λ^k *converges to the solution* $\bar{\lambda}$ *of (6.5).*
(ii) If problem (6.5) is not dual degenerate, then there is k *such that* $\bar{\lambda} = \lambda^k$*.*
(iii) If $\Gamma \geq \kappa(\theta_{\lambda\lambda})^{1/2}$*, where* κ *denotes the spectral condition number, then there is* k *such that* $\bar{\lambda} = \lambda^k$*.*

Step (ii) of Algorithm GPS may be implemented by means of the conjugate gradient method. The most simple implementation of this step starts from $y^0 = \lambda^k$ and generates the conjugate gradient iterations $y^1, y^2, \ldots$ for $\min\{\theta(y) : y \in \mathcal{W}_J,\ J = \mathcal{A}(y^0)\}$ until y^i is found that is not feasible or not proportional or minimizes $\theta(\lambda)$ subject to $\lambda_I \geq -\widetilde{\lambda_I}$. If y^i is feasible, then we put $\lambda^{k+1} = y^i$, otherwise $y^i = y^{i-1} - \alpha^i p^i$ is not feasible and we can find $\widetilde{\alpha}^i$ so that $\lambda^{k+1} = y^i - \widetilde{\alpha}^i p^i$ is feasible and $\mathcal{A}(\lambda^k) \not\subseteq \mathcal{A}(\lambda^{k+1})$. We shall call the resulting algorithm *feasible proportioning* [10].

An obvious drawback of feasible proportioning is that the algorithm is usually unable to add more than one index to the active set in one iteration. A simple but efficient alternative is to replace the feasibility condition by $\theta(Py^{i+1}) \leq \theta(Py^i)$, where Py denotes the projection on the set $\Omega = \{y : y_i \geq -\widetilde{\lambda_i} \ \text{ for } \ i \in I\}$. If the conjugate gradient iterations are interrupted on condition $\theta(Py^{i+1}) > \theta(Py^i)$, then a new iteration is defined by $\lambda^{k+1} = Py^i$. Resulting modification of the feasible proportioning algorithm is called *monotone proportioning* [10].

The algorithm uses the single parameter Γ. Apparently, a good choice is $\Gamma \approx 1$, as it seems reasonable to change the face when the error in the chopped gradient dominates that in the free gradient as the conjugate gradient method typically reduces only the latter.

The performance of the algorithm depends essentially on the rate of convergence of the conjugate gradient method that minimizes θ in faces. In our case, the optimality results (5.7) and (5.8) suggest that examination of faces will be carried out efficiently.

7 Numerical Experiments

In this section, we illustrate the practical behavior of our algorithm on solution of the model problem of Section 2 with $f(x, y) = -3$ for $(x, y) \in (0, 1) \times [0.75, 1)$, $f(x, y) = 0$ for $(x, y) \in (0, 1) \times (0, 0.75)$, $f(x, y) = -1$ for $(x, y) \in (1, 2) \times (0, 0.25)$ and $f(x, y) = 0$ for $(x, y) \in (1, 2) \times (0.25, 1)$. The model problem was discretized by regular grids defined by the stepsize $h = 1/n$ with $n + 1$ nodes in each direction per subdomain Ω^i, $i = 1, 2$. Each subdomain Ω^i was decomposed into $n_x \times n_y$ identical rectangles of the dimensions $H_x = 1/n_x$ and $H_y = 1/n_y$. Solution of the model problem for regular decomposition is in Figure 4.

The model problem was solved for $h \in \{1/64, 1/128, 1/256, 1/512\}$ with op-

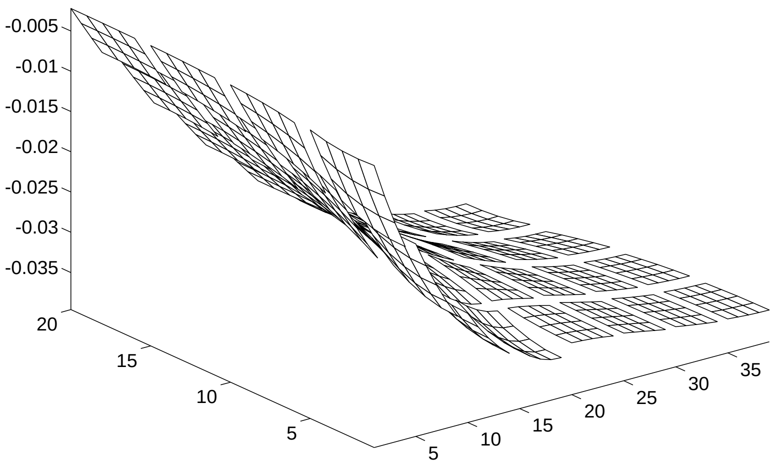

Figure 4: Solution of the model problem with $h = 1/16$ and $H_x = H_y = 1/4$

Table 1.
Convergence with the decomposition into strips for $h = 1/128$.

H_x	dual dim.	Outer iter.	cg iter.	M − V prod.	time sec.
1	129	5	45	95	62.9
1/2	387	5	66	137	84.3
1/4	903	5	63	131	82.0
1/8	1935	6	91	188	114.6
1/16	3999	5	147	299	177.9
1/32	8127	5	258	521	292.5

tional secondary decompositions in order to test experimentally the dependance of the rate of convergence on the discretization, decomposition and penalization parameters. In all cases, we use the stopping criterium

$$||g^P(\lambda, \mu, 0)|| \leq 10^{-5}||d|| \quad \text{and} \quad ||G\lambda|| \leq 10^{-5}||f||.$$

The results are summarized in Tables 1-5, that give several discretization and decomposition parameters together with the number of iterations in the outer loop of the augmented Lagrangian algorithm, the number of the conjugate gradient iterations in the inner loops of the algorithm for the solution of bound constrained QP problems, the number of multiplications by the matrix F that dominates the cost in each inner iteration, and the time spent by the serial implementation, in seconds. If not specified explicitly, we use the values of parameters suggested at Algorithms 6.1 and 6.2, $\mu^0 = 0$ and $\lambda^0 = -\widetilde{\lambda}$.

In particular, Table 1 shows the convergence with decomposition into strips for the fixed discretization parameter $h = 1/128$. Horizontal dimension H_x of vertical strips is in the first column. The primal dimension of the problems ranges from 33153 to 41151. We can observe that the performance deteriorates with an increasing number of subdomains and resulting increase of the aspect ratio. The penalty parameter ρ was not updated during the solution.

The effect of the aspect ratio may be observed in more detail in Table 2. These results are comparable with the results reported for linear problems [20].

Table 3 shows the near optimal performance of the algorithm with the natural coarse space with respect to the ratio of the decomposition parameter H and the discretization parameter h. The numbers are given for regular decompositions that are characterized by $H_x = H_y$ for some variants of the discretization parameter h.

Table 2.
Effect of the subdomain aspect ratio
for $h = 1/128$.

H_x	H_y	dual dim.	Outer iter.	cg iter.	M-V prod.	time sec.
$1/16$	1	3999	5	147	299	177.9
$1/16$	$1/2$	4286	5	91	187	50.0
$1/16$	$1/4$	4860	5	59	123	19.9
$1/16$	$1/8$	6008	4	51	106	19.1
$1/16$	$1/16$	8304	4	76	157	81.6

Table 3.
Near optimality of regular decompositions
with $H = H_x = H_y$.

h	H	primal dim.	dual dim.	Outer iter.	cg iter.	$\mathrm{M} - \mathrm{V}$ prod.	time sec.
$1/64$	$1/2$	8646	326	4	38	80	2.94
$1/128$	$1/4$	34716	1692	5	51	107	15.4
$1/256$	$1/8$	139128	7544	6	74	154	89.5
$1/512$	$1/16$	557040	31728	6	107	221	586.0

Table 4.
Effect of the initial penalty parameter ρ_0
for $H_x = H_y = 1/8$ and $h = 1/128$.

ρ_0	final ρ	Outer iter.	cg iter.	$\mathrm{M-V}$ prod.	time sec.
10	10^4	7	84	175	17.5
10^2	10^4	6	67	140	14.1
10^3	10^4	5	62	129	13.1
10^4	10^4	4	60	124	12.6
10^5	10^5	5	75	157	15.8
10^6	10^6	5	86	178	17.9
10^7	10^7	5	98	203	20.4

These results are completed by Table 4, which illustrates the sensitivity of the algorithm on the initial penalization parameter. Though the linear theory [11] predicts a little sensitivity of the performance on the penalty parameter in exact arithmetics, we can observe that large penalty values may cause problems.

We have also implemented our algorithm with preconditioning, using the modified lumped precoditioner in the form $C^{-1} = PBKB^TP + (1/\rho)Q$, in an attempt to reduce the effective condition number of the Hessian of the augmented Lagrangian. The untransformed scheme [3] for preconditioning in faces after the proportioning step has been used. Comparing the last two columns from Tables 3 and 5, we can notice that, on average, the number of multiplications by matrix F was reduced by 29.9%, while the time spent by the algorithm was reduced by 27.6%. These preliminary results suggest that, at least for this specific value of the ratio H/h, it is indeed possible to improve the performance of our algorithm by preconditioning.

All of the experiments were run on a SUN sparc Ultra 1 computer, under SunOS 5.5.1, using the `f77` (version 4.0) FORTRAN compiler and double precision. The auxiliary problems were solved using the `QUACON` routine [4], developed at the Institute of Mathematics, Statistics and Scientific Computation of Unicamp.

8 Comments and conclusions

We have described a new algorithm for the solution of variational inequalities arising from the elliptic boundary value problems with boundary conditions that include inequalities. The algorithm gives directly multipliers for the boundary inequalities.

The approach combines a variant of the domain decomposition method of the

Table 5.
Effect of preconditioning for regular decompositions
with $H = H_x = H_y$.

h	H	primal dim.	dual dim.	Outer iter.	cg iter.	$\mathrm{M-V}$ prod.	time sec.
1/64	1/2	8646	326	4	26	56	2.11
1/128	1/4	34716	1692	5	34	73	10.6
1/256	1/8	139128	7544	7	51	109	64.6
1/512	1/16	557040	31728	6	76	158	448.6

Neumann-Neumann type based on the duality theory of quadratic programming with new algorithms for the solution of the quadratic programming problems with simple bounds and equalities. A new feature of the algorithm is the combination of the preconditioning by the natural coarse subspace with the adaptive control of precision of the solution of auxiliary problems with effective application of the projections and penalty technique that preserve the optimal rate of convergence of the conjugate gradient iterations in faces.

The implementation of the algorithm deals separately with each subdomain, so that the algorithm is suitable for parallel implementation. The convergence results have been reported. In particular, it has been shown that the rate of convergence of the iterative method for the solution of auxiliary problems is not hindered by the penalty term in the augmented Lagrangian and does not depend either on the penalty parameter of the augmented Lagrangian or on the grid parameter, provided that the ratio between the subdomain and grid parameters are kept constant. Numerical experiments confirm predicted numerical scalability that is similar to related methods for linear problems. Numerical experiments also indicate that the performance of the algorithms may be further improved by suitable implementation of preconditioners for related type of domain decomposition methods.

References

[1] O. Axelsson, A class of iterative methods for finite element equations. Comp. Meth. in Appl. Mech. and Engng., 9(1976)127-137.

[2] O. Axelsson, Iterative Solution Methods, Cambridge Univ. Press, Cambridge, 1995.

[3] O. Axelsson, V. A. Barker, Finite Element Solution of Boundary Value Problems, Academic Press, London 1984.

[4] R.H. Bielschowsky, A. Friedlander, F.A.M. Gomes, J.M.Martínez, M. Raydan, An adaptive algorithm for bound constrained quadratic minimization, Technical Report, Institute of Mathematics, State University of Campinas, 1995. To appear in Investigación Operativa.

[5] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer Verlag, New York 1991.

[6] A.R. Conn, N.I.M.Gould, Ph.L. Toint, A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds, SIAM J. Num. Anal. 28(1991) 545-572.

[7] Z. Dostál, Conjugate gradient method with preconditioning by projector, Int. J. Comp. Math. 23(1988),315-324.

[8] Z. Dostál, Duality based domain decomposition with inexact subproblem solver for contact problems. Contact Mechanics II, eds. M.H. Alibiadi, C. Alessandri, Wessex Inst. of Technology, Southampton (1995) 461-468.

[9] Z. Dostál, Duality based domain decomposition with proportioning for the solution of free boundary problems, J. Comp. Appl. Math. 63(1995), 203-208.

[10] Z. Dostál, Box constrained quadratic programming with proportioning and projections, SIAM J.Opt. 7 (1997), 871-887.

[11] Z. Dostál, On preconditioning and penalized matrices, to appear in Num.Lin.Alg.Appl.

[12] Z. Dostál, A. Friedlander, S.A. Santos, Augmented Lagrangians with adaptive precision control for quadratic programming with simple bounds and equality constraints, Technical Report RP 74/96, IMECC-UNICAMP, University of Campinas, October 1996.

[13] Z. Dostál, A. Friedlander, S.A. Santos, Solution of coercive and semicoercive contact problems by FETI domain decomposition, Contemporary Math. 218(1998), 82-93.

[14] Z. Dostál, A. Friedlander, S.A. Santos, Augmented Lagrangians with adaptive precision control for quadratic programming with equality constraints, to appear in Comp.Opt.Appl. 1998.

[15] Z. Dostál, V. Vondrák, Duality based solution of contact problems with Coulomb friction, Arch.Mech. 49(1997), 453-460.

[16] G. Duvant, J. L. Lions, Inequalities in Mechanics and Physics, Springer Verlag, Berlin 1976.

[17] C. Farhat, M. Gérardin, On the general solution by a direct method of a large scale singular system of linear equations: application to the analysis of floating structures, Int. J. Num. Met. Eng. 40(1997)

[18] C. Farhat, F.-X. Roux, An unconventional domain decomposition method for an efficient parallel solution of large-scale finite element systems, SIAM J. Sci. Stat. Comput. 13(1992), 379-396.

[19] C. Farhat, P. Chen, F.-X. Roux, The dual Schur complement method with well posed local Neumann problems, SIAM J. Sci. Stat. Comput. 14(1993), 752-759.

[20] C. Farhat, J. Mandel, F.-X. Roux, Optimal convergence properties of the FETI domain decomposition method, Comput. Methods Appl. Mech. Eng. 115 (1994), 365-385.

[21] A. Friedlander, J.M. Martínez, S.A. Santos, A new trust region algorithm for bound constrained minimization, Applied Mathematics & Optimization 30 (1994), 235-266.

[22] A. Friedlander, J.M. Martínez, On the maximization of concave quadratic function with box constraints, SIAM J. Opt. 4(1994) 177-192.

[23] A. Friedlander, J.M. Martínez, M. Raydan, A new method for large scale box constrained quadratic minimization problems, Optimization Methods and Software 5(1995) 57-74.

[24] S. Fučík, A. Kufner, Nonlinear Differential Equations, Elsevier, Amsterdam 1980.

[25] R. Glowinski, J. L. Lions, R. Trémolieres, Analyse Numérique des Inéquations Variationnelles, Tom 1-2, Dunod, Paris 1976.

[26] G. H. Golub, C. F. Van Loan, Matrix Computation, The John Hopkins Univ. Press, Baltimore 1985.

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

[28] N. Kikuchi, J.T. Oden, Contact Problems in Elasticity, SIAM, Philadelphia, 1988.

[29] F.-X. Roux, Spectral analysis of interface operator, Proceedings of the 5th Int. Symp. on Domain Decomposition Methods for Partial Differential Equations, ed. D. E. Keyes at al., SIAM, Philadelphia 1992, pp. 73-90.

[30] F.-X. Roux et al., Efficient parallel coarse grid preconditioner to the FETI method, to appear in Contemporary Math. 218(1998).