Applications of Mathematics

Mark Ainsworth; L. Angela Mihai

A comparison of solvers for linear complementarity problems arising from large-scale masonry structures

Applications of Mathematics, Vol. 51 (2006), No. 2, 93–128

Persistent URL: http://dml.cz/dmlcz/134633

51 (2006) APPLICATIONS OF MATHEMATICS No. 2, 93–128

A COMPARISON OF SOLVERS FOR LINEAR COMPLEMENTARITY PROBLEMS ARISING FROM LARGE-SCALE MASONRY STRUCTURES*

Mark Ainsworth, L. Angela Mihai, Glasgow

(Received September 26, 2005)

Abstract. We compare the numerical performance of several methods for solving the discrete contact problem arising from the finite element discretisation of elastic systems with numerous contact points. The problem is formulated as a variational inequality and discretised using piecewise quadratic finite elements on a triangulation of the domain. At the discrete level, the variational inequality is reformulated as a classical linear complementarity system. We compare several state-of-art algorithms that have been advocated for such problems. Computational tests illustrate the use of these methods for a large collection of elastic bodies, such as a simplified bidimensional wall made of bricks or stone blocks, deformed under volume and surface forces.

Keywords: linear elasticity, equilibrium problems, variational inequality, complementarity problems, masonry structures

MSC 2000: 74B10, 74G15, 49J40, 90C33, 74L99

1. Introduction

An important problem arising in practical engineering applications involves a collection of linearly elastic bodies that are deformed due to volume and surface forces, but cannot penetrate each other. The work presented in this paper is motivated by our interest in masonry structures. We assume that they can be modelled satisfactorily as a linear elasticity system assembled from a large number of elastic components situated at nonnegative distance from one another. Our present objective is to com-

* This work was supported by the Engineering and Physical Science Research Council of Great Britain under grant GR/S35101, and the first author was supported by a fellowship from the Royal Society of Edinburgh.

pare several state-of-art algorithms that have been advocated for the solution of the linear complementarity problem that arise when such problems are discretised.

In Section 2, we describe the model problem in terms of classical partial differential equations of linear elasticity with contact conditions. The problem is formulated as a variational inequality and discretised using piecewise quadratic finite elements on a triangulation of the domain. The treatment of variational inequalities and their applications in continuum mechanics is discussed, for example, in Fichera [4], Duvaut & Lions [3], Glowinski et al. [6], Hlaváček et al. [11], Kikuchi & Oden [13]. In Section 3, at the discrete level, the variational inequality is reformulated as a classical linear complementarity system. In Section 4, we discuss several iterative solvers for the discrete constrained system. The solvers we consider are: the successive over-relaxation with projection, cf. e.g. Glowinski et al. [6], the linear least squares with nonnegativity constraints, cf. Lawson & Hanson [14], the primal-dual active-set method, cf. Hintermüller et al. [9], the primal-dual predictor-corrector method, cf. e.g. Wright [18], and the principal pivoting simplex method, cf. Graves [7]. In Section 5, numerical experiments are presented to illustrate the use of these solvers for a large collection of elastic bodies, such as a simplified bidimensional wall made of bricks or stone blocks, deformed under volume and surface forces. Concluding remarks are addressed in Section 6.

2. Formulation and discretisation of the model problem

We introduce the model contact problem in both the strong and weak forms and discuss the finite element approximation of the problem expressed as a variational inequality. Both the primal formulation of the problem (i.e. in terms of displacements only) and the primal-dual formulation (i.e. in terms of displacements and stresses) will be needed in view of the fact that the different solvers we consider are sometimes viewed more naturally in terms of the primal or dual problem.

The mathematical model. We consider an elastic system consisting of a finite, but possibly large, number of elastic bodies situated at nonnegative distance from one another (Fig. 1). Each body occupies a Lipschitz domain $\Omega^k \subset \mathbb{R}^d$, $d = 2$ or 3, $k = 1, \ldots, n_b$, and the domain occupied by the overall system is defined as $\Omega = \Omega^1 \cup \ldots \cup \Omega^{n_b}$. Let $\partial\Omega = \partial\Omega^1 \cup \ldots \cup \partial\Omega^{n_b}$ represent the global boundary of Ω. We denote by Γ_C the potential contact surface between the elastic bodies, and by $\Gamma_B = \partial\Omega \setminus \Gamma_C$ the exterior boundary of the overall system.

Let $u(x) = (u_1(x), \ldots, u_d(x))$, $x \in \Omega$, denote the vector field of displacements of the elastic system, and let e_{kl} $(k, l = 1, \ldots, d)$ represent the corresponding linearised

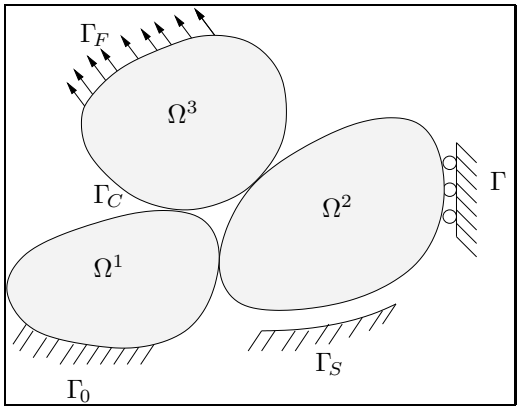

Figure 1. The assembled domain $\Omega \subset \mathbb{R}^2$, with $n_b = 3$.

tensor field of strain,

$$e_{kl} = \frac{1}{2}\Big(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\Big).$$

The forces acting in the interior of the elastic system are characterised by the stress tensor σ, given by the usual constitutive relation (Hooke's law):

$$\sigma_{ij}(u) = a_{ijkl} e_{kl}(u), \quad i, j = 1, \ldots, d.$$

In the above equations, the summation convention has been used. The moduli of elasticity a_{ijkl} are bounded, measurable functions of $x \in \Omega$ which satisfy the usual symmetry conditions:

$$a_{ijkl} = a_{klij} = a_{jikl}, \tag{2.1}$$

and the ellipticity condition that there exists a positive constant $c_0 > 0$ such that for all symmetric ξ_{ij},

$$a_{ijkl}(x)\xi_{ij}\xi_{kl} \geqslant c_0 \xi_{ij}\xi_{ij}. \tag{2.2}$$

In the special case when the elastic bodies are made of a material which is homogeneous (i.e. the material properties are independent of position) and isotropic (i.e. the behaviour of the material is the same in all directions), the stress tensor can be written explicitly as

$$\sigma_{ij}(u) = \lambda \delta_{ij} e_{ll}(u) + 2\mu e_{ij}(u), \quad i, j = 1, \ldots, d,$$

where $0 \leqslant \lambda$ and $0 < \mu$ are the Lamé parameters which represent physical properties of the given material. Instead of the Lamé coefficients, the Young modulus (elasticity

modulus) E and the Poisson ratio ν can be used. These are defined implicitly by

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{2(1+\nu)}.$$

We recall that Poisson's ratio takes its values within the theoretical interval $[-1, 0.5]$, with the value 0.5 corresponding to perfectly incompressible materials. Most practical engineering materials have Poisson's ratio between 0.0 and 0.5. Thus,

$$\sigma_{ij}(u) = \lambda\delta_{ij}\frac{\partial u_l}{\partial x_l} + \mu\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big), \quad i,j = 1,\dots,d.$$

Let $f = (f_1,\dots,f_d)$ be the prescribed volume force densities in the elastic bodies, i.e. $f_i = \varrho F_i$, where ϱ represents density and F_i is a body force, most usually due to gravity, and $f_i \in L^2(\Omega)$ $(i = 1,\dots,d)$. Then the displacement u of the elastic system is modelled by the following equations of mechanical equilibrium:

$$\mu\Delta u_i + (\lambda+\mu)\frac{\partial^2 u_j}{\partial x_i \partial x_j} + f_i = 0 \quad \text{in } \Omega, \qquad i = 1,\dots,d.$$

These equations can be written equivalently as

$$\frac{\partial}{\partial x_j}\sigma_{ij} + f_i = 0 \quad \text{in } \Omega, \qquad i = 1,\dots,d. \tag{2.3}$$

We assume that, in the initial stage, the bodies are two by two in contact and that the overall zone of contact is equal to Γ_C. We also suppose that the unknown final contact zone after deformation will be included in Γ_C. In order to express the conditions on $\partial\Omega$, for every $k = 1,\dots,n_b$, we define

$$\begin{cases} u_N^k = u^k \cdot n^k, \quad u_T^k = u^k - u_N^k n^k, \\ \sigma_N^k = \sigma_{ij}^k n_i^k n_j^k, \quad (\sigma_T^k)_i = \sigma_{ij}^k n_j^k - \sigma_N^k n_i^k, \quad i,j = 1,\dots,d, \end{cases} \tag{2.4}$$

where $n^k = (n_1^k,\dots,n_d^k)$ represents the outward unit normal to $\partial\Omega^k$, $u^k = u|_{\Omega^k}$, and $\sigma^k = \sigma|_{\Omega^k}$. Then, we set the normal

$$n = \begin{cases} n^k \text{ or } n^l & \text{on } \partial\Omega^k \cap \partial\Omega^l, \quad k,l = 1,\dots,n_b, \ k \neq l, \\ n^k & \text{on } \partial\Omega^k \cap \partial\Gamma_B, \quad k = 1,\dots,n_b, \end{cases} \tag{2.5}$$

where $n^l = -n^k$. Correspondingly, we denote

$$u_N = \begin{cases} u_N^k \text{ or } u_N^l & \text{on } \partial\Omega^k \cap \partial\Omega^l, \quad k,l = 1,\dots,n_b, \ k \neq l, \\ u_N^k & \text{on } \partial\Omega^k \cap \partial\Gamma_B, \quad k = 1,\dots,n_b, \end{cases}$$

etc.

On Γ_C, the conditions are as follows:

- only nonpenetrative displacements are allowed, i.e.

$$[u_N] = u_N^k + u_N^l \leqslant 0 \quad \text{on } \partial\Omega^k \cap \partial\Omega^l, \quad k, l = 1, \ldots, n_b, \ k \neq l; \tag{2.6}$$

- the action and reaction principle is satisfied, i.e.

$$\sigma_N = \sigma_N^k = \sigma_N^l \quad \text{on } \partial\Omega^k \cap \partial\Omega^l, \quad k, l = 1, \ldots, n_b, \ k \neq l; \tag{2.7}$$

- the following complementarity condition holds:

$$\sigma_N [u_N] = 0, \tag{2.8}$$

  such that where two bodies have a contact point, they do not attract each other, i.e.

$$[u_N] = 0 \Rightarrow \sigma_N \leqslant 0;$$

- for the restricted class of frictionless problems, the tangential (shear, frictional) stress is equal to zero, i.e.

$$\sigma_T = \sigma_T^k = 0, \quad k = 1, \ldots, n_b. \tag{2.9}$$

Finally, the outer boundary Γ_B is partitioned as $\Gamma_B = \Gamma_0 \cup \Gamma_D \cup \Gamma_F \cup \Gamma_S$, according to the following

- Dirichlet conditions:

$$\begin{aligned} &u_i = 0 \quad \text{on } \Gamma_0, \quad i = 1, \ldots, d \quad \text{(body is clamped)}, \\ &u_N = 0, \quad \sigma_T = 0 \ \text{ on } \Gamma_D \quad \text{(data are prescribed)}. \end{aligned} \tag{2.10}$$

- Neumann conditions:

$$\sigma_{ij} n_j = \theta_i \quad \text{on } \Gamma_F, \quad i = 1, \ldots, d \quad \text{(surface forces are acting)}, \tag{2.11}$$

  where $\theta_i \in L^2(\Gamma_F)$ represent traction forces.

- Signorini conditions:

$$u_N \leqslant g, \quad \sigma_N \leqslant 0, \quad \sigma_N (u_N - g) = 0, \quad \sigma_T = 0 \quad \text{on } \Gamma_S, \tag{2.12}$$

where g is a bounded, measurable initial gap between the body's surface and a rigid obstacle.

The variational form. To be able to describe the behaviour of the given elastic system, first, we state the model problem (2.3)–(2.12) (the strong form) in the equivalent variational form (the weak form), then show existence and uniqueness for the solution of this problem. For more comprehensive details and proofs we refer, for instance, to Hlaváček et al. [11] and Kikuchi & Oden [13]. For the subsequent analysis, important function spaces are the following ones:

- $L^2(D)$, $L^\infty(D)$, and $H^1(D)$, where $D \subset \mathbb{R}^d$, are the usual Hilbert spaces.
- $\mathcal{H}^1(\Omega) = \{v = (v^1, \dots, v^{n_b})\colon\ v^k \in [H^1(\Omega^k)]^d,\ k = 1, \dots, n_b\}$ is the space of displacement functions with finite energy, endowed with the broken norm

$$|||v|||_{1,\Omega} = \left(\sum_{k=1}^{n_b} \|v^k\|^2_{1,\Omega^k} \right)^{1/2},$$

  where $\|\cdot\|_{1,\Omega^k}$ represents the norm in $[H^1(\Omega^k)]^d$.
- $\mathcal{H}^1(\Gamma) = \{v_N\colon\ v \in \mathcal{H}^1(\Omega)\}$ is the range of $\mathcal{H}^1(\Omega)$ by the normal trace operator on $\Gamma = \Gamma_C \cup \Gamma_S$.
- $\mathcal{V} = \{v \in \mathcal{H}^1(\Omega)\colon\ v = 0 \text{ on } \Gamma_0,\ v_N^k = 0 \text{ on } \Gamma_D^k,\ k = 1, \dots, n_b\}$ is the space of virtual displacements.

We define the bilinear form

$$a(u, v) = \int_\Omega a_{ijkl} e_{ij}(u) e_{kl}(v)\, \mathrm{d}x, \quad \forall\, u, v \in \mathcal{V},$$

and the linear form

$$b(v) = \int_\Omega f_i v_i\, \mathrm{d}x + \int_{\Gamma_F} \theta_i v_i\, \mathrm{d}\gamma, \quad \forall\, v \in \mathcal{V},$$

where $\mathrm{d}\gamma$ is the element of area on Γ_F (the arc length along Γ_F, when this is a one-dimensional set).

Let

$$\mathcal{K} = \{v \in \mathcal{V}\colon\ [v_N] \leqslant 0 \text{ on } \Gamma_C,\ v_N \leqslant g \text{ on } \Gamma_S\}$$

be the closed convex set of admissible displacements. Then the displacement in the model system is the solution of the primal variational inequality: find $u \in \mathcal{K}$ such that

$$a(u, v - u) \geqslant b(v - u), \quad \forall\, v \in \mathcal{K}. \tag{2.13}$$

The equivalence between (2.3)–(2.12) and (2.13) can be proved by arguments analogous to those given, for example, in Hlaváček et al. [11, Chapter 2.1], or Kikuchi & Oden [13, Chapter 2].

Next, we define the functional of total potential energy as

$$\mathcal{L}(v) = \frac{1}{2}a(v,v) - b(v), \quad \forall\, v \in \mathcal{K}.$$

Then the variational inequality (2.13) can be rewritten as the minimisation problem: find the solution $u \in \mathcal{K}$ to

$$\mathcal{L}(u) = \min_{v \in \mathcal{K}} \mathcal{L}(v), \tag{2.14}$$

or equivalently: find $u \in \mathcal{K}$ such that

$$\mathcal{L}(u) \leqslant \mathcal{L}(v), \quad \forall\, v \in \mathcal{K}.$$

Under the previous assumptions on the moduli of elasticity, the functional $\mathcal{L}(v)$ is convex (Kikuchi & Oden [13, p. 113]), i.e.

$$\mathcal{L}(\alpha u + (1-\alpha)v) \leqslant \alpha\mathcal{L}(u) + (1-\alpha)\mathcal{L}(v), \quad \forall\, u, v \in \mathcal{K}, \ \forall \alpha \in [0,1],$$

and lower semicontinuous, i.e.

$$\mathcal{L}(u) \leqslant \liminf_{n\to\infty} \mathcal{L}(u_n) \quad \text{for} \quad u_n \to u.$$

We also assume that $\mathcal{L}(v)$ is proper, i.e. $\mathcal{L}(u) > -\infty$ for all $u \in V$, and $\mathcal{L} \neq +\infty$.

For the problem (2.14), in the case considered by our numerical examples (see Section 5), existence and uniqueness of the solution are established through the following results.

Let

$$\mathcal{R} = \prod_{k=1}^{n_b} \mathcal{R}_k, \quad \text{where } \mathcal{R}_k = \{(0, r_k)\colon\ r_k \in \mathbb{R}\}$$

and

$$\mathcal{R}^+ = \prod_{k=1}^{n_b} \mathcal{R}_k^+, \quad \text{where } \mathcal{R}_k^+ = \{(0, r_k)\colon\ r_k \geqslant 0\}.$$

Lemma 2.1. *If the compatibility condition*

$$b(v) < 0, \quad \forall\, v \in \mathcal{K} \cap \mathcal{R}^+ \setminus \{0\} \tag{2.15}$$

is satisfied, then for every elastic body that occupies a domain Ω^k, $k = 1, \ldots, n_b$, there exists at least one point of contact with a body that occupies a domain Ω^l situated below Ω^k, or with the rigid obstacle on Γ_S, i.e. $[u_N(\gamma)] = 0$ for some $\gamma \in \partial\Omega^k \cap \partial\Omega^l$, or $u_N(\gamma) = g$ for some $\gamma \in \Gamma_S$.

P r o o f. From the total potential energy, by contradiction. □

Theorem 2.2. *If the compatibility condition* (2.15) *is satisfied, then there exists at most one solution* $u \in \mathcal{K}$ *to* (2.14).

P r o o f. Let u and w be two solutions of (2.14). Then $u-w = r$, where $r \in \mathcal{V} \cap \mathcal{R}$, and, by Lemma 2.1, $r = 0$. Thus $u = w$. □

Let $r_0 = (r_0^1, \dots, r_0^{n_b})$, where $r_0^k = (0, -g)$ for the bodies which can settle on Γ_S and $r_0^k = (0,0)$ for the remaining bodies, $k = 1, \dots, n_b$. Then $\mathcal{K} = \mathcal{K}_0 + r_0$, where $\mathcal{K}_0$ is a cone such that $\mathcal{K}_0 \cap \mathcal{R} = \mathcal{K} \cap \mathcal{R}^+$.

Theorem 2.3. *If the compatibility condition* (2.15) *is satisfied, then the functional of total potential energy* $\mathcal{L}(v)$ *is coercive on* $\mathcal{K}$ *and there exists a solution of the problem* (2.14).

P r o o f. See Hlaváček and Nedoma [12, Theorem 4.1]. □

Next, we define

$$\mathcal{M} = \Big\{ \Lambda \in \mathcal{H}^1(\Gamma) \colon \int_\Gamma \Lambda\psi \,\mathrm{d}\gamma \leqslant 0, \quad \forall\, \psi \in \mathcal{H}^1(\Gamma), \quad \psi \leqslant 0 \ \text{ on } \ \Gamma \Big\}$$

as the closed convex cone of admissible Lagrangian multipliers. Then, we consider the primal-dual problem: find $(u, \Sigma) \in \mathcal{V} \times \mathcal{M}$ such that

$$\begin{cases} a(u,v) + m(v,\Sigma) = b(v), & \forall\, v \in \mathcal{V}, \\ m(u, \Lambda - \Sigma) \leqslant 0, & \forall\, \Lambda \in \mathcal{M}, \end{cases} \tag{2.16}$$

where

$$m(v,\Lambda) = \int_{\Gamma_C} \Lambda[v_N] \,\mathrm{d}\gamma + \int_{\Gamma_S} \Lambda(v_N - g) \,\mathrm{d}\gamma.$$

The equivalence between (2.13) and (2.16) can be shown by arguments analogous to those given, for example, in Hlaváček et al. [11, Chapter 2.4], or Kikuchi & Oden [13, Chapter 3].

Let

$$\bar{\mathcal{L}}(v,\Lambda) = \frac{1}{2} a(v,v) + m(v,\Lambda) - b(v), \quad \forall\, (v,\Lambda) \in \mathcal{V} \times \mathcal{M}.$$

Then (2.16) can be written equivalently as the saddle-point problem: find the solution $(u, \Sigma) \in \mathcal{V} \times \mathcal{M}$ to

$$\bar{\mathcal{L}}(u,\Sigma) = \min_{v \in \mathcal{V}} \max_{\Lambda \in \mathcal{M}} \bar{\mathcal{L}}(v,\Lambda), \tag{2.17}$$

or equivalently: find $(u, \Sigma) \in \mathcal{V} \times \mathcal{M}$ such that

$$\bar{\mathcal{L}}(u,\Lambda) \leqslant \bar{\mathcal{L}}(u,\Sigma) \leqslant \bar{\mathcal{L}}(v,\Sigma), \quad \forall\, v \in \mathcal{V}, \ \forall \Lambda \in \mathcal{M}.$$

The finite element approximation. We approximate the given contact problem by the h-version piecewise quadratic finite element method with nodal (Lagrange) basis as follows. For every $k = 1, \ldots, n_b$, we consider a standard finite element partition $\mathcal{T}_h^k = \{\tau_h^k\}$ of the subdomain $\Omega^k \subset \Omega$, as described in Brenner & Scott [2, Chapter 3]. We choose the degrees of freedom to be the vertices and the midpoints of the edges of each element τ_h^k, and require the triangulation $\mathcal{T}_h^k$ to be consistent with the boundary conditions on $\partial\Omega^k$ in the following sense: a node of an element $\tau_h^k \in \mathcal{T}_h^k$ lying on $\partial\Omega^k \cap \partial\Omega^l \subseteq \Gamma_C$ $(k \neq l)$ must also be a node of an element $\tau_h^l \in \mathcal{T}_h^l$, and each point on $\partial\Omega^k \cap \Gamma_B$ at which the boundary condition changes must be a node of an element $\tau_h^k \in \mathcal{T}_h^k$. We denote by Φ_h^k the set of degrees of freedom of the triangulation $\mathcal{T}_h^k$ lying on Γ. Let $\mathcal{P}_2(\tau_h^k)$ be the set of quadratic polynomials (i.e. polynomials of degree less than or equal to 2 globally with respect to all space variables) defined on the element τ_h^k, with the element nodal basis. We define the piecewise quadratic finite element subspace of $[H^1(\Omega^k)]^d$ associated with $\mathcal{T}_h^k$ as

$$\mathcal{S}_h(\Omega^k) = \{v^k \in [H^1(\Omega^k)]^d \colon\ v^k|_{\tau_h^k} \in [\mathcal{P}_2(\tau_h^k)]^d, \quad \forall\, \tau_h^k \in \mathcal{T}_h^k\}.$$

We define $\mathcal{T}_h = \mathcal{T}_h^1 \cup \ldots \cup \mathcal{T}_h^{n_b}$ to be the triangulation of the domain $\Omega \subset \mathbb{R}^d$ and denote by Φ_h the set of degrees of freedom of $\mathcal{T}_h$ lying on Γ, $\Phi_h = \Phi_h^1 \cup \ldots \cup \Phi_h^{n_b}$. Then, we set

$$\mathcal{S}_h(\Omega) = \{v \in \mathcal{H}^1(\Omega) \colon\ v|_{\tau_h} \in [\mathcal{P}_2(\tau_h)]^d,\ \ \forall\, \tau_h \in \mathcal{T}_h\}$$

to be the piecewise quadratic finite element subspace of $\mathcal{H}^1(\Omega)$ associated with $\mathcal{T}_h$. In order to express the contact constraints on Γ, we also set

$$\mathcal{S}_h(\Gamma) = \{v_N \colon\ v \in \mathcal{S}_h(\Omega)\}$$

to be the range of $\mathcal{S}_h(\Omega)$ under the normal trace operator on Γ.

Next, we define

$$\mathcal{V}_h = \{v \in \mathcal{V} \colon\ v|_{\tau_h} \in [\mathcal{P}_2(\tau_h)]^d,\ \ \forall\, \tau_h \in \mathcal{T}_h\}$$

to be the discrete space of virtual displacements,

$$\mathcal{K}_h = \{v \in \mathcal{V}_h \colon\ [v_N](\xi) \leqslant 0,\ \ \forall\, \xi \in \Phi_h \cap \Gamma_C,\ \ v_N(\xi) \leqslant g(\xi),\ \ \forall\, \xi \in \Phi_h \cap \Gamma_S\}$$

to be the closed convex set of discrete admissible displacements, and

$$\mathcal{M}_h = \left\{\Lambda \in \mathcal{S}_h(\Gamma) \colon\ \int_\Gamma \Lambda\psi \,\mathrm{d}\gamma \leqslant 0,\ \ \forall\, \psi \in \mathcal{S}_h(\Gamma),\ \ \psi(\xi) \leqslant 0,\ \ \forall\, \xi \in \Phi_h\right\}$$

to be the closed convex cone of discrete admissible Lagrangian multipliers. Note that, in the definitions of $\mathcal{K}_h$ and $\mathcal{M}_h$, the inequalities hold only at the mesh points and at the midpoints of element edges, which will lead to convex constraints of linear type.

Now, the finite element formulation of the problem (2.13) is: find $u_h \in \mathcal{K}_h$ such that

$$a(u_h, v_h - u_h) \geqslant b(v_h - u_h), \quad \forall\, v_h \in \mathcal{K}_h, \tag{2.18}$$

i.e. find the solution $u_h \in \mathcal{K}_h$ to

$$\mathcal{L}(u_h) = \min_{v_h \in \mathcal{K}_h} \mathcal{L}(v_h).$$

Equivalently, the finite element formulation of the problem (2.16) is: find $(u_h, \Sigma_h) \in \mathcal{V}_h \times \mathcal{M}_h$ such that

$$\begin{cases} a(u_h, v_h) + m(v_h, \Sigma_h) = b(v_h), & \forall\, v_h \in \mathcal{V}_h, \\ m(u_h, \Lambda_h - \Sigma_h) \leqslant 0, & \forall\, \Lambda_h \in \mathcal{M}_h, \end{cases} \tag{2.19}$$

i.e. find the solution $(u_h, \Sigma_h) \in \mathcal{V}_h \times \mathcal{M}_h$ to

$$\bar{\mathcal{L}}(u_h, \Sigma_h) = \min_{v_h \in \mathcal{V}_h} \max_{\Lambda_h \in \mathcal{M}_h} \bar{\mathcal{L}}(v_h, \Lambda_h).$$

Approximations of contact problems by the finite element method are discussed, for example, in Hlaváček et al. [11, Chapter 2.3], while more general results on the finite element approximations of variational inequalities can be found in Kikuchi & Oden [13, Chapter 4].

3. Matrix formulations

In this section we reformulate the above discrete problem as a classical linear complementarity system and show how the given problem can be approximated by regularised problems with a symmetric positive definite Hessian. Here we switch to matrix-vector formulation to describe the algorithms.

Let $\{\varphi_j\}_{j=1}^m$ represent the basis functions of $\mathcal{S}_h(\Omega)$. Inserting the expression

$$u_h(x) = \sum_{j=1}^{m} u_j \varphi_j(x), \quad x \in \Omega,$$

into (2.18) generates the equivalent quadratic program with (selected) inequality constraints (QIP): find the minimiser $u \in \mathbb{R}^{n\times 1}$ for

$$L(v) = \frac{1}{2} v^T A v - v^T b, \tag{3.1}$$

defined for all $v \in \mathbb{R}^{n\times 1}$, such that the following componentwise inequality between vectors holds:

$$Cv \geqslant c, \tag{3.2}$$

where A is a symmetric and positive semidefinite $n \times n$ matrix, b is an $n \times 1$ vector, C is a rectangular $n_c \times n$ matrix, and c is an $n_c \times 1$ vector, with n equal to the number of variables and n_c equal to the number of inequality constraints assigned to the nodes in Γ ($n_c \leqslant n$). We denote by S the set of indices corresponding to the constraints (3.2). Then each row $i \in S$ of the matrix C contains at least one nonzero entry and at most $2d$ nonzero entries corresponding to the components of the normal vector, while each column of C contains at most d such entries. For example, in the two-dimensional case, for every index $i \in S$, let $u_k^i = (u_{k_1}^i, u_{k_2}^i)$ and $u_l^i = (u_{l_1}^i, u_{l_2}^i)$ be two opposite nodes on the potential contact zone such that the constraint $[u_N^i] = u_k^i \cdot n_k^i + u_l^i \cdot n_l^i \leqslant 0$ holds, where $n_k^i = (n_{k_1}^i n_{k_2}^i)$ and $n_l^i = (n_{l_1}^i, n_{l_2}^i)$ are the outward unit normals to the contact zone at u_k^i and u_l^i, respectively. In this case,

$$C_{ij} = \begin{cases} n_{k_1}^i, & \text{if } j = k_1, \\ n_{k_2}^i, & \text{if } j = k_2, \\ n_{l_1}^i, & \text{if } j = l_1, \\ n_{l_2}^i, & \text{if } j = l_2, \\ 0, & \text{otherwise,} \end{cases}$$

and $c_i = 0$. The extension to the three-dimensional case is obvious.

We note that, since A is symmetric and positive semidefinite, the null space of A, denoted by $\mathrm{null}(A)$, is of dimension greater than one, i.e. there exists $r \in \mathbb{R}^{n\times 1} \setminus \{0\}$ such that $Ar = 0$. On the other hand, $L(u) = L(u + r) - r^T b$ for all $r \in \mathrm{null}(A)$. In this case, the existence and uniqueness of the minimiser u for (3.1)–(3.2) are guaranteed by Theorem 2.2 and Theorem 2.3.

A change of variables. We define $\tilde{C}$ to be the $n \times n$ matrix

$$\tilde{C}_{i,:} = \begin{cases} C_{i,:}, & \text{if } i \in S, \\ I_{i,:}, & \text{if } i \in \{1, \ldots, n\} \setminus S, \end{cases}$$

where I is the identity matrix of order n and the index notation '$i,:$' stands for all the entries in the row i of the matrix. We also define the n vector

$$\tilde{c}_i = \begin{cases} c_i, & \text{if } i \in S, \\ 0, & \text{if } i \in \{1, \ldots, n\} \setminus S. \end{cases}$$

Let $\tilde{A} = \tilde{C}^{-T} A \tilde{C}^{-1}$ and $\tilde{b} = \tilde{C}^{-T} b - \tilde{C}^{-T} A \tilde{C}^{-1} \tilde{c}$. Then the matrix $\tilde{A}$ has the same properties as the matrix A, i.e. it is symmetric and positive semidefinite. If we apply the change of variables

$$\tilde{v} = \tilde{C} v - \tilde{c} \Leftrightarrow v = \tilde{C}^{-1}(\tilde{v} + \tilde{c}) \tag{3.3}$$

and replace A by $\tilde{A}$ and b by $\tilde{b}$ in the above QIP, then we obtain an equivalent formulation of this problem (Hlaváček et al. [11, pp. 160–161]). Henceforth, for convenience, we shall drop the tilde from our notation (i.e. set $A = \tilde{A}$ and $b = \tilde{b}$).

Thus, by the change of variables (3.3), we obtain that (3.1)–(3.2) is equivalent to the quadratic program with (selected) nonnegativity constraints (QNP): find the minimiser $u \in \mathbb{R}^{n \times 1}$ for

$$L(v) = \frac{1}{2} v^T A v - v^T b, \tag{3.4}$$

defined for all $v \in \mathbb{R}^{n \times 1}$, such that $v_S \geqslant 0$.

When $n_c < n$, by carrying out a reordering of the unknowns, we may assume that $S = \{1, \ldots, n_c\}$ and $E = \{n_c + 1, \ldots, n\}$. Then the matrix A can be written as

$$A = \begin{bmatrix} A_{SS} & A_{SE} \\ A_{SE}^T & A_{EE} \end{bmatrix},$$

where A_{SS} is the block submatrix associated with the nodes in Γ and A_{EE} is the block submatrix associated with the nodes in $\bar{\Omega} \setminus \Gamma$.

Now, the above QNP can be rewritten as the linear problem with (selected) nonnegativity constraints (LNP): for the vector $u \in \mathbb{R}^{n \times 1}$ solve the componentwise inequality system:

$$\begin{cases} A_{SS} u_S + A_{SE} u_E \geqslant b_S, \\ A_{SE}^T u_S + A_{EE} u_E = b_E \end{cases} \tag{3.5}$$

such that

$$u_S \geqslant 0 \quad \text{and} \quad (A_{SS} u_S + A_{SE} u_E - b_S)^T u_S = 0. \tag{3.6}$$

Equally well, we may rewrite (3.4) as the linear complementarity problem (LCP): for the vector $(u, \Sigma) \in \mathbb{R}^{n\times 1} \times \mathbb{R}^{n\times 1}$ solve the linear system

$$Au + \Sigma = b \tag{3.7}$$

such that

$$u_S \geqslant 0, \quad \Sigma_S \leqslant 0, \quad \Sigma_S^T u_S = 0, \quad \text{and} \quad \Sigma_i = 0, \ \forall\, i \notin S. \tag{3.8}$$

When $n_c = n$, the LNP (3.5)–(3.6) reduces to: for the vector $u \in \mathbb{R}^{n\times 1}$ solve the componentwise inequality system

$$Au \geqslant b \tag{3.9}$$

such that:

$$u \geqslant 0 \quad \text{and} \quad (Au - b)^T u = 0. \tag{3.10}$$

Analogously, the LCP (3.7)–(3.8) reduces to: for $(u, \Sigma) \in \mathbb{R}^{n\times 1} \times \mathbb{R}^{n\times 1}$ solve the linear system

$$Au + \Sigma = b, \tag{3.11}$$

such that

$$u \geqslant 0, \quad \Sigma \leqslant 0, \quad \text{and} \quad \Sigma^T u = 0. \tag{3.12}$$

The reduced Schur complement problem. When $n_c < n$ and A_{EE} is nonsingular, we can define $\hat{A} = A_{SS} - A_{SE}A_{EE}^{-1}A_{SE}^T$ to be the Schur complement (SC) of A_{EE} in A, and $\hat{b} = b_S - A_{SE}A_{EE}^{-1}b_E$. We note that, for A symmetric and positive semidefinite (definite), $\hat{A}$ is also symmetric and positive semidefinite (definite) (Prasolov [17, p. 151]).

After eliminating u_E in (3.5)–(3.6), we obtain the (reduced SC) LNP: for the vector $\hat{u} \in \mathbb{R}^{n_c\times 1}$ solve the componentwise inequality system

$$\hat{A}\hat{u} \geqslant \hat{b} \tag{3.13}$$

such that

$$\hat{u} \geqslant 0 \quad \text{and} \quad (\hat{A}\hat{u} - \hat{b})^T \hat{u} = 0. \tag{3.14}$$

Then, for (3.5)–(3.6), $u_S = \hat{u}$ and $u_E = A_{EE}^{-1}(b_E - A_{SE}^T\hat{u})$.

Similarly, the LCP (3.7)–(3.8) is: find $(u_S, \Sigma_S) \in \mathbb{R}^{n_c \times 1} \times \mathbb{R}^{n_c \times 1}$ and $u_E \in \mathbb{R}^{(n-n_c)\times 1}$ such that

$$\begin{cases} A_{SS}u_S + A_{SE}u_E + \Sigma_S = b_S, \\ A_{SE}^T u_S + A_{EE}u_E = b_E, \end{cases} \tag{3.15}$$

subject to

$$u_S \geqslant 0, \quad \Sigma_S \leqslant 0, \quad \text{and} \quad \Sigma_S^T u_S = 0. \tag{3.16}$$

Again, after eliminating u_E in (3.15), we obtain the (reduced SC) LCP: find $(\hat{u}, \hat{\Sigma}) \in \mathbb{R}^{n_c \times 1} \times \mathbb{R}^{n_c \times 1}$ such that

$$\hat{A}\hat{u} + \hat{\Sigma} = \hat{b}, \tag{3.17}$$

subject to

$$\hat{u} \geqslant 0, \quad \hat{\Sigma} \leqslant 0, \quad \text{and} \quad \hat{\Sigma}^T \hat{u} = 0. \tag{3.18}$$

Then, for (3.15)–(3.16), $u_S = \hat{u}$ and $u_E = A_{EE}^{-1}(b_E - A_{SE}^T \hat{u})$.

The regularised approximation. The fact that many bodies are free to move under rigid body motions means that the matrix A has a large null space. Many algorithms for the solution of the discrete problem require the matrix to be invertible. Hence it will be useful to consider a regularisation of the problem by perturbing the matrix A as follows. Let $A^\varepsilon = A + \varepsilon I$, where I is the identity matrix of order n and $\varepsilon > 0$ is a regularisation parameter. If A is symmetric positive semidefinite, then A^ε is symmetric positive definite.

The regularised (perturbed) LNP: for the vector $u^\varepsilon \in \mathbb{R}^{n\times 1}$ solve the componentwise inequality system:

$$\begin{cases} A_{SS}^\varepsilon u_S^\varepsilon + A_{SE}^\varepsilon u_E^\varepsilon \geqslant b_S, \\ (A_{SE}^\varepsilon)^T u_S^\varepsilon + A_{EE}^\varepsilon u_E^\varepsilon = b_E, \end{cases} \tag{3.19}$$

such that

$$u_S^\varepsilon \geqslant 0 \quad \text{and} \quad (A_{SS}^\varepsilon u_S^\varepsilon + A_{SE}^\varepsilon u_E^\varepsilon - b_S)^T u_S^\varepsilon = 0 \tag{3.20}$$

represents an approximation of (3.5)–(3.6).

Equivalently, the LCP: for the vector $(u^\varepsilon, \Sigma^\varepsilon) \in \mathbb{R}^{n\times 1} \times \mathbb{R}^{n\times 1}$ solve the linear system

$$A^\varepsilon u^\varepsilon + \Sigma^\varepsilon = b \tag{3.21}$$

such that

$$u_S^\varepsilon \geqslant 0, \quad \Sigma_S^\varepsilon \leqslant 0, \quad (\Sigma_S^\varepsilon)^T u_S^\varepsilon = 0, \quad \text{and} \quad \Sigma_i^\varepsilon = 0, \ \forall\, i \notin S, \tag{3.22}$$

represents an approximation of (3.7)–(3.8).

The following result relates the solutions of the original problem to those of the regularised one. In particular, as we shall show later, it means that if the regularisation parameter $\varepsilon > 0$ is chosen appropriately, then we may use information about the active set for the regularised problem to deduce information about the active set for the original problem. The active set represents the contact part between elastic components, where the stress is compressive.

Theorem 3.1. *Let u be the solution of the* QNP (3.4) *such that the compatibility conditions in Theorem* 2.2 *and Theorem* 2.3 *are satisfied. For $\varepsilon > 0$ and B a positive definite $n \times n$ matrix, we define*

$$L^\varepsilon(v) = \frac{1}{2} v^T A v + \frac{1}{2} \varepsilon v^T B v - v^T b \tag{3.23}$$

for all $v \in \mathbb{R}^{n\times 1}$ such that $v_S \geqslant 0$, and $u^\varepsilon \in \mathbb{R}^{n\times 1}$ such that $u_S^\varepsilon \geqslant 0$, to be its minimiser. Then

$$\lim_{\varepsilon \to 0} u^\varepsilon = u.$$

P r o o f. First, we observe that

$$L(u^\varepsilon) = L(u) + (u^\varepsilon - u)^T A u - (u^\varepsilon - u)^T b + \frac{1}{2}(u^\varepsilon - u)^T A (u^\varepsilon - u).$$

Hence, since

$$(u^\varepsilon - u)^T A u \geqslant (u^\varepsilon - u)^T b$$

and A is positive semidefinite, we deduce that

$$L(u^\varepsilon) - L(u) \geqslant \frac{1}{2}(u^\varepsilon - u)^T A (u^\varepsilon - u) \geqslant 0. \tag{3.24}$$

Moreover, by hypothesis,

$$\begin{aligned} L(u^\varepsilon) - L(u) &= L^\varepsilon(u^\varepsilon) - L^\varepsilon(u) + \frac{1}{2}\varepsilon u^T B u - \frac{1}{2}\varepsilon (u^\varepsilon)^T B u^\varepsilon \\ &\leqslant \frac{1}{2}\varepsilon u^T B u - \frac{1}{2}\varepsilon (u^\varepsilon)^T B u^\varepsilon. \end{aligned} \tag{3.25}$$

Therefore, by (3.24) and the inequaliy in (3.25) we obtain

$$(u^\varepsilon)^T B u^\varepsilon \leqslant u^T B u \quad \text{and} \quad \frac{1}{2}(u^\varepsilon - u)^T A (u^\varepsilon - u) \leqslant \frac{1}{2}\varepsilon u^T B u. \tag{3.26}$$

The first estimate in (3.26) shows that the sequence $\{(u^\varepsilon)^T B u^\varepsilon\}_{\varepsilon>0}$ is uniformly bounded. Thus, there exists a convergent subsequence of $\{u^\varepsilon\}_{\varepsilon>0}$, denoted also by $\{u^\varepsilon\}_{\varepsilon>0}$, with a limit $u^* \in \mathbb{R}^{n\times 1}$, such that $u^*_S \geqslant 0$. The second estimate implies that $u^* - u \in \text{null}(A)$. Hence, $u = u^* + r$, where $r \in \text{null}(A)$. Finally, since

$$L(u) = L(u^* + r) = L(u^*) - r^T b$$

and, by Theorem 2.2, $r^T b \leqslant 0$, with equality iff $r = 0$, we deduce that

$$L(u^*) \leqslant L(u), \quad u^* \in \mathbb{R}^{n\times 1}, \ \ u^*_S \geqslant 0,$$

with equality iff $r = 0$. By the uniqueness of the minimiser u, we conclude that $r = 0$. Hence $u^* = u$. Since the choice of $\varepsilon > 0$ was arbitrary, it follows that $\lim\limits_{\varepsilon\to 0} u^\varepsilon = u$. □

Remark 3.2. By taking B to be the identity matrix of order n, an immediate consequence of the above result is that the non-contact set for the solution u of the original LNP (3.5)–(3.6) is contained within the non-contact set for the solution u^ε of the regularised LNP (3.19)–(3.20), i.e. there exists $\bar\varepsilon > 0$ such that, for all $0 < \varepsilon \leqslant \bar\varepsilon$,

$$u_i > 0 \Rightarrow u_i^\varepsilon > 0, \quad \forall\, i \in S.$$

This can be shown as follows. By Theorem 3.1, $\lim\limits_{\varepsilon\to 0} u_i^\varepsilon = u_i$ for all $i \in S$. Hence, if $u_i > 0$, then there exists $\varepsilon_i > 0$ such that $u_i^\varepsilon > 0$ for all $0 < \varepsilon \leqslant \varepsilon_i$. Thus we can take $0 < \bar\varepsilon \leqslant \min\{\varepsilon_i \colon\ u_i > 0,\ \ i \in S\}$.

In practice, we wish to use information about the non-contact set for u^ε to make deductions about the non-contact set for u. Therefore, when performing numerical tests, the finite machine precision $0 < \varepsilon_M \ll 1$ has to be taken into account. We have the following consequence of Theorem 3.1: there exists $\bar\varepsilon > 0$ such that, for all $0 < \varepsilon \leqslant \bar\varepsilon$,

$$u_i^\varepsilon > \varepsilon_M \Rightarrow u_i > 0, \quad \forall\, i \in S.$$

To see this, we choose $\bar\varepsilon$ sufficiently small, such that if $u_i^\varepsilon > \varepsilon_M$, where $i \in S$, then $|u_i - u_i^\varepsilon| < \varepsilon_M$, hence $u_i > u_i^\varepsilon - \varepsilon_M > 0$ for all $0 < \varepsilon \leqslant \bar\varepsilon$.

4. Iterative solvers

In this section, we describe the salient features and give a brief account of the convergence behaviour of several popular iterative methods for the numerical solution of the LNP (3.5)–(3.6), or equivalently, of the LCP (3.7)–(3.8). For extensive convergence and complexity issues we shall refer to the relevant references.

First, we present three algorithms for the solution of the regularised problem. These algorithms are: the successive over-relaxation with projection, cf. e.g. Glowinski et al. [6], the linear least squares with nonnegativity constraints, cf. Lawson & Hanson [14], and the primal-dual active-set method, cf. Hintermüller et al. [9]. In all cases, the existing convergence theory for the algorithms depends on the assumption that the matrix A is positive definite. However, this is not the case in general, nor in the particular situation we consider. For many practical purposes, the solution to the regularised problem may serve quite well as an approximation to the solution of the original problem. Next, we discuss a primal-dual predictor-corrector method, cf. e.g. Wright [18], where the positive semidefinite Hessian can be used directly, but the complementarity condition in (3.8) is approximated by a sequence of duality measures that converges monotonically to zero. Finally, we present the principal pivoting simplex method, cf. Graves [7], for the solution of the reduced SC problem. The previous methods can also be employed to solve this problem.

Successive Over-Relaxation with Projection (PSOR). This is a standard projected relaxation method for solving the regularised LNP (3.19)–(3.20) (Glowinski et al. [6, pp. 66–67, 589]). In this primal method, all basic vectors u are feasible in the sense that they satisfy the constraint $u_S \geqslant 0$.

Algorithm: PSOR(A^ε, b, u^0, n, n_c, ω, *tol*). We choose u^0 such that $u^0_S \geqslant 0$, to be an initial approximation to the vector solution u^ε of (3.19)–(3.20) (without restricting the generality we can assume the starting approximation to be zero), and *tol* to be a tolerance for zero. The steps of the PSOR procedure are as follows.

Step 1: Set $u = u^0$, $S = \{1, \ldots, n_c\}$, $error = tol$, and $iter = 0$.

Step 2: If $error < tol$, then Stop.

Step 3: Set $iter = iter + 1$. For all $i \in S$, compute

$$u_i = \frac{1}{A^\varepsilon_{ii}} \left(b_i - \sum_{j=1}^{i-1} A^\varepsilon_{ij} u_j - \sum_{j=i+1}^{n} A^\varepsilon_{ij} u^0_j \right),$$

then set $u_i = \max\{(1-\omega)u^0_i + \omega u_i, 0\}$.

Step 4: If $n_c < n$, then set $E = \{n_c + 1, \ldots, n\}$ and compute

$$u_E = (A^{\varepsilon}_{EE})^{-1}(b_E - A^{\varepsilon}_{ES} u_S).$$

Step 5: Set $error = \|u^0 - u\|_2 / \|u\|_2$ and $u^0 = u$, then go to Step 2.

The convergence of this algorithm is guaranteed by the following result (Glowinski et al. [6, pp. 68, 589–590]).

Theorem 4.1. *Let u^{ε} denote the solution vector of the regularised* LNP (3.19)–(3.20). *If $0 < \omega < 2$ and u^{iter} is the solution vector at iteration $iter$ of the* PSOR *algorithm, then*

$$\lim_{iter \to +\infty} u^{iter} = u^{\varepsilon}.$$

Linear Least Squares with Selected Nonnegativity Constraints (LSSNN). This is a single principal pivoting technique for the solution of the regularised LCP (3.21)–(3.22) (Lawson & Hanson [14, pp. 160–165], also Gill et al. [5, pp. 180–182]). In this method, all iterates u are feasible in the sense that $u_S \geqslant 0$. Moreover, the complementarity condition $\Sigma_S^T u_S = 0$ is always satisfied, and $\Sigma_i = 0$ for all $i \notin S$. When a feasible dual vector Σ is also found, in the sense that $\Sigma_S \leqslant 0$, the method terminates.

Algorithm: LSSNN(A^{ε}, b, u^0, n, n_c). We choose u^0 such that $u^0_S \geqslant 0$, to be an initial approximation to the basic vector solution u^{ε} of (3.21)–(3.22) (without restricting the generality we can assume the starting approximation to be zero). First, we compute the Cholesky factor R for A^{ε}, i.e. $A^{\varepsilon} = R^T R$. The steps of the LSSNN procedure are as follows.

Step 1: Set $u = u^0$, $S = \{1, \ldots, n_c\}$, $SZ = S$, $SP = \emptyset$, and $iter = 0$.

Step 2: Compute $\Sigma = b - A^{\varepsilon} u$.

Step 3: If $SZ = \emptyset$ or $\Sigma_{SZ} \leqslant 0$, then Stop.

Step 4: Set $iter = iter + 1$. Find $t \in SZ$ such that $\Sigma_t = \max\limits_{i \in SZ} \Sigma_i$, then set $SP = SP \cup \{t\}$ and $SZ = SZ \setminus \{t\}$.

Step 5: Set $\overline{R}$ as the matrix defined by

$$\overline{R}_{:,i} = \begin{cases} R_{:,i}, & \text{if } i \in \{1, \ldots, n\} \setminus SZ, \\ 0, & \text{if } i \in SZ, \end{cases}$$

where the index notation ':, i' stands for all entries in the column i of the matrix. Compute the vector $\bar{u}$ as the solution of the least squares problem:

$$\min_{v} \|\overline{R}v - R^{-T}b\|_2.$$

Set $\bar{u}_i = 0$ for $i \in SZ$.

Step 6: If $\bar{u}_{SP} > 0$, then set $u = \bar{u}$ and go to Step 2.

Step 7: Find all the critical steps, along the search direction $\bar{u} - u$, where the constraints become binding, then take

$$\alpha = \max\Big\{ \frac{u_i}{\bar{u}_i - u_i} \colon\ \bar{u}_i \leqslant 0,\ \ i \in SP \Big\}.$$

Set

$$u = u + \alpha(\bar{u} - u),$$

$SP = SP \setminus \{i \in S \colon\ u_i = 0\}$ and $SZ = SZ \cup \{i \in S \colon\ u_i = 0\}$, then go to Step 5.

On termination, the sets of indices SP and SZ form a partition of S, the vector u satisfies

$$u_{SZ} = 0, \quad u_{SP} > 0,$$

and is a solution vector for the quadratic program with (selected) equality constraints (QEP):

$$\min_{v_{SZ}=0} \Big(\frac{1}{2} v^T A^{\varepsilon} v - v^T b \Big).$$

The dual vector $\Sigma = b - A^{\varepsilon}u$ satisfies

$$\Sigma_{SZ} \leqslant 0, \quad \Sigma_{SP} = 0, \quad \text{and} \quad \Sigma_i = 0,\ \ \forall\, i \notin S.$$

The fact that this algorithm terminates in a finite number of iterations can be shown as follows (Lawson & Hanson [14, p. 164]).

Theorem 4.2. *In the* LSSNN *algorithm, if* $(u^{iter}, \Sigma^{iter})$ *is the basic complementarity solution at Step 2 of the iteration iter, then the sequence of residual norms* $\{\|Ru^{iter} - R^{-T}b\|_2\}$ *strictly decreases with iter.*

The Primal-Dual Active-Set Method (PDAS). This is a block principal pivoting technique for solving the regularised LCP (3.21)–(3.22) (Hintermüller et al. [8], [9], [10]). This primal-dual technique combines two complementary ideas that

lead to rapid convergence. On the one hand, as an active-set strategy, it requires a reduced amount of work at each iteration, where an equality-constrained quadratic problem is solved, with the constraints on the actual estimate of the active set. On the other hand, as a semismooth Newton technique, it achieves superlinear local convergence. An essential feature of the active-set method is that all the basic vectors (u, Σ), except the final one, are infeasible in the sense that u_S and/or Σ_S might change signs several times during the algorithm, although they always satisfy the complementarity condition $\Sigma_S^T u_S = 0$, and $\Sigma_i = 0$ for all $i \notin S$. When a complementarity feasible basic solution is obtained, the method terminates.

Algorithm: PDAS(A^ε, b, u^0, n, n_c). The method requires A^ε to be a P-matrix, that is, all its principal minors to be positive (Berman & Plemmons [1, Chapter 10.2]). It also requires an artificial parameter $\tau > 0$, which is chosen arbitrary. We choose u^0 such that $u_S^0 \geqslant 0$, to be an initial approximation to the basic vector solution u^ε of (3.21)–(3.22) (without restricting the generality we can assume the starting approximation to be zero). The steps of the PDAS procedure are as follows.

Step 1: Set $u = u^0$, $S = \{1, \ldots, n_c\}$, and $iter = 0$.

Step 2: Compute $\Sigma = b - A^\varepsilon u$, then set:

$$SZ = \{i \in S\colon\ u_i + \tau\Sigma_i < 0\} \quad \text{(the active set)},$$
$$SP = \{i \in S\colon\ u_i + \tau\Sigma_i \geqslant 0\} \quad \text{(the inactive set)}.$$

Step 3: If $iter \geqslant 1$, $u_{SZ} = 0$, and $\Sigma_{SP} = 0$, then Stop.

Step 4: Solve (directly, e.g. by Gaussian elimination) for u:

$$\begin{cases} (A^\varepsilon u)_i = b_i, & \forall\, i \in \{1, \ldots, n\} \setminus SZ, \\ u_i = 0, & \forall\, i \in SZ, \end{cases}$$

where $(A^\varepsilon u)_i = A^\varepsilon_{i1} u_1 + \ldots + A^\varepsilon_{in} u_n$.

Step 5: Set $iter = iter + 1$, then go to Step 2.

On termination, the sets of indices SP and SZ form a partition of S, the vector u satisfies

$$u_{SZ} = 0, \quad u_{SP} \geqslant 0,$$

and is a solution vector for the QEP

$$\min_{v_{SZ}=0} \Big(\frac{1}{2} v^T A^\varepsilon v - v^T b \Big).$$

The dual vector $\Sigma = b - A^\varepsilon u$ satisfies:

$$\Sigma_{SZ} < 0, \quad \Sigma_{SP} = 0, \quad \text{and} \quad \Sigma_i = 0, \ \forall\, i \notin S.$$

The local convergence of this algorithm is shown by the following result (Hintermüller et al. [8, Theorem 3.1]).

Theorem 4.3. *In the* PDAS *algorithm, if* (u^0, Σ^0) *is sufficiently close to the complementarity solution* $(u^\varepsilon, \Sigma^\varepsilon)$ *of the regularised* LCP (3.21)–(3.22)*, and* $(u^{iter}, \Sigma^{iter})$ *is the basic complementarity solution at iteration* $iter$*, then the sequence* $\{(u^{iter}, \Sigma^{iter})\}$ *converges superlinearly to* $(u^\varepsilon, \Sigma^\varepsilon)$*.*

The projection procedure. In the algorithms PSOR, LSSNN, and PDAS, the (regularised, perturbed) positive definite matrix A^ε has to be used in place of the positive semidefinite matrix A. As a result, for the primal algorithm, $\{u^{iter}\}$ converges to the solution u^ε of (3.19)–(3.20), or equivalently, for the primal-dual algorithms, $\{(u^{iter}, \Sigma^{iter})\}$ converges to the complementarity solution $(u^\varepsilon, \Sigma^\varepsilon)$ of (3.21)–(3.22). Of course, Theorem 3.1 shows that if $\varepsilon \to 0$, then $\{u^\varepsilon\}$ converges to the solution u of the LNP (3.5)–(3.6), or equivalently, $\{(u^\varepsilon, \Sigma^\varepsilon)\}$ converges to the complementarity solution (u, Σ) of the LCP (3.7)–(3.8). In many applications, the solution to the regularised problem may, for suitably small $\varepsilon > 0$, be regarded as an acceptable approximation to the solution of the original problem. However, if an exact solution is required, this can be obtained by a two-stage process, as follows.

At the first stage, one of the schemes described above is applied to solve the regularised problem. Let u^ε be the final basic vector obtained at this first stage, and let $\Sigma^\varepsilon = b - A^\varepsilon u^\varepsilon$ be the corresponding dual vector. We define

$$SZ^\varepsilon = \{i \in S\colon\ u_i^\varepsilon = 0 \ \text{ and } \ b_i - (Au^\varepsilon)_i - \varepsilon u_i^\varepsilon < 0\},$$
$$SP^\varepsilon = \{i \in S\colon\ u_i^\varepsilon \geqslant 0 \ \text{ and } \ b_i - (Au^\varepsilon)_i - \varepsilon u_i^\varepsilon = 0\}.$$

The second stage consists of a procedure that projects $(u^\varepsilon, \Sigma^\varepsilon)$ onto the nearest point (u, Σ) that satisfies (3.7)–(3.8). The projection algorithm is the PDAS scheme described above, where the matrix A is used instead of A^ε. The starting iterate for this procedure is $u^0 = u^\varepsilon$. Let

$$SZ^0 = \{i \in S\colon\ u_i^0 + \tau(b_i - (Au^0)_i) < 0\},$$
$$SP^0 = \{i \in S\colon\ u_i^0 + \tau(b_i - (Au^0)_i) \geqslant 0\}$$

be the partitioning of S at the beginning of the second stage. It is easy to see that $SZ^\varepsilon \subseteq SZ^0$ and $SP^\varepsilon \subseteq SP^0$ for all $\tau > 0$. Thus $SZ^\varepsilon = SZ^0$ and $SP^\varepsilon = SP^0$.

Hence, if this is the correct partitioning of S corresponding to (u, Σ) (see Remark 3.2), then the projection algorithm successfully terminates in one iteration and as such constitutes a negligible additional cost to the overall procedure.

The Primal-Dual Predictor-Corrector Method (PDPC). This is a Newton approach to solving the LCP (3.7)–(3.8) (Wright [18], Portugal et al. [16]). In this approach, all the basic vectors (u, Σ) are infeasible in the sense that $\Sigma_S^T u_S \neq 0$. After a suitable approximation has been derived by this method, a projection procedure can also be applied to obtain the complementarity solution of (3.7)–(3.8) (Wright [18, pp. 145–149]).

Algorithm: PDPC(A, b, u^0, Σ^0, n, n_c, tol_1, tol_2). We choose an initial basic vector (u^0, Σ^0) such that $u_S^0 > 0$, $\Sigma_S^0 < 0$, and $\Sigma_i^0 = 0$ for all $i \notin S$. We also set tol_1 and tol_2 to be two tolerances for zero. The steps of the PDPC procedure are as follows.

Step 1: Set $iter = 0$, $u = u^0$, $\Sigma = \Sigma^0$, $S = \{1, \ldots, n_c\}$, and $E = \{n_c + 1, \ldots, n\}$.

Step 2: Set $error_1 = |\Sigma_S^T u_S|/n_c$ and $error_2 = \|b - Au - \Sigma\|_2$. If $error_1 < tol_1$ and $error_2 < tol_2$, then Stop.

Step 3: Set $iter = iter + 1$, set $\mu = |\Sigma_S^T u_S|/n_c$. If $iter$ is an odd number, then set $\sigma = 0$, else set $\sigma = 0.99$.

Step 4: Solve (directly, e.g. by Gaussian elimination) for the vector δu, of size n, and the vector $\delta\Sigma$, of size n_c, the linear system

$$\begin{cases} (A\delta u)_i + \delta\Sigma_i = b_i - (Au)_i - \Sigma_i, & \forall\, i \in S, \\ (A\delta u)_i = b_i - (Au)_i, & \forall\, i \in E, \\ \Sigma_i \delta u_i + u_i \delta\Sigma_i = \Sigma_i u_i + \sigma\mu, & \forall\, i \in S, \end{cases}$$

where $(A\delta u)_i = A_{i1}\delta u_1 + \ldots + A_{in}\delta u_n$, and similarly for $(Au)_i$.

Step 5: Take

$$\alpha_1 = \min\Big\{-\frac{u_i}{\delta u_i} :\ \delta u_i < 0,\ \ i \in S\Big\},$$
$$\alpha_2 = \min\Big\{-\frac{\Sigma_i}{\delta\Sigma_i} :\ \delta\Sigma_i > 0,\ \ i \in S\Big\},$$

then set $\alpha_1 = \min\{\alpha_1, 1\}$ and $\alpha_2 = \min\{\alpha_2, 1\}$. If $\delta\Sigma_S^T \delta u_S < 0$, then set

$$\alpha_3 = 0.99(1 - \sigma)\, \frac{\Sigma_S^T u_S}{\delta\Sigma_S^T \delta u_S}.$$

Take

$$\alpha = \begin{cases} \min\{\alpha_1, \alpha_2, \alpha_3\}, & \text{if } \delta\Sigma_S^T \delta u_S < 0, \\ \min\{\alpha_1, \alpha_2\}, & \text{otherwise.} \end{cases}$$

Set $u = u + \alpha\delta u$, $\Sigma = \Sigma + \alpha\delta\Sigma$, and $\Sigma_i = 0$ for all $i \in E$, then go to Step 2.

On termination, the basic vector (u, Σ) satisfies

$$\|Au + \Sigma - b\|_2 < tol_2$$

such that

$$u_S > 0, \quad \Sigma_S < 0, \quad |\Sigma_S^T u_S|/n_c < tol_1, \quad \text{and} \quad \Sigma_i = 0, \ \forall\, i \notin S.$$

We remark that the linear system at Step 4 in the PDPC algorithm has a unique solution. In order to show this, it is sufficient to prove that if $(\delta u, \delta\Sigma)$ is a solution of the corresponding homogeneous system, then $\delta u = 0$ and $\delta\Sigma = 0$. We replace the right-hand side of this system by zeros, to obtain the homogenised system

$$\begin{cases} (A\delta u)_i + \delta\Sigma_i = 0, & \forall\, i \in S, \\ (A\delta u)_i = 0, & \forall\, i \in E, \\ \Sigma_i \delta u_i + u_i \delta\Sigma_i = 0, & \forall\, i \in S. \end{cases} \tag{4.1}$$

In (4.1), when we multiply the first n equations from the left by δu^T, we obtain

$$\delta u^T A \delta u + \delta u_S^T \delta\Sigma = 0.$$

As A is positive semidefinite, this implies

$$\delta u_S^T \delta\Sigma = -\delta u^T A \delta u \leqslant 0. \tag{4.2}$$

Then, in the last n_c equations of (4.1), when we multiply, for every $i \in S$, the corresponding equation by δu_i, we obtain

$$\Sigma_i \delta u_i \delta u_i + u_i \delta\Sigma_i \delta u_i = 0.$$

As $\Sigma_i < 0$, for all $i \in S$, this implies

$$0 \leqslant -\Sigma_i \delta u_i \delta u_i = u_i \delta\Sigma_i \delta u_i. \tag{4.3}$$

Next, since $u_i > 0$ for all $i \in S$, we also deduce

$$0 \leqslant \delta\Sigma_i \delta u_i, \quad \forall\, i \in S. \tag{4.4}$$

When we sum the above inequalities with respect to $i \in S$, we obtain $\delta u_S^T \delta\Sigma \geqslant 0$, which together with (4.2) implies

$$\delta u_S^T \delta\Sigma = 0. \tag{4.5}$$

On the other hand, in (4.3), when we take the maximum of u_i and of Σ_i, respectively, over $i \in S$, by (4.4) we deduce

$$0 \leqslant -\big(\max_{i\in S} \Sigma_i\big)\delta u_i \delta u_i \leqslant \big(\max_{i\in S} u_i\big)\delta\Sigma_i \delta u_i, \quad \forall\, i \in S.$$

After we sum the above inequalities, with respect to $i \in S$, by (4.5) we obtain

$$0 \leqslant -\big(\max_{i\in S} \Sigma_i\big)\delta u_S^T \delta u_S \leqslant \big(\max_{i\in S} u_i\big)\delta u_S^T \delta\Sigma = 0.$$

This implies $\delta u_S^T \delta u_S = 0$ or equivalently, $\delta u_i = 0$, for all $i \in S$. Finally, by replacing these in the first n equations of (4.1), assuming that the restriction of A to the set of indices E is nonsingular (we have previously denoted this submatrix by A_{EE}), we obtain, first, $u_i = 0$ for all $i \in E$, then $\delta\Sigma_i = 0$ for all $i \in S$.

We also note that, at Step 5,

$$\begin{aligned} \alpha_1 &= \max\{\alpha \in [0,1]\colon\ u_i + \alpha\delta u_i \geqslant 0,\ \ i \in S\}, \\ \alpha_2 &= \max\{\alpha \in [0,1]\colon\ \Sigma_i + \alpha\delta\Sigma_i \leqslant 0,\ \ i \in S\}. \end{aligned}$$

Furthermore, when $\delta\Sigma_S^T \delta u_S < 0$, taking

$$\alpha_3 < (1-\sigma)\,\frac{\Sigma_S^T u_S}{\delta\Sigma_S^T \delta u_S}$$

ensures that $error_1$ and $error_2$ are strictly decreasing. Now, the convergence of the PDPC algorithm can be shown as follows.

Theorem 4.4. *In the* PDPC *algorithm, if* $(u^{iter}, \Sigma^{iter})$ *is the basic solution at iteration iter, then the sequence of duality measures* $\{|(\Sigma^{iter})_S^T u_S^{iter}|/n_c\}$ *and the sequence of residual norms* $\{\|b - Au^{iter} - \Sigma^{iter}\|_2\}$ *strictly decrease to zero with every iter.*

We conclude this section with a discussion on how to solve the reduced SC problem. First, we remark that, like the Hessian A, the SC matrix $\hat{A}$ can be approximated by the symmetric positive definite matrix $\hat{A}^\varepsilon = \hat{A} + \varepsilon I$, where I is the identity matrix of order n_c and $\varepsilon > 0$ is a regularisation parameter.

In this case, (3.13)–(3.14) is approximated by the regularised LNP: for the vector $\hat{u}^\varepsilon \in \mathbb{R}^{n_c \times 1}$ solve the componentwise inequality system

$$\hat{A}^\varepsilon \hat{u}^\varepsilon \geqslant \hat{b} \tag{4.6}$$

such that

$$\hat{u}^\varepsilon \geqslant 0 \quad \text{and} \quad (\hat{A}^\varepsilon \hat{u}^\varepsilon - \hat{b})^T \hat{u}^\varepsilon = 0. \tag{4.7}$$

Equivalently, (3.17)–(3.18) is approximated by the regularised LCP: for the vector $(\hat{u}^\varepsilon, \hat{\Sigma}^\varepsilon) \in \mathbb{R}^{n_c \times 1} \times \mathbb{R}^{n_c \times 1}$ solve the linear system

$$\hat{A}^\varepsilon \hat{u}^\varepsilon + \hat{\Sigma}^\varepsilon = \hat{b} \tag{4.8}$$

such that

$$\hat{u}^\varepsilon \geqslant 0, \quad \hat{\Sigma}^\varepsilon \leqslant 0, \quad \text{and} \quad (\hat{\Sigma}^\varepsilon)^T \hat{u}^\varepsilon = 0. \tag{4.9}$$

Thus the algorithms PSOR, LSSNN, and PDAS can be applied to solve the regularised LNP (4.6)–(4.7), or equivalently, the regularised LCP (4.8)–(4.9). Alternatively, the algorithm PDPC can be used to solve the LCP (3.17)–(3.18). Once the desired approximation to the solution of (3.17)–(3.18) has been derived, the projection procedure can be applied to solve this problem exactly. The exact solution can also be obtained directly by the following simplex algorithm.

The principal pivoting simplex method (Graves). This is a simplex method for solving the reduced LCP (3.17)–(3.18) (Graves [7], also Murty [15, Chapter 4]). An advantage of this method is that it produces the exact solution to the problem. Moreover, if (3.17)–(3.18) has no solution, then this algorithm will indicate so. This is achieved by using single or double principal pivots and a vector-valued function which decreases lexicographically at each iteration, i.e. the difference between the vector-value at one iteration and the vector-value at the next iteration is lexico-positive. A nonzero vector is lexico-positive (lexico-negative) if its first nonzero component is positive (negative) (Berman et al. [1, p. 273]).

Algorithm: Graves($\hat{A}$, $\hat{b}$, n_c). The method requires an artificial $n_c \times n_c$ matrix B that is nonsingular and all its rows are initially lexico-positive but it is otherwise arbitrary. For simplicity, we choose B as the identity matrix I of order n_c. The Graves procedure is as follows.

Step 1: Set $A = \hat{A}$, $b = -\hat{b}$, $S = \{1, \ldots, n_c\}$, $SZ = S$, $SP = \emptyset$, $B = I$, and $iter = 0$.

Step 2: If $b \geqslant 0$, then set the solution vector:

$$u_{SZ} = 0, \qquad u_{SP} = b_{SP},$$

the dual vector:

$$\Sigma_{SZ} = -b_{SZ}, \quad \Sigma_{SP} = 0,$$

and Stop.

Step 3: Set $iter = iter + 1$. Find t such that

$$\frac{B_{t,:}}{b_t} = \max_{b_i<0}(B_{i,:}/b_i),$$

where max represents the maximum in lexicographical order relation. If $t \in SZ$, then set $SP = SP \cup \{t\}$ and $SZ = SZ \setminus \{t\}$, else set $SP = SP \setminus \{t\}$ and $SZ = SZ \cup \{t\}$.

Step 4: Case 1. When $A_{tt} \neq 0$, set:

$$\begin{aligned}
\bar{A}_{tt} &= \frac{1}{A_{tt}}, \\
\bar{A}_{tj} &= -\frac{A_{tj}}{A_{tt}}, \quad \forall\, j \in \{1, \ldots, n_c\}, \ j \neq t, \\
\bar{A}_{it} &= \frac{A_{it}}{A_{tt}}, \quad \forall\, i \in \{1, \ldots, n_c\}, \ i \neq t, \\
\bar{A}_{ij} &= A_{ij} - A_{it}\frac{A_{tj}}{A_{tt}}, \quad \forall\, i, j \in \{1, \ldots, n_c\}, \ i, j \neq t,
\end{aligned}$$

and

$$\begin{aligned}
\bar{b}_t &= -\frac{b_t}{A_{tt}}, \\
\overline{B}_{tj} &= -\frac{B_{tj}}{A_{tt}}, \quad \forall\, j \in \{1, \ldots, n_c\}, \\
\bar{b}_i &= b_i - A_{it}\frac{b_t}{A_{tt}}, \quad \forall\, i \in \{1, \ldots, n_c\}, \ i \neq t, \\
\overline{B}_{ij} &= B_{ij} - A_{it}\frac{B_{ij}}{A_{tt}}, \quad \forall\, i, j \in \{1, \ldots, n_c\}, \ i \neq t.
\end{aligned}$$

Case 2. When $A_{tt} = 0$, if $A_{it} \geqslant 0$ for all i, then indicate that the problem has no solution and Stop, else find s such that

$$(B_{s,:} - b_s(B_{t,:}/b_t))/A_{st} = \max_{A_{it}<0}(B_{i,:} - b_i(B_{t,:}/b_t))/A_{it},$$

then set:

$$\begin{aligned}
\bar{A}_{sj} &= -\Big(A_{sj} - \frac{A_{ss}}{A_{ts}}A_{tj}\Big)\frac{1}{A_{st}}, \quad \forall\, j \in \{1,\ldots,n_c\},\\
\bar{A}_{ij} &= A_{ij} - \frac{A_{is}}{A_{ts}}A_{tj} - \frac{A_{it}}{A_{st}}\Big(A_{sj} - \frac{A_{ss}}{A_{ts}}A_{tj}\Big), \quad \forall\, i,j \in \{1,\ldots,n_c\},\ \ i \neq s,t,\\
\bar{A}_{tj} &= -\frac{A_{tj}}{A_{ts}}, \quad \forall\, j \in \{1,\ldots,n_c\},
\end{aligned}$$

and

$$\begin{aligned}
\bar{b}_s &= -\Big(b_s - \frac{A_{ss}}{A_{ts}}b_t\Big)\frac{1}{A_{st}},\\
\overline{B}_{sj} &= -\Big(B_{sj} - \frac{A_{ss}}{A_{ts}}B_{tj}\Big)\frac{1}{A_{st}}, \quad \forall\, j \in \{1,\ldots,n_c\},\\
\bar{b}_i &= b_i - \frac{A_{is}}{A_{ts}}b_t - \frac{A_{it}}{A_{st}}\Big(b_s - \frac{A_{ss}}{A_{ts}}b_t\Big), \quad \forall\, i \in \{1,\ldots,n_c\},\ \ i \neq s,t,\\
\overline{B}_{ij} &= B_{ij} - \frac{A_{is}}{A_{ts}}B_{tj} - \frac{A_{it}}{A_{st}}\Big(B_{sj} - \frac{A_{ss}}{A_{ts}}B_{tj}\Big), \quad \forall\, i,j \in \{1,\ldots,n_c\},\ \ i \neq s,t,\\
\bar{b}_t &= -\frac{b_t}{A_{ts}},\\
\overline{B}_{tj} &= -\frac{B_{tj}}{A_{ts}}, \quad \forall j \in \{1,\ldots,n_c\}.
\end{aligned}$$

Step 5: Set $A = \bar{A}$, $b = \bar{b}$, and $B = \overline{B}$, then go to Step 2.

On termination, the sets of indices SP and SZ form a partition of S, the vector u satisfies

$$u_{SZ} = 0, \quad u_{SP} \geqslant 0,$$

and the dual vector $\Sigma = \hat{b} - \hat{A}u$ satisfies

$$\Sigma_{SZ} \leqslant 0, \quad \Sigma_{SP} = 0.$$

This method can be interpreted as follows. Let $\hat{u} = 0$ and $\hat{\Sigma} = \hat{b}$ be a trial solution for the reduced LCP (3.17)–(3.18). If all of the entries of $\hat{b}$ are nonpositive, then this is an acceptable solution, else the trial solution fails to satisfy the second condition in (3.18). In the latter case, another trial solution may be chosen as follows. First, some positive component of $\hat{\Sigma}$ is set to zero, then, the new trial solution is expressed in terms of the corresponding component of $\hat{u}$ and the remaining components of $\hat{\Sigma}$. When this is not possible, two components of $\hat{\Sigma}$ are set to zero, then the new trial solution is expressed in terms of the corresponding components of $\hat{u}$ and the remaining components of $\hat{\Sigma}$. In Graves [7], it is shown that, when $\hat{A}$ is a

positive semidefinite matrix, either one of the two choices mentioned above is always possible, or the problem has no solution. The termination of this algorithm in a finite number of iterations can be proved as follows (Graves [7, Theorem 3]).

Theorem 4.5. *In the Graves algorithm, the vector function* $\max_{b_i<0}(B_{i,:}/b_i)$ *defined at Step* 4 *strictly decreases lexicographically with every iteration iter.*

5. Numerical examples

The aim of this section is to provide a comparison of the algorithms described in Section 4, with respect to the CPU time until successful termination, when solving the (large-scale) LNP (3.5)–(3.6), or equivalently, the LCP (3.7)–(3.8). The information about the test problems is contained in Tabs. 1 and 2, while the performances of the selected algorithms when dealing with these problems are displayed in Tabs. 3 to 7. The implementations of these algorithms have been coded in Matlab 7 and have been tested on a Sun Java Workstation W2100z. No special attempt has been made to optimise the implementation such as to exploit the sparsity.

In the numerical tests, the closure $\bar{\Omega}$ of the domain $\Omega \subset \mathbb{R}^2$, occupied by the given structure, is considered to be the unit square $[0,1]\times[0,1]$. Three different geometries are considered. For a given $H \in (0,1)$, the 'stack bond' system is assembled from squares of size $H \times H$ (Fig. 2 up-left), the 'running bond' system is assembled from rectangular components ('bricks') of size $2H \times H$ (Fig. 2 up-right), and the 'laminae' system is assembled from (long and thin) rectangular domains of size $1 \times H$ (Fig. 2 down).

On the exterior boundary $\Gamma_B = (\{0,1\}\times[0,1]) \cup ([0,1]\times\{0,1\})$, the following conditions are satisfied:

$$\begin{cases} u_N = 0, \;\; \sigma_T = 0 & \text{on } \Gamma_D, \\ \sigma_{ij} n_j = 0 & \text{on } \Gamma_F, \\ u_N \leqslant g, \;\; \sigma_N \leqslant 0, \;\; \sigma_N(u_N - g) = 0, \;\; \sigma_T = 0 & \text{on } \Gamma_S, \end{cases}$$

where the constant $g \in (0, H)$, $\Gamma_D = (\{0,1\}\times[0,1]) \cup ([0.5,1]\times\{0\})$, $\Gamma_F = [0,1]\times\{1\}$, and $\Gamma_S = [0, 0.5)\times\{0\}$. We represent these boundary conditions graphically in Fig. 3 and note that this kind of conditions typically arise in the case of structural masonry undergoing (vertical) settlement of the ground (due to subsidence).

In the finite element approximation, initially, a coarse triangulation is considered, of mesh size $h = H$, as follows: for the 'stack bond' configuration, each square component is partitioned into two triangles; for the 'running bond' geometry, each 'brick' is first partitioned into two squares, then each square is further partitioned

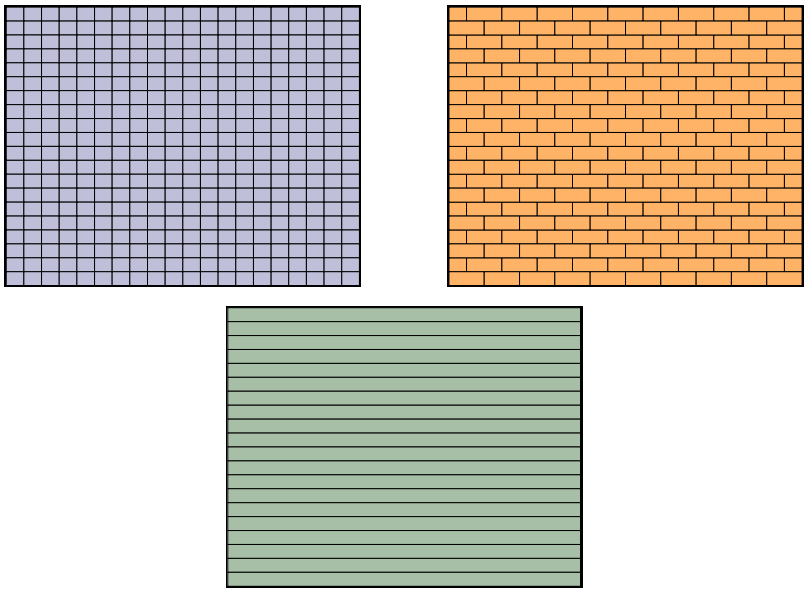

Figure 2. The assembled domain $\Omega \subset \mathbb{R}^2$ with $n_b = 400$, $n_b = 210$, and $n_b = 20$, respectively.

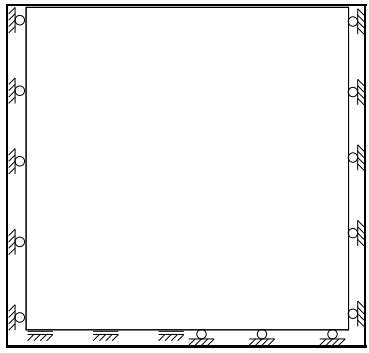

Figure 3. The conditions on the exterior boundary Γ_B of the domain $\Omega \subset \mathbb{R}^2$.

into two triangles, as in the previous case; for 'laminae', each component is first partitioned into $1/H$ squares, then each square is partitioned into two triangles, as before. In Fig. 4, a uniform coarse mesh for each of the three systems represented in Fig. 2 is illustrated. The number n_b of elastic components, with the associated mesh size, is specified in Tab. 1, while the corresponding dimensions n (of A) and n_c (of S) are listed in Tab. 2. The given data are $E = 4 \cdot 10^3$, $\nu = 0.3$, and $f = (0, -1)$.

We derive the exact solution of the LCP (3.7)–(3.8) in two stages, as follows. First, we apply one of the algorithms PSOR, LSSNN, or PDAS to solve the LNP (3.19)–(3.20), or equivalently, the LCP (3.21)–(3.22), with a regularisation parameter $\varepsilon \in$

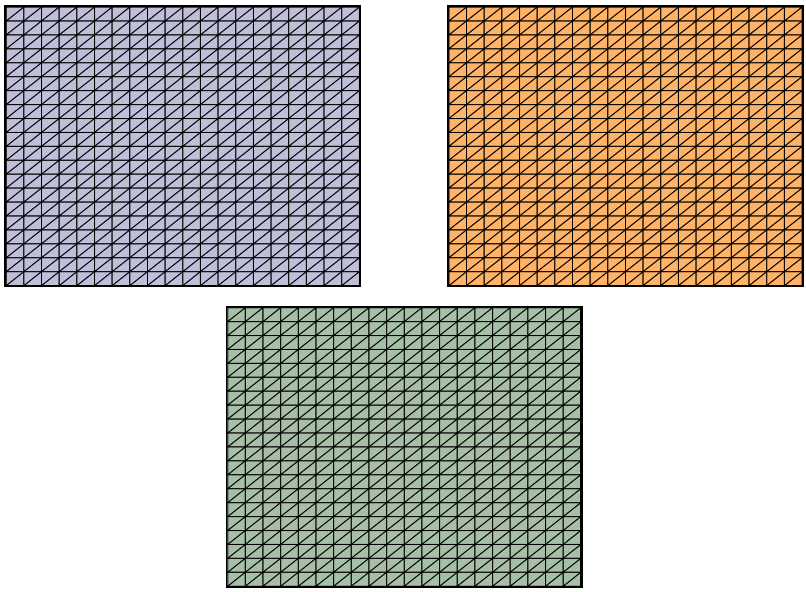

Figure 4. The triangulation of $\Omega \subset \mathbb{R}^2$ with $h = 1/20$, for $n_b = 400$, $n_b = 210$, and $n_b = 20$, respectively.

Stack Bond (n_b)	Running Bond (n_b)	Laminae (n_b)	Mesh Size (h)
25,000	15,000	5	1/5
100,000	55,000	10	1/10
225,000	120,000	15	1/15
400,000	210,000	20	1/20
625,000	325,000	25	1/25
2,500	1,275	50	1/50
5,625	2,850	75	1/75
10,000	5,050	100	1/100
15,625	7,875	125	1/125
22,500	11,325	150	1/150
30,625	15,400	175	1/175
40,000	20,100	200	1/200
50,625	25,425	225	1/225

Table 1. The size of the test problems.

$(10^{-6}, 10^{-4})$. The performances of these algorithms when the initial value is $u^0 = 0$ are recorded in Tabs. 3, 4, and 5, respectively. At the second stage, we employ the projection procedure, which successfully terminates in one iteration. For PSOR, we take $\omega = 2/(1+\sin(\pi/n))$ and $tol \approx n\varepsilon_M$, where $\varepsilon_M \approx 10^{-16}$ represents the machine epsilon. For PDAS and the projection procedure, we choose $\tau \approx 10^{-3}$.

Mesh Size (h)	Number of unknowns	Stack Bond	Running Bond	Laminae
1/5	n	411,000	355,000	298,000
	n_c	126,000	96,000	49,000
1/10	n	1,725	1,461	1,198
	n_c	555,000	417,000	199,000
1/15	n	3,936	3,317	2,698
	n_c	1,281	963,000	449,000
1/20	n	7,050	5,924	4,798
	n_c	2,310	1,735	799,000
1/25	n	11,061	9,280	7,498
	n_c	3,636	2,731	1,249
1/50	n	44,625	37,311	29,998
	n_c	14,775	11,087	4,999
1/75	n	100,686	84,092	67,498
	n_c	33,411	25,068	11,249
1/100	n	179,250	149,624	119,998
	n_c	59,550	44,675	19,999
1/125	n	280,311	233,905	187,498
	n_c	93,186	69,906	31,249
1/150	n	403,875	336,936	269,998
	n_c	134,325	100,762	44,999
1/175	n	549,936	458,717	367,498
	n_c	182,961	137,243	61,249
1/200	n	718,500	599,249	479,998
	n_c	239,100	179,350	79,999
1/225	n	909,561	758,530	607,498
	n_c	302,736	227,081	101,249

Table 2. The number of unknowns for the test problems.

Mesh Size	Stack Bond	Running Bond	Laminae
1/5	0.610	4.920	1.360
1/10	8.750	147.230	25.240
1/15	60.270	1,187.390	151.380
1/20	264.270	6,344.700	587.780

Table 3. PSOR-CPU time (in seconds).

Mesh Size	Stack Bond	Running Bond	Laminae
1/5	0.010	0.100	0.030
1/10	0.050	1.790	0.240
1/15	0.140	9.700	0.890
1/20	0.260	36.050	2.550
1/25	0.450	92.250	5.170
1/50	4.840	2,974.000	77.090

Table 4. LSSNN-CPU time (in seconds).

Mesh Size	Stack Bond	Running Bond	Laminae
1/5	0.030	0.010	0.010
1/10	0.070	0.130	0.040
1/15	0.190	0.470	0.090
1/20	0.460	0.800	0.200
1/25	0.900	1.670	0.320
1/50	7.000	17.850	2.360
1/75	30.460	74.910	27.490
1/100	78.730	180.030	70.120
1/125	180.870	524.650	155.270
1/150	358.260	951.230	303.130
1/175	685.340	2,407.420	486.240
1/200	1,193.680	4,266.140	900.730
1/225	2,080.830	6,232.460	1,429.850

Table 5. PDAS-CPU time (in seconds).

Mesh Size	Stack Bond	Running Bond	Laminae
1/5	0.220	0.250	0.120
1/10	1.020	1.200	0.580
1/15	2.530	3.290	1.370
1/20	4.260	5.920	2.650
1/25	6.970	10.980	4.410
1/50	32.240	69.650	21.440
1/75	84.140	194.670	61.040
1/100	152.460	459.270	123.760
1/125	259.070	873.510	194.220
1/150	419.730	1,421.100	333.700
1/175	621.930	2,275.980	434.530
1/200	869.020	3,365.440	818.480
1/225	1,334.010	4,634.230	971.180

Table 6. PDPC-CPU time (in seconds).

We also apply the algorithm PDPC to solve the LCP (3.7)–(3.8). In this case we note that, due to round-off errors, the system at Step 4 may be found unsolvable. To avoid this, we apply diagonal scaling such that all the entries on the diagonal of A become equal to 1. The performance of PDPC is illustrated in Tab. 6. For this algorithm, the initial values are $u^0 \approx \varepsilon e_n$ and $\Sigma_S^0 \approx \varepsilon e_{n_c}$, where $\varepsilon \in (10^{-6}, 10^{-2})$ and e_n and e_{n_c} are two vectors of size n and n_c, respectively, with all entries equal to 1. The tolerances are $tol_1 \approx n_c \varepsilon_M$ and $tol_2 \approx n\sqrt{\varepsilon_M}$.

Next, we solve the (reduced SC) LCP (3.17)–(3.18) in one stage, using the Graves algorithm, or in two stages, using the algorithms LSSNN or PDAS at the first stage, and the (one iteration) projection scheme at the second stage. We note that, at

Step 4 of the Graves procedure, Case 2 may incorrectly be entered, due to round-off errors (cycling of the algorithm, which never occurs in theory, may be observed in this case). To avoid this, before we construct the SC test problems, we use the diagonal scaling for the original problem as mentioned above. When we apply LSSNN or PDAS to solve the scaled SC problems, we choose the regularisation parameter $\varepsilon \approx n_c \varepsilon_M$. The respective performances of these algorithms are presented in Tab. 7. The performance of the PDPC procedure applied to the SC problem is also recorded there. In this case, $u^0 \approx \varepsilon e_{n_c}$, $\Sigma_S^0 \approx \varepsilon e_{n_c}$, where $\varepsilon \in (10^{-6}, 10^{-2})$, $tol_1 \approx n_c \varepsilon_M$, and $tol_2 \approx n_c \sqrt{\varepsilon_M}$.

The numerical solutions for the test problems with $H = 1/20$ and $H = 1/50$ are graphically illustrated in Figs. 5–6 and 7, respectively.

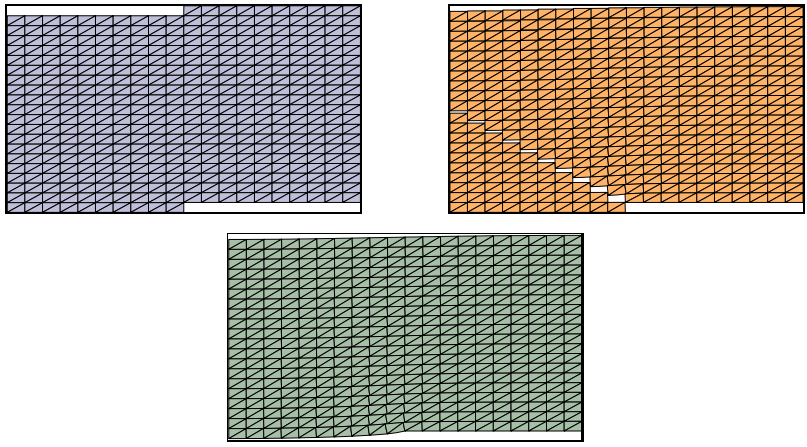

Figure 5. The deformed structure using a mesh of size $h = 1/20$, for $n_b = 400$, $n_b = 210$, and $n_b = 20$, respectively.

Discussion. From Tables 3 to 7, we deduce that all of the methods presented here are very sensitive to the number of physical components n_b. This is expected, as the size of the discrete problem increases with n_b. We also note the sensitivity of these methods to the geometry of the given structure, as every algorithm tends to take longer to achieve successful termination when applied to the 'running bond' configuration than when applied to the other two configurations. This is due to the different contact conditions between the elastic components, specific to each of the three types of structures. We have noticed a similar behaviour when solving the contact problems, on the given structures, with other exterior-boundary conditions.

The PDAS and PDPC techniques are quite efficient in terms of CPU time needed to find the desired solution. For the PDAS algorithm, this is reflected in Tab. 5, as well as in Tab. 7. For the PDPC algorithm, this is indicated in Tab. 6, while

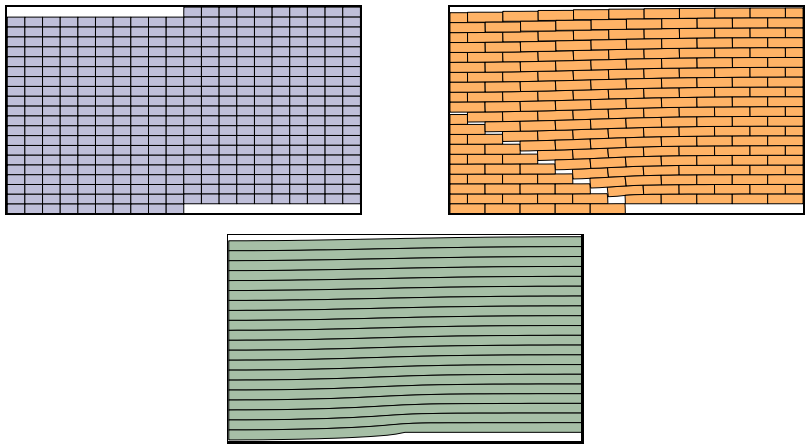

Figure 6. The solution to the test problems with $n_b = 400$, $n_b = 210$, and $n_b = 20$, respectively.

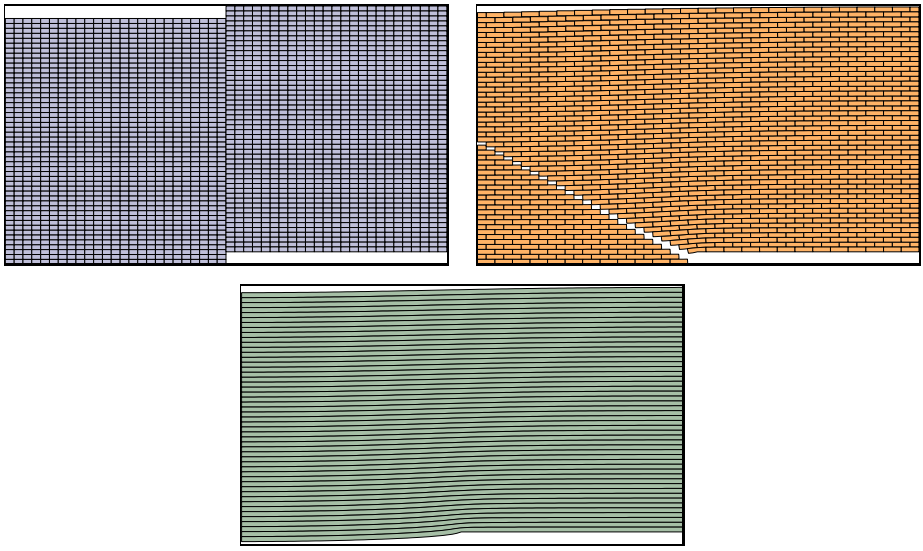

Figure 7. The solution to the test problems with $n_b = 2,500$, $n_b = 1,275$, and $n_b = 50$, respectively.

in Tab. 7 we see that, in its current form, PDPC is unsuited to solve the reduced SC problem, where the SC matrix is dense. In our numerical tests, we have also observed that for the PDAS and PDPC procedures, the number of iterations until successful termination increases only slightly or remains constant when n_b increases, and increases slightly when n_b is fixed and the mesh is refined. However, more iterations are required when solving the problems with the 'running bond' geometry than with the other two types of geometries.

Method	Mesh Size	Stack Bond	Running Bond	Laminae
Graves	1/5	1.190	0.260	0.010
	1/10	1.170	28.060	0.320
	1/15	10.180	530.760	2.500
	1/20	41.850	5,445.910	11.530
LSSNN with projection	1/5	0.010	0.030	0.010
	1/10	0.070	0.750	0.030
	1/15	0.620	10.430	0.140
	1/20	2.250	74.340	0.500
PDAS with projection	1/5	0.020	0.010	0,000
	1/10	0.010	0.040	0.020
	1/15	0.030	0.180	0.010
	1/20	0.020	0.620	0.020
PDPC	1/5	1.280	0.330	0.070
	1/10	115.860	25.400	1.960
	1/15	1,664.700	407.920	27.550
	1/20	8,155.700	1,547.820	175.370

Table 7. CPU time (in seconds) when solving the SC problem.

Tabs. 3 and 4 indicate that LSSNN takes less time to find the desired solution than PSOR, although both these methods are much too slow when n_b is large. Tab. 7 shows that the practical efficiency of the Graves algorithm does not match those of LSSNN and PDAS. However, in contrast with all the other methods discussed here, Graves technique requires no solution to be computed until the final iteration. This particular feature offers an advantage, especially when the coefficient matrix is dense.

6. Conclusions

We conclude that the primal-dual active-set (PDAS) and primal-dual predictor-corrector (PDPC) methods are quite appropriate to solve contact problems of the form (2.3)–(2.12) when the number of elastic components is large, while the successive over-relaxation with projection (PSOR), linear least squares with selected nonnegativity constraints (LSSNN), and principal pivoting simplex (Graves) methods are suitable to solve only the smallest of these problems.

References

[1] *A. Berman, R. J. Plemmons*: Nonnegative Matrices in the Mathematical Sciences. Computer Science and Scientific Computing Series. Academic Press, New York, 1979. Zbl 0484.15016

[2] *S. C. Brenner, L. R. Scott*: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics, Vol. 15. Springer-Verlag, New York, 1994. Zbl 0804.65101

[3] *G. Duvaut, J.-L. Lions*: Inequalities in Mechanics and Physics. Grundlehren der mathematischen Wissenschaften, Vol. 219. Springer-Verlag, Berlin-Heidelberg-New York, 1976. Zbl 0331.35002

[4] *G. Fichera*: Encyclopedia of physics. Existence Theorems in Elasticity-Boundary Value Problems of Elasticity with Unilateral Constraints, Volume VI a/2 (S. Flügge, ed.). Springer-Verlag, Berlin, 1972, pp. 347–427.

[5] *P. E. Gill, W. Murray, and M. H. Wright*: Practical Optimization. Academic Press, London, 1981. Zbl 0503.90062

[6] *R. Glowinski, J.-L. Lions, and R. Trémolières*: Numerical Analysis of Variational Inequalities. Studies in Mathematics and its Applications, Vol. 8. North-Holland, Amsterdam-New York-Oxford, 1981, English version edition. Zbl 0463.65046

[7] *R. L. Graves*: A principal pivoting simplex algorithm for linear and quadratic programming. Oper. Res. *15* (1967), 482–494. Zbl 0154.19604

[8] *M. Hintermüller, K. Ito, and K. Kunisch*: The primal-dual active set strategy as a semismooth Newton method. SIAM J. Optim. *13* (2003), 865–888.

[9] *M. Hintermüller, V. A. Kovtunenko, and K. Kunisch*: The primal-dual active set method for a crack problem with non-penetration. IMA J. Appl. Math. *69* (2004), 1–26. Zbl pre 02109616

[10] *M. Hintermüller, V. A. Kovtunenko, and K. Kunisch*: Generalized Newton methods for crack problems with nonpenetration condition. Numer. Methods Partial Differential Equations *21* (2005), 586–610. Zbl pre 02173394

[11] *I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek*: Solution of Variational Inequalities in Mechanics. Applied Mathematical Sciences, Vol. 66. Springer-Verlag, Berlin-Heidelberg-New York, 1988. Zbl 0654.73019

[12] *I. Hlaváček, J. Nedoma*: On a solution of a generalized semi-coercive contact problem in thermo-elasticity. Math. Comput. Simul. *60* (2002), 1–17. Zbl 1021.74030

[13] *N. Kikuchi, J. T. Oden*: Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods. Studies in Applied Mathematics, Vol. 8. SIAM, Philadelphia, 1988. Zbl 0685.73002

[14] *C. L. Lawson, R. J. Hanson*: Solving Least Squares Problems. Series in Automatic Computation. Prentice-Hall, Englewood Cliffs, 1974. Zbl 0860.65028

[15] *K. G. Murty*: Complementarity, Linear and Nonlinear Programming. Heldermann-Verlag, Berlin, 1988. Zbl 0634.90037

[16] *L. F. Portugal, J. J. Judice, and L. N. Vicente*: A comparison of block pivoting and interior-point algorithms for linear least squares problems with nonnegative variables. Math. Comput. *63* (1994), 625–643. Zbl 0812.90124

[17] *V. V. Prasolov*: Problems and Theorems in Linear Algebra. Translations of Mathematical Monographs, Vol. 134. AMS, Providence, 1994. Zbl 0803.1501

[18] *S. J. Wright*: Primal-Dual Interior-Point Methods. SIAM, Philadelphia, 1997. Zbl 0863.65031

Authors' address: *M. Ainsworth*, *L. A. Mihai*, University of Strathclyde, 26 Richmond Street, Glasgow, G1 1XH, Scotland, UK, e-mails: `m.ainsworth@strath.ac.uk`, `angela.mihai@strath.ac.uk`.