Solution methods for quasi variational inequalities

Simone Sagratella

May 29, 2013

PhD thesis in Operations Research
Sapienza University of Rome

Advisor: Prof. Francisco Facchinei

Introduction

Quasi variational inequalities (QVIs) were introduced by Bensoussan and Lions in a series of papers [4, 5, 6] in connection with the study of impulse control problems and soon they turned out to be **a powerful modeling tool** capable of describing complex equilibrium situations that can appear in such different fields as generalized Nash games (see e.g. [3, 32, 36, 58]), mechanics (see e.g. [2, 7, 33, 38, 51, 52]), economics (see e.g. [36, 62]), statistics (see e.g. [37]), transportation (see e.g. [8, 11]), and biology (see e.g. [31]). We refer the reader to the basic monographs of Mosco [44] and Baiocchi and Capelo [2] for a more comprehensive analysis of QVIs.

Although in the literature hundreds of papers were devoted to various and very important aspects of QVIs, like existence, uniqueness and stability of its solutions, in this thesis we concentrate our work only in one of the most challenging ones: the numerical solution of finite-dimensional QVIs.

State-of-the-art in the numerical solution of QVIs

In spite of their modeling power, relatively **few studies have been devoted to the numerical solution of finite-dimensional QVIs**; a topic which, beside being of great interest in its own, also forms the backbone of solution methods for infinite-dimensional QVIs.

Motivated by earlier research on the implicit complementarity problem [44, 54, 55], Chan and Pang introduced in [9] what is probably the first globally convergent algorithm for a QVI. In this seminal paper, the authors use a fixed point argument to prove convergence of a projection-type algorithm **only in the case in which the QVI falls within the so-called "moving set" class**.

It is safe to say that practically all subsequent works, where globally convergent algorithms are analyzed, consider variants or extensions of the basic setting proposed in [9] and then, still following a fixed point approach, although they are successful in extending the field of "moving set" QVIs for which globally convergence can be garanteed, however **fail to detect new numerical tractable classes of QVIs**, see e.g. [45, 46, 48, 59, 61] and references therein.

In a departure from this setting, Pang and Fukushima [58] proposed a sequential penalty approach to general QVIs. The method in [58] reduces the solution of a QVI to the solution of a sequence of Variational Inequalities (VIs); however, even if this approach is very interesting and promising, its global convergence properties are in jeopardy since they ultimately hinge of the capability of solving a sequence of possibly very challenging VIs.

More recently, Fukushima [27] studied a class of gap functions for QVIs, reducing the solution of a QVI to the global minimization of a nondifferentiable gap function, but no algorithms are explicitly proposed in [27] (see [39] for a further and more detailed application of this approach in a specialized game setting).

This essentially completes the picture of globally convergent proposals for the solution of QVIs. We also mention that Outrata and co-workers studied some interesting local

Newton methods, see [51, 52, 53], but the globalization of these methods is not discussed.

In Section 1 we briefly describe more in detail some important methods (previously cited) in order to compare them to our algorithm.

Main contributions of this thesis

In this thesis we propose a totally different approach to the solution of a QVI. Given the blanket assumption that the feasible set mapping can be described by a finite number of parametric inequalities, we consider the Karush-Kuhn-Tucker (KKT) conditions of the QVI, reformulate them as a system of constrained equations and then apply a powerful interior-point method (which was introduced in [43]). It turns out that the convergence properties of the resulting algorithm depend essentially on the nonsingularity of a certain Jacobian matrix JH. Our main contributions are both **theoretical**:

- an in-depth analysis of the nonsingularity of JH, showing that **global convergence of our method can be obtained in the "moving set" case, but also in many other situations covering a wide array of new and significant settings**, thus enlarging considerably the range of QVIs that it is possible to solve with theoretical guarantees,
- a discussion of the **boundedness of the sequence generated by the algorithm**,

and **numerical**:

- a **collection of test problems** from diverse sources which, being the largest test set for QVIs considered so far, gives a uniform basis on which algorithms for the solution of QVIs can be tested and compared,
- an **implementation** of our new method and a **numerical testing demonstrating its effectiveness and robustness** even if compared to the best solver for complementarity systems: the PATH solver [26].

Organization of the thesis

This thesis is divided into three parts. In **Part I** we formally define the QVI problem and briefly describe some important solution methods available in the literature (**Section 1**). In **Part II** we describe our method and establish global convergence results for many interesting instances of QVIs, vastly broadening the class of problems that can be solved with theoretical guarantees. In particular in **Section 2** we present the Potential Reduction Algorithm, in **Section 3** we define classes of QVIs for which our algorithm globally converges, in **Section 4** we give some further assumptions on the QVI in order to garantee boundedness of the sequence generated by our algorithm and in **Section 5** we specialize results developed in previous sections for generalized Nash equilibrium problems. **Part III** is devoted to numerical issues: in **Section 6** we present the collection of test problems and its Matlab interface and in **Section 7** we describe an implementation and report numerical testing results of our algorithm and its comparison with PATH.

Further information

The material presented in this thesis led to the following publications:

- Francisco Facchinei, Christian Kanzow and **Simone Sagratella**: Solving quasi-variational inequalities via their KKT conditions, **Math. Prog.** Ser. A DOI 10.1007/s10107-013-0637-0 (2013),

- Axel Dreves, Francisco Facchinei, Christian Kanzow and **Simone Sagratella**: On the Solution of the KKT Conditions of Generalized Nash Equilibrium Problems, **SIAM J. Optim.** 21 (3), pp 1082-1108 (2011),

- Francisco Facchinei, Christian Kanzow and **Simone Sagratella**: QVILIB: A Library of Quasi-Variational Inequality Test Problems, **Pac. J. Optim.** (to appear).

In particular Sections 2-4 are extracted from the Math-Prog paper, Section 5 from the SIAM-J-Optim paper and Section 6 from the Pac-J-Optim paper. The author is grateful to all co-authors for their precious work and in particular to its advisor for the priceless guide.

The following are other publications on complementarity issues that do not find a place in this thesis:

- Francisco Facchinei and **Simone Sagratella**: On the computation of all solutions of jointly convex generalized Nash equilibrium problems, **Optim. Lett.** 5 (3), pp 531-547 (2011),

- Francisco Facchinei, Lorenzo Lampariello and **Simone Sagratella**: Recent Advancements in the Numerical Solution of Generalized Nash Equilibrium Problems, in: Recent Advances in Nonlinear Optimization and Equilibrium Problems: a Tribute to Marco D'Apuzzo, V. De Simone, D. di Serafino, and G. Toraldo (eds.), **Quaderni di Matematica**, Dipartimento di Matematica della Seconda Universit di Napoli, vol. 27, pp. 137-174 (2012).

Notations

$\mathbb{R}_+$: set of nonnegative numbers;

$\mathbb{R}_{++}$: set of positive numbers;

$\mathbb{M}_{m\times n}$: set of $m \times n$ matrices;

$\|\cdot\|$: Euclidean norm operator for vectors; spectral norm for matrices, i.e. the norm induced by the Euclidean vector norm (we recall that $M \in \mathbb{M}_{m\times n}$, $\|M\| = \max\{\sqrt{\lambda} \mid \lambda$ is an eigenvalue of $M^T M\}$; the spectral norm is compatible with the Euclidean norm in the sense that $\|Mv\| \leq \|M\|\|v\|$, $v \in \mathbb{R}^n$);

$\Pi_K(\cdot)$: Euclidean projector on K;

$JF(x)$: Jacobian of a differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ at x;

$\nabla F(x)$: transposed Jacobian of a differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ at x;

$J_y F(y, x)$: partial Jacobian with respect to y of a differentiable mapping $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ at (y, x);

$\nabla_y F(y, x)$: transposed partial Jacobian with respect to y of a differentiable mapping $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ at (y, x);

$\mu_m(M)$: minimum eigenvalue of $M \in \mathbb{M}_{n\times n}$ symmetric;

$\mu_m^+(M)$: minimum positive eigenvalue of $M \in \mathbb{M}_{n\times n}$ symmetric;

$\mu_m^s(M)$: minimum eigenvalue of the symmetric part of $M \in \mathbb{M}_{n\times n}$ (*i*.e. $\mu_m(\frac{1}{2}(M^T + M))$);

$\circ$: Hadamard (componentwise) product operator (*i*.e. $x, y \in \mathbb{R}^n$, $x \circ y = (x_i y_i)_{i=1}^n \in \mathbb{R}^n$);

$(\cdot)^{-1}$: componentwise inverse operator (*i*.e. $x \in \mathbb{R}^n$, $x^{-1} = (\frac{1}{x_i})_{i=1}^n \in \mathbb{R}^n$);

$\operatorname{int} K$: interior of K;

$\lfloor\cdot\rfloor$: floor-function;

M_{ij} : $(i, j)^{\text{th}}$ entry of $M \in \mathbb{M}_{m\times n}$, $1 \leq i \leq m$, $1 \leq j \leq n$;

M_{i*} : i^{th} row of $M \in \mathbb{M}_{m\times n}$, $1 \leq i \leq m$;

M_{*j} : j^{th} column of $M \in \mathbb{M}_{m\times n}$, $1 \leq j \leq n$;

$\operatorname{diag}(v)$: (square) diagonal matrix whose diagonal entries are the elements of $v \in \mathbb{R}^n$;

I : (square) identity matrix.

Contents

Part I
Preliminaries

1 Quasi Variational Inequalities

In this section we formally describe the quasi variational inequality problem and, under some assumptions, we derive its KKT conditions, which are fundamental for our subsequent analysis. Then we briefly describe some important solution methods available in the literature in order to make a theoretical comparison of our algorithm. Moreover we report some existence result (with an algorithmic genesis) for QVIs, among the multitude obtainable from the literature, because we recall them in the sequel.

1.1 Problem Definition

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a (point-to-point) continuous mapping and $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ a point-to-set mapping with closed and convex images. The quasi variational inequality QVI (K, F) is the problem of finding a point $x^* \in K(x^*)$ such that the following holds

$$F(x^*)^T(y - x^*) \geq 0, \qquad \forall y \in K(x^*). \tag{1}$$

For sake of simplicity, we always assume that all functions involved are defined over $\mathbb{R}^n$, even if this request could easily be weakened. A particularly well known and studied case occurs when $K(x)$ is actually independent of x, so that, for all x, $K(x) = K$ for some closed convex set K. In this case, the QVI becomes the Variational Inequality VI (K, F), that is the problem of finding $x^* \in K$ such that $F(x^*)^T(y - x^*) \geq 0$, $\forall y \in K$. For this latter problem, an extensive theory exists, see for example [23].

In most practical settings, the point-to-set mapping K is defined through a parametric set of equality and inequality constraints:

$$K(x) := \{y \in \mathbb{R}^n \mid M(x)y + v(x) = 0,\, g(y, x) \leq 0\}, \tag{2}$$

where $M : \mathbb{R}^n \to \mathbb{M}_{p\times n}$, $v : \mathbb{R}^n \to \mathbb{R}^p$ and $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$. We will use the following assumption

Assumption 1.1 *$g_i(\cdot, x)$ is convex and continuously differentiable on $\mathbb{R}^n$, for each $x \in \mathbb{R}^n$ and for each $i = 1, \dots, m$.*

The convexity of $g_i(\cdot, x)$ is obviously needed in order to guarantee that $K(x)$ be convex, while we require the differentiability assumption to be able to write down the KKT conditions of the QVI. We say that a point $x \in \mathbb{R}^n$ satisfies the KKT conditions if multipliers $\zeta \in \mathbb{R}^p$ and $\lambda \in \mathbb{R}^m$ exist such that

$$\begin{aligned} & F(x) + M(x)^T\zeta + \nabla_y g(x, x)\lambda \,=\, 0, \\ & M(x)x + v(x) = 0, \\ & 0 \,\leq\, \lambda \,\perp\, g(x, x) \,\leq\, 0. \end{aligned} \tag{3}$$

Note that $M(x)x + v(x) = 0$ and $g(x, x) \leq 0$ mean that $x \in K(x)$ and recall that $\nabla_y g(x, x)$ indicates the partial Jacobian of $g(y, x)$ with respect to y evaluated at $y = x$. These KKT conditions parallel the classical KKT conditions for a VI, see [23], and it is quite easy to show the following result.

Theorem 1.2 *Suppose Assumption 1.1 holds. If a point x, together with two suitable vectors $\zeta \in \mathbb{R}^p$ and $\lambda \in \mathbb{R}^m$ of multipliers, satisfies the KKT system* (3)*, then x is a solution of the QVI (K, F). Vice versa, if x is a solution of the QVI (K, F) and the constraints $g(\cdot, x)$ satisfy any standard constraint qualification, then multipliers $\zeta \in \mathbb{R}^p$ and $\lambda \in \mathbb{R}^m$ exist such that (x, ζ, λ) satisfies the KKT conditions* (3)*.*

Proof. It is well known that x is a solution of the QVI if and only if x is a global solution of the following optimization problem:

$$\min_y F(x)^T y, \quad \text{s.t. } M(x)y + v(x) = 0, \ g(y, x) \leq 0. \tag{4}$$

By the convexity of $g(\cdot, x)$, the optimization problem (4) is convex in y. Therefore, under the assumption that a constraint qualification holds, the KKT conditions (3) are necessary and sufficient optimality conditions for the problem (4) and then for the QVI. □

In the theorem above, by "any standard constraint qualification" we mean any classical optimization constraint qualification for $g(\cdot, x)$ at $y = x$ such as the linear independence of the active constraints, the Mangasarian-Fromovitz constraint qualification, Slater's one and so on.

The KKT conditions (3) are central to the approach described in this thesis as our potential reduction algorithm aims at finding KKT points of the QVI (K, F). In view of Theorem 1.2, the solution of these KKT conditions is essentially equivalent to the solution of the underlying QVI and, in any case whenever we can find a solution of the KKT conditions, we are sure that the corresponding x-part solves the QVI itself.

For simplicity, in Part II of this thesis, we suppose that the set K of the QVI is defined only through the inequalities g (we recall that it is always possible to rewrite $M(x)y + v(x) = 0$ as $M(x)y + v(x) \leq 0$ and $-M(x)y - v(x) \leq 0$). In this case the KKT conditions are the following

$$\begin{aligned} F(x) + \nabla_y g(x, x)\lambda &= 0, \\ 0 \leq \lambda \perp g(x, x) &\leq 0. \end{aligned} \tag{5}$$

1.2 Existence

Existence analysis on QVIs goes beyond the scope of this thesis. In this subsection we report only some simple existence (and uniqueness) results from the literature on QVIs that stems from algorithmic frameworks, because we recall them in the sequel of the thesis.

The following theorem proves the existence of at least one solution of a QVI whose feasible set satisfies some boundedness and nonemptiness conditions. It is an adaptation of Corollary 3.1 in [9] to QVIs.

Theorem 1.3 *Suppose that there exists a nonempty compact convex set C such that:*

(i) $\cup_{x \in C} K(x) \subseteq C$,

(ii) $K(x) \neq \emptyset$, $\forall\, x \in C$,

then there exists a solution of the QVI (K, F).

In [9] the previous result is proved for generalized quasi variational inequalities (GQVI for short) that is a more complex problem that we do not consider in this thesis. Moreover always in [9] authors showed a way to slightly weaken the previous assumption (i) with some technicalities.

Another existence (and uniqueness) result (improving the original one in [49]) is given in Corollary 2 in [45], we report it below.

Theorem 1.4 *Suppose that the following assumptions hold:*

(i) Operator F *is Lipschitz continuous and strongly monotone with constants* L *and* $\sigma > 0$ *respectively.*

(ii) There exists $\alpha < \frac{\sigma}{L}$ *such that*

$$\|\Pi_{K(x)}(z) - \Pi_{K(y)}(z)\| \leq \alpha \|x - y\|, \quad \forall x, y, z \in \mathbb{R}^n,$$

(we recall that Π_K *is the Euclidean projector on* K*).*

Then the QVI (K, F) *has a (unique) solution.*

Following this way other papers has been devoted to give (slightly different) existence (and uniqueness) results for QVIs, see for example [47].

In section 4 we give some new existence results and compare their assumptions with those of theorems in this subsection to show that they are in some sense weaker.

1.3 Solution Methods

In this subsection we briefly report and analyze global convergence properties for some solution methods for QVIs available in the literature. Methods described in this section consider QVIs in their general form, however we must mention that other very important algorithms with strong global convergence properties for QVIs with specific structures were proposed for example in [30, 52].

1.3.1 Projection Methods

One of the first algorithm with global convergence properties was proposed by Chan and Pang [9]. Their method consists in reformulating the QVI as a fixed point problem by using a projection operator.

Algorithm 1: Projection Algorithm for QVIs

`(S.0)` : Choose $x^0 \in \mathbb{R}^n$, $\rho > 0$, $\epsilon > 0$ and set $k = 0$.

`(S.1)` : If $\|x^k - \Pi_{K(x^k)}(x^k - \rho F(x^k))\| \leq \epsilon$: STOP.

`(S.2)` : Set

$$x^{k+1} := \Pi_{K(x^k)}(x^k - \rho F(x^k)),$$

$k \leftarrow k + 1$, and go to `(S.1)`.

Convergence result proposed by Chan and Pang is based on Brouwer fixed point theorem by giving conditions on the QVI to obtain the projection operator to be a contraction. However they only prove convergence for "moving set" problems, that is QVIs in which the feasible mapping $K(\cdot)$ is defined by a nonempty closed convex set $Q \subseteq \mathbb{R}^n$ and a "trajectory" described by $c : \mathbb{R}^n \to \mathbb{R}^n$ according to:

$$K(x) = c(x) + Q. \tag{6}$$

Theorem 1.5 *Consider the QVI (F, K) in which the set K is defined by* (6)*. Suppose that both F and c are Lipschitz continuous and strongly monotone with moduli L, σ, α and τ respectively. Suppose that*

$$\rho^2 L^2 + 2\rho(\alpha L - \sigma) - 2(\tau - \alpha) < 0. \tag{7}$$

Then the operator $\Pi_{K(\cdot)}((\cdot) - \rho F(\cdot))$ is a contraction and thus has a fixed point $\bar{x}_\rho$. Moreover, Algorithm 1 globally converges to $\bar{x}_\rho$ which is the (unique) solution of the QVI.

Where the proof of Theorem 1.5 is based on the fact that:

$$\Pi_{K(x)}(y) = c(x) + \Pi_Q(y - c(x)).$$

It is not difficult to see that condition (7) and $\rho > 0$ together means that

$$-(\alpha L - \sigma) + \left|\sqrt{(\alpha L - \sigma)^2 + 2L^2(\tau - \alpha)}\right| > 0, \tag{8}$$

which is the condition for the solvability of the QVI by using Algorithm 1.

1.3.2 Sequential VI Methods

To the best of our knowledge, the method that improves more than any other Chan and Pang's convergence result was proposed by Nesterov and Scrimali in [45]. This method consists in solving a sequence of VIs generated by sequentially parameterizing at each iteration the feasible set of the QVI by using the solution of the VI at the previous iterate:

Algorithm 2: Sequential VI method for QVIs

`(S.0)` : Choose $x^0 \in \mathbb{R}^n$ such that $K(x^0) \neq \emptyset$ and set $k = 0$.

`(S.1)` : If x^k is a solution of QVI (F, K): STOP.

`(S.2)` : Set x^{k+1} equal to the solution of VI $(F, K(x^k))$,
$k \leftarrow k + 1$, and go to `(S.1)`.

Clearly Algorithm 2 is well defined only if each VI generated during its iterations has a solution. In [45] authors proved global convergence for moving set QVIs:

Theorem 1.6 *Consider the QVI (F, K) in which the set K is defined by* (6)*. Suppose that F is Lipschitz continuous and strongly monotone with moduli L and σ respectively, and that c is Lipschitz continuous with modulus α. Suppose that*

$$\alpha < \frac{\sigma}{L}, \tag{9}$$

then Algorithm 2 globally converges to the (unique) solution of the QVI.

Note that, under the assumptions of Theorem 1.6, Algorithm 2 is well defined, in fact by (9) each VI generated by the algorithm has a (unique) solution.

It is not difficult to see that condition (8) implies condition (9), in fact it sufficies to note that:

$$\begin{aligned} 0 &< -(\alpha L - \sigma) + \left|\sqrt{(\alpha L - \sigma)^2 + 2L^2(\tau - \alpha)}\right| \\ &\leq -(\alpha L - \sigma) + \left|\sqrt{(\alpha L - \sigma)^2}\right| \\ &= -(\alpha L - \sigma) + \left|\alpha L - \sigma\right|, \end{aligned}$$

that is $(\alpha L - \sigma) < 0$. Then we can conclude that global convergence conditions for Algorithm 2 are weaker than those for Algorithm 1 in the moving set case.

Another sequential VI approach to general QVIs is described in [58]. This method reduces the solution of a QVI to the solution of a sequence of VIs by using a penalization strategy; however, even if this approach is very interesting and promising, its global convergence properties are in jeopardy since they ultimately hinge of the capability of solving a sequence of possibly very challenging VIs.

1.3.3 Gap Function Methods

An attempt to reformulate a QVI by using a gap function is imputable to Fukushima in [27]. In his paper, the gap function is defined as follows:

$$f(x) := -\inf\left\{\varphi(x, y) \mid y \in \Gamma(x)\right\},$$

where

$$\varphi(x, y) := F(x)^T(y - x) + \frac{1}{2}(y - x)^T G(y - x),$$

with a positive definite symmetric matrix G, and

$$\Gamma(x) := \left\{y \in \mathbb{R}^n \mid g_i(x, x) + \nabla_y g_i(x, x)^T(y - x) \leq 0,\ i = 1, \ldots, m\right\}.$$

To solve the QVI, the method proposed by Fukushima consists in finding a global solution (with zero value) of the following optimization problem:

$$\min\ f(x),\ \text{s.t.}\ g(x, x) \leq 0. \tag{10}$$

Despite its practical utility, from the theoretical point of view this method does not improve global convergence results for QVIs. In fact Fukushima only proves that f is continuous and directionally differentiable, then finding a global solution of problem (10) is not easy in general.

1.3.4 KKT Methods

KKT methods try to solve the QVI (1) by solving its KKT system (5). Different approches are possible, in fact the KKT system may be reformulated as:

- a semismooth system of equations by using complementarity functions like the Fischer-Burmeister one (see [13] for a description of this method for generalized Nash equilibrium problems): it has been showed in [13] that these methods require stronger assumptions for global convergence than those required by the potential reduction algorithm proposed in this thesis;
- a differentiable constrained system of equations (the approach proposed in this thesis, see Part II for details);
- a nonmonotone VI (the basis for the solution of a QVI by using the PATH solver [12], see Part III for details): despite its poor global convergence properties, from the practical point of view this method has been widely used to numerically solve QVIs; however in Part III we will show the effectiveness of our potential reduction algorithm and its robustness even if compared to the PATH solver.

Motivated by considerations developed in [13] for Nash games, in this thesis we will show that the potential reduction algorithm described in Section 2 is (so far) the best method for the solution of QVIs from both the theoretical (Part II) and the numerical (Part III) point of view.

Part II
Theoretical Results

2 Potential Reduction Algorithm

As we already mentioned, we propose to solve the KKT conditions (5) by an interior-point method designed to solve constrained systems of equations. In order to reformulate system (5) as a constrained system of equations (CE for short), we introduce slack variables $w \in \mathbb{R}^m$ and consider the CE system

$$H(z) = 0, \quad z = (x, \lambda, w) \in Z \tag{11}$$

with

$$H(x, \lambda, w) := \begin{pmatrix} L(x, \lambda) \\ h(x) + w \\ \lambda \circ w \end{pmatrix}$$

and where

$$L(x, \lambda) := F(x) + \nabla_y g(x, x)\lambda, \qquad h(x) := g(x, x) \tag{12}$$

and

$$Z := \{z = (x, \lambda, w) \mid x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m_+, w \in \mathbb{R}^m_+\}.$$

It is clear that the couple (x, λ) solves system (5) if and only if (x, λ), together with a suitable w, solves the CE (11). From now on, we will aim at solving the CE (11) by the interior-point method described next.

Let $r : \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++} \to \mathbb{R}$ be the function

$$r(u, v) := \zeta \log(\|u\|^2 + \|v\|^2) - \sum_{i=1}^{2m} \log(v_i), \quad (u, v) \in \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}, \quad \zeta > m,$$

and let

$$\psi(z) := r(H(z))$$

be the potential function of the CE, which is defined for all

$$z \in Z_I := H^{-1}(\mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}) \cap \operatorname{int} Z,$$

(we reacall that $\operatorname{int} Z$ denotes the interior of the set Z). In order to be able to define our potential reduction interior-point method we need some further differentiability conditions.

Assumption 2.1 *$F(x)$, $h(x)$ and $\nabla_y g(x, x)$ are continuously differentiable on $\mathbb{R}^n$.*

The following algorithm is precisely the interior-point method from [43]; see also [23] for further discussion and [13] for an inexact version.

Algorithm 3: Potential Reduction Algorithm (PRA) for CEs

`(S.0)` : Choose $z^0 \in Z_I, \beta, \gamma \in (0,1)$, and set $k := 0, a^T = (0_n^T,\, 1_{2m}^T)$.

`(S.1)` : If $H(z^k) = 0$: STOP.

`(S.2)` : Choose $\rho_k \in [0,1)$ and find a solution d^k of the linear system

$$JH(z^k)d^k = -H(z^k) + \rho_k \frac{a^T H(z^k)}{\|a\|^2} a. \tag{13}$$

`(S.3)` : Compute a stepsize $t_k := \max\left\{\beta^\ell \mid \ell = 0, 1, 2, \dots\right\}$ such that

$$z^k + t_k d^k \in Z_I$$

and

$$\psi(z^k + t_k d^k) \le \psi(z^k) + \gamma t_k \nabla\psi(z^k)^T d^k. \tag{14}$$

`(S.4)` : Set $z^{k+1} := z^k + t_k d^k, k \leftarrow k+1$, and go to `(S.1)`.

Algorithm 3 will generate a sequence $\{z^k\} := \{(x^k, \lambda^k, w^k)\}$ with $\lambda^k > 0$ and $w^k > 0$ for every k. The core of this approach is the calculation of a Newton-type direction for the system $H(z) = 0$. According to standard procedures in interior point methods, the Newton direction is "bent" in order to follow the central path. Operatively this means that the search direction used in this method is the solution of the system (13). Once this direction has been calculated, a line-search is performed by using the potential function ψ.

Algorithm 3 is well-defined as long as the Jacobians $JH(z^k)$ in (13) are nonsingular. Actually, the following theorem, which can be found in [43] and [23], shows that this condition also guarantees that every limit point of the sequence generated by the algorithm is a solution.

Theorem 2.2 *Suppose that Assumptions 1.1 and 2.1 hold. Assume that $JH(z)$ is nonsingular for all $z \in Z_I$, and that the sequence $\{\rho_k\}$ from (S.2) of Algorithm 3 satisfies the condition $\limsup_{k\to\infty} \rho_k < 1$. Let $\{z^k\}$ be any sequence generated by Algorithm 3. Then the following statements hold:*

(a) the sequence $\{H(z^k)\}$ is bounded;

(b) any accumulation point of $\{z^k\}$ is a solution of CE (11).

In view of Theorem 2.2, the main question we must answer in order to make our approach viable is: for which classes of QVIs can we guarantee that the Jacobian matrices $JH(z)$ are nonsingular for all $z \in Z_I$? A related, albeit practically less crucial, question is whether we can guarantee that the sequence $\{z^k\}$ generated by Algorithm 3 is bounded. This obviously would guarantee that the algorithm actually has at least a limit point and therefore that a solution is certainly found. The first question will be answered in detail in Section 3, whereas the second question will be dealt with in Section 4.

3 Solvability: Nonsingularity Conditions

As noted before, the main topic in order to guarantee global convergence of Algorithm 3 to a solution of CE (11) (and then the QVI) is the nonsingularity of $JH(z)$. The structure of this Jacobian is given by

$$JH(x, \lambda, w) = \begin{pmatrix} J_x L(x, \lambda) & \nabla_y g(x, x) & 0 \\ J_x h(x) & 0 & I \\ 0 & \mathrm{diag}(w) & \mathrm{diag}(\lambda) \end{pmatrix}.$$

This section is devoted entirely to the study of classes of QVIs for which the nonsingularity of JH can be established. It is not too difficult to give conditions that guarantee the nonsingularity of JH, what is less obvious is how we can establish *sensible* and *significant* conditions for interesting classes of QVIs. This we achieve in two stages: in the next subsection we give several sufficient or necessary and sufficient conditions for the nonsingularity of JH which are then used in the following subsections to analyze various classes of QVIs. In particular, we will discuss and establish nonsingularity results for the following classes of QVIs:

- Problems where $K(x) = c(x) + Q$ (the so called "moving set" case, already mentioned in the introduction and described in Section 1);

- Problems where $K(x)$ is defined by a linear system of inequalities with a variable right-hand side;

- Problems where $K(x)$ is defined by box constraints with parametric upper and lower bounds;

- Problems where $K(x)$ is defined by "binary constraints", i.e. parametric inequalities $g(x, y) \leq 0$ with each g_i actually depending only on two variables: x_j and y_j;

- Problems where $K(x)$ is defined by bilinear constraints.

While we refer the reader to the following subsections for a more accurate description of the problem classes, we underline that, as far as we are aware of and with the exception of the moving set case, these problem classes are all new and we can establish here for the first time convergence results, according to Theorem 2.2.

3.1 General Nonsingularity Conditions

The results in this subsection do not make explicit reference to a specific structure of the QVI and, in particular, of the feasible set mapping K. However, they are instrumental in proving the more specific results in the following subsections. The first result we present is a necessary and sufficient condition for the nonsingularity of JH.

Theorem 3.1 *Suppose that Assumptions 1.1 and 2.1 hold. Let* $(x, \lambda, w) \in \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}$ *be given. Then the matrix*

$$N(x, \lambda, w) := J_x L(x, \lambda) + \nabla_y g(x, x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x h(x) \tag{15}$$

is nonsingular if and only if $JH(x, \lambda, w)$ *is nonsingular.*

Proof. We first prove the only-if-part. Let $q = \left(q^{(1)}, q^{(2)}, q^{(3)}\right)$ be a suitably partitioned vector such that $JH(x, \lambda, w)q = 0$. This equation can be rewritten in partitioned form as

$$J_x L(x, \lambda) q^{(1)} + \nabla_y g(x, x) q^{(2)} = 0, \tag{16}$$

$$J_x h(x) q^{(1)} + q^{(3)} = 0, \tag{17}$$

$$\operatorname{diag}(w) q^{(2)} + \operatorname{diag}(\lambda) q^{(3)} = 0. \tag{18}$$

Solving (18) for $q^{(3)}$ gives

$$q^{(3)} = -\operatorname{diag}\left(\lambda^{-1} \circ w\right) q^{(2)}. \tag{19}$$

Inserting this expression into (17) yields

$$J_x h(x) q^{(1)} - \operatorname{diag}\left(\lambda^{-1} \circ w\right) q^{(2)} = 0$$

which, in turn, gives

$$q^{(2)} = \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x h(x) q^{(1)}. \tag{20}$$

Substituting this expression into (16) finally yields

$$\left[J_x L(x, \lambda) + \nabla_y g(x, x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x h(x)\right] q^{(1)} = 0.$$

However, the matrix in brackets is precisely the matrix $N(x, \lambda, w)$ from (15) and, therefore, nonsingular. Hence, it follows that $q^{(1)} = 0$ which then also implies $q^{(2)} = 0$ and $q^{(3)} = 0$.

Now, to prove the if-part, we show that if $N(x, \lambda, w)$ is singular, then $JH(x, \lambda, w)$ is singular, too. If $N(x, \lambda, w)$ is singular, there exists a nonzero vector $q^{(1)}$ such that

$$\left[J_x L(x, \lambda) + \nabla_y g(x, x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x h(x)\right] q^{(1)} = 0.$$

Now, let $q^{(2)}$ and $q^{(3)}$ be vectors defined by (20) and (19), respectively. Then (16)–(18) hold, and hence $JH(x, \lambda, w)q = 0$ for $q = \left(q^{(1)}, q^{(2)}, q^{(3)}\right) \neq 0$. This shows that $JH(x, \lambda, w)$ is singular and, therefore, completes the proof. □

We next state a simple consequence of Theorem 3.1.

Corollary 3.2 *Suppose that Assumptions 1.1 and 2.1 hold and let* $(x, \lambda, w) \in \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}$ *be given. Suppose that* $J_x L(x, \lambda, w)$ *is positive definite and one of the following conditions holds:*

(a) $\nabla_y g(x, x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x h(x)$ *is positive semidefinite, or*

(b) $\nabla_y g(x,x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x g(x,x)$ *is positive semidefinite.*

Then $JH(x,\lambda,w)$ *is nonsingular.*

Proof. In view of Theorem 3.1, it suffices to show that the matrix $N(x,\lambda,w)$ from (15) is nonsingular. Since $J_x L(x,\lambda)$ is positive definite by assumption, the statement is trivially satisfied under condition (a). Hence, suppose that (b) holds. Since $h(x) = g(x,x)$, we have $J_x h(x) = J_y g(x,x) + J_x g(x,x)$. This implies

$$\begin{aligned} N(x,\lambda,w) &= J_x L(x,\lambda) + \nabla_y g(x,x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x h(x) \\ &= J_x L(x,\lambda) + \nabla_y g(x,x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_y g(x,x) \\ &\quad + \nabla_y g(x,x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x g(x,x). \end{aligned}$$

Now, the first term $J_x L(x,\lambda)$ in the last expression is positive definite by assumption, the second term is obviously positive semidefinite since $\lambda, w > 0$, and the third term is positive semidefinite by condition (b). Consequently, $N(x,\lambda,w)$ is positive definite, hence nonsingular. □

Note that the previous proof actually shows that condition (b) from Corollary 3.2 implies condition (a) which, therefore, is a weaker assumption in general, whereas condition (b) might be easier to verify in some situations.

We now state another consequence of Theorem 3.1.

Corollary 3.3 *Suppose that Assumptions 1.1 and 2.1 hold and let* $(x,\lambda,w) \in \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}$ *be given. Suppose that* $J_x L(x,\lambda)$ *is nonsingular and*

$$\bar{M}(x,\lambda) := J_x h(x) J_x L(x,\lambda)^{-1} \nabla_y g(x,x)$$

is a P_0*-matrix. Then* $JH(x,\lambda,w)$ *is nonsingular.*

Proof. For notational simplicity, let us write

$$A(x,\lambda,w) := J_x L(x,\lambda)^{-1} \nabla_y g(x,x) \operatorname{diag}\left(w^{-1} \circ \lambda\right) J_x h(x).$$

We note that $\operatorname{diag}\left(w^{-1} \circ \lambda\right)$ is a positive definite diagonal matrix and can therefore be written as a product DD, where D is another positive definite diagonal matrix.

We have that the matrix $N(x,\lambda,w)$ is nonsingular if and only if $I + A(x,\lambda,w)$ is nonsingular. In turn, recalling that μ is an eigenvalue of $A(x,\lambda,w)$ if and only if $1+\mu$ is an eigenvalue of $I + A(x,\lambda,w)$, we see that $N(x,\lambda,w)$ is surely nonsingular if $A(x,\lambda,w)$ has all real eigenvalues nonnegative. But it is well known that, given two square matrices A, B, the matrix product AB has the same eigenvalues as the matrix product BA, see [35, Theorem 1.3.20], hence it follows that $A(x,\lambda,w)$ has the same eigenvalues as $DJ_x h(x) J_x L(x,\lambda)^{-1} \nabla_y g(x,x) D$ which is exactly the matrix $D\bar{M}(x,\lambda)D$. By assumption, we have that $\bar{M}(x,\lambda)$ is a P_0 matrix, hence $D\bar{M}(x,\lambda)D$ is also a P_0 matrix since D is diagonal and positive definite, and then it has all real eigenvalues nonnegative, see [10, Theorem 3.4.2]. This completes the proof. □

The remaining part of this section specializes the previous results to deal with specific constraint structures.

3.2 The Moving Set Case

As we mentioned in the introduction, this is the most studied class of problems in the literature and (variants and generalizations apart) essentially the only class of problems for which clear convergence conditions are available (see Section 1). In this class of problems, the feasible mapping $K(\cdot)$ is defined by a closed convex set $Q \subseteq \mathbb{R}^n$ and a "trajectory" described by $c : \mathbb{R}^n \to \mathbb{R}^n$ according to:

$$K(x) = c(x) + Q.$$

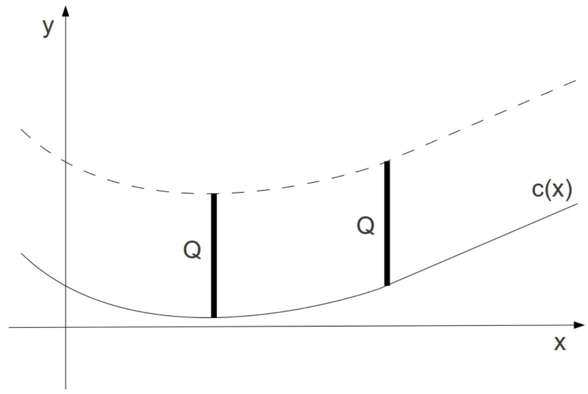

Figure 1: Example of moving set: the set K is between the dashed and the continuous curve lines.

In order to proceed in our analysis, we suppose that Q is defined by a set of convex inequalities:

$$Q \;=\; \{x \in \mathbb{R}^n \,|\, q(x) \leq 0\},$$

where $q : \mathbb{R}^n \to \mathbb{R}^m$ and each q_i is convex on $\mathbb{R}^n$. It is easy to see that, in this setting, we have

$$K(x) = \{y \in \mathbb{R}^n \mid q(y - c(x)) \leq 0\}. \tag{21}$$

By exploiting this structure, we can prove the following theorem.

Theorem 3.4 *Let $K(x)$ be defined as in (21), with q_i convex for every $i = 1, \dots, m$. Let a point $x \in \mathbb{R}^n$ be given and assume that around x it holds that F and c are C^1 and q is C^2. Suppose further that $JF(x)$ is nonsingular and that*

$$\|Jc(x)\| \leq \frac{\mu_m^s(JF(x)^{-1})}{\|JF(x)^{-1}\|}. \tag{22}$$

Then $JH(x, \lambda, w)$ is nonsingular for all positive λ and w.

Proof. We are going to show that the conditions from Theorem 3.1 are satisfied. First of all note that the hypotheses imply Assumptions 1.1 and 2.1. Taking into account (21), we have, using the notation in (2) and (12),

$$g(y,x) = q(y - c(x)), \quad h(x) = q(x - c(x))$$

and, hence,

$$\nabla_y g(x,x) \;=\; \nabla q(x - c(x)), \qquad J_x h(x) \;=\; Jq(x - c(x))(I - Jc(x)).$$

Therefore we can write

$$N(x,\lambda,w) \;=\; JF(x) + \bar{S}(I - Jc(x)),$$

where

$$\bar{S} \;=\; \sum_{i=1}^{m} \lambda_i \nabla^2 q_i(x - c(x)) + \nabla q(x - c(x))\,\mathrm{diag}(w^{-1} \circ \lambda) Jq(x - c(x)).$$

Note that, for any positive λ and w, $\bar{S}$ is positive semidefinite and symmetric. Therefore, we can write $\bar{S} = SS^T$ for some suitable square matrix S. Recalling that $JF(x)$ is nonsingular by assumption, we have that the matrix $N(x,\lambda,w)$ is nonsingular if and only if

$$I + JF(x)^{-1}SS^T(I - Jc(x))$$

is nonsingular. In turn, since μ is an eigenvalue of $JF(x)^{-1}SS^T(I - Jc(x))$ if and only if $1+\mu$ is an eigenvalue of $I + JF(x)^{-1}SS^T(I - Jc(x))$, we see that $N(x,\lambda,w)$ is surely nonsingular if $JF(x)^{-1}SS^T(I - Jc(x))$ has all real eigenvalues nonnegative. But, similar to the proof of Corollary 3.3, it follows that $JF(x)^{-1}SS^T(I - Jc(x))$ has the same eigenvalues as $S^T(I - Jc(x))JF(x)^{-1}S$. If we can show that $(I - Jc(x))JF(x)^{-1}$ is positive semidefinite, we obviously also have that $S^T(I - Jc(x))JF(x)^{-1}S$ is positive semidefinite and, therefore, has all the real eigenvalues (if any) nonnegative. Hence, to complete the proof, we only need to show that (22) implies that $(I - Jc(x))JF(x)^{-1}$ is positive semidefinite. In order to see this, it is sufficient to observe that for any $v \in \mathbb{R}^n$ we can write

$$v^T Jc(x)JF(x)^{-1}v \;\le\; \|Jc(x)\|\|JF(x)^{-1}\|\|v\|^2 \;\le\; \mu_m^s(JF(x)^{-1})\|v\|^2 \;\le\; v^T JF(x)^{-1}v,$$

where the second inequality follows from (22). From this chain of inequalities the positive semidefiniteness of $(I - Jc(x))JF(x)^{-1}$ follows readily and this concludes the proof. $\square$

Note that in assumptions of Theorem 3.4 the request to have $JF(x)$ nonsingular and the (22) implicitly imply that $JF(x)$ must be positive definite. Moreover note that (22) is a condition purely in terms of F and c, neither q nor the values of λ and w are involved. The fact that q is not involved simply indicates that the nonsingularity of N is not related to the "shape" of the set Q, but only to the trajectory the moving set follows. More precisely, as will also become more clear in the following corollary, (22) requires the trajectory described by c to be not "too steep", where the exact meaning of "too steep" is given by (22). The following corollary shades some further light on this condition.

Remark 3.5 In part (d) of the following Corollary, and in the rest of this section we freely use some notation and definitions for Lipschitz and monotonicity constants that are fully explained and discussed at length in the Appendix A.

Corollary 3.6 *Assume the setting of Theorem 3.4 and consider the following conditions:*

(a) The matrix $N(x, \lambda, w)$ *is nonsingular on* $\mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}$;

(b) Condition (22) holds for all $x \in \mathbb{R}^n$;

(c) It holds that

$$\sup_{x \in \mathbb{R}^n} \|Jc(x)\| \leq \inf_{x \in \mathbb{R}^n} \frac{\mu^s_m(JF(x)^{-1})}{\|JF(x)^{-1}\|};$$

(d) c *is Lipschitz continuous on* $\mathbb{R}^n$ *with Lipschitz modulus* α*,* F *is Lipschitz continuous on* $\mathbb{R}^n$ *and strongly monotone on* $\mathbb{R}^n$*, the moduli of Lipschitz continuity and strong monotonicity of* F^{-1} *are* L_{-1} *and* σ_{-1}*, respectively, and*

$$\alpha \leq \frac{\sigma_{-1}}{L_{-1}}. \tag{23}$$

Then it holds that

$$(d) \implies (c) \implies (b) \implies (a).$$

Proof. The only implication that needs a proof is $(d) \implies (c)$. By Proposition A.2 (a) in the Appendix, we have $\alpha = \sup_{x \in \mathbb{R}^n} \|Jc(x)\|$. We now recall that since F is strongly monotone on $\mathbb{R}^n$, its range is $\mathbb{R}^n$, see [50, Theorem 5.4.5]. Therefore, by Proposition A.2 in the Appendix and taking into account that $JF^{-1}(F(x)) = JF(x)^{-1}$, we can write

$$\frac{\sigma_{-1}}{L_{-1}} = \frac{\inf_{y \in \mathbb{R}^n} \mu^s_m(JF^{-1}(y))}{\sup_{y \in \mathbb{R}^n} \|JF^{-1}(y)\|} = \frac{\inf_{x \in \mathbb{R}^n} \mu^s_m(JF(x)^{-1})}{\sup_{x \in \mathbb{R}^n} \|JF(x)^{-1}\|} \leq \inf_{x \in \mathbb{R}^n} \frac{\mu^s_m(JF(x)^{-1})}{\|JF(x)^{-1}\|}.$$

This completes the proof. □

Although the sufficient condition (23) is the strongest one among those we analyzed, it gives a clear geometric picture of the kind of conditions we need in order to guarantee nonsingularity. Note that Lipschitz continuity and strong monotonicity of F imply that also the inverse of F enjoys the same properties, see Proposition A.4 in the Appendix A, so that L_{-1} and σ_{-1} are well defined. Furthermore, observe that $(\sigma_{-1}/L_{-1}) \leq 1$ (this is obvious from the very definition of these constants, see Appendix A). Therefore (23) stipulates that $c(x)$ is rather "flat" and consequently, $K(x)$ varies "slowly", in some sense.

Remark 3.7 Reference [45] is one of the most interesting papers where the moving set structure has been used in order to show convergence of some algorithms for QVIs. It is shown in [45] that if $\alpha \leq \frac{\sigma}{L}$, where α and L are the Lipschitz moduli of c and F, respectively, and σ is the strong monotonicity modulus of F, then a certain gradient projection type method converges to the unique solution of the QVI (see Section 1). It is then of interest to contrast this condition to our condition $\alpha \leq \frac{\sigma_{-1}}{L_{-1}}$ in Corollary 3.6 (d) (which is the strongest among the conditions we considered). If the function F is a gradient mapping, then Proposition A.5 in the Appendix implies that $\sigma/L = \sigma_{-1}/L_{-1}$, so that our condition and that in [45] are exactly the same. However, in general $\sigma_{-1}/L_{-1} < \sigma/L$ and $\sigma_{-1}/L_{-1} > \sigma/L$ can both occur. In fact, consider the function

$$F(x) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} x.$$

It is easy to see that $\sigma(\mathbb{R}^n, F) = 1 - \frac{1}{\sqrt{2}}$ and $L(\mathbb{R}^n, F) \simeq 1.8019$. Moreover, we have

$$F^{-1}(x) = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} x.$$

Again, it is easy to see that $\sigma(\mathbb{R}^n, F^{-1}) = \frac{1}{2}$ and $L(\mathbb{R}^n, F^{-1}) \simeq 2.2470$. Therefore, we have

$$\frac{\sigma(\mathbb{R}^n, F)}{L(\mathbb{R}^n, F)} \simeq 0.1625 < 0.2225 \simeq \frac{\sigma(\mathbb{R}^n, F^{-1})}{L(\mathbb{R}^n, F^{-1})},$$

and then for this function our condition is less restrictive than that in [45]. But it is sufficient to switch the function with its inverse to get exactly the opposite. Therefore there is no one condition that dominates the other one in general. □

The following example shows how condition (c) in Corollary 3.6 simplifies in certain situations and the way it can be used (i) to show how interesting classes of problems can be analyzed and (ii) to easily check whether this condition is actually satisfied in a concrete situation.

Example 3.8 The discretization of many (elliptic) infinite-dimensional QVIs involving suitable partial differential operators often leads to linear mappings of the form $F(x) = Ax + b$ for some positive definite matrix A, see e.g. [29, 30]. Furthermore, in many application in mechanics an implicit-obstacle-type constraint described by the set $K(x) := \{y \mid y \leq c(x)\}$ for some smooth mapping c is present, see [38]. In these cases $K(x)$ belongs to the class of moving sets with q being the identity mapping in (21). Taking into account that $JF(x) = A$, we can easily calculate $\frac{\mu_m^s(JF(x)^{-1})}{\|JF(x)^{-1}\|}$ which is obviously a positive constant. It actually turns out that there are intereseting applications where A is symmetric. Furthermore the minimum and maximum eigenvalues of A, here denoted by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ respectively, are even known analytically in some cases, e.g., if A corresponds to a standard finite difference-discretization of the two-dimensional Laplace operator on the unit square $(0,1) \times (0,1)$. In this setting we can write

$$\frac{\mu_m^s(JF(x)^{-1})}{\|JF(x)^{-1}\|} = \frac{\lambda_{\min}(A^{-1})}{\|A^{-1}\|} = \frac{1/\lambda_{\max}(A)}{1/\lambda_{\min}(A)} = \frac{\lambda_{\min}(A)}{\lambda_{\max}(A)} > 0.$$

Hence condition (c) in Corollary 3.6 holds provided that $\|Jc(x)\|$ is less or equal to this positive constant, i.e. provided that c is Lipschitz continuous with a sufficiently small Lipschitz constant. □

3.3 Linear Constraints with Variable Right-hand Side

We now pass to consider the case in which the feasible set $K(x)$ is given by

$$K(x) \ = \ \{y \in \mathbb{R}^n \mid g(y,x) := Ey - b - c(x) \leq 0\}, \tag{24}$$

where $E \in \mathbb{R}^{m\times n}$ is a given matrix, $c : \mathbb{R}^n \to \mathbb{R}^m$ and $b \in \mathbb{R}^m$. In this class of QVIs, the feasible set is defined by linear inequalities in which the right-hand side depends on x.

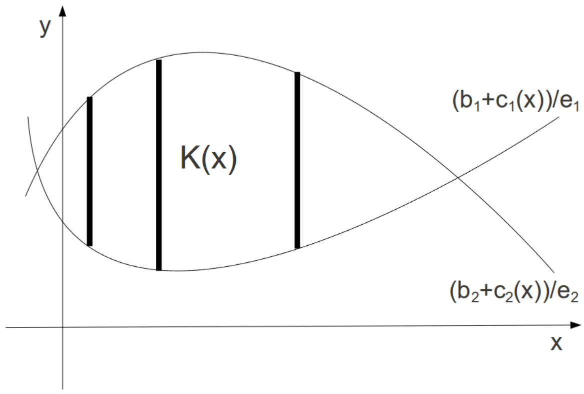

Figure 2: Example of linear constraints with variable right-hand side: here $E = (e_1 \ e_2)^T$ and $e_1 < 0$, $e_2 > 0$.

Theorem 3.9 *Let g be defined as in (24), let $x \in \mathbb{R}^n$ be a given point, and assume that F and c are C^1 around x. Suppose further that $JF(x)$ is positive definite and that*

$$\|Jc(x)\| \ \leq \ \frac{\mu_m^+(x)}{\|JF(x)^{-1}\|\|E\|}, \tag{25}$$

where

$$\mu_m^+(x) = \min\{\mu_m^+(A) \mid A \text{ is a principal submatrix of } \frac{1}{2}E(JF(x)^{-1} + JF(x)^{-T})E^T\},$$

$\mu_m^+(A)$ denotes the minimum positive eigenvalue of the matrix A, and A^{-T} is the transpose of the inverse of A. Then $JH(x,\lambda,w)$ is nonsingular for all positive λ and w.

Proof. We will show that the assumptions from Corollary 3.3 hold. First of all note that the hypotheses imply Assumptions 1.1 and 2.1. Taking into account (24), we have

$\nabla_y g(x,x) = E^T$ and $J_x h(x) = E - Jc(x)$. Since $JF(x)$ is nonsingular by assumption, we can write

$$\bar{M} := \bar{M}(x,\lambda) = (E - Jc(x))JF(x)^{-1}E^T.$$

In view of Corollary 3.3, we need to show that $\bar{M}$ is a P_0 matrix.

To this end, we first observe that the rank of $\bar{M}$ is obviously less or equal to n (the rank of $JF(x)$). Hence each principal minor of $\bar{M}$ with dimension greater than n is equal to zero. Therefore, it suffices to show that each principal minor of $\bar{M}$ with size less or equal to n is nonnegative.

A generic principal submatrix of $\bar{M}$ with dimension $s \leq n$ is defined by

$$(E_{i*} - Jc(x)_{i*})_{i\in I_s} JF(x)^{-1}(E_{i*})^T_{i\in I_s}$$

where I_s is a subset of $\{1,\dots,m\}$ with cardinality equal to s. Therefore, each of these subsets of indices defines a principal submatrix of $\bar{M}$. Now we have two cases: $E_{I_s} := (E_{i*})_{i\in I_s}$ has full row rank or not. If not, the principal minor corresponding to I_s is equal to zero. Otherwise, if E_{I_s} has full row rank, then we can prove that the principal submatrix corresponding to I_s is positive semidefinite. In fact, we can write

$$\begin{aligned} v^T E_{I_s} JF(x)^{-1} E^T_{I_s} v &\geq \mu_m^+(x)\|v\|^2 \\ &\overset{(25)}{\geq} \|Jc(x)\|\,\|JF(x)^{-1}\|\,\|E\|\,\|v\|^2 \\ &\geq \|Jc(x)_{I_s}\|\,\|JF(x)^{-1}\|\,\|E_{I_s}\|\,\|v\|^2 \\ &\geq v^T Jc(x)_{I_s} JF(x)^{-1} E_{I_s} v, \quad \forall v \in \mathbb{R}^n, \end{aligned}$$

where the third inequality follows from the fact that the spectral norm of a submatrix is less or equal to the spectral norm of the matrix itself. Then we have

$$v^T(E_{I_s} - Jc(x)_{I_s})JF(x)^{-1}E^T_{I_s} v \geq 0, \quad \forall v \in \mathbb{R}^n.$$

Hence $\bar{M}$ is a P_0 matrix, and using Corollary 3.3, we have the thesis. □

By the inclusion principle (see, for example, [35, Theorem 4.3.15]) and recalling condition (25), it is clear that if the matrix E has full row rank, then we have

$$\mu_m^+(x) = \mu_m^s(EJF(x)^{-1}E^T).$$

This allows us to state the following immediate corollary.

Corollary 3.10 *Let g be defined as in (24), let $x \in \mathbb{R}^n$ be a given point, and assume that F and c are C^1 around x. Moreover, suppose that E has full row rank. Suppose that $JF(x)$ is positive definite and that*

$$\|Jc(x)\| \leq \frac{\mu_m^s(EJF(x)^{-1}E^T)}{\|JF(x)^{-1}\|\|E\|}.$$

Then $JH(x,\lambda,w)$ is nonsingular for all positive λ and w.

Technicalities apart, the meaning of Theorem 3.9 is that $c(x)$ should not vary "too quickly".

The following result parallels Corollary 3.6 and gives stronger, but more expressive conditions for the nonsingularity of JH.

Corollary 3.11 *Assume the same setting as in Theorem 3.9 and consider the following conditions:*

(a) The matrix $N(x, \lambda, w)$ is nonsingular on $\mathbb{R}^n \times \mathbb{R}^m_{++}$;

(b) For all $x \in \mathbb{R}^n$, $JF(x)$ is positive definite and condition (25) holds;

(c) For all x, $JF(x)$ is positive definite and it holds that

$$\|Jc(x)\| \leq \frac{\mu_m^s(JF(x)^{-1})}{\|JF(x)^{-1}\|} \frac{\mu_m^+}{\|E\|},$$

where $\mu_m^+ = \min\{\mu_m^+(A) \mid A$ is a principal submatrix of $EE^T\}$;

(d) The Jacobian $JF(x)$ is positive definite for all $x \in \mathbb{R}^n$, and it holds that

$$\sup_{x \in \mathbb{R}^n} \|Jc(x)\| \leq \inf_{x \in \mathbb{R}^n} \frac{\mu_m^s(JF(x)^{-1})}{\|JF(x)^{-1}\|} \frac{\mu_m^+}{\|E\|};$$

(e) c is Lipschitz continuous on $\mathbb{R}^n$ with Lipschitz modulus α, F is Lipschitz continuous on $\mathbb{R}^n$ and strongly monotone on $\mathbb{R}^n$, the moduli of Lipschitz continuity and strong monotonicity of F^{-1} are L_{-1} and σ_{-1}, respectively, and

$$\alpha \leq \frac{\sigma_{-1}}{L_{-1}} \frac{\mu_m^+}{\|E\|},$$

where μ_m^+ is defined as before.

Then the following implications hold:

$$(e) \implies (d) \implies (c) \implies (b) \implies (a).$$

Proof. We only prove the implication $(c) \implies (b)$, the other ones are very similar to those of Corollary 3.6, hence they are left to the reader.

In order to verify the implication $(c) \implies (b)$, we have to show that

$$\mu_m^+(x) \geq \mu_m^s(JF(x)^{-1})\mu_m^+, \quad \forall x \in \mathbb{R}^n \tag{26}$$

holds. Take an arbitrary x, and let I_s^* be a set of indices such that $\frac{1}{2}E_{I_s^*}(JF(x)^{-1} + JF(x)^{-T})E_{I_s^*}^T$ is a submatrix of $\frac{1}{2}E(JF(x)^{-1} + JF(x)^{-T})E^T$ for which one obtains the minimum positive eigenvalue $\mu_m^+(x)$ for the given x (where as before $E_{I_s^*} := (E_{i*})_{i \in I_s^*}$). Let

$\bar{v}$ be an eigenvector of the matrix $\frac{1}{2}E_{I_s^*}(JF(x)^{-1} + JF(x)^{-T})E_{I_s^*}^T$ associated to $\mu_m^+(x)$; we may assume without loss of generality that $\|\bar{v}\| = 1$. Then we have

$$\bar{v}^T E_{I_s^*} JF(x)^{-1} E_{I_s^*}^T \bar{v} = \frac{1}{2}\bar{v}^T E_{I_s^*}(JF(x)^{-1} + JF(x)^{-T})E_{I_s^*}^T \bar{v} = \mu_m^+(x)\|\bar{v}\|^2 = \mu_m^+(x). \quad (27)$$

Since the eigenvectors corresponding to different eigenvalues of a symmetric matrix are orthogonal to each other, we have $\bar{v} \perp \operatorname{null}\big(\frac{1}{2}E_{I_s^*}(JF(x)^{-1} + JF(x)^{-T})E_{I_s^*}^T\big)$. However, it is easy to see that, for any positive definite (not necessarily symmetric) matrix A, the two matrices $E_{I_s^*}AE_{I_s^*}^T$ and $E_{I_s^*}E_{I_s^*}^T$ have the same null space. Hence we also have $\bar{v} \perp \operatorname{null}\big(E_{I_s^*}E_{I_s^*}^T\big)$. Now, assuming that $E_{I_s^*}E_{I_s^*}^T$ is an $s \times s$-matrix, let $E_{I_s^*}E_{I_s^*}^T = QDQ^T$ with $Q \in \mathbb{R}^{s\times s}$ orthogonal and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_s)$ be the spectral decomposition of $E_{I_s^*}E_{I_s^*}^T$, i.e. λ_i are the eigenvalues with corresponding eigenvectors v_i being the i-th column of Q. Suppose further that the null space of this matrix has dimension $r \leq s$ and that the eigenvalues are ordered such that $\lambda_1 \leq \dots \leq \lambda_s$. Then $\lambda_1 = \dots = \lambda_r = 0$ (and $\lambda_{r+1} \geq \mu_m^+$ in our notation) and the eigenvectors $v_1, \dots, v_r$ form a basis of the null space of $E_{I_s^*}E_{I_s^*}^T$. We therefore have $\bar{v}^T v_i = 0$ for all $i = 1, \dots, r$. Consequently, $w_i = 0$ for all $i = 1, \dots, r$, where $w := Q^T\bar{v}$. It therefore follows that

$$\begin{aligned}\bar{v}^T E_{I_s^*}E_{I_s^*}^T\bar{v} &= \bar{v}^T QDQ^T\bar{v} = w^T Dw = \sum_{i=1}^{s}\lambda_i w_i^2 \\ &= \sum_{i=r+1}^{s}\lambda_i w_i^2 \geq \mu_m^+ \sum_{i=r+1}^{s} w_i^2 = \mu_m^+ \sum_{i=1}^{s} w_i^2 = \mu_m^+\|w\|^2 = \mu_m^+\|\bar{v}\|^2 = \mu_m^+.\end{aligned}$$

Combining this inequality with (27), we obtain

$$\mu_m^+(x) = \bar{v}^T E_{I_s^*} JF(x)^{-1} E_{I_s^*}^T\bar{v} \geq \mu_m^s(JF(x)^{-1})\bar{v}^T E_{I_s^*}E_{I_s^*}^T\bar{v} \geq \mu_m^s(JF(x)^{-1})\mu_m^+,$$

and this shows that (26) holds. □

We illustrate the previous result by the following example which comes from a realistic model described in [52], and which is also used as a test problem in Section 7 (test problems OutKZ31 and OutKZ41).

Example 3.12 Consider the problem of an elastic body in contrast to a rigid obstacle. In particular assume that Coulomb friction is present. After discretization, this class of QVIs is characterized by a linear function $F(x) := Bx - g$ with a positive definite matrix B, and by the following constraints:

$$a_i - y_i \leq 0, \quad a_i = \begin{cases} l, & \text{if } i \in \{1, ..., n\} \text{ is even,} \\ \phi\, x_{i+1}, & \text{if } i \in \{1, ..., n\} \text{ is odd,} \end{cases}$$

$$y_i - b_i \leq 0, \quad b_i = \begin{cases} u, & \text{if } i \in \{1, ..., n\} \text{ is even,} \\ -\phi\, x_{i+1}, & \text{if } i \in \{1, ..., n\} \text{ is odd,} \end{cases}$$

with $l < u \leq 0$ and where $\phi \in \mathbb{R}$ is the friction coefficient. Let $x^* \in \mathbb{R}^n$ be a solution of the described QVI, then odd elements of x^* are interpreted as tangential stress components on the rigid obstacle in different points of such obstacle, while even elements are interpreted as outer normal stress components. This example fits into the framework of this subsection with

$$E := \begin{pmatrix} -I \\ I \end{pmatrix}, \quad \|Jc(x)\| = \sqrt{2}\phi, \quad \|E\| = \sqrt{2},$$

$$\mu_m^+ = \min\left\{ \mu_m^+(A) : \ A \text{ is a principal submatrix of } \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} \right\} = 1.$$

According to Corollary 3.11 (c)→(a), we can say that if

$$\phi \le \frac{1}{\sqrt{2}} \frac{\mu_m^s(B^{-1})}{\|B^{-1}\|} \frac{\mu_m^+}{\|E\|} = \frac{1}{\sqrt{2}} \frac{\mu_m^s(B^{-1})}{\|B^{-1}\|} \frac{1}{\sqrt{2}} \left(\le \frac{1}{2} \right),$$

then we are sure that $JH(x, \lambda, w)$ is nonsingular for all λ and w positive. Note that this condition holds for all sufficiently small friction coefficients ϕ. □

So far, in this subsection we have considered only QVIs that are linear in the y-part. This restriction has allowed us to give conditions that do not depend on the multipliers λ. However, we can extend the results we have obtained to the more general setting in which

$$K(x) \;=\; \{y \in \mathbb{R}^n \,|\, g(y, x) := q(y) - c(x) \le 0\}, \tag{28}$$

where both q and c are functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. We can prove the following theorem, in which nonsingularity conditions now also depend on the Lagrange multiplier λ. The proof follows lines identical to those of Theorem 3.9 and is therefore omitted.

Theorem 3.13 *Let g be defined as in (28), let a point $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m_{++}$ be given and assume that F and c are C^1 while q is C^2 around x. Suppose further that $J_x L(x, \lambda)$ is positive definite and that*

$$\|Jc(x)\| \;\le\; \frac{\mu_m^+(x, \lambda)}{\|J_x L(x, \lambda)^{-1}\| \|Jq(x)\|},$$

where $\mu_m^+(x, \lambda) = \min\{\mu_m^+(A) \mid A$ is a principal submatrix of $\frac{1}{2} Jq(x)(J_x L(x, \lambda)^{-1} + J_x L(x, \lambda)^{-T})$ $Jq(x)^T\}$ and $\mu_m^+(A)$ denotes once again the minimum positive eigenvalue of a symmetric matrix A. Then $JH(x, \lambda, w)$ is nonsingular for all positive w.

We conclude by considering a particular structure of the constraints of the QVI that is a subclass of that studied in this section. Suppose that

$$g(y, x) := \begin{pmatrix} l - y \\ y - u \\ I^{\pm}\,(y - c(x)) \end{pmatrix} \le 0, \tag{29}$$

where $I^{\pm}$ is a diagonal matrix with elements equal to 1 or -1, that is there are box constraints for y with lower bounds l and upper bounds u, and n special linear constraints with variable right-hand side.

Theorem 3.14 *Let g be defined as in (29), let a point $x \in \mathbb{R}^n$ be given and assume that around x it holds that F and c are C^1. Suppose that $JF(x)$ and $I - Jc(x)$ are row diagonally dominant with positive diagonal entries. Then $JH(x, \lambda, w)$ is nonsingular for all positive λ and w.*

Proof. Let

$$D := \begin{pmatrix} D_1 & 0 & 0 \\ 0 & D_2 & 0 \\ 0 & 0 & D_3 \end{pmatrix} := \mathrm{diag}(w^{-1} \circ \lambda),$$

where $D_1, D_2, D_3 \in \mathbb{R}^{n \times n}$ and $w = (w^1, w^2, w^3), \lambda = (\lambda^1, \lambda^2, \lambda^3)$ denote the slack variables and Lagrange multipliers corresponding to the three blocks in the definition of the inequality constraints from (29), respectively. Then we can write

$$\begin{aligned} N(x, \lambda, w) &= JF(x) + \begin{pmatrix} -I & I & I^{\pm} \end{pmatrix} D \begin{pmatrix} -I \\ I \\ I^{\pm}(I - Jc(x)) \end{pmatrix} \\ &= JF(x) + D_1 + D_2 + D_3 \left(I - Jc(x)\right). \end{aligned}$$

Note that $D_3 (I - Jc(x))$ is a row diagonally dominant matrix with positive diagonal entries for all λ and w positive. Hence $N(x, \lambda, w)$ is a strictly row diagonally dominant matrix for all λ and w positive since it is the sum of two row diagonally dominant matrices with positive diagonal entries ($JF(x)$ and $D_3 (I - Jc(x))$) and two strictly row diagonally dominant matrices with positive diagonal entries (D_1 and D_2). Recalling that every strictly row diagonally dominant matrix is nonsingular, we obtain the thesis. □

It is possible to generalize constraints (29) by imposing that lower or upper bounds may not exist for every variable and that the number of special linear constraints with variable right-hand side may be less or greater than n:

$$g(y, x) := \begin{pmatrix} (l_i - y_i)_{i \in L} \\ (y_i - u_i)_{i \in U} \\ \left(y_i - c_i^j(x)\right)_{i \in S_+}^{j \in C(i)} \\ \left(-y_i + d_i^j(x)\right)_{i \in S_-}^{j \in D(i)} \end{pmatrix} \leq 0, \tag{30}$$

where $L, U, S_-, S_+ \subseteq \{1, \ldots, n\}$ and for any $i \in S_+$, $C(i) \subseteq \{1, 2, \ldots\}$ and for any $i \in S_-$, $D(i) \subseteq \{1, 2, \ldots\}$. For QVIs with these constraints a result similar to Theorem 3.14 can be given. The proof of this theorem is akin to that of Theorem 3.14 and hence it is left to the reader.

Theorem 3.15 *Let g be defined as in (30), let a point $x \in \mathbb{R}^n$ be given and assume that around x it holds that F, c and d are continuously differentiable. Suppose that $JF(x)$ is row diagonally dominant with positive diagonal entries and such that for every $i \notin L \cup U$*

it holds that $JF(x)_{ii} > \sum_{k=1,\ldots,n,\,k\neq i} |JF(x)_{ik}|$. *Suppose further that for all* $i \in S_+$ *and all* $j \in C(i)$ *it holds that*

$$1 - \frac{\partial c_i^j(x)}{\partial x_i} \geq \sum_{k=1,\ldots,n,\,k\neq i} \left| \frac{\partial c_i^j(x)}{\partial x_k} \right|,$$

and that for all $i \in S_-$ *and all* $j \in D(i)$ *it holds that*

$$1 - \frac{\partial d_i^j(x)}{\partial x_i} \geq \sum_{k=1,\ldots,n,\,k\neq i} \left| \frac{\partial d_i^j(x)}{\partial x_k} \right|.$$

Then $JH(x, \lambda, w)$ *is nonsingular for all positive* λ *and* w.

3.4 Box Constraints and "Binary Constraints"

We now consider the situation where each component g_i of the constraint function from (2) depends only on a single pair (y_j, x_j) for some index $j \in \{1, \ldots, n\}$. In particular, this includes the case of bounds having parametric bound constraints. We use the terminology "binary constraints" for this class of problems. The following result shows how the nonsingularity Theorem 3.1 can be applied.

Theorem 3.16 *Let* $x \in \mathbb{R}^n$ *and* $\lambda > 0$ *be given. Suppose that each constraint* $g_i(\cdot, \cdot)$ $(i = 1, \ldots, m)$ *depends only on a single couple* $(y_{j(i)}, x_{j(i)})$ *for some* $j(i) \in \{1, \ldots, n\}$ *and that Assumptions 1.1 and 2.1 hold. Assume further that one of the following conditions holds:*

(a) $J_x L(x, \lambda)$ *is a* P*-matrix and* $\nabla_{y_{j(i)}} g_i(x_{j(i)}, x_{j(i)}) \nabla_{x_{j(i)}} h_i(x_{j(i)}) \geq 0$ *for all* i*, or*

(b) $J_x L(x, \lambda)$ *is a* P_0*-matrix and* $\nabla_{y_{j(i)}} g_i(x_{j(i)}, x_{j(i)}) \nabla_{x_{j(i)}} h_i(x_{j(i)}) > 0$ *for all* i*.*

Then $JH(x, \lambda, w)$ *is nonsingular for all positive* w.

Proof. We verify the statement under condition (a) since the proof under (b) is essentially identical.

We assume without loss of generality that the constraints g are ordered in such a way that the first m_1 constraints depend on the pair (y_1, x_1) only, the next m_2 constraints depend on the couple (y_2, x_2) only, and so on, with the last m_n constraints depending on (y_n, x_n) only. Note that m_i might be equal to zero for some of the indices $i \in \{1, \ldots, n\}$, and that we have $m_1 + m_2 + \ldots + m_n = m$. Taking this ordering into account, it is not difficult to see that

$$J_x h(x) = \begin{pmatrix} \nabla_{x_1} h_1(x_1) & 0 & \cdots & 0 \\ \nabla_{x_1} h_2(x_1) & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \nabla_{x_1} h_{m_1}(x_1) & 0 & \cdots & 0 \\ 0 & \nabla_{x_2} h_{m_1+1}(x_2) & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \nabla_{x_2} h_{m_1+m_2}(x_2) & & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & 0 & \nabla_{x_n} h_{m-m_n+1}(x_n) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \nabla_{x_n} h_m(x_n) \end{pmatrix},$$

whereas $\nabla_y g(x, x)$ is given by

$$\begin{pmatrix} \nabla_{y_1} g_1(x_1, x_1) & \cdots & \nabla_{y_1} g_{m_1}(x_1, x_1) & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & \ddots & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & \nabla_{y_n} g_{m-m_n+1}(x_n, x_n) & \cdots & \nabla_{y_n} g_m(x_n, x_n) \end{pmatrix}.$$

Then, an easy calculation shows that the matrix $N(x, \lambda, w)$ from (15) is given by

$$N(x, \lambda, w) = J_x L(x, \lambda) + D$$

with the diagonal matrix

$$D := \begin{pmatrix} \sum_{i=1}^{m_1} \frac{\lambda_i}{w_i} \nabla_{y_1} g_i(x_1, x_1) \nabla_{x_1} h_i(x_1) & & 0 \\ & \ddots & \\ 0 & & \sum_{i=m-m_n+1}^{m} \frac{\lambda_i}{w_i} \nabla_{y_n} g_i(x_n, x_n) \nabla_{x_n} h_i(x_n) \end{pmatrix}.$$

In view of assumption (a) together with $\lambda, w > 0$, it follows that $J_x L(x, \lambda)$ is a P-matrix and the diagonal matrix D is positive semidefinite. This implies that $N(x, \lambda, w)$ is nonsingular for all positive w, and then from Theorem 3.1 we obtain the thesis. □

We give below a specialization which deals with the most important case of Theorem 3.16: the case in which the constraints are bound constraints of the type

$$u_i(y_i, x_i) := y_i - a_i x_i \leq c_i \quad \forall i = 1, \ldots, n \quad \text{and} \tag{31}$$
$$l_i(y_i, x_i) := -y_i + b_i x_i \leq d_i \quad \forall i = 1, \ldots, n. \tag{32}$$

For this class of QVIs, Theorem 3.16 easily gives the following corollary.

Corollary 3.17 *Let $x \in \mathbb{R}^n$ be given, and consider a QVI whose feasible set is defined by the constraints* (31) *and* (32) *and suppose that F is C^1 around x. Assume that one of the following conditions hold:*

(a) $JF(x)$ is a P_0-matrix and $a_i < 1, b_i < 1$ for all $i = 1, \ldots, n$, or

(b) $JF(x)$ is a P-matrix and $a_i \leq 1, b_i \leq 1$ for all $i = 1, \ldots, n$.

Then $JH(x, \lambda, w)$ is nonsingular for all positive λ and w.

In principle, QVIs with box constraints can be viewed as a subclass of QVIs with linear constraints and variable right-hand sides, see (30). However, the conditions we got here are somewhat weaker. Note in particular that the conditions in Theorem 3.15 require JF to be diagonally dominant with positive diagonal elements, which implies that JF must be P_0, while a P_0 matrix is not necessarily diagonally dominant.

3.5 Bilinear Constraints

We conclude this section on nonsingularity results for JH by considering the case of bilinear constraints which can be considered as a natural variant of the case of (linear) constraints with variable right-hand side in which the right-hand sides are fixed, but the coefficients of the linear part vary. Specifically, we consider a QVI in which the feasible set is defined by some convex "private" constraints $q_i(y) \leq 0$ (that depend only on y) and some bilinear constraints of the form

$$x^T Q_i y - c_i \leq 0$$

in which each matrix Q_i is symmetric and positive semidefinite. Hence we consider constraints of the form

$$g(y, x) := \begin{pmatrix} q_1(y) \\ \vdots \\ q_p(y) \\ x^T Q_1 y - c_1 \\ \vdots \\ x^T Q_b y - c_b \end{pmatrix} \leq 0. \tag{33}$$

In order to deal with these constraints we give a preliminary result on QVIs in which the feasible set satisfies the condition

$$\nabla_x h(x) = \nabla_y g(x, x)\, D_+, \tag{34}$$

where D_+ is a diagonal matrix with nonnegative entries. Although this is a technical result, it is the key to the analysis of QVIs with bilinear constraints.

Theorem 3.18 *Suppose that Assumptions 1.1 and 2.1 hold. Let $x \in \mathbb{R}^n$ and $\lambda > 0$ be given. Assume that g and h satisfy equation* (34) *in x, and that $J_x L(x, \lambda)$ is a positive definite matrix. Then $JH(x, \lambda, w)$ is nonsingular for all positive w.*

Proof. It is easy to see that the matrix $\bar{M}$ from Corollary 3.3 is given by

$$\bar{M}(x, \lambda) = D_+\, \nabla_y g(x, x)^T J_x L(x, \lambda)^{-1} \nabla_y g(x, x),$$

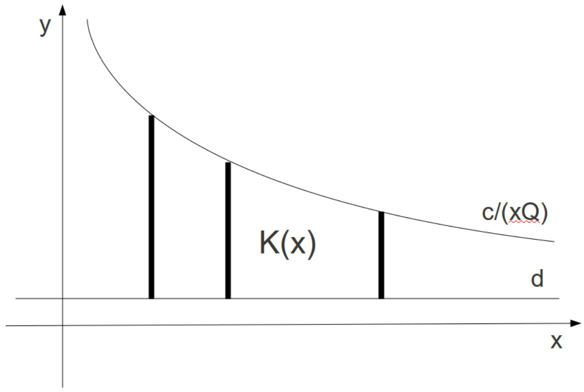

Figure 3: Example of bilinear constraints: here $y \geq d > 0$, $c > 0$ and $Q > 0$.

which is the product of a diagonal matrix with nonnegative entries and a positive semidefinite matrix. It is well known that a matrix with this form is P_0 (see Proposition B.12 in Appendix B), and then by Corollary 3.3 the thesis holds. □

Now, it is not difficult to see that the constraints (33) satisfy condition (34) with D_+ having the first p entries equal to 1 and the last b entries equal to 2. Therefore, the nonsingularity of JH follows immediately from Theorem 3.18.

Corollary 3.19 *Consider the constraints* (33)*, with each* q_i*,* $i = 1, \ldots, p$*, convex and* C^2 *and each* Q_j*,* $j = 1, \ldots, b$*, positive semidefinite and symmetric and suppose that* F *is* C^1*. Let* $x \in \mathbb{R}^n$ *and* $\lambda > 0$ *be given and assume that* $J_x L(x, \lambda)$ *is a positive definite matrix. Then* $JH(x, \lambda, w)$ *is nonsingular for all positive* w*.*

Note that $J_x L(x, \lambda)$ is certainly positive definite if either F is strongly monotone, or at least one q_i is strongly convex or at least one Q_j is positive definite.

4 Existence: Boundedness Conditions

In this section, we consider conditions guaranteeing that a sequence generated by Algorithm 3 is bounded and, therefore, has an accumulation point. We first discuss a general result and then its application to the moving set case. Application of the general result to the remaining settings considered before does not require any specific investigation, so we conclude the section with a few more examples and general considerations.

4.1 General Boundedness Conditions

We begin with a general result that shows that under a sort of coercivity condition ((a1) below) and constraint qualification ((a2) below) we can guarantee boundedness of the sequence generated by Algorithm 3. We recall that we assume that $K(x)$ is defined by (2) and that $h(x) := g(x, x)$.

Theorem 4.1 *Let the setting and the assumptions of Theorem 2.2 be satisfied and suppose, in addition, that*

(a1) $\lim_{\|x\|\to\infty} \|\max\{0, h(x)\}\| = \infty$,

(a2) *for all* $x \in \mathbb{R}^n$ *there exist a* d *such that* $\nabla_y g_i(x, x)^T d < 0$ *for all* $i \in \{i : h_i(x) \geq 0\}$.

Then any sequence generated by Algorithm 3 remains bounded, and any accumulation point is a solution of the QVI.

Proof. By Theorem 2.2 (a), it is enough to show that $\|H(x, \lambda, w)\|$ has bounded level sets over Z_I. To this end, suppose that a sequence $\{(x^k, \lambda^k, w^k)\} \subseteq Z_I$ exists such that $\lim_{k\to\infty} \|(x^k, \lambda^k, w^k)\| = \infty$. We will show that $\|H(x^k, \lambda^k, w^k)\| \to \infty$ as $k \to \infty$.

We first claim that the sequence $\{x^k\}$ is bounded. Assume we have $\|x^k\| \to \infty$. Then condition (a1) would imply $\| \max\{0, h(x^k)\}\| \to \infty$. Hence there would exist an index $j \in \{1, \dots, m\}$ such that, on a suitable subsequence, $h_j(x^k) \to +\infty$, and therefore also $\|h(x^k) + w^k\| \to \infty$ since $w^k > 0$. But this would imply $\|H(x^k, \lambda^k, w^k)\| \to \infty$ and gives the desired contradiction. Hence it remains to consider the case in which $\|(\lambda^k, w^k)\| \to \infty$ and $\{x^k\}$ is bounded.

Suppose that $\|w^k\| \to \infty$ and $\{x^k\}$ is bounded. Then $\{h(x^k)\}$ is also bounded due to the continuity of h. We therefore obtain $\|h(x^k) + w^k\| \to \infty$. This, in turn, implies $\|H(x^k, \lambda^k, w^k)\| \to \infty$ which, again, is a contradiction. Thus we have to consider only the case where $\|\lambda^k\| \to \infty$ and $\{(x^k, w^k)\}$ is bounded.

For $\|\lambda^k\| \to \infty$, let J_λ be the set of indices such that $\{\lambda^k_j\} \to \infty$, whereas, subsequencing if necessary, we may assume that the remaining components stay bounded. Without loss of generality, we may also assume that $x^k \to \bar{x}$ and $w^k \to \bar{w}$. If, for some $j \in J_\lambda$, we have $\bar{w}_j > 0$, it follows that $\lambda^k_j w^k_j \to +\infty$ and, therefore, again $\|H(x^k, \lambda^k, w^k)\| \to \infty$. Consequently, it remains to consider the case where $\bar{w}_j = 0$ for all $j \in J_\lambda$.

Since (x^k, λ^k, w^k) belongs to Z_I, we have $h_j(x^k)+w_j^k > 0$ which implies $h_j(\bar{x}) \geq 0$ for all $j \in J_\lambda$. Hence we can apply condition (a2) and obtain a vector d such that $\nabla_y g_j(\bar{x}, \bar{x})^T d < 0$, $\forall j \in J_\lambda$. This implies

$$\begin{aligned}\lim_{k\to\infty} L(x^k, \lambda^k)^T d &= \lim_{k\to\infty}\left(F(x^k)^T d + \sum_{j\notin J_\lambda} \lambda_j^k \nabla_y g_j(x^k, x^k)^T d\right) + \\ &\quad + \lim_{k\to\infty}\left(\sum_{j\in J_\lambda} \lambda_j^k \nabla_y g_j(x^k, x^k)^T d\right) = -\infty\end{aligned}$$

since the first term is bounded (because $\{x^k\} \to \bar{x}$ and the functions F and $\nabla_y g$ are continuous, and because all sequences $\{\lambda_j^k\}$ for $j \notin J_\lambda$ are bounded by definition of the index set J_λ), whereas the second term is unbounded since $\lambda_j^k \to +\infty$ and $\nabla_y g_j(\bar{x}, \bar{x})^T d < 0$ for all $j \in J_\lambda$. Using the Cauchy-Schwarz inequality, we therefore obtain

$$\|L(x^k, \lambda^k)\| \, \|d\| \geq |L(x^k, \lambda^k)^T d| \to \infty$$

for $k \to \infty$. Since d is a fixed vector, this implies $\|L(x^k, \lambda^k)\| \to \infty$ which, in turn, implies $\|H(x^k, \lambda^k, w^k)\| \to \infty$ for $k \to \infty$. This contradiction, together with Theorems 1.2 and 2.2, completes the proof. □

Note that condition (a1) in Theorem 4.1 guarantees boundedness of the x- and w-parts, whereas (a2) is needed for the λ-part. In principle, if we knew an upper bound for the multipliers value, we could add this bound to the constrained equation reformulation of the KKT system and dispense with assumption (a2) altogether; we do not elaborate further on this idea for lack of space.

Condition (a1) is a mild coercivity condition that implies neither the boundedness of $K(x)$ for any x nor the existence of a compact set K such that $K(x) \subseteq K$ for all x, as one might think at first sight.

Example 4.2 Consider a problem with $K(x) = \{y \in \mathbb{R} \,|\, y + x^2 \leq 1\}$. In this case, for every x, the set $K(x)$ is unbounded and yet (a1) is easily seen to hold, since $h(x) = x + x^2 - 1$. □

Example 4.3 Consider a problem with $K(x) = \{y \in \mathbb{R}^2 \,|\, \|y + x\|^2 \leq 1\}$. In this case, for every x, the set $K(x)$ is a ball of radius 1 and center in $-x$. We have $\cup_{x\in\mathbb{R}^n} K(x) = \mathbb{R}^n$. But $h(x) = 4\|x\|^2 - 1$ and so (a1) holds. □

However, uniform boundedness of $K(x)$ implies (a1) if some, very common and natural, further structure is assumed. So suppose $K(x) = K \cap K'(x)$, i.e. suppose that $K(x)$ is given by the intersection of a fixed set K and a point to set mapping K'. Analytically, this simply means that if $K(x) = \{y \in \mathbb{R}^n \,|\, g(x, y) \leq 0\}$, then some of the g_i actually only depend on y. Obviously, if K is bounded, $K(x)$ is uniformly bounded when x varies. In the proposition below we assume for simplicity that K is a bounded polyhedron (a quite common case, but see the remark following the proposition for a simple generalization).

Proposition 4.4 *Let $K(x)$ be defined by* (2) *with g continuous and convex in y for every $x \in \mathbb{R}^n$. Suppose that the first p inequalities of g are of the form $Ay \leq b$ and that the polyhedron defined by these inequalities is bounded. Then (a1) in Theorem 4.1 holds.*

Proof. Denote by K the bounded polyhedron defined by the inequalities $Ay \leq b$. By Hoffman's error bound, we know there exists a positive constant c such that for every $x \in \mathbb{R}^n$ we have $\text{dist}(x, K) \leq c \,\| \max\{0, Ax - b\}\|$. Since K is bounded, this shows that $\lim_{\|x\| \to \infty} \|\max\{0, Ax - b\}\| = \infty$. But then (a1) in Theorem 4.1 follows readily. □

Remark 4.5 It is clear that the polyhedrality of the set K is only used to deduce that an error bound holds. Therefore, it is straightforward to generalize the above result in the following way: Suppose that the first inequalities of g define a bounded set $K = \{g_i(y) \leq 0, i = 1, \ldots, p\}$ and that an error bound holds for this system of p inequalities. Then (a1) in Theorem 4.1 holds. The literature on error bounds is vast and there are many conditions that ensure the error bound condition, polyhedrality is just one of them. We refer the interested reader to [23, 56]. □

Condition (a2) in Theorem 4.1 is a very mild constraint qualification. It is related to the well-known Extended Mangasarian-Fromovitz Constraint Qualification (EMFCQ) for a system of inequalities.

Definition 4.6 *We say that a system of continuously differentiable inequalities $f(x) \leq 0$, with $f : \mathbb{R}^n \to \mathbb{R}^m$, satisfies the EMFCQ if, for all $x \in \mathbb{R}^n$, there exists a vector $d \in \mathbb{R}^n$ such that $\nabla f_i(x)^T d < 0$, for all i such that $f_i(x) \geq 0$.*

For each given x, the set $K(x)$ is defined by the system of inequalities $g(y, x) \leq 0$. It is then clear that condition (a2) is the requirement that an EMFCQ-like conditions holds just at the point $y = x$ and that this is a much weaker requirement than requiring the EMFCQ to hold for the system $g(y, x) \leq 0$. We give an example to clarify this point.

Example 4.7 Consider a problem with $K(x) = \{y \in \mathbb{R} \mid y^2 + x^2 - 1 \leq 0\}$. We have $\nabla_y g(x, x) = 2x$. It is clear that we can find a $d \in \mathbb{R}$ such tht $2xd < 0$ at any point except at $x = 0$. Therefore, (a2) holds because we have $h(0) < 0$. The EMFCQ, instead, is not satisfied for the set $K(1)$. In fact $\nabla_y g(y, 1) = 2y$ and for $y = 0$ it is immediate to verify that the EMFCQ fails. It is interesting to observe that for this problem also (a1) clearly holds.

Furthermore, if x is greater than 1, we have $K(x) = \varnothing$, hence this simple example also shows that (a1) and (a2) together do not imply $K(x) \neq \varnothing$ for all x. The latter is a condition often encountered in papers dealing with algorithms for the solution of QVIs. □

Armed with the developments so far, we can now study the applicability of Theorem 4.1 to the moving set case, which is the only setting, among those considered in the previous section, for which an additional analysis is useful.

4.2 The Moving Set Case

Consider the QVI structure defined in subsection 3.2:

$$K(x) := c(x) + Q = \{y \in \mathbb{R}^n \,|\, q(y - c(x)) \leq 0\}, \quad Q := \{y \in \mathbb{R}^n \,|\, q(y) \leq 0\}.$$

We recall that, in the previous section, we have given sufficient conditions for nonsingularity of JH. Such conditions presuppose that $\|Jc(x)\| \leq 1$ for all $x \in \mathbb{R}^n$. In the next proposition, we show that if the constraints $q(x) \leq 0$ define a full-dimensional bounded set and $\|Jc(x)\|$ is uniformly bounded away from 1, then conditions (a1) and (a2) of Theorem 4.1 hold.

Proposition 4.8 *In the setting described above, assume that c and q are continuously differentiable. Suppose that:*

(b1) $\|Jc(x)\| \leq \alpha < 1$ *for all* $x \in \mathbb{R}^n$*;*

(b2) *Q is compact and the system $q(y) \leq 0$ satisfies Slater's condition, i.e. there exists $\bar{y}$ such that $q(\bar{y}) < 0$.*

Then conditions (a1) and (a2) of Theorem 4.1 hold.

Proof. Since $\|Jc(x)\| \leq \alpha$ for all $x \in \mathbb{R}^n$, the Cauchy-Schwarz inequality implies $y^T Jc(x) y \leq \alpha \|y\|^2$ for all $x, y \in \mathbb{R}^n$. Therefore, $y^T(I - Jc(x))y \geq (1 - \alpha)\|y\|^2$ for all $x, y \in \mathbb{R}^n$, hence the function $x - c(x)$ is strongly monotone on $\mathbb{R}^n$ and, consequently, $\lim_{\|x\| \to \infty} \|x - c(x)\| = \infty$. Now, since q_i is convex for all components i, it follows that $\max\{0, q_i\}$ and, therefore, also $\|\max\{0, q(z)\}\|$ is convex. Hence, the corresponding level sets are bounded for all levels if and only if at least one level set is bounded. But the level set with level zero is precisely the set Q which was assumed to be compact. It therefore follows that all level sets of the function $z \mapsto \|\max\{0, q(z)\}\|$ are bounded. But then $\lim_{\|x\| \to \infty} \|x - c(x)\| = \infty$ implies $\lim_{\|x\| \to \infty} \|\max\{0, q(x - c(x))\}\| = \infty$, hence condition (a1) holds.

To show that also (a2) is satisfied, we first note that $\nabla_y g(x, x) = \nabla q(x - c(x))$. Therefore, taking $d := \bar{y} - (x - c(x))$, with $\bar{y}$ being the Slater point from assumption (b2), the convexity of q_i implies

$$0 > q(\bar{y}) \geq q_i(x - c(x)) + \nabla q_i(x - c(x))^T \big(\bar{y} - (x - c(x)\big)$$

for all components i such that $h_i(x) = q_i(x - c(x)) \geq 0$. But this immediately gives $\nabla_y g_i(x, x)^T d < 0$ for all i with $h_i(x) \geq 0$. □

4.3 Final Examples and Comments

We complete our discussion by giving a few additional examples on which we apply the results of both this and the previous section in order to show the ability of our algorithm to solve problems that are not solvable by other methods. This will also give us the opportunity to discuss very briefly some existence implications of the results obtained so far.

An often used assumption in the analysis of algorithms and also in many existence proofs is that either $K(x)$ is nonempty for all $x \in \mathbb{R}^n$ or that there exists a convex compact set $T \subset \mathbb{R}^n$ such that $K(T) \subseteq T$ and $K(x)$ is nonempty for all $x \in T$, see Theorem 1.3. The following example shows that this assumption is not implied by our conditions.

Example 4.9 Consider a one dimensional QVI with $F(x) = x^3$ and $K(x) = \{y \in \mathbb{R}^n \,|\, y^2 + x^2 + x^4 - 1 \leq 0\}$. First of all note that $K(x) = \varnothing$ if $x \notin [-a, a]$, where $a \approx 0.7862$ is the only positive root of the equation $x^2 + x^4 = 1$. Furthermore, it is not difficult to see that there cannot exist a convex compact set T (which would be a closed interval in our case) such that $K(T) \subseteq T$ holds and $K(x)$ is nonempty for all $x \in T$. In fact, it should be $T \subseteq [-a, a]$ since otherwise $K(x)$ is empty for some $x \in T$. Furthermore, 0 can not belong to T, otherwise $K(0) = [-1, 1] \not\subseteq T$. Then T should be an interval of either all negative or all positive numbers. But if nonempty, $K(x)$ always contains both positive and negative points.

Nevertheless, we can show that the conditions of Theorem 3.16 (a) are satisfied. We have $h(x) = 2x^2 + x^4 - 1$, so that $\nabla_y g(x, x) \nabla_x h(x) = (2x)(4x + 4x^3) = 8(x^2 + x^4) \geq 0$. Furthermore $J_x L(x, \lambda) = 3x^2 + 2\lambda$ which, for every x and positive λ, is positive. So Theorem 3.16 (a) tells us that $JH(x, \lambda, w)$ is nonsingular for any x and positive λ and w.

We next verify that also the assumptions of Theorem 4.1 are met. Condition (a1) is obvious from the expression of $h(x)$, so we consider (a2). We have $\nabla_y g(x, x) = 2x$, and if $x \neq 0$, it is sufficient to take $d = -x$ to have $\nabla_y g(x, x) d < 0$. If $x = 0$, this is not possible, but in this case we also have $h(x) < 0$ so that (a2) is satisfied. We can then conclude that every sequence generated by our interior-point method will be bounded and that every limit point is a solution of the QVI. Note that this also gives an algorithmic proof of the existence of a solution. We do not know any method that could provably solve this example. Also proving existence by using other known results seems not obvious. □

As far as we are aware of, all methods for which convergence to a solution can be proved make assumptions that imply the existence of a (at least locally) unique solution and require the function F to be strongly monotone. In the following example, we present a problem with a monotone, but not strongly monotone F, that has infinitely many solutions and for which we can prove convergence of our method.

Example 4.10 Consider again a one dimensional problem with

$$F(x) = \begin{cases} -(x+1)^4 & \text{if } x \leq -1, \\ 0 & \text{if } x \in [-1, 0], \\ x^4 & \text{if } x \geq 0 \end{cases}$$

and $K(x) = \{y \in \mathbb{R} \,|\, -10 \leq y \leq -2x\}$. The function F is monotone, but not strongly monotone, and the solutions of the problem are all points in $[-1, 0]$. The assumptions of

Corollary 3.17 are easily checked to be satisfied; in fact $a_1 = -2$, $b_1 = 0$ and since F is monotone, its Jacobian is positive semidefinite. Also condition (a1) in Theorem 4.1 holds trivially. Consider then (a2) in the same theorem. We have $h(x) = (-x - 10, 3x)^T$, so that it is clear that at most one component of h can be positive at any point, a fact that easily permits to check that also (a2) is satisfied. We conclude that the interior-point method is able to find a solution of this problem which admits infinitely many solutions. □

We remarked already several times that, when it comes to algorithms, the most studied QVIs are those with a moving set type of constraints. One of the most interesting papers in this category is [45] where, among other things, a wider class of problems is studied under a condition, subsequently used also by other authors, which is implied by the moving set structure (which actually constitutes the main case in which the condition below can be verified). This condition is

$$\|\Pi_{K(x)}(z) - \Pi_{K(y)}(z)\| \leq \alpha\|x - y\|, \quad \alpha < 1, \quad \forall x, y, z \in \mathbb{R}^n, \tag{35}$$

where Π_K denotes the Euclidean projector on K and α is a positive constant whose exact definition is immaterial here (see also Theorem 1.4). Roughly speaking, condition (35) is a strenghtening of a contraction property of the point-to-set mapping $K(\cdot)$. The following example shows that our assumptions do not imply condition (35).

Example 4.11 Consider the same problem as in the previous example and, in particular, its feasible set mapping $K(x) = \{y \in \mathbb{R} \,|\, -10 \leq y \leq -2x\}$. Then

$$\|\Pi_{K(0)}(1) - \Pi_{K(1)}(1)\| = \|0 - (-2)\| = 2 \leq \alpha\|0 - 1\| = \alpha$$

implies $\alpha \geq 2$, so that condition (35) does not hold, whereas we have already mentioned in Example 4.10 that our method provably solves this example. □

5 Generalized Nash Equilibrium Problems

In this section we specialize results developed in Section 2, 3 and 4 for one of the most important and studied applications for QVIs: the generalized Nash equilibrium problem (GNEP for short). In a GNEP each player ν ($\nu = 1, \dots, N$) controls $x^\nu \in \mathbb{R}^{n_\nu}$ and tries to solve the optimization problem

$$\min_{x^\nu} \; \theta_\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad g^\nu(x^\nu, x^{-\nu}) \leq 0 \tag{36}$$

with given $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$ and $g^\nu : \mathbb{R}^n \to \mathbb{R}^{m_\nu}$. Here, $n := n_1 + \ldots + n_N$ denotes the total number of variables, $m := m_1 + \ldots + m_N$ will be the total number of (inequality) constraints, and $(x^\nu, x^{-\nu})$ is a short-hand notation for the full vector $\mathbf{x} := (x^1, x^2, \dots, x^N)$, so that $x^{-\nu}$ subsumes all the block vectors x^μ with $\mu \neq \nu$. A vector $\mathbf{x} = (x^1, \dots, x^N)$ is called feasible for the GNEP if it satisfies the constraints $g^\nu(\mathbf{x}) \leq 0$ for all players $\nu = 1, \dots, N$. A feasible point $\bar{\mathbf{x}}$ is a solution of the GNEP if, for all players $\nu = 1, \dots, N$, we have

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \;\leq\; \theta_\nu(x^\nu, \bar{x}^{-\nu}) \quad \forall x^\nu : g^\nu(x^\nu, \bar{x}^{-\nu}) \;\leq\; 0,$$

i.e. if, for all players ν, $\bar{x}^\nu$ is the solution of the ν-th player's problem, when the other players set their variables to $\bar{x}^{-\nu}$.

In the literature on GNEPs it is typical to assume that the following blanket assumptions always hold:

Assumption 5.1 *$\theta_\nu(\cdot, x^{-\nu})$ and $g_i^\nu(\cdot, x^{-\nu})$ are convex for every $x^{-\nu}$, and for every $\nu = 1, \dots, N$ and $i = 1, \dots, m_\nu$.*

Assumption 5.2 *θ_ν and g^ν are C^2 functions for every $\nu = 1, \dots, N$.*

This is a very general form of a GNEP, and finding a solution of such a GNEP is a very hard problem, see [17, 24] for a detailed discussion. In fact, the solution of a GNEP in this general form is still little investigated. Due to its daunting difficulty, only very few results are available for the solution of a GNEP at the level of generality described above, see [14, 16, 18, 22, 28, 57, 58] for some different approaches. Some subclasses, in particular jointly convex Nash equilibrium problems (where $g^1 = g^2 = \ldots = g^N$ are the same convex functions, defining the same joint constraints for all players) and pure Nash equilibrium problems (where g^ν depends on x^ν alone for all $\nu = 1, \dots, N$), have been more widely investigated and some reasonably efficient methods for the solution of these latter problems have been proposed, see [17, 25].

Aim of this section is to study and give convergence results based on the use of the KKT conditions of the general GNEP (36). This has been done previously in [16, 57], where the authors were mainly interested in the local convergence behaviour of suitable Newton-type methods. In particular, it is shown in [16] that one has to expect difficulties in solving the KKT system due to some singularity problems, hence local fast convergence cannot be obtained in many standard situations. Apart from these papers, the KKT approach is also part of the folklore in the engineering world, but in spite of this, there is still a lack of any serious analysis dealing with the solution of this peculiar KKT-like system. In fact, the study of this system is not trivial at all, and deriving convergence results for methods based on the solution of the KKT system turns out to be a rather involved issue.

Here we fill this gap and provide sound results establishing the viability of the KKT approach, both at the theoretical and numerical level, and with a special emphasis on the global behaviour of the methods. In particular, we provide conditions under which the global convergence is guaranteed. These conditions are reasonable and, to the best of our knowledge, they are the first set of explicit conditions on a general GNEP under which global convergence can be established. Global convergence results are also discussed in [22], where a penalty technique for the solution of a general GNEP is proposed. Although the results in [22] are of great interest, global convergence for genuine GNEPs can only be established under restrictive conditions. These conditions depend on the unknown value of a (penalty) parameter and so their application appears to be problematic in practice.

Let $\bar{\mathbf{x}}$ be a solution of the GNEP (36). Assuming any standard constraint qualification holds, it is well known that the following KKT conditions will be satisfied for every player $\nu = 1, \dots, N$:

$$\begin{array}{rcl} \nabla_{x^\nu}\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) + \sum_{i=1}^{m_\nu} \lambda_i^\nu \nabla_{x^\nu} g_i^\nu(\bar{x}^\nu, \bar{x}^{-\nu}) & = & 0, \\ \lambda_i^\nu \geq 0, \qquad g_i^\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq 0, \qquad \lambda_i^\nu g_i^\nu(\bar{x}^\nu, \bar{x}^{-\nu}) & = & 0 \quad \forall i = 1, \dots, m_\nu, \end{array} \tag{37}$$

where λ^ν is the vector of Lagrange multipliers of player ν. Vice versa, recalling that the player's problems are convex (see Assumption 5.1), we have that if a point $\bar{\mathbf{x}}$ satisfies, together with a suitable vector of multipliers $\boldsymbol{\lambda} := (\lambda^1, \lambda^2, \dots, \lambda^N)$, the KKT conditions (37) for every ν, then $\bar{\mathbf{x}}$ is a solution of the GNEP. It then seems rather natural to try to solve the GNEP by solving the system obtained by concatenating the N systems (37). In order to use a more compact notation, we introduce some further definitions.

We denote by $L^\nu(\mathbf{x}, \lambda^\nu) := \theta_\nu(x^\nu, x^{-\nu}) + \sum_{i=1}^{m_\nu} \lambda_i^\nu g_i^\nu(x^\nu, x^{-\nu})$ the Lagrangian of player ν. If we set $\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}) := (\nabla_{x^\nu} L^\nu(\mathbf{x}, \lambda^\nu))_{\nu=1}^N$ and $\mathbf{g}(\mathbf{x}) := (g^\nu(\mathbf{x}))_{\nu=1}^N$, the concatenated KKT system can be written as

$$\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}) = 0, \qquad \boldsymbol{\lambda} \geq 0, \qquad \mathbf{g}(\mathbf{x}) \leq 0, \qquad \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = 0. \tag{38}$$

By introducing slack variables $\mathbf{w} := (w^\nu)_{\nu=1}^N$, where $w^\nu \in \mathbb{R}^{m_\nu}$, and setting $\boldsymbol{\lambda} \circ \mathbf{w} := \left(\lambda_1^1 w_1^1, \dots, \lambda_{m_N}^N w_{m_N}^N\right)^T$, we can define

$$H(z) := H(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{w}) := \begin{pmatrix} \mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}) \\ \mathbf{g}(\mathbf{x}) + \mathbf{w} \\ \boldsymbol{\lambda} \circ \mathbf{w} \end{pmatrix} \tag{39}$$

and

$$Z := \{z = (\mathbf{x}, \boldsymbol{\lambda}, \mathbf{w}) \mid \mathbf{x} \in \mathbb{R}^n, \boldsymbol{\lambda} \in \mathbb{R}_+^m, \mathbf{w} \in \mathbb{R}_+^m\}. \tag{40}$$

It is immediate to verify that a point $(\mathbf{x}, \boldsymbol{\lambda})$ solves the KKT system (38) if and only if this point, together with a suitable $\mathbf{w}$, solves the constrained equation defined by (39) and (40). Then we can apply Algorithm 3 to solve the CE (H, Z) and then the GNEP.

For the sake of notational simplicity, it is useful to introduce the matrix

$$E(\mathbf{x}) := \begin{pmatrix} \nabla_{x^1} g^1(\mathbf{x}) & & 0 \\ & \ddots & \\ 0 & & \nabla_{x^N} g^N(\mathbf{x}) \end{pmatrix} \quad \text{with} \quad \nabla_{x^\nu} g^\nu(\mathbf{x}) \in \mathbb{R}^{n_\nu \times m_\nu}. \tag{41}$$

5.1 Nonsingularity Conditions

We recall that the critical issue in applying Theorem 2.2 is establishing the nonsingularity of

$$JH(z) := \begin{pmatrix} J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}) & E(\mathbf{x}) & 0 \\ J_{\mathbf{x}}\mathbf{g}(\mathbf{x}) & 0 & I \\ 0 & \operatorname{diag}(\mathbf{w}) & \operatorname{diag}(\boldsymbol{\lambda}) \end{pmatrix}, \tag{42}$$

for all $\boldsymbol{\lambda}$ and $\mathbf{w}$ positive.

The following theorem gives a sufficient condition for the nonsingularity of JH in Z. This condition is interesting because it gives a quantitative insight into what is necessary to guarantee the nonsingularity of JH.

Theorem 5.3 *Let $z = (\mathbf{x}, \boldsymbol{\lambda}, \mathbf{w}) \in \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}$ be given such that $J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})$ is nonsingular and*

$$\mu_m^s \left(E(\mathbf{x})^T J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} E(\mathbf{x}) \right) \geq \| J_{\mathbf{x}}\mathbf{g}(\mathbf{x}) - E(\mathbf{x})^T \|_2 \left\| J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} \right\|_2 \| E(\mathbf{x}) \|_2 .$$

Then the Jacobian $JH(z)$ is nonsingular.

Proof. For all $u \in \mathbb{R}^m$ we have

$$\begin{aligned} u^T E(\mathbf{x})^T J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} E(\mathbf{x}) u &= \frac{1}{2} u^T \left(E(\mathbf{x})^T \left(J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} + J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1^T} \right) E(\mathbf{x}) \right) u \\ &\geq \mu_m^s \left(E(\mathbf{x})^T J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} E(\mathbf{x}) \right) \|u\|_2^2 \\ &\geq \| J_{\mathbf{x}}\mathbf{g}(\mathbf{x}) - E(\mathbf{x})^T \|_2 \left\| J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} \right\|_2 \| E(\mathbf{x}) \|_2 \|u\|_2^2 \\ &\geq | u^T \left(J_{\mathbf{x}}\mathbf{g}(\mathbf{x}) - E(\mathbf{x})^T \right) J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} E(\mathbf{x}) u | \\ &\geq -u^T \left(J_{\mathbf{x}}\mathbf{g}(\mathbf{x}) - E(\mathbf{x})^T \right) J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} E(\mathbf{x}) u. \end{aligned}$$

Using the matrix $\bar{M}(\mathbf{x}, \boldsymbol{\lambda})$ (defined for QVIs in Corollary 3.3), this implies that $u^T \bar{M}(\mathbf{x}, \boldsymbol{\lambda}) u = u^T J_{\mathbf{x}}\mathbf{g}(\mathbf{x}) \; J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})^{-1} E(\mathbf{x}) u \geq 0$ for all $u \in \mathbb{R}^m$. Therefore $\bar{M}(\mathbf{x}, \boldsymbol{\lambda})$ is positive semidefinite, hence a P_0-matrix, and Corollary 3.3 guarantees nonsingularity of $JH(z)$. $\square$

It should be pointed out that in general, in Theorem 5.3, we do not need the matrix $J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda})$ to be positive (semi-) definite. This is illustrated by the following example.

Example 5.4 Consider a GNEP with two players, each controlling a single variable. The problem is given by

$$\begin{aligned}
&\text{Player 1:} \quad \min_{x_1} \ \frac{1}{2}x_1^2 - 2x_1 && \text{s.t.} \quad x_1^2 + x_2 \le 0,\\
&\text{Player 2:} \quad \min_{x_2} \ \frac{1}{2}x_2^2 + (2 - x_1^2)x_2 && \text{s.t.} \quad x_2 \in \mathbb{R}.
\end{aligned}$$

It is easy to see that $J_{\mathbf{x}}\mathbf{F}(\mathbf{x},\lambda) = \begin{pmatrix} 1+2\lambda & 0 \\ -2x_1 & 1 \end{pmatrix}$ is nonsingular for all $x \in \mathbb{R}^2$ and all $\lambda > 0$ but it is not positive semidefinite everywhere. However, since a simple calculation shows that $J_{\mathbf{x}}\mathbf{g}(\mathbf{x})J_{\mathbf{x}}\mathbf{F}(\mathbf{x},\lambda)^{-1}E(\mathbf{x}) = 8x_1^2/(1+2\lambda) \ge 0$, it follows that the conditions from Theorem 5.3 are satisfied.

As it is well known, it is possible to reformulate a GNEP as a QVI. Doing this, we can give new results for GNEPs improving on the ones exposed so far (and which are taken from [13]).

We consider GNEPs where each player solves a problem whose feasible set is defined by a system of linear inequalities with variable right-hand side, i.e., player ν ($\nu = 1, \dots, N$) controls $x^\nu \in \mathbb{R}^{n_\nu}$ and tries to solve the optimization problem

$$\min_{x^\nu} \ \theta_\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad E^\nu x^\nu - b^\nu - c^\nu(x^{-\nu}) \le 0 \tag{43}$$

with given $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$, $E^\nu \in \mathbb{R}^{m_\nu \times n_\nu}$ and $c^\nu : \mathbb{R}^{n-n_\nu} \to \mathbb{R}^{m_\nu}$, $b^\nu \in \mathbb{R}^{m_\nu}$. It is well known that a solution of the GNEP (43) can be computed by solving the following QVI:

$$\begin{aligned}
&\text{Find } \bar{\mathbf{x}} \in \{\mathbf{x} \in \mathbb{R}^n : E^\nu x^\nu - b^\nu - c^\nu(x^{-\nu}) \le 0, \nu = 1, \dots, N\} \text{ such that}\\
&(\nabla_{x^\nu}\theta_\nu(\bar{\mathbf{x}}))_{\nu=1}^{N\ T}(\mathbf{y} - \bar{\mathbf{x}}) \ge 0, \quad \forall \mathbf{y} \in \mathbb{R}^n : E^\nu y^\nu - b^\nu - c^\nu(\bar{x}^{-\nu}) \le 0,\ \nu = 1, \dots, N.
\end{aligned} \tag{44}$$

To simplify the notation, we write

$$F(\mathbf{x}) := (\nabla_{x^\nu}\theta_\nu(\mathbf{x}))_{\nu=1}^N, \quad E := \begin{pmatrix} E^1 & & 0 \\ & \ddots & \\ 0 & & E^N \end{pmatrix}, \quad c(\mathbf{x}) := \begin{pmatrix} c^1(x^{-1}) \\ \vdots \\ c^N(x^{-N}) \end{pmatrix}. \tag{45}$$

Note that the QVI (44) belongs to the class of QVIs whose constraints are defined by (24). This fact allows us to rewrite Theorem 3.9 for the GNEP (43).

Theorem 5.5 *Consider a GNEP in which each player tries to solve (43). Recalling the notation in (45), let a point $\mathbf{x} \in \mathbb{R}^n$ be given and assume that F and c are C^1 around x. Suppose further that $JF(\mathbf{x})$ is positive definite and that*

$$\|Jc(\mathbf{x})\| \le \frac{\mu_m^+(\mathbf{x})}{\|JF(\mathbf{x})^{-1}\|\|E\|}, \tag{46}$$

where $\mu_m^+(\mathbf{x}) = \min\{\mu_m^+(A) \mid A$ is a principal submatrix of $\frac{1}{2}E(JF(\mathbf{x})^{-1} + JF(\mathbf{x})^{-T})E^T\}$, and $\mu_m^+(A)$ is again the minimum positive eigenvalue of the matrix A. Then $JH(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{w})$ is nonsingular for all positive $\boldsymbol{\lambda}$ and $\mathbf{w}$.

Theorem 5.5 gives weaker conditions than those of Theorem 5.3, in fact it is clear that

$$\mu_m^s(EJF(\mathbf{x})^{-1}E^T) \le \mu_m^+(\mathbf{x}).$$

The following example describes a GNEP that satisfies conditions of Theorem 5.5, but violates those from Theorem 5.3 for all $\mathbf{x}$.

Example 5.6 Consider a GNEP in which there are two players controlling one variable each one, x^1 and x^2 respectively. The optimization problems of the players are the following:

$$\begin{array}{ll|ll} \min\limits_{x^1} & (x^1-2)^2 & \min\limits_{x^2} & (x^2-2)^2 \\ \text{s.t.} & x^1+\frac{1}{2}x^2 \le 1, & \text{s.t.} & x^2+\frac{1}{2}x^1 \le 1, \\ & x^1 \ge 0, & & x^2 \ge 0. \end{array}$$

This GNEP has only one equilibrium in $(\frac{2}{3}, \frac{2}{3})$. Referring to the notation in (45), we write

$$F(\mathbf{x}) = \begin{pmatrix} 2x^1-4 \\ 2x^2-4 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix}, \quad c(\mathbf{x}) = \begin{pmatrix} -\frac{1}{2}x^2 \\ 0 \\ -\frac{1}{2}x^1 \\ 0 \end{pmatrix}.$$

Then

$$JF(\mathbf{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \succ 0, \quad JF(\mathbf{x})^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \quad Jc(\mathbf{x}) = \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & 0 \\ -\frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}.$$

Since $\|JF(\mathbf{x})^{-1}\| = \frac{1}{2}$, $\|E\| = \sqrt{2}$, $\|Jc(\mathbf{x})\| = \frac{1}{2}$, and

$$EJF(\mathbf{x})^{-1}E^T = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

conditions of Theorem 5.3 do not hold since $\mu_m^s(EJF(\mathbf{x})^{-1}E^T) = 0$ for all $\mathbf{x}$ and therefore $\frac{1}{2} \not< 0$. However, condition (46) holds because, recalling the notation of Theorem 5.5, $\mu_m^+(\mathbf{x}) = \frac{1}{2}$ for all $\mathbf{x}$, and then we have $\frac{1}{2} < \frac{\frac{1}{2}}{\frac{1}{2}\sqrt{2}} = \frac{1}{\sqrt{2}}$. □

Now we consider games with totally different structure and still get convergence results. In particular we consider GNEPs that can be reformulated as QVIs with moving sets (see Subsection 3.2). Suppose that each player ν has to solve the following optimization problem

$$\min_{x^\nu} \; \theta_\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad q^\nu(x^\nu - c^\nu(x^{-\nu})) \le 0 \tag{47}$$

with given $\theta_\nu : \mathbb{R}^n \to \mathbb{R}$, $q^\nu : \mathbb{R}^{n_\nu} \to \mathbb{R}^{m_\nu}$ and $c^\nu : \mathbb{R}^{n-n_\nu} \to \mathbb{R}^{n_\nu}$. The feasible region of player ν is therefore a "moving set" whose position depends on the variables of all other players. The GNEP can be reformulated as a QVI (F, K) with

$$F(\mathbf{x}) := (\nabla_{x^\nu}\theta_\nu(\mathbf{x}))_{\nu=1}^N, \quad K(\mathbf{x}) := c(\mathbf{x}) + Q, \tag{48}$$

$$c(\mathbf{x}) := \begin{pmatrix} c^1(x^{-1}) \\ \vdots \\ c^N(x^{-N}) \end{pmatrix}, \quad Q := \prod_{\nu=1}^N \{y^\nu \in \mathbb{R}^{n_\nu} \mid q^\nu(y^\nu) \le 0\}, \tag{49}$$

that is a QVI with a moving set to which the nonsingularity results in Section 3.2 can readily be applied.

Theorem 5.7 *Consider a GNEP in which each player tries to solve (47). Recalling the notation in (48) and (49), let a point $\mathbf{x} \in \mathbb{R}^n$ be given and assume that F and c are C^1 around x. Suppose further that $JF(\mathbf{x})$ is positive definite and that*

$$\|Jc(\mathbf{x})\| \le \frac{\mu_m^s(JF(\mathbf{x})^{-1})}{\|JF(\mathbf{x})^{-1}\|}, \tag{50}$$

then $JH(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{w})$ is nonsingular for all positive $\boldsymbol{\lambda}$ and $\mathbf{w}$.

5.2 Boundedness Conditions

The following theorem is about boundedness of Algorith 3 in the solution of KKT system of GNEPs. Since its proof is similar to that of Theorem 4.1 it is left to the reader.

Theorem 5.8 *Suppose that $JH(z)$ is nonsingular for all $z \in Z_I$. If it holds that*

(a1) $\lim_{\|\mathbf{x}\|\to\infty} \|\max\{0, \mathbf{g}(\mathbf{x})\}\| = \infty$,

(a2) *The Extended Mangasarian-Fromovitz Constraint Qualification (EMFCQ) holds for each player, i.e., for all $\nu = 1, \dots, N$ and for all $\mathbf{x} \in \mathbb{R}^n$,*

$$\exists d^\nu \in \mathbb{R}^{n_\nu} : \ \nabla_{\mathbf{x}^\nu} g_i^\nu(\mathbf{x})^T d^\nu < 0 \ \ \forall i \in I_\ge^\nu(\mathbf{x}),$$

where $I_\ge^\nu(\mathbf{x}) := \{i \in \{1, \dots, m_\nu\} \mid g_i^\nu(\mathbf{x}) \ge 0\}$ denotes the set of active or violated constraints for player ν.

Then any sequence generated by Algorithm 3 remains bounded, and any accumulation point is a solution of the GNEP.

Part III
Applications and Numerical Results

6 Applications

In this section we present a collection of 55 QVI test problems. The problems range from small (few variables and constraints) to large (several thousands of variables and constraints). They include academic problems, problems arising from real-world applications (e.g. Walrasian equilibrium problems) and problems resulting from the discretization of infinite-dimensional QVIs modelling diverse engineering and physical problems. For each problem we provide a succinct, but complete description, along with Matlab files which allow the user to easily obtain function values and derivatives and that can be easily incorporated in any solution routine developed in order to solve QVIs. It is hoped that the availability of this collection, which we plan to maintain and enlarge, will stimulate the development of new solution methods and will permit a uniform and fair comparison of existing and future algorithms. The collection can freely be obtained by writing to `facchinei@dis.uniroma1.it`, `kanzow@mathematik.uni-wuerzburg.de` or `sagratella@dis.uniroma1.it`.

6.1 Classification of Test Problems

Each QVI, see (1), is defined by the function F and the point-to-set mapping $K(x)$. We assume that $K(x)$ is defined as the intersection of a fixed set $\bar{K}$ and a set $\tilde{K}(x)$ that depends on the point x:

$$K(x) = \bar{K} \cap \tilde{K}(x).$$

The sets $\bar{K}$ and $\tilde{K}(x)$ are described by inequalities and equalities:

$$\begin{aligned}\bar{K} &:= \{y \in \mathbb{R}^n \mid g^I(y) \leq 0,\, M^I y + v^I = 0\},\\ \tilde{K}(x) &:= \{y \in \mathbb{R}^n \mid g^P(y,x) \leq 0,\, M^P(x)y + v^P(x) = 0\}.\end{aligned}$$

The constraints defining the set $\bar{K}$ are individual constraints that are independent of x, hence we use the superscript "I" in our notation (for individual/independent of x). On the other hand, the constraints defining $\tilde{K}(x)$ are parametric due to their dependence on x, therefore, we use the superscript "P" (for parametric). According to Assumption 1.1, we assume that $g^I(\cdot)$ is a vector of convex functions and that each component function of $g_i^P(\cdot, x)$ is convex for all x. When we refer to the whole set of inequality or (linear) equality constraints, we use the notation

$$g(y,x) := \begin{pmatrix} g^I(y) \\ g^P(y,x) \end{pmatrix}, \qquad M(x)y + v(x) := \begin{pmatrix} M^I \\ M^P(x) \end{pmatrix} y + \begin{pmatrix} v^I \\ v^P(x) \end{pmatrix}.$$

For each test problem, we therefore report F and the functions defining $K(x)$ along with few more information concerning origin of the problem, known characteristics (for example monotonicity of F, uniqueness of the solution, etc). Furthermore, in some cases we also give some more details on the construction of the test problem (for example in the case of a discretization of an infinite-dimensional problem).

Each problem in the test set comes with a simple problem classification which we explain below. A problem is classified by the label [XXX/XX/n-m_I-p_I-m_P-p_P]. The first character in the label defines the type of the operator F. Possible values are:

L : F is linear (L = linear)

N : F is nonlinear (N = nonlinear).

The second character in the label defines the type of constraints g^I of the problem. Possible values are:

A : there are no constraints g^I (A = absent)

B : g^I defines only bounds on the variables (B = box/bound)

L : g^I is linear (L = linear)

Q : g^I is quadratic (Q = quadratic)

N : g^I is nonlinear (N = nonlinear).

The third character in the label defines the type of constraints g^P of the problem. The classification below is based on the classes of constraints analyzed in Section 3, which are briefly recalled below. Possible values are:

A : there are no constraints g^P (A = absent)

B : g^P defines separable box (in the y-part) constraints only: a generic constraint has the form $ay_i + b(x) - c \leq 0$ (B = box/bound)

L : g^P defines separable linear (in the y-part) constraints only: a generic constraint has the form $a^T y + b(x) - c \leq 0$ (L = linear)

O : g^P defines constraints different from any of the above (O = other).

Characters immediately following the first slash indicate the primary origin and/or interest of the problem (one or two characters are allowed here). Possible values are:

A : the problem is academic, that is, has been constructed specifically by researchers to test one or more algorithms (A = academic)

R : the problem models some real problem: economic, physical, etc. (R = real)

D : the problem is the discretization of an infinite-dimensional problem (D = discretized).

The numbers after the second slash indicate the "dimensions" of the problem, in particular:

- n is the number of variables;

- m_I is the number of inequality constraints defining $\bar{K}$;

- p_I is the number linear equalities in $\bar{K}$;
- m_P is the number of inequality constraints defining $\tilde{K}(x)$;
- p_P is the number of equalities in the definition of $\tilde{K}(x)$.

In Table 1 we report the list of all problems currently in the library, with the corresponding labels.

6.2 Description of Matlab Functions

Each QVI test problem described below is distributed as a single Matlab M-file function, whose name is the same as that of the problem. For some of the larger problems, a data file is also necessary which, again, has the same name as the problem (see below). All these files are contained in a folder called QVILIB. For each problem, and given two points y and x, the quantities given in Table 2 can be computed.

Let us give some more explanations. To this end, consider a generic problem whose name is `<QVI_name>`; in the folder QVILIB one can find the M-file `<QVI_name>.m` and, for some of the problems, a second data file `<QVI_name>.dat`. The function `<QVI_name>` can have up to three input arguments. The first input argument of `<QVI_name>` is a mandatory flag and it is used by the user to select the behavior of the function as detailed in the previous list.

In this list, it is also shown how many additional input arguments should be used in correspondence to each admissible value for the flag `i`, which can take any integer value between 0 and 11. If the flag value `i` is out of range or if the number of input arguments is not correct an exception will be thrown. Note that, if present, the second and third input argument of `<QVI_name>` must be column vectors with `nVar` elements; otherwise, a corresponding exception will be thrown.

When the function is called with the first argument equal to 0, some preliminary operations are performed, in particular in the scope of the function some global variables are initialized. This set of global variables always contains the positive integer `nVar`, i.e. the number of variables, the nonnegative integer `nIneq`, i.e. the total number of inequality constraints, the nonnegative integer `nEq`, i.e. the total number of equality constraints, the nonnegative integer `nIneqInd`, i.e. the number of inequality constraints independent of x, and the nonnegative integer `nEqInd`, i.e. the number of equality constraints that do not depend on x. In order to make these quantities available to the user's calling function, one should define them as global also in the user's calling function(s). When the function is called with the first argument equal to 0, other global variables might be defined that store data used by the function when invoked with other flags. All these further variables begin with the string "QVItest" and therefore it is better to avoid the use of any quantities which includes this string in the user's functions. `<QVI_name>(0)` must be called before any other function call. If this rule is not respected, an exception will be thrown. If it is called more than one time, a warning will be displayed since unnecessary operations are performed.

Table 1: Problem list.

Academic problems	
Problem name	**Label**
OutZ40 - OutZ41	[LBB/A/2-4-0-2-0]
OutZ42	[LBB/A/4-4-0-4-0]
OutZ43 - OutZ44	[LAB/A/4-0-0-4-0]
MovSet1A - MovSet1B - MovSet2A - MovSet2B	[LAO/A/5-0-0-1-0]
MovSet3A1 - MovSet3B1	[LAO/A/1000-0-0-1-0]
MovSet3A2 - MovSet3B2	[LAO/A/2000-0-0-1-0]
MovSet4A1 - MovSet4B1	[LAO/A/400-0-0-801-0]
MovSet4A2 - MovSet4B2	[LAO/A/800-0-0-1601-0]
Box1A - Box1B	[LAB/A/5-0-0-10-0]
Box2A - Box2B - Box3A - Box3B	[NAB/A/500-0-0-1000-0]
BiLin1A - BiLin1B	[LBO/A/5-10-0-3-0]
RHS1A1 - RHS1B1 - RHS2A1 - RHS2B1	[LAL/A/200-0-0-199-0]
Problems from real-world models	
Problem name	**Label**
WalEq1	[LBO/R/18-18-1-5-0]
WalEq2	[NBO/R/105-105-1-20-0]
WalEq3	[LBO/R/186-186-1-30-0]
WalEq4	[NBO/R/310-310-1-30-0]
WalEq5	[NBO/R/492-492-1-40-0]
Wal2	[NLO/A/105-107-0-20-0]
Wal3	[LLO/R/186-188-0-30-0]
Wal5	[NLO/A/492-494-0-40-0]
LunSS1	[NBA/R/501-1002-0-0-6]
LunSS2	[NBA/R/1251-2502-0-0-6]
LunSS3	[NBA/R/5001-10002-0-0-6]
LunSSVI1	[NBA/R/501-1002-1-0-0]
LunSSVI2	[NBA/R/1251-2502-1-0-0]
LunSSVI3	[NBA/R/5001-10002-1-0-0]
Discretized problems	
Problem name	**Label**
Scrim11	[LBA/RD/2400-2400-0-0-1200]
Scrim12	[LBA/RD/4800-4800-0-0-2400]
Scrim21	[LBL/RD/2400-2400-0-2400-0]
Scrim22	[LBL/RD/4800-4800-0-4800-0]
OutKZ31	[LBB/RD/62-62-0-62-0]
OutKZ41	[LBB/RD/82-82-0-82-0]
KunR11 - KunR21 - KunR31	[LAO/RD/2500-0-0-2500-0]
KunR12 - KunR22 - KunR32	[LAO/RD/4900-0-0-4900-0]

When the function is called with the first argument equal to 11, then all global variables in the scope of the function will be cleared. If used, it must be the last function call.

The function `<QVI_name>.m` can have one output, or no output at all, depending on the value of the flag `i`. When present, the output can be either a column vector, a sparse matrix or a cell array of sparse matrices. Table 3 summarizes in detail all possible output formats.

Table 2: Description of all possible calls to a generic QVI function in the library.

Call	Description
`<QVI_name>(0)`	initializes the data that are used when invoking `<QVI_name>` with other flags; does not return anything. In particular sets, as global variables, the following "dimensions": `nVar`: number of variables `nIneq`: total number of inequality constraints `nEq`: total number of equality constraints `nIneqInd`: number of inequality constraints independent of x `nEqInd`: number of equality constraints independent of x
`<QVI_name>(1,x)`	returns $F(x)$
`<QVI_name>(2,x)`	returns $JF(x)$, the Jacobian of F at x
`<QVI_name>(3,x,y)`	returns $g(y, x)$
`<QVI_name>(4,x,y)`	returns $J_y g(y, x)$, the partial Jacobian of g with respect to y
`<QVI_name>(5,x)`	returns $Jh(x)$, the Jacobian of $h(x) := g(x, x)$
`<QVI_name>(6,x)`	returns $Js_i(x)$ for all i, the Jacobian of all functions $s_i(x) := J_y g_i(y, x)\|_{y=x}$
`<QVI_name>(7,x,y)`	returns $M(x)y + v(x)$
`<QVI_name>(8,x)`	returns $M(x)$
`<QVI_name>(9,x)`	returns $Jt(x)$, the Jacobian of $t(x) := M(x)x + v(x)$
`<QVI_name>(10,x)`	returns $J(M_{i*}(x)^T)$ for all i, where $M_{i*}(x)$ denotes the ith row of the matrix $M(x)$
`<QVI_name>(11)`	clears all problem data from memory; does not return anything

We already observed that, in order to help the user debugging, some simple checks are performed when the `<QVI_name>.m` function is invoked. If these checks fail, a corresponding error message is provided by throwing an exception or a warning. Some of these have been mentioned already; we report the complete list in Table 4.

For users convenience short string aliases for the mandatory flag `i` are enabled. The complete list of aliases, which are case insensitive, is reported in Table 5.

The library also includes an M-file `startingPoints.m` that can be used by the user to get the starting points of each test problem. If the function `startingPoints` is called without any input arguments, it displays a list of all test problems available with a brief description of their starting points. The function `startingPoints` returns the number of starting points available for one specific test problem by accepting a string of characters, corresponding to the test problem name, as the only input argument. Finally the function

Table 3: Description of outputs of a generic QVI function in the library.

Input flag	Output
`i`= 0 or `i`= 11	no output
`i`= 1	column vector of dimension `nVar`
`i`= 2	sparse square matrix of dimensions `nVar` $\times$ `nVar`
`i`= 3	column vector of dimension `nIneq` if `nIneq`= 0 the output is the empty matrix
`i`= 4 or `i`= 5	sparse matrix of dimension `nIneq` $\times$ `nVar` if `nIneq`= 0 the output is the empty matrix
`i`= 6	cell array of dimension `nIneq`, each cell in the array contains a sparse square matrix of dimension `nVar` $\times$ `nVar` (the matrices contained in the cells are the evaluations of $Js_i(x)$) if `nIneq`= 0 the output is the empty cell array
`i`= 7	column vector of dimension `nEq` if `nEq`= 0 the output is the empty matrix
`i`= 8 or `i`= 9	sparse matrix of dimension `nEq` $\times$ `nVar` if `nEq`= 0 the output is the empty matrix
`i`= 10	cell array of dimension `nEq`, each cell in the array contains a sparse square matrix of dimension `nVar` $\times$ `nVar` (the matrices contained in the cells are the evaluations of $J(M_{i*}(x)^T))$ if `nEq`= 0 the output is the empty cell array

Table 4: Description of exceptions.

Event	Exception/warning thrown
`i` is not an integer between 0 and 11	`QVItest:BadFlagInput`
`QVI_name` is invoked with `i` between 1 and 10 before invoking it with `i` = 0	`QVItest:DataNotInitialized`
`QVI_name`(0) is invoked more than once	`QVItest:MultipleDataInitialization` (warning)
`QVI_name` is invoked with a wrong number of arguments	`QVItest:BadInputNumber`
the second or third argument of `QVI_name` have wrong dimensions	`QVItest:BadInputArgument`

`startingPoints` returns a starting point for a test problem by accepting in input two arguments: a string of characters corresponding to the test problem name and a positive integer which selects the desired starting point of such problem. Table 6 summarizes all possible utilizations of the function `startingPoints`.

Table 5: String aliases for flag i.

String alias	**Corresponding value for `i`**
Init	0
F	1
JF	2
Ineq	3
JyIneq	4
JhIneq	5
JsIneq	6
Eq	7
JyEq	8
JhEq	9
JsEq	10
Clear	11

Table 6: Possible utilizations of the M-file `startingPoints.m`.

Call	**Description**
`startingPoints`	displays a list of all test problems available with a brief description of their starting points
`startingPoints(QVIname)`	returns the number of starting points available for the test problem `QVIname` (the input argument (`QVIname` must be a string of characters)
`startingPoints(QVIname,number)`	returns the `number`-th starting point for the test problem `QVIname` (the input argument `QVIname` must be a string of characters, while the input argument `number` must be an integer)

Finally the library also contains the M-file `solution.m` that can be used by the user to get one solution for each test problem. The function `solution` returns a solution for a test problem by accepting in input a string of characters corresponding to the test problem name. Note that, except for OutZ40, OutZ41 and OutZ43, all solutions are approximated.

6.3 Test Problems Description

In this subsection we report the test problems. The subsection is divided into three subparts. Subsubsection 6.3.1 contains pure academic problems, subsubsection 6.3.2 contains QVIs modelling some real problems, while subsubsection 6.3.3 contains discretization of infinite dimensional problems.

6.3.1 Academic Problems

OutZ40 [LBB/A/2-4-0-2-0]

source: [53, p. 13]

description:

$$\begin{aligned}
F(x) &:= \begin{pmatrix} 2 & 8/3 \\ 5/4 & 2 \end{pmatrix} x - \begin{pmatrix} 34 \\ 24.25 \end{pmatrix}, \\
g^I(y) &:= \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} y - \begin{pmatrix} 0 \\ 11 \\ 0 \\ 11 \end{pmatrix}, \\
g^P(y,x) &:= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} y + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} x - \begin{pmatrix} 15 \\ 15 \end{pmatrix}
\end{aligned}$$

JF: positive definite (everywhere)

comments: this problem was built so that it does not satisfy the assumptions for the local convergence of the Newton method discussed in [53] at the known solution listed below

known solution: $x^* = (5, 9)^T$

OutZ41 [LBB/A/2-4-0-2-0]

source: [53, Example 4.1]

description:

$$\begin{aligned}
F(x) &:= \begin{pmatrix} 2 & 8/3 \\ 5/4 & 2 \end{pmatrix} x - \begin{pmatrix} 100/3 \\ 22.5 \end{pmatrix}, \\
g^I(y) &:= \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} y - \begin{pmatrix} 0 \\ 11 \\ 0 \\ 11 \end{pmatrix}, \\
g^P(y,x) &:= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} y + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} x - \begin{pmatrix} 15 \\ 20 \end{pmatrix}
\end{aligned}$$

JF: positive definite (everywhere)

comments: a variant of the OutZ40 that satisfies the assumptions for the local convergence of the Newton method discussed in [53] at the known solution listed below

known solution: $x^* = (10, 5)^T$

OutZ42 [LBB/A/4-4-0-4-0]

source: [53, Example 4.2]

description:

$$\begin{aligned} F(x) &:= \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \\ g^I(y) &:= y, \\ g^P(y,x) &:= \begin{pmatrix} -y_1 - 2.5 + x_1 + x_1^2 \\ \vdots \\ -y_4 - 2.5 + x_4 + x_4^2 \end{pmatrix} \end{aligned}$$

JF: positive definite (everywhere)

known solution: $x^* \approx (-1.291, -1.5811, -1.5811, -1.291)^T$

OutZ43 [LAB/A/4-0-0-4-0]

source: [53, Example 4.3]

description:

$$\begin{aligned} F(x) &:= \text{same as for problem OutZ42}, \\ g^P(y,x) &:= -y - \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} x - \begin{pmatrix} 1.5 \\ 1.5 \\ 1.5 \\ 1.5 \end{pmatrix} \end{aligned}$$

JF: positive definite (everywhere)

comments: this problem satisfies conditions of Theorem 3.14

known solution: $x^* = (-0.9, -1.2, -1.2, -0.9)^T$

OutZ44 [LAB/A/4-0-0-4-0]

source: [53, Example 4.4]

description:

$$\begin{aligned} F(x) &:= \text{same as for problem OutZ42} \\ g^P(y,x) &:= -y - 1.5 \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} x - \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} \\ &\quad +0.25 \begin{pmatrix} (2x_1 - x_2 + 1)^2 \\ (-x_1 + 2x_2 - x_3 + 1)^2 \\ (-x_2 + 2x_3 - x_4 + 1)^2 \\ (-x_3 + 2x_4 + 1)^2 \end{pmatrix} \end{aligned}$$

JF: positive definite (everywhere)

known solution: $x^* \approx (-1.0021, -1.36, -1.36, -1.0021)^T$

Moving set problems

This is the most studied class of QVIs (see Section 1 and 3), namely problems where the set $\tilde{K}(x)$ is given by

$$\tilde{K}(x) := c(x) + Q$$

for some function $c : \mathbb{R}^n \to \mathbb{R}^n$ and a fixed set $Q \subseteq \mathbb{R}^n$. Assuming that Q has a representation of the form

$$Q \ := \ \{x \in \mathbb{R}^n \mid q(x) \leq 0\},$$

for some function $q : \mathbb{R}^n \to \mathbb{R}^{m_P}$, it is easy to see that $\tilde{K}(x)$ can be rewritten as

$$\tilde{K}(x) := \{y \in \mathbb{R}^n \mid q(y - c(x)) \leq 0\}$$

which corresponds to the general setting considered in this paper of a QVI with $g^P : \mathbb{R}^n \to \mathbb{R}^{m_P}$ being defined by

$$g^P(y, x) \ := \ q(y - c(x)). \tag{51}$$

MovSet1A - MovSet1B [LAO/A/5-0-0-1-0]

source: this thesis

description:

$$\begin{aligned} F(x) &:= Ax + b, \\ g^P(y, x) &:= \|y - \alpha x\|^2 - 0.5 \end{aligned}$$

with

$$A := \begin{pmatrix} 19.8699 & 0.5369 & 2.9482 & 0.3358 & 7.1239 \\ 4.1819 & 16.3484 & -5.2030 & 5.4332 & 2.7143 \\ -5.6554 & 0.9422 & 19.0981 & 7.1556 & -7.3810 \\ -1.8770 & 0.1918 & -5.3596 & 18.3565 & -7.8847 \\ -6.0303 & -3.6171 & -1.4658 & 4.6238 & 15.4085 \end{pmatrix}, \quad b := \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{pmatrix}, \tag{52}$$

and $\alpha := 0.1$ for MovSet1A and $\alpha := 10$ for MovSet1B

JF: positive definite (everywhere)

comments: referring to the general description (51) of g^P: $q(z) := \|z\|^2 - 0.5$, $c(x) := \alpha x$. Note that MovSet1A satisfies conditions of Theorem 3.4, while MovSet1B does not

MovSet2A - MovSet2B [LAO/A/5-0-0-1-0]

source: this thesis

description:

$$\begin{aligned} F(x) &:= Ax+b, \\ g^P(y,x) &:= \|y-\alpha\big(\cos(x_i)\big)_{i=1}^n\|^2 - 0.5 \end{aligned}$$

with A and b as in (52) and $\alpha := 0.1$ for MovSet2A and $\alpha := 10$ for MovSet2B

JF: positive definite (everywhere)

comments: referring to the general description (51) of g^P: $q(z) := \|z\|^2 - 0.5$, $c(x) := \alpha\big(\cos(x_i)\big)_{i=1}^n$. Note that MovSet2A satisfies conditions of Theorem 3.4, while MovSet2B does not

MovSet3A1 - MovSet3B1 [LAO/A/1000-0-0-1-0]

MovSet3A2 - MovSet3B2 [LAO/A/2000-0-0-1-0]

source: this thesis

description:

$$\begin{aligned} F(x) &:= Ax+b, \\ g^P(y,x) &:= (y-Mx)^T Q(y-Mx) - d \end{aligned}$$

where A, b, M, Q and d are available in the corresponding Matlab functions (MovSet3A1 and MovSet3A2 differ from MovSet3B1 and MovSet3B2, respectively, only in the matrix M)

JF: positive definite (everywhere)

comments: referring to the general description (51) of g^P: $q(z) := z^T Qz - d$, $c(x) := Mx$. Note that MovSet3A1 and MovSet3A2 satisfy conditions of Theorem 3.4, while MovSet3B1 and MovSet3B2 do not

MovSet4A1 - MovSet4B1 [LAO/A/400-0-0-801-0]

MovSet4A2 - MovSet4B2 [LAO/A/800-0-0-1601-0]

source: this thesis

description:

$$\begin{aligned} F(x) &:= Ax+b, \\ g^P(y,x) &:= \begin{pmatrix} -y+Mx \\ y-Mx-1 \\ \mathbf{1}_n^T y - \mathbf{1}_n^T Mx - \frac{n}{2} \end{pmatrix} \end{aligned}$$

where A, b and M are available in the corresponding Matlab functions (MovSet4A1 and MovSet4A2 differ from MovSet4B1 and MovSet4B2, respectively, only in the matrix M)

JF: positive definite (everywhere)

comments: referring to the general description (51) of g^P:

$$q(z) := \begin{pmatrix} -z \\ z-1 \\ \mathbf{1}_n^T z - \frac{n}{2} \end{pmatrix}, \quad c(x) := Mx.$$

Note that MovSet4A1 and MovSet4A2 satisfy conditions of Theorem 3.4, while MovSet4B1 and MovSet4B2 do not

Problems with box constraints

This class of QVIs have a set $\tilde{K}(x)$ defined by constraints of the form

$$g^P(y,x) := \begin{pmatrix} (y_i - s_i x_i - c_i)_{i=1}^n \\ (-y_i + t_i x_i - d_i)_{i=1}^n \end{pmatrix}. \tag{53}$$

We call this a QVI with box constraints since, given a fixed vector x, the feasible set describes box constraints for the variables y (see Section 3). The particular values of the box constraints for a variable y_i, however, varies with x_i.

Box1A - Box1B [LAB/A/5-0-0-10-0]

source: this thesis

description:

$$\begin{aligned} F(x) &:= Ax + b, \\ g^P(y,x) &:= \begin{pmatrix} (y_i - \alpha x_i - c_i)_{i=1}^n \\ (-y_i + \alpha x_i - c_i)_{i=1}^n \end{pmatrix} \end{aligned}$$

where A and b are defined as in (52),

$$c := \begin{pmatrix} 0.1202 \\ 1.7418 \\ 2.7064 \\ 2.0502 \\ 4.4616 \end{pmatrix}$$

and $\alpha := 0.1$ for Box1A and $\alpha := 2$ for Box1B

JF: positive definite (everywhere)

comments: Box1A satisfies conditions of Corollary 3.17, while Box1B does not

Box2A - Box2B [NAB/A/500-0-0-1000-0]

source: this thesis

description:

$$F(x) \;:=\; Ax + b + (\exp(x_i))_{i=1}^{500},$$

g^P is defined as in (53), where A, b, s, t, c and d are available in the corresponding Matlab functions (Box2A differs from Box2B only in the vectors s and t)

JF: positive definite (everywhere)

comments: Box2A satisfies conditions of Corollary 3.17, while Box2B does not

Box3A - Box3B [NAB/A/500-0-0-1000-0]

source: this thesis

description: these problems are identical to Box2 except the function F:

$$F(x) \;:=\; Ax + b + M\left(x_i^3\right)_{i=1}^{500},$$

A, b, M, s, t, c and d are available in the corresponding Matlab functions (Box3A differs from Box3B only in the vectors s and t)

JF: positive definite (everywhere)

comments: Box3A satisfies conditions of Corollary 3.17, while Box3B does not

Problems with bilinear constraints

In these problems, the set $\tilde{K}(x)$ is defined by the following inequality constraints only (see Section 3):

$$g^P(y,x) := \begin{pmatrix} x^T Q_1 y - c_1 \\ \vdots \\ x^T Q_q y - c_q \end{pmatrix}.$$

BiLin1A - BiLin1B [LBO/A/5-10-0-3-0]

source: this thesis

description:

$$\begin{aligned} F(x) &:= Ax + b, \\ g^I(y) &:= \begin{pmatrix} l - y \\ y - u \end{pmatrix}, \\ g^P(y,x) &:= \begin{pmatrix} x^T Q_1 y - c_1 \\ \vdots \\ x^T Q_3 y - c_3 \end{pmatrix}, \end{aligned}$$

where A and b are defined as in (52),

$$l := \begin{pmatrix} -0.1202 \\ -1.7418 \\ -2.7064 \\ -2.0502 \\ -4.4616 \end{pmatrix}, \quad u := -l, \quad c := \begin{pmatrix} 0.3070 \\ 1.1186 \\ 2.6149 \end{pmatrix},$$

$$Q_1 := \begin{pmatrix} 1.9073 & 0.2403 & 0.2352 & -0.4903 & -0.2651 \\ 0.2403 & 1.1319 & 1.2087 & -0.3268 & 0.2540 \\ 0.2352 & 1.2087 & 1.6862 & 0.2941 & 0.6732 \\ -0.4903 & -0.3268 & 0.2941 & 1.8258 & 0.1363 \\ -0.2651 & 0.2540 & 0.6732 & 0.1363 & 1.5527 \end{pmatrix} + \alpha \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

$$Q_2 := \begin{pmatrix} 2.7307 & 0.5988 & 1.5728 & 1.4072 & -0.3082 \\ 0.5988 & 2.2435 & 0.7546 & 1.3632 & 1.5852 \\ 1.5728 & 0.7546 & 2.3809 & 1.2625 & 1.0403 \\ 1.4072 & 1.3632 & 1.2625 & 1.7612 & 0.3071 \\ -0.3082 & 1.5852 & 1.0403 & 0.3071 & 2.6305 \end{pmatrix} + \alpha \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

$$Q_3 := \begin{pmatrix} 2.5189 & 2.1947 & 1.7697 & 2.2753 & 1.9885 \\ 2.1947 & 3.8143 & 1.3839 & 1.5636 & 1.8451 \\ 1.7697 & 1.3839 & 3.3655 & 1.6441 & 1.9946 \\ 2.2753 & 1.5636 & 1.6441 & 3.6885 & 2.3272 \\ 1.9885 & 1.8451 & 1.9946 & 2.3272 & 2.2883 \end{pmatrix} + \alpha \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

and $\alpha := 0$ for BiLin1A and $\alpha := 10$ for BiLin1B

JF: positive definite (everywhere)

comments: BiLin1A satisfies conditions of Corollary 3.19, while BiLin1B does not

Problems with linear constraints with variable right-hand side

In this class of problems, the feasible set $\tilde{K}(x)$ is defined by

$$g^P(y, x) := Ey - d + c(x),$$

where $E \in \mathbb{R}^{m \times n}$ is a given matrix, $c : \mathbb{R}^n \to \mathbb{R}^{m_P}$ and $d \in \mathbb{R}^{m_P}$. In this class of QVIs, the feasible set is defined by linear inequalities in which the right-hand side depends on x (see Section 3).

RHS1A1 - RHS1B1 [LAL/A/200-0-0-199-0]

source: this thesis

description:

$$\begin{aligned} F(x) &:= Ax + b, \\ g^P(y, x) &:= Ey - d + C\,(\sin(x_i))_{i=1}^n \end{aligned}$$

where A, b, E, d and C are available in the corresponding Matlab functions (RHS1A1 differs from RHS1B1 only in the matrix C)

JF: positive definite (everywhere)

comments: RHS1A1 satisfies conditions of Theorem 3.9, while RHS1B1 does not

RHS2A1 - RHS2B1 [LAL/A/200-0-0-199-0]

source: this thesis

description:

$$F(x) := Ax + b,$$
$$g^P(y,x) := Ey - d + Cx$$

where A, b, E, d and C are available in the corresponding Matlab functions (RHS2A1 differs from RHS2B1 only in the matrix C)

JF: positive definite (everywhere)

comments: RHS2A1 satisfies conditions of Theorem 3.9, while RHS2B1 does not

6.3.2 Problems from Real-World Models

Walrasian equilibrium problems

Problems described here are QVI reformulations of a Walrasian pure exchange economy with utility function without production whose general structure is described [15]; the specific data used here are taken from [19]. In this model there are C agents, whose preferences are given by a utility function u_i, exchanging P goods. Each agent controls a variable $x^i \in \mathbb{R}^P$ (representing quantity of goods) and has an initial endowment of $\xi^i \in \mathbb{R}^P$. There is also one extra, 0-th "player" controlling a vector $x^0 \in \mathbb{R}^P$ representing prices. Therefore, the vector of variables is $x = (x^i)_{i=0}^C \in \mathbb{R}^{(C+1)P}$. The dimensions and the description of the QVI model depend on the parameters C and P:

$$n := P(C+1), \quad m_I := P(C+1), \quad p_I := 1, \quad m_P := C, \quad p_P := 0,$$

$$F(x) := \begin{pmatrix} \sum_{i=1}^C \xi^i - x^i \\ \nabla_{x^1} u_1(x^1) \\ \vdots \\ \nabla_{x^C} u_C(x^C) \end{pmatrix}, \quad g^I(y) := -y, \quad M^I := (\mathbf{1}_P^T \quad \mathbf{0}_{PC}^T), \tag{54}$$

$$v^I := -1, \quad g^P(y,x) := \begin{pmatrix} \sum_{j=1}^P x_j^0 (y_j^1 - \xi_j^1) \\ \vdots \\ \sum_{j=1}^P x_j^0 (y_j^C - \xi_j^C) \end{pmatrix}.$$

The utility functions u of the agents, as well as the parameters C and P and the endowment ξ, are specified for each test problem.

WalEq1 [LBO/R/18-18-1-5-0]

source: model from [15], data from [19]

description: the general description is (54) where $C := 5$, $P := 3$, the utility functions are quadratic and convex:

$$u_i(x^i) \;:=\; \frac{1}{2}(x^i)^T Q^i x^i - (b^i)^T x^i, \quad i = 1, \ldots, 5,$$

$$Q^i \;:=\; \begin{pmatrix} 6 & -2 & 5 \\ -2 & 6 & -7 \\ 5 & -7 & 20 \end{pmatrix}, \quad b^i := \begin{pmatrix} 32+i \\ 32+i \\ 32+i \end{pmatrix}, \quad i = 1, 2,$$

$$Q^i \;:=\; \begin{pmatrix} 6 & 1 & 0 \\ 1 & 7 & -5 \\ 0 & -5 & 7 \end{pmatrix}, \quad b^i := \begin{pmatrix} 30+(i+2)*2 \\ 30+(i+2)*2 \\ 30+(i+2)*2 \end{pmatrix}, \quad i = 3, 4, 5,$$

and

$$\xi^i \;:=\; \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \quad i = 1, 2, \qquad \xi^i := \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix}, \quad i = 3, 4, 5$$

JF: P_0 (everywhere) but never positive semidefinite

WalEq2 [NBO/R/105-105-1-20-0]

source: model from [15], data from [19]

description: the general description is (54) where $C := 20$, $P := 5$, the utility functions are of logarithmic type:

$$u_i(x^i) \;:=\; -\sum_{k=1}^{5}(a_k + i + 4)\log(x^i_k + b_k + 2(i+4)), \qquad i = 1, \ldots, 10,$$

$$u_i(x^i) \;:=\; -\sum_{k=1}^{5}(c_k + i + 4)\log(x^i_k + d_k + i + 4), \qquad i = 11, \ldots, 20,$$

$$a \;:=\; \begin{pmatrix} 1 \\ 2 \\ 4 \\ 6 \\ 8 \end{pmatrix}, \quad b := \begin{pmatrix} 20 \\ 30 \\ 30 \\ 40 \\ 50 \end{pmatrix}, \quad c := \begin{pmatrix} 10 \\ 6 \\ 4 \\ 10 \\ 1 \end{pmatrix}, \quad d := \begin{pmatrix} 50 \\ 40 \\ 30 \\ 20 \\ 20 \end{pmatrix},$$

and

$$\xi^i \;:=\; \begin{pmatrix} 2 \\ 3 \\ 4 \\ 1 \\ 6 \end{pmatrix}, \quad i = 1, \ldots, 10, \qquad \xi^i := \begin{pmatrix} 6 \\ 5 \\ 4 \\ 3 \\ 2 \end{pmatrix}, \quad i = 11, \ldots, 20$$

JF: P_0 (everywhere) but never positive semidefinite

WalEq3 [LBO/R/186-186-1-30-0]

source: model from [15], data from [19]

description: the general description is (54) where $C := 30$, $P := 6$, the utility functions are quadratic and convex:

$$u_i(x^i) \;:=\; \frac{1}{2}(x^i)^T Q^i x^i - (b^i)^T x^i, \quad i \;=\; 1, \ldots, 30,$$

$$Q^i \;:=\; A,\; b^i := \begin{pmatrix} 56+i \\ 66+i \\ 76+i \\ 66+i \\ 66+i \\ 56+i \end{pmatrix},\; i = 1, \ldots, 15,$$

$$Q^i \;:=\; B, \quad b^i := \begin{pmatrix} 50+2*(i+6) \\ 60+2*(i+6) \\ 50+2*(i+6) \\ 70+2*(i+6) \\ 70+2*(i+6) \\ 60+2*(i+6) \end{pmatrix},\; i = 16, \ldots, 30,$$

and

$$\xi^i \;:=\; \begin{pmatrix} 2 \\ 3 \\ 4 \\ 1 \\ 6 \\ 1 \end{pmatrix}, \quad i = 1, \ldots, 15, \qquad\qquad \xi^i := \begin{pmatrix} 6 \\ 5 \\ 4 \\ 3 \\ 2 \\ 8 \end{pmatrix}, \quad i = 16, \ldots, 30.$$

Set A equal to

68.22249416536778	12.12481199690621	-8.35496210217478	-6.81177486915109	-4.66752803051747	3.64100170417482
12.12481199690621	53.51450780426463	-21.77618227261339	-15.00376305863444	-0.11788350473544	2.03354709400720
-8.35496210217478	-21.77618227261339	35.44033408387684	4.35160649036518	19.17472558234163	-3.40090742729160
-6.81177486915109	-15.00376305863444	4.35160649036518	52.25155022199242	-5.99490328518247	20.40443259092577
-4.66752803051747	-0.11788350473544	19.17472558234163	-5.99490328518247	23.32798561358070	-3.58535668529727
3.64100170417482	2.03354709400720	-3.40090742729160	20.40443259092577	-3.58535668529727	10.21258119890765

and B equal to

61.74633559943146	-23.83006225091380	16.78581949473039	14.42073900860500	-2.75188745616575	13.44307656650567
-23.83006225091380	37.64246654306209	-3.76510322128227	16.32022449045404	-39.90743633716275	11.38657250296817
16.78581949473039	-3.76510322128227	53.34843665848310	4.60388415537161	-23.04611587657949	-25.31392346426841
14.42073900860500	16.32022449045404	4.60388415537161	40.69699687713468	-30.78019133996427	17.08866411420883
-2.75188745616575	-39.90743633716275	-23.04611587657949	-30.78019133996427	66.22678445157413	-12.28091080313848
13.44307656650567	11.38657250296817	-25.31392346426841	17.08866411420883	-12.28091080313848	41.37849544246254

JF: P_0 (everywhere) but never positive semidefinite

WalEq4 [NBO/R/310-310-1-30-0]

source: model from [15], data from [19]

description: the general description is (54) where $C := 30$, $P := 10$, the utility functions are of logarithmic type:

$$u_i(x^i) \;:=\; -\sum_{k=1}^{10}(a_k + i + 6)\log(x_k^i + b_k + 2(i+6)), \qquad i \;=\; 1, \ldots, 15,$$

$$u_i(x^i) \;:=\; -\sum_{k=1}^{10}(c_k + i + 6)\log(x_k^i + d_k + i + 6), \qquad i \;=\; 16, \ldots, 30,$$

$$a := \begin{pmatrix} 1\\2\\4\\6\\8\\7\\8\\10\\1\\5 \end{pmatrix}, \quad b := \begin{pmatrix} 50\\60\\70\\60\\60\\50\\50\\80\\60\\70 \end{pmatrix}, \quad c := \begin{pmatrix} 10\\6\\4\\10\\1\\2\\6\\4\\9\\4 \end{pmatrix}, \quad d := \begin{pmatrix} 50\\60\\50\\70\\70\\60\\50\\50\\80\\50 \end{pmatrix},$$

and

$$\xi^i := \begin{pmatrix} 2\\3\\4\\1\\6\\1\\3\\6\\2\\10 \end{pmatrix}, \quad i = 1,\ldots,15, \qquad \xi^i := \begin{pmatrix} 6\\5\\4\\3\\2\\8\\4\\6\\2\\0 \end{pmatrix}, \quad i = 16,\ldots,30$$

JF: P_0 (everywhere) but never positive semidefinite

WalEq5 [NBO/R/492-492-1-40-0]

source: model from [15], data from [19]

description: the general description is (54) where $C := 40$, $P := 12$, the utility functions are of logarithmic type:

$$u_i(x^i) := -\sum_{k=1}^{12}(a_k + i + 7)\log(x_k^i + b_k + 2(i+7)), \qquad i = 1,\ldots,20,$$

$$u_i(x^i) := -\sum_{k=1}^{12}(c_k + i + 7)\log(x_k^i + d_k + i + 7), \qquad i = 21,\ldots,40,$$

$$a := \begin{pmatrix} 1\\2\\4\\6\\8\\7\\8\\10\\1\\5\\2\\4 \end{pmatrix}, \quad b := \begin{pmatrix} 50\\60\\70\\60\\60\\50\\50\\80\\60\\70\\70\\80 \end{pmatrix}, \quad c := \begin{pmatrix} 10\\6\\4\\10\\1\\2\\6\\4\\9\\4\\5\\1 \end{pmatrix}, \quad d := \begin{pmatrix} 50\\60\\50\\70\\70\\60\\50\\50\\80\\50\\60\\70 \end{pmatrix},$$

and

$$\xi^i := \begin{pmatrix} 2\\3\\4\\1\\6\\1\\3\\6\\2\\10\\3\\4 \end{pmatrix}, \quad i = 1,\ldots,20, \qquad \xi^i := \begin{pmatrix} 6\\5\\4\\3\\2\\8\\4\\6\\2\\0\\6\\0 \end{pmatrix}, \quad i = 21,\ldots,40$$

JF: P_0 (everywhere) but never positive semidefinite

Wal2 [NLO/A/105-107-0-20-0]

source: model from [15], data from [19]

description: the general description is (54), except for

$$m_I \;:=\; P(C+1)+2, \quad p_I := 0,$$

$$g^I(y) \;:=\; \begin{pmatrix} -y \\ \sum_{i=0}^{P} y_i^0 - 1 \\ 1 - \sum_{i=0}^{P} y_i^0 \end{pmatrix},$$

where $C := 20$, $P := 5$ and the utility functions are of logarithmic type:

$$u_i(x^i) \;:=\; \sum_{k=1}^{5} (a_k + k + 4)\log(x_k^i + b_k + 2(i+4)), \qquad i \;=\; 1,\dots,10,$$

$$u_i(x^i) \;:=\; \sum_{k=1}^{5} (c_k + k + 4)\log(x_k^i + d_k + i + 4), \qquad i \;=\; 11,\dots,20,$$

a, b, c, d and ξ are the same as for WalEq2

JF: never P_0

comments: this QVI arises from an implementation mistake, in fact it differs from WalEq2 essentially only in the sign and in one parameter of the u_i functions; furthermore the equality constraint $\sum_{i=1}^{P} y_i^0 = 1$ is substituted by a double inequality. Since the problem proved challenging, we kept it in the library.

Wal3 [LLO/R/186-188-0-30-0]

source: model from [15], data from [19]

description: the general description is (54), except for

$$m_I \;:=\; P(C+1)+2, \quad p_I := 0,$$

$$g^I(y) \;:=\; \begin{pmatrix} -y \\ \sum_{i=0}^{P} y_i^0 - 1 \\ 1 - \sum_{i=0}^{P} y_i^0 \end{pmatrix},$$

where all parameters and functions are the same as for WalEq3

JF: P_0 (everywhere) but never positive semidefinite

comments: this QVI differs from WalEq3 only for the fact that the equality constraint $\sum_{i=1}^{P} y_i^0 = 1$ is substituted by a double inequality.

Wal5 [NLO/A/492-494-0-40-0]

source: model from [15], data from [19]

description: the general description is (54), except for

$$m_I \;:=\; P(C+1)+2, \quad p_I := 0,$$

$$g^I(y) \;:=\; \begin{pmatrix} -y \\ \sum_{i=0}^{P} y_i^0 - 1 \\ 1 - \sum_{i=0}^{P} y_i^0 \end{pmatrix},$$

where $C := 40$, $P := 12$, the utility functions are of logarithmic type:

$$u_i(x^i) \;:=\; \sum_{k=1}^{12} (a_k + k + 7) \log(x_k^i + b_k + 2(i+7)), \qquad i \;=\; 1, \dots, 20,$$

$$u_i(x^i) \;:=\; \sum_{k=1}^{12} (c_k + k + 7) \log(x_k^i + d_k + i + 7), \qquad i \;=\; 21, \dots, 40,$$

a, b, c, d and ξ are the same as for WalEq5

JF: never P_0

comments: this QVI arises from an implementation mistake, in fact it differs from WalEq2 essentially only in the sign and in one parameter of the u_i functions; furthermore the equality constraint $\sum_{i=1}^{P} y_i^0 = 1$ is substituted by a double inequality. Since the problem proved challenging, we kept it in the library.

Generalized Nash equilibrium problems

It is well known that finding an equilibrium of a generalized Nash game is equivalent to solving a QVI problem, see [20]. This QVI model of an energy market Nash equilibrium is taken from [41]. Let N agents owns l plants each one to generate electric energy for sale. We denote as x_j^i the energy produced by agent i in the j-th plant. The unitary energy price in the market depends on the total amount of energy produced by all agents, it is modelled by a quadratic concave function of one variable. Then the profit of each agent depends on the generation level of the other agents in the market. In turn, each generation level is constrained by technological limitations of the power plants. The coordination, or regulation, of the market is done by the Independent System Operator (ISO), whose actions in the market are considered as those of an additional player. Accordingly, letting the ISO be player number 0, the ISO tries to maximize the social welfare by encouraging all other agents to satisfy the total market demand.

$$n := Nl+1, \quad m_I := 2(Nl+1), \quad p_P := N+1,$$

$$F(x) := \begin{bmatrix} 0 \\ c\left(\sum_{i=1}^N \sum_{j=1}^l x_j^i\right)^2 - 120 + 2c\left(\sum_{i=1}^N \sum_{j=1}^l x_j^i\right)\left(\sum_{j=1}^l x_j^1\right) \\ \vdots \\ c\left(\sum_{i=1}^N \sum_{j=1}^l x_j^i\right)^2 - 120 + 2c\left(\sum_{i=1}^N \sum_{j=1}^l x_j^i\right)\left(\sum_{j=1}^l x_j^1\right) \\ \vdots \\ c\left(\sum_{i=1}^N \sum_{j=1}^l x_j^i\right)^2 - 120 + 2c\left(\sum_{i=1}^N \sum_{j=1}^l x_j^i\right)\left(\sum_{j=1}^l x_j^N\right) \\ \vdots \\ c\left(\sum_{i=1}^N \sum_{j=1}^l x_j^i\right)^2 - 120 + 2c\left(\sum_{i=1}^N \sum_{j=1}^l x_j^i\right)\left(\sum_{j=1}^l x_j^N\right) \end{bmatrix} + \begin{bmatrix} 0 \\ A_1^1 x_1^1 \\ \vdots \\ A_l^1 x_l^1 \\ \vdots \\ A_1^N x_1^N \\ \vdots \\ A_l^N x_l^N \end{bmatrix} + \begin{bmatrix} 120 \\ b_1^1 \\ \vdots \\ b_l^1 \\ \vdots \\ b_1^N \\ \vdots \\ b_l^N \end{bmatrix}, \tag{55}$$

$$g^I(y) := \begin{bmatrix} -y^0 \\ -y^1 \\ \vdots \\ -y^N \\ y^0 - u^0 \\ y^1 - u^1 \\ \vdots \\ y^N - u^N \end{bmatrix}, \qquad M^P(x) := \begin{pmatrix} 1 & & & \\ & \mathbf{1}_l^T & & \\ & & \ddots & \\ & & & \mathbf{1}_l^T \end{pmatrix},$$

$$v^P(x) := \begin{pmatrix} 0 & \mathbf{1}_l^T & & \mathbf{1}_l^T \\ \mathbf{1}_l^T & \mathbf{0}_l^T & & \mathbf{1}_l^T \\ & & \ddots & \\ \mathbf{1}_l^T & \mathbf{1}_l^T & & \mathbf{0}_l^T \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ \vdots \\ x^N \end{pmatrix} - \begin{pmatrix} d \\ d \\ \vdots \\ d \end{pmatrix},$$

where $c \in \mathbb{R}, A \in \mathbb{R}^{Nl}, b^{Nl}, u \in \mathbb{R}^{Nl}$ and $d \in \mathbb{R}$.

LunSS1 [NBA/R/501-1002-0-0-6]

LunSS2 [NBA/R/1251-2502-0-0-6]

LunSS3 [NBA/R/5001-10002-0-0-6]

source: [41]

description: these problems are described by (55), where $N := 5$ and $l := 100$ for LunSS1, $l := 250$ for LunSS2, $l := 1000$ for LunSS3. c, A, b, u and d are available for all these problems in the corresponding Matlab functions.

It is possible to compute some equilibria (not all in general) of a jointly convex Nash problem by solving a variational inequality instead. These points are called variational equilibria and have some properties from the economic point of view, see [20]. The next 3 problems pursue this goal. Note that those problems are pure VIs in which the parametric set $\tilde{K}(x)$ vanishes and $K(x) = \bar{K}$.

LunSSVI1 [NBA/R/501-1002-1-0-0]

LunSSVI2 [NBA/R/1251-2502-1-0-0]

LunSSVI3 [NBA/R/5001-10002-1-0-0]

source: [41]

description: F and g^I are taken from (55) while $M^I := \mathbf{1}_n^T$, and $v^I := -d$, where $N := 5$ and $l := 100$ for LunSSVI1, $l := 250$ for LunSSVI2, $l := 1000$ for LunSSVI3. c, A, b, u and d are available for all these problems in the corresponding Matlab functions. Note that $p_P := 0$.

comments: these problems are pure VIs

6.3.3 Discretized Problems

Here we consider finite dimensional QVIs obtained by making a discretization procedure on infinite dimensional QVIs. This series of problems stemmed from different fields.

Transportation problems

In the modeling of competition on networks in [60] it is assumed that users either behave following the Wardrop equilibrium or the Nash equilibrium concept. In the time-dependent network model shared by two types of users: group users (Nash players) and individual users (Wardrop players), both classes of users choose the paths to ship their jobs so as to minimize their costs, but they apply different optimization criteria. The source of interaction of users is represented by the travel demand, which is assumed to be elastic with respect to the equilibrium solution. Thus, the equilibrium distribution is proved to be equivalent to the solution of an appropriate time-dependent quasi-variational inequality problem. This example taken from [60] is relative to a simple network with 4 nodes and 7 edges, in which there are two users: one Nash user and one Wardrop user. The time interval considered is $[0, N]$, and in particular it is discretized so that the time instants are $0, \ldots, N$. Then the solution of the following QVI contains the flows on the paths of the network at the equilibrium in the instants $1, \ldots, N$ for the two users. The dimensions and the description of the following two discretized models depend on the parameter N:

$$n := 4N, \quad m^I := 4N, \quad p^P := 2N,$$

$$
\begin{aligned}
F(x) &:= \begin{pmatrix} A & & 0 \\ & \ddots & \\ 0 & & A \end{pmatrix} x + \begin{pmatrix} b \\ \vdots \\ b \end{pmatrix}, \\
g^I(y) &:= -y, \quad M^P(x) := \begin{pmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{pmatrix}, \\
v^P(x) &:= \begin{pmatrix} -E & & 0 \\ & \ddots & \\ 0 & & -E \end{pmatrix} x + \begin{pmatrix} -d_1 - d_2(1-1)/(N-1) \\ \vdots \\ -d_1 - d_2(N-1)/(N-1) \end{pmatrix},
\end{aligned} \tag{56}
$$

where

$$
A := \begin{pmatrix} 4 & 2 & 0 & 0 \\ 2 & 10 & 0 & 4 \\ 0 & 0 & 2 & 1 \\ 0 & 2 & 1 & 5 \end{pmatrix}, \quad b := \begin{pmatrix} 40 \\ 30 \\ 40 \\ 30 \end{pmatrix}, \quad C := \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix},
$$

$$
E := \begin{pmatrix} 2/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \end{pmatrix}, \quad d1 := \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad d2 := \begin{pmatrix} 10 \\ 4 \end{pmatrix}
$$

Scrim11 [LBA/RD/2400-2400-0-0-1200]

Scrim12 [LBA/RD/4800-4800-0-0-2400]

source: [60]

description: the general description is (56) with $N := 600$ for Scrim11 and $N := 1200$ for Scrim12

JF: positive definite (everywhere)

Scrim21 [LBL/RD/2400-2400-0-2400-0]

Scrim22 [LBL/RD/4800-4800-0-4800-0]

source: [60]

description: the general description is (56), but here

$$
\begin{aligned}
m^P &:= 4N, \quad p^P := 0, \\
g^P(y,x) &:= \begin{pmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \\ -C & & 0 \\ & \ddots & \\ 0 & & -C \end{pmatrix} y + \begin{pmatrix} -E & & 0 \\ & \ddots & \\ 0 & & -E \\ E & & 0 \\ & \ddots & \\ 0 & & E \end{pmatrix} x + \begin{pmatrix} -d_1 - d_2(1-1)/(N-1) \\ \vdots \\ -d_1 - d_2(N-1)/(N-1) \\ d_1 + d_2(1-1)/(N-1) \\ \vdots \\ d_1 + d_2(N-1)/(N-1) \end{pmatrix},
\end{aligned}
$$

with $N := 600$ for Scrim21 and $N := 1200$ for Scrim22

JF: positive definite (everywhere)

comments: Problems Scrim21 and Scrim22 are essentially the same as problems Scrim11 and Scrim22, respectively, except that each equality constraint has been rewritten as two inequalities. In particular, the standard linear independence constraint qualification is therefore violated for problems Scrim21 and Scrim22.

Contact problems with Coulomb friction

This is the problem of an elastic body in contact to a rigid obstacle. In particular, this is the most realistic model in which Coulomb friction is present (in this problem $\phi \in \mathbb{R}$ is the friction coefficient). The problem is taken from Example 11.1 in [52]. Let $x^* \in \mathbb{R}^n$ be a solution of the QVI, then odd elements of x^* are interpreted as tangential stress components on the rigid obstacle in different points of such obstacle, while even elements are interpreted as outer normal ones. We consider different instances of this problem which derive from different discretizations generating different fragmentations of the obstacle in identical segments. In particular, the case in which the obstacle is divided into N segments involves $2(N+1)$ variables in the model (since there are $N+1$ extreme segment points and having to consider both tangential and outer normal stress components for all of them). The dimensions and the description of the following two discretized models depend on the parameter N:

$$
\begin{aligned}
n &:= 2(N+1), \quad m^I := 2(N+1), \quad m^P := 2(N+1), \\
F(x) &:= Ax - b, \quad g^I(y) := \begin{bmatrix} (-y_{2i} - 10)_{i=1}^{N+1} \\ (y_{2i})_{i=1}^{N+1} \end{bmatrix}, \qquad (57) \\
g^P(y,x) &:= \begin{bmatrix} \left(-y_{(2i-1)} + \phi x_{2i}\right)_{i=1}^{N+1} \\ \left(y_{(2i-1)} + \phi x_{2i}\right)_{i=1}^{N+1} \end{bmatrix},
\end{aligned}
$$

where the positive definite square matrix A and the vector b depend on N and are available in the library for $N := 30$ and $N := 40$ (data for these problems have been kindly provided by J.V. Outrata, M. Kočvara and J. Zowe).

OutKZ31 [LBB/RD/62-62-0-62-0]

OutKZ41 [LBB/RD/82-82-0-82-0]

source: [52]

description: the general description is (57) with friction coefficient $\phi := 10$ and the fragmentation granularity $N := 30$ for OutKZ31 and $N := 40$ for OutKZ41

JF: positive definite (everywhere)

QVIs with gradient constraints

The problems considered here are taken from [40] (see also [34]) and represent a stationary model for the magnetization of type-II superconductors.

Specifically, let $\Omega \subseteq \mathbb{R}^2$ be an open and convex domain, let $W := W^{1,2}(\Omega)$ be the corresponding Sobolev space, and let j_c be a nonnegative continuous function. Then the infinite-dimensional problem from [40] (using $p = 2$) is to find a solution $u \in K(u)$ satisfying

$$\int_\Omega \nabla u(\xi)^T \nabla(v-u)(\xi) d\xi \geq 0 \quad \forall v \in K(u), \tag{58}$$

where the feasible set $K(u)$ is defined by

$$K(u) := \left\{ v \in W \mid v\big|_{\partial\Omega} = u_1,\ \|\nabla v(\xi)\| \leq j_c\big(|u(\xi)|\big) \text{ a.e. in } \Omega \right\} \tag{59}$$

for a given function u_1.

In our realizations of this problem, we always take $\Omega = (0,1) \times (0,1)$ and $j_c(t) := t$. To discretize the problem, we choose a number $N \in \mathbb{N}$, a stepsize $h := \frac{1}{N+1}$, and the discrete points

$$\xi_i^{(1)} := ih = \frac{i}{N+1}, \quad \xi_j^{(2)} := jh = \frac{j}{N+1} \quad \forall\, i,j = 0,1,\ldots,N+1.$$

Furthermore, let

$$u_{i,j} := u(\xi_i^{(1)}, \xi_j^{(2)}), \quad v_{i,j} := v(\xi_i^{(1)}, \xi_j^{(2)}) \quad \forall\, i,j = 0,1,\ldots,N+1$$

and note that the values of $u_{i,j}, v_{i,j}$ are known for $i,j \in \{0, N+1\}$ due to the given boundary condition. Therefore, the discrete unknowns are the components $u_{i,j}$ for $i,j \in \{1,\ldots,N\}$. We approximate the partial gradients of u and v by using forward finite differences. Moreover, an integral of the form $\int_\Omega f(\xi)d\xi$ for a suitable function f is approximated by a piecewise constant function in such a way that we get

$$\int_\Omega f(\xi)d\xi \approx h^2 \sum_{i,j=0}^{N} f_{i,j},$$

where $f_{i,j} := f\big(\xi_i^{(1)}, \xi_j^{(2)}\big)$. We then reorder the unknows $u_{i,j}$ and define

$$\text{vec}(u) := (u_{1,1}, u_{2,1}, \ldots, u_{N,1}, u_{1,2}, u_{2,2}, \ldots, u_{N,2}, \ldots, u_{N,N})^T \in \mathbb{R}^n, \quad n := N^2.$$

In a similar way, we also define $\text{vec}(v)$. To get back to our standard notation, we finally set

$$x := \text{vec}(u) \quad \text{and} \quad y := \text{vec}(v).$$

Altogether, this results in a QVI with a linear function F of the form

$$F(x) := A^T(Ax + a) + C^T(Cx + c) \tag{60}$$

for certain matrices $A, C \in \mathbb{R}^{(n+N)\times n}$, and vectors $a, c \in \mathbb{R}^{n+N}$ (these vectors take into account the boundary conditions). Furthermore, the constraints take the form

$$g_k^P(y, x) := (A \cdot y + a)^2_{k+1+\lfloor (k-1)/N \rfloor} + (C \cdot y + c)^2_{N+k} - h^2 x_k^2 \tag{61}$$

for all $k = 1, \ldots n$, where $\lfloor \cdot \rfloor$ denotes the floor-function. The precise data of A, C, a, c are given in the corresponding Matlab files. Different instances of the discretized problems arise from different choices of the discretization parameter $N \in \mathbb{N}$ and the boundary function u_1.

KunR11 - KunR21 - KunR31 [LAO/RD/2500-0-0-2500-0]

KunR12 - KunR22 - KunR32 [LAO/RD/4900-0-0-4900-0]

source: [40]

description: These problems arise from the general description with F and g^P described in (60) and (61). We took $N = 50$ for problems KunR11, KunR21 and KunR31, and $N = 70$ for KunR12, KunR22 and KunR32. The boundary function is $u_1(\xi^{(1)}, \xi^{(2)}) := 1 + \xi^{(1)} + \xi^{(2)}$ for problems KunR11 and KunR12, $u_1(\xi^{(1)}, \xi^{(2)}) := 1 - \frac{\sin(2\pi\xi^{(1)}) + \cos(2\pi\xi^{(2)})}{10}$ for problems KunR21 and KunR22, and $u_1(\xi^{(1)}, \xi^{(2)}) := e^{\xi^{(1)} + \xi^{(2)}}$ for problems KunR31 and KunR32. Matrices A, C and vectors a, c can be found in the corresponding Matlab source files.

JF: positive definite (everywhere)

7 Numerical Results

In this section we report the results obtained by an implementation of the Potential Reduction Algorithm (analyzed so far) on the set of test problems described in Section 6. These results are intended to show the viability of our approach and to give the reader a concrete feel for the practical behavior of PRA on QVI problems. All the computations in this thesis were done using Matlab 7.6.0 on a Ubuntu 10.04 64 bits PC with Intel(R) Core(TM) i7 CPU 870 and 7.8 GiB of RAM.

7.1 Implementation Details

The implemented algorithm corresponds to the theoretical scheme given in Algorithm 3. However now we consider the framework of Section 6 in which also equality constraints are present. This means that, at step `(S.2)`, Algorithm 3 calls for the solution of an $n + 2m + p$ square linear system in order to determine the search direction d^k (we recall that n corresponds to the number of variables, m to the number of inequalities and p to that of equalities). More precisely, at each iteration k we must find a solution $(\bar{d}_1^k, \bar{d}_2^k, \bar{d}_3^k, \bar{d}_4^k) \in \mathbb{R}^{n+2m+p}$ of the following linear system

$$\begin{pmatrix} A^k & \nabla_y g(x^k, x^k) & 0 & M(x^k)^T \\ J_x h(x^k) & 0 & I & 0 \\ 0 & \operatorname{diag}(w^k) & \operatorname{diag}(\lambda^k) & 0 \\ S^k & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix} = \begin{pmatrix} b_1^k \\ b_2^k \\ b_3^k \\ b_4^k \end{pmatrix}, \tag{62}$$

where

$$\begin{aligned}
A^k &:= JF(x^k) + \sum_{i=1}^m \lambda_i^k \, J_x \left[\nabla_y g_i(x^k, x^k)\right] + \sum_{j=1}^p \zeta_j^k \, J_x \left[(M(x^k)_{j*})^T\right], \\
h(x) &:= g(x, x), \\
S^k &:= M(x^k) + \sum_{i=1}^n x_i^k J_x \left[M(x^k)_{*i}\right] + J_x v(x^k), \\
b_1^k &:= -F(x^k) - \nabla_y g(x^k, x^k)\lambda^k - M(x^k)^T \zeta^k, \\
c^k &:= \frac{\rho^k}{2m} \sum_{i=1}^m \left(g_i(x^k, x^k) + w_i^k + \lambda_i^k w_i^k\right), \\
b_2^k &:= -g(x^k, x^k) - w^k + c^k \mathbf{1}_m, \\
b_3^k &:= -\lambda^k \circ w^k + c^k \mathbf{1}_m, \\
b_4^k &:= -M(x^k)x^k - v(x^k),
\end{aligned}$$

$w^k \in \mathbb{R}^m$ are slack variables, $\lambda^k \in \mathbb{R}^m$ are the lagrangian parameters of inequality constraints and $\zeta^k \in \mathbb{R}^p$ are those of equality constraints computed at iteration k.

All data are stored in sparse matrices and full vectors. In order to perform the linear algebra involved, we used Matlab's linear system solver *mldivide* (note that, by the reference manual of mldivide, since the input matrix is sparse the solver actually used is the efficient code CHOLMOD written by Timothy A. Davis).

In what follows we give some implementation details.

Linear System Reduction

As said above, step `(S.2)` of Algorithm 3 consist in the solution of the $n+2m+p$ square linear system (62). However, this system is very structured and some simple manipulations permit to reduce its solution to that of a linear system of dimension $n+p$.

It is easy to verify, by substitution and by the fact that $w^k > 0$, that if we compute $(\bar{d}_1^k, \bar{d}_4^k)$ as solution of

$$\begin{pmatrix} A^k + G^k & M(x^k)^T \\ S^k & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_4 \end{pmatrix} = \begin{pmatrix} r^k \\ b_4^k \end{pmatrix}, \tag{63}$$

where

$$\begin{aligned} G^k &:= \nabla_y g(x^k, x^k)\,\mathrm{diag}((w^k)^{-1} \circ \lambda^k) J_x h(x^k), \\ r^k &:= b_1^k - \nabla_y g(x^k, x^k)\,\mathrm{diag}(w^k)^{-1} b_3^k + \nabla_y g(x^k, x^k)\,\mathrm{diag}((w^k)^{-1} \circ \lambda^k) b_2^k, \end{aligned}$$

and then we compute $\bar{d}_2^k$, $\bar{d}_3^k$ by $\bar{d}_3^k = b_2^k - J_x h(x^k)\bar{d}_1^k$ and $\bar{d}_2^k = \mathrm{diag}(w^k)^{-1} b_3^k - \mathrm{diag}((w^k)^{-1} \circ \lambda^k)\bar{d}_3^k$, respectively, this is indeed a solution of (62). This shows clearly that the main computational burden in solving the linear system (62) is actually the solution of an $(n+p) \times (n+p)$ square linear system. Note that $A^k + G^k$ is exactly the matrix $N(x^k, \lambda^k, w^k)$ defined in (15).

The procedure just described has the advantage of reducing the dimension of the system as much as possible; however this might not always be the best strategy, since sparsity patterns could be lost. For example it might be more convenient, from this point of view, to eliminate just the d_2 variables and then solve the resulting $n+m+p$ system in d_1, d_3 and d_4. Or solve directly the original $n+2m+p$ system and leave to the solver the task of exploiting sparsity. This, as well as the choice of the most suitable linear solver, along with numerical procedures to deal with singularities (discussed below), are very important issues that can have huge practical impact. We are currently investigating on these topics and will report on this research elsewhere, however, about the solution of the linear system (62), our Matlab results seem to indicate that reducing as much as possible the linear system is the best choice (but in a C++ environment the situation may change).

Linear System Perturbation

In practice it is possible that the rows of matrix $M(x^k)$ or of matrix S^k are linear dependent (this occurs for example in test problems LunSS1-2-3 in the library), in this case the reduced linear system (63) is difficult to solve since the condition number of its coefficient matrix is rather big. Therefore to overcome this drawback we apply a perturbation procedure to obtain a linear system better conditioned. In particular the procedure starts if the lower bound for the 1-norm condition number of the coefficient matrix in (63) computed by using the Matlab function `condest` is greater than 10^{17}. In this case the coefficient matrix of (63) is modified by adding a diagonal matrix D to its last block:

$$\begin{pmatrix} A^k + G^k & M(x^k)^T \\ S^k & D \end{pmatrix}, \tag{64}$$

where entries of D are equal to $\pm 10^{-3}$, in particular the i-th entry is negative if

$$(S^k)_{i*}\,(M(x^k)^T)_{*i} \geq 0$$

and positive otherwise. This procedure is partially motivated by results on 1-rank matrices in Appendix B, however we are still working on it. For the time being it must be considered only as an heuristic useful to handle some challenging test problems.

Centering Parameter Computation and Stepsize Failure

During all computations the value of ρ^k is always set to 0.1. This is only changed if in the previous iteration the step size t_{k-1} is not greater than 0.1, in particular ρ^k is iteratively increased by 0.1 until we obtain $t_{k-1} > 0.1$. Whenever $t_{k-1} > 0.1$ the value of ρ^k is reset to 0.1 (its default value). Also if ρ^k reach the value of 0.9 and $t_{k-1} \leq 0.1$ occurs the value of ρ^k is reset to 0.1, but in this case a flag f is activated. This flag f remains active until $t_{k-1} > 0.1$ and if occurs that $t_k < 10^{-6}$ during f active, the algorithm fails (stepsize failure).

Line Search Description

In the line search at step `(S.3)` of Algorithm 3, we take $\beta = 0.5$, $\gamma = 10^{-2}$ and $\xi = 2m$. In order to stay in Z_I we preliminarily rescale $d^k = (d_1^k, d_2^k, d_3^k, d_4^k)$. First we analytically compute a positive constant α such that $\lambda^k + \alpha d_2^k$ and $w^k + \alpha d_3^k$ are greater than 10^{-10}. This ensures that the middle two blocks in $z^k + \alpha d^k$ are in the interior of $\mathbb{R}^{2m}_+$. Then, if necessary, we further reduce this α until $h(x^k + \alpha d_1^k) + w^k + \alpha d_3^k \geq 10^{-10}$ thus finally guaranteeing that $z^k + \alpha d^k$ belongs to Z_I. In this latter phase, an evaluation of h is needed for each bisection. At the end of this process, we set $d^k \leftarrow \alpha d^k$ and then perform the Armijo line search.

Stopping Criteria

The main stopping criterion is based on an equation reformulation of the KKT conditions which uses the Fischer-Burmeister function that, we recall, is defined by $\phi(a,b) = \sqrt{a^2+b^2} - (a+b)$ and has the property that $\phi(a,b) = 0$ if and only if $a \geq 0, b \geq 0, ab = 0$. The equation reformulation is then defined by

$$V(x,\lambda,\zeta) := \begin{pmatrix} F(x) + M(x)^T\zeta + \nabla_y g(x,x)\lambda \\ (\phi(\lambda_i, -g_i(x,x)))_{i=1}^m \\ M(x)y + v(x) \end{pmatrix}.$$

The main termination criterion, declaring the success of the run, is $\|V(x^k, \lambda^k, \zeta^k)\|_\infty \leq 10^{-4}$. The iterations are stopped with a failure if the number of iterations exceeds 1000 or the running time exceeds one hour or a stepsize failure (described above) occurs.

7.2 Computational Results

The following is a list of problems for which the key nonsingularity assumption of JH can be guaranteed based on the results in Section 3 (see Section 6 for details): OutZ43, MovSet1A, MovSet2A, MovSet3A1, MovSet3A2, MovSet4A1, MovSet4A2, Box1A, Box2A, Box3A, BiLin1A, RHS1A1, RHS2A1.

Note that the structure of problems OutKZ31 and OutKZ41 is the one analyzed in Example 3.12 and therefore these problems are nonsingular if the friction coefficient ϕ is small enough. However in the test problem we used, we took the friction coefficient large to make the problems more difficult. More in general, we included many problems whose nonsingularity is not guaranteed (i.e. we actually do not know whether nonsingularity is satisfied or not) in order to test the robustness of the method.

In Table 7 we report the numerical results of our algorithm on all test problems of the library (described in Section 6). For each problem we list

- the x-part of the starting point (the number reported is the value of all components of the x-part of the starting point) [x^0];
- the number of iterations, which is equal to the number of evaluations of JH [It/JH];
- the number of evaluations of the constraints vectors [g];
- the number of evaluations of F, which is equal to the number of evaluations of the gradients of the constraints vectors [$F/\nabla g$];
- the value of the KKT violation measure $\|V(x, \lambda, \zeta)\|_\infty$ at termination [$\|V\|_\infty$].

Note that for the (λ, w)-part of the starting vector, we always used $\lambda^0 = 5$ and further set $w^0 = \max(5, 5 - h(x^0))$, so as to ensure that the starting point is "well inside" Z_I. While for the ζ-part of the starting vector (the equality parameters), we always used $\zeta^0 = 0$.

We see that overall the algorithm seems efficient and reliable and able to solve a wide array of different problems. Note that for six runs (LunSS1, LunSS3, LunSSVI1, LunSSVI3) $\|V\|_\infty$ is slightly greater than 10^{-4}, however, since the numerical solution of these problems is very challenging and we can easily improve the solution given by using fast local methods, we do not consider these runs as failures. The five failures reported in Table 7 deserve a few more comments. The failures on KunR31 and KunR32 are due to the limit on computing time (3600 seconds), but the algorithm actually appears to be converging in both cases, in fact $\|V\|_\infty$ is 1.1*10^{-3} and 1.3*10^{-3} respectively. In the case of Box1B, Box3A and Box3B instead, difficulties arise because of an almost singularity of the linear system giving the search direction; this leads to a stepsize failure (we are currently working on a procedure able to recover from such situations).

Table 7: Potential Reduction Algorithm numerical results for QVIs.

Problem	x^0	It / JH	g	$F/\nabla g$	$\|V\|_\infty$	Problem	x^0	It / JF	g	$F/\nabla g$	$\|V\|_\infty$
OutZ40	0	8	9	9	7.4853e-05	Wal3	0	48	84	82	4.2859e-05
OutZ41	0	18	19	19	9.7789e-05	Wal3	10	63	110	110	3.4127e-05
OutZ42	0	8	9	9	1.4467e-05	Wal5	0	46	80	47	4.6612e-05
OutZ43	0	8	9	9	1.8955e-05	Wal5	10	42	43	43	6.4139e-05
OutZ44	0	8	9	9	2.9380e-05	LunSS1	0	33	34	34	8.7775e-05
MovSet1A	0	10	12	12	1.4767e-05	LunSS1	10	47	48	48	1.2273e-04
MovSet1B	0	16	27	27	3.8251e-05	LunSS2	0	28	29	29	9.6952e-05
MovSet2A	0	12	15	15	2.0701e-05	LunSS2	10	34	35	35	8.8951e-05
MovSet2B	0	36	93	93	2.1019e-05	LunSS3	0	39	40	40	8.3032e-04
MovSet3A1	0	11	12	12	2.7945e-05	LunSS3	10	45	46	46	5.3789e-04
MovSet3B1	0	11	12	12	1.8449e-05	LunSSVI1	0	33	34	34	8.7775e-05
MovSet3A2	0	11	12	12	5.6040e-05	LunSSVI1	10	47	48	48	1.2272e-04
MovSet3B2	0	11	12	12	3.6660e-05	LunSSVI2	0	28	29	29	9.6952e-05
MovSet4A1	0	12	13	13	7.1662e-05	LunSSVI2	10	34	35	35	8.8950e-05
MovSet4B1	0	12	13	13	4.5120e-05	LunSSVI3	0	39	40	40	8.3032e-04
MovSet4A2	0	12	13	13	7.1632e-05	LunSSVI3	10	45	46	46	5.3789e-04
MovSet4B2	0	12	13	13	7.1343e-05	Scrim11	0	10	11	11	2.7489e-05
Box1A	10	9	10	10	1.6652e-05	Scrim11	10	9	10	10	5.1617e-05
Box1B	10	Failure				Scrim12	0	10	11	11	2.7475e-05
Box2A	10	167	187	187	7.7965e-06	Scrim12	10	9	10	10	5.1600e-05
Box2B	10	195	220	220	2.3443e-06	Scrim21	0	17	18	18	2.5190e-05
Box3A	10	Failure				Scrim21	10	19	20	20	8.9428e-05
Box3B	10	Failure				Scrim22	0	17	18	18	2.5188e-05
BiLin1A	0	13	14	14	1.2465e-05	Scrim22	10	19	20	20	8.9414e-05
BiLin1B	0	10	11	11	1.2308e-05	OutKZ31	0	18	19	19	2.4473e-05
RHS1A1	10	19	20	20	3.5596e-05	OutKZ31	10	17	18	18	1.6132e-05
RHS1B1	10	23	29	24	2.2245e-05	OutKZ41	0	20	21	21	4.6573e-05
RHS2A1	10	19	20	20	3.5592e-05	OutKZ41	10	20	21	21	3.5913e-05
RHS2B1	10	19	20	20	2.4006e-05	KunR11	0	14	15	15	7.9623e-05
WalEq1	0	15	17	17	1.9814e-05	KunR11	10	24	40	40	8.3369e-05
WalEq1	10	17	24	24	6.5298e-05	KunR12	0	22	35	35	7.7460e-05
WalEq2	0	22	33	23	4.2337e-06	KunR12	10	25	43	43	9.0344e-05
WalEq2	10	17	19	18	4.9068e-06	KunR21	0	21	35	35	6.1531e-05
WalEq3	0	35	53	48	4.1380e-05	KunR21	10	22	33	33	4.6800e-05
WalEq3	10	60	95	95	1.9673e-05	KunR22	0	23	40	40	7.6296e-05
WalEq4	0	25	29	26	3.4682e-05	KunR22	10	23	40	40	8.9724e-05
WalEq4	10	22	26	23	3.6072e-05	KunR31	0	154	764	764	4.6276e-05
WalEq5	0	24	27	25	6.2267e-05	KunR31	10	Failure			
WalEq5	10	30	54	31	3.8680e-05	KunR32	0	168	807	807	2.7577e-05
Wal2	0	34	59	35	2.5898e-05	KunR32	10	Failure			
Wal2	10	47	95	48	5.0181e-05						

In order to better gauge the robustness of our algorithm we also solved all the problems using a C version of the PATH solver with a Matlab interface downloaded from `http://pages.cs.wisc.edu/~ferris/path/` and whose detailed description can be found in [12, 26]. PATH is a well-established and mature software implementing a stabilized Josephy-Newton method for the solution of Mixed Complementarity Problems and it is well known that it can be also used to solve the KKT conditions of a QVI, although with no theoretical guarantee of convergence in our setting. We used the same (x, ζ)-part for the starting point as we used in the testing of our method. For the λ-part, we considered two options. In the first one we took $\lambda^0 = 5$, therefore using exactly the same starting point we used in the testing of the interior-point algorithm. In the second option we set $\lambda^0 = 0$; this latter alternative was considered because the choice of $\lambda^0 = 5$ is geared towards our interior-point method, while $\lambda^0 = 0$ seems more natural for PATH. It might be useful to remark that we run PATH with its default settings and the stopping criteria using by PRA and PATH are marginally different. In spite of this, the precision at the computed solution, measured in terms of $\|V(x, \lambda, \zeta)\|_\infty$, is consistently comparable. In both the tested cases, PATH was not able to solve ten problems (the failures are given in Table 8, where under the heading PATH (5) we report the timings for PATH with $\lambda^0 = 5$ and analogously under PATH (0) we have the timings for PATH with $\lambda^0 = 0$). These results seem to indicate that our method has the potential to become a very robust solver for the solution of the KKT conditions arising from QVIs.

The comparison of CPU times is somewhat more problematic. In fact one should take into account that, although the main computational burden in our algorithm is given by the solution of linear systems, a task very efficiently performed by the Matlab built-in function *mldivide*, we did use Matlab, an interpreted language, and furthermore, what we implemented is a straightforward version of our algorithm, with none of all those crash and recover techniques that are to be found in a well developed software as PATH. In spite of this, having current CPU times would still be of interest, and so we report them in Table 8. Note that we do not report major, minor and crash iterations for PATH. In fact PRA and PATH are very unlike, and the meaning of "iteration" is so different in the two methods that we feel that, besides the number of failures, CPU time is the only other meaningful parameter to compare. These results show that even the current prototypical Matlab implementation of the interior-point method compares well to PATH, also in terms of computing times. The development of a more sophisticated C++ version of our method, fully exploiting its potential for parallelism, is currently under the way and more extended and detailed numerical results, along with more accurate comparisons, will be reported elsewhere.

Table 8: CPU times in seconds and failures (F).

Problem	x^0	PRA	PATH (5)	PATH (0)	Problem	x^0	PRA	PATH (5)	PATH (0)
OutZ40	0	< 0.1	< 0.1	< 0.1	Wal3	0	0.3	0.1	0.1
OutZ41	0	< 0.1	< 0.1	< 0.1	Wal3	10	0.4	0.4	0.1
OutZ42	0	< 0.1	< 0.1	< 0.1	Wal5	0	0.7	F	F
OutZ43	0	< 0.1	< 0.1	< 0.1	Wal5	10	0.6	F	1.4
OutZ44	0	< 0.1	< 0.1	< 0.1	LunSS1	0	3.2	4.7	3.4
MovSet1A	0	< 0.1	< 0.1	< 0.1	LunSS1	10	4.6	4.9	5.6
MovSet1B	0	< 0.1	< 0.1	F	LunSS2	0	18.0	106.6	62.0
MovSet2A	0	< 0.1	< 0.1	< 0.1	LunSS2	10	22.6	108.6	68.0
MovSet2B	0	< 0.1	< 0.1	F	LunSS3	0	722.0	F	F
MovSet3A1	0	2.2	5.4	4.8	LunSS3	10	858.9	F	F
MovSet3B1	0	2.2	5.8	5.3	LunSSVI1	0	3.3	4.7	1.6
MovSet3A2	0	10.1	37.5	35.2	LunSSVI1	10	4.8	4.6	5.6
MovSet3B2	0	9.9	46.8	38.1	LunSSVI2	0	19.2	109.5	50.8
MovSet4A1	0	0.4	7.5	0.4	LunSSVI2	10	24.0	110.8	62.9
MovSet4B1	0	0.3	7.6	0.4	LunSSVI3	0	732.6	F	F
MovSet4A2	0	1.3	92.0	2.0	LunSSVI3	10	867.6	F	F
MovSet4B2	0	1.3	96.2	2.0	Scrim11	0	0.3	2.6	< 0.1
Box1A	10	< 0.1	< 0.1	< 0.1	Scrim11	10	0.3	2.6	< 0.1
Box1B	10	F	< 0.1	< 0.1	Scrim12	0	0.6	17.7	0.1
Box2A	10	4.0	5.7	2.7	Scrim12	10	0.5	17.7	0.1
Box2B	10	4.7	7.2	8.7	Scrim21	0	0.2	4.7	< 0.1
Box3A	10	F	16.0	4.6	Scrim21	10	0.2	4.7	2.1
Box3B	10	F	F	34.3	Scrim22	0	0.3	29.2	0.1
BiLin1A	0	< 0.1	< 0.1	< 0.1	Scrim22	10	0.3	29.3	10.1
BiLin1B	0	< 0.1	< 0.1	< 0.1	OutKZ31	0	< 0.1	< 0.1	< 0.1
RHS1A1	10	0.6	0.3	0.4	OutKZ31	10	< 0.1	< 0.1	< 0.1
RHS1B1	10	0.4	0.5	0.8	OutKZ41	0	< 0.1	< 0.1	< 0.1
RHS2A1	10	0.7	0.3	0.5	OutKZ41	10	< 0.1	< 0.1	< 0.1
RHS2B1	10	0.6	0.3	0.3	KunR11	0	26.2	110.5	37.7
WalEq1	0	0.2	< 0.1	0.1	KunR11	10	53.8	130.0	50.0
WalEq1	10	0.1	< 0.1	0.1	KunR12	0	177.1	1382.8	161.8
WalEq2	0	0.2	0.8	3.9	KunR12	10	209.0	1280.6	196.3
WalEq2	10	0.1	1.7	0.7	KunR21	0	46.7	91.7	42.8
WalEq3	0	0.4	0.3	0.4	KunR21	10	47.0	119.1	82.2
WalEq3	10	0.7	0.3	0.7	KunR22	0	192.5	845.4	114.1
WalEq4	0	0.6	F	5.6	KunR22	10	193.6	1028.0	218.3
WalEq4	10	0.6	3.9	0.4	KunR31	0	637.9	168.7	1225.4
WalEq5	0	0.9	F	5.9	KunR31	10	F	185.3	F
WalEq5	10	1.2	2.3	0.6	KunR32	0	2531.6	552.3	F
Wal2	0	0.1	0.2	0.3	KunR32	10	F	817.4	F
Wal2	10	0.2	F	0.1					

Conclusions

We presented a detailed convergence theory for an interior-point method for the solution of the KKT conditions of a general QVI. We could establish convergence for a wide array of different classes of problems including QVIs with the feasible set given by "moving sets", linear systems with variable right-hand sides, box constraints with variable bounds, and bilinear constraints. These results surpass by far existing convergence analyses, the latter all having a somewhat limited scope. In our view, the results in this thesis constitute an important step towards the development of theoretically reliable and numerically efficient methods for the solution of QVIs.

Moreover we presented a big collection of test problems from diverse sources which is the largest test set for QVIs considered so far and gives a uniform basis on which algorithms for the solution of QVIs can be tested and compared.

Future works on this thesis topics are the development and the implementation of a C++ solver for the solution of QVIs based on our method.

A Appendix on Monotonicity and Lipschitz Properties

In this appendix we recall some well-known definitions and discuss some related results. Although the latter are also mostly well-known, in some cases we could not find in the literature the exact versions we needed. Therefore, for completeness we also report the proofs of these results.

We begin by recalling the definitions of several classes of functions.

Definition A.1 Let $D \subseteq \mathbb{R}^n$ and $F : D \to \mathbb{R}^n$ be a given function. Then

(a) F is *strongly monotone* on D with constant σ if $\sigma > 0$ and

$$(x-y)^T\big(F(x)-F(y)\big) \;\geq\; \sigma\|x-y\|^2, \quad \forall x,y \in D;$$

The largest σ for which such a relation holds is termed the *monotonicity modulus* of F on D:

$$\sigma(D,F) \;:=\; \inf_{x\neq y, x,y\in D} \frac{(x-y)^T\big(F(x)-F(y)\big)}{\|x-y\|^2}.$$

(b) F is *co-coercive* on D with constant ξ if $\xi > 0$ and

$$(x-y)^T\big(F(x)-F(y)\big) \;\geq\; \xi\|F(x)-F(y)\|^2, \quad \forall x,y \in D;$$

(c) F is *Lipschitz continuous* on D with constant $L \geq 0$ if

$$\|F(x)-F(y)\| \;\leq\; L\|x-y\|, \quad \forall x,y \in D.$$

The smallest L for which such relation holds is termed the *Lipschitz modulus* of F on D:

$$L(D,F) \;:=\; \sup_{x\neq y, x,y\in D} \frac{\|F(x)-F(y)\|}{\|x-y\|}.$$

(d) F is a *homeomorphism* of D onto $F(D)$ if F is one-to-one on D (that is $F(x) \neq F(y)$ whenever $x, y \in D, x \neq y$, or, in other words, F has a single-valued inverse F^{-1} defined on $F(D)$), and F and F^{-1} are continuous on D and $F(D)$, respectively. □

Characterizations of the Lipschitz and strong monotonicity moduli are given in the following result.

Proposition A.2 *Let $D \subseteq \mathbb{R}^n$ be an open, convex subset of $\mathbb{R}^n$ and let $F : D \to \mathbb{R}^n$ be a continuously differentiable function. Then the following statements hold:*

(a) F is Lipschitz continuous on D with constant L if and only if $\|JF(x)\| \leq L$ for all $x \in D$; consequently

$$L(D,F) \;=\; \sup_{x\in D} \|JF(x)\|,$$

provided the sup *on the right hand side is finite.*

(b) F is strongly monotone on D with constant σ if and only if $h^T JF(x)h \geq \sigma\|h\|^2$ for all $x \in D$, $h \in \mathbb{R}^n$; consequently

$$\sigma(D, F) \;=\; \inf_{x\in D} \mu_m^s(JF(x)),$$

provided the inf *on the right hand side is positive.*

Proof. (a) From Theorem 3.2.3 in [50], if $\|JF(x)\| \leq L$ then L is a Lipschitz constant for F on D. Conversely, assume that

$$\big\|F(x) - F(y)\big\| \leq L\|x - y\|, \quad \forall x, y \in D \tag{65}$$

holds. Applying the differential mean value theorem to each component function F_i of F, it follows that, for any given $x, y \in D$, we can find suitable points $\xi^{(i)} \in (x, y)$ such that

$$F_i(x) - F_i(y) = \nabla F_i(\xi^{(i)})^T(x - y) \quad \forall i = 1, \ldots, n.$$

Setting

$$G(\xi) := \begin{pmatrix} \nabla F_1(\xi^{(1)})^T \\ \vdots \\ \nabla F_n(\xi^{(n)})^T \end{pmatrix} \in \mathbb{R}^{n\times n},$$

this can be rewritten in a compact way as

$$F(x) - F(y) = G(\xi)(x - y). \tag{66}$$

Now, let $x \in D$ be fixed, and note that

$$G(\xi) \to JF(x) \tag{67}$$

for any sequence $y \to x$ in view of the continuous differentiability of F. We now consider a particular sequence $y = x + td$ with a fixed (but arbitrary) vector $d \in \mathbb{R}^n \setminus \{0\}$ and a sequence $t \downarrow 0$. Then (65) and (66) together imply

$$\big\|G(\xi)td\big\| = \big\|F(x) - F(x + td)\big\| \leq L\|td\|.$$

Dividing by t and subsequently letting $t \downarrow 0$ (note that ξ still depends on t), we obtain

$$\big\|JF(x)d\big\| \leq L\|d\|$$

in view of (67). Since d was taken arbitrarily, this implies $\|JF(x)\| \leq L$, and this inequality is true for any vector $x \in D$.
(b) See [50, Theorem 5.4.3]. □ □

The following result gives a relation between the Lipschitz constants etc. of a given mapping F and its inverse F^{-1}.

Proposition A.3 *Let a function $F : D \to \mathbb{R}^n$ be given where D is an open subset of $\mathbb{R}^n$. Assume that two positive constants ℓ and L exist such that*

$$\ell\|x - y\| \leq \|F(x) - F(y)\| \leq L\|x - y\|, \quad \forall x, y \in D. \tag{68}$$

Then F is a homeomorphism from D to $F(D)$ (which is an open set) and

$$\frac{1}{L}\|a - b\| \leq \|F^{-1}(a) - F^{-1}(b)\| \leq \frac{1}{\ell}\|a - b\|, \quad \forall a, b \in F(D), \tag{69}$$

in particular, F and F^{-1} are Lipschitz continuous on D and $F(D)$, respectively.

Proof. The first inequality from (68) implies that F is one-to-one on D (therefore the inverse F^{-1} exists on $F(D)$), and that, setting $a = F(x)$ and $b = F(y)$, the second inequality in (69) holds. In particular, this implies that F^{-1} is Lipschitz continuous on $F(D)$, hence continuous, so that $F(D) = (F^{-1})^{-1}(D)$, being the pre-image of a continuous map of the open set D, is also an open set. Finally, let $a, b \in F(D)$ be arbitrarily given. Setting $x = F^{-1}(a), y = F^{-1}(b)$, we obtain from the second inequality in (68) that

$$\|F^{-1}(a) - F^{-1}(b)\| = \|x - y\| \geq \frac{1}{L}\|F(x) - F(y)\| = \frac{1}{L}\|a - b\|,$$

and this completes the proof. □ □

The next result considers a strongly monotone and Lipschitz continuous mapping and provides suitable bounds for the moduli of Lipschitz continuity and strong monotonicity of the corresponding inverse function. We stress, however, that the constant of strong monotonicity of the inverse function provided by this result is really just an estimate and typically not exact. It seems difficult to find a stronger bound in the general context discussed here. In a more specialized situation, much better results can be obtained, see Proposition A.5 below.

Proposition A.4 *Let $D \subseteq \mathbb{R}^n$ be an open set and $F : D \to \mathbb{R}^n$ be strongly monotone with modulus σ and Lipschitz continuous with modulus L on D. Then F is co-coercive with constant $\frac{\sigma}{L^2}$. Furthermore, it holds that the inverse F^{-1} exists on $F(D)$, is Lipschitz with constant $\frac{1}{\sigma}$ and strongly monotone with constant $\frac{\sigma}{L^2}$.*

Proof. We can write

$$\big(F(x) - F(y)\big)^T (x - y) \geq \sigma\|x - y\|^2, \quad \forall x, y \in D$$

and

$$\|F(x) - F(y)\|^2 \leq L^2\|x - y\|^2, \quad \forall x, y \in D$$

by assumption. A combination of these two inequalities yields

$$\|F(x) - F(y)\|^2 \le L^2 \|x - y\|^2 \le \frac{L^2}{\sigma} \big(F(x) - F(y)\big)^T (x - y), \quad \forall x, y \in D. \tag{70}$$

Hence F is co-coercive with constant $\frac{\sigma}{L^2}$.

By Proposition A.3 we know that F is a homeomorphism from D to $F(D)$ and F^{-1} is Lipschitz continuous with constant $\frac{1}{\sigma}$. Finally writing $a = F(x), b = F(y)$ in (70) gives

$$\big(F^{-1}(a) - F^{-1}(b)\big)^T (a - b) \ge \frac{\sigma}{L^2} \|a - b\|^2, \quad \forall a, b \in F(D).$$

This completes the proof. □ □

The following result gives an exact estimate of the Lipschitz and strong monotonicity moduli of the inverse of a function under the assumption that the mapping F itself is a *gradient mapping*, i.e. that $F = \nabla f$ for a differentiable real-valued function f.

Proposition A.5 *Let $D \subseteq \mathbb{R}^n$ be an open convex set and $F : D \to \mathbb{R}^n$ be a gradient mapping. Assume that F is strongly monotone with modulus σ and Lipschitz continuous with modulus L on D. Then the inverse function F^{-1} exists on $F(D)$, is Lipschitz with modulus $\frac{1}{\sigma}$ and strongly monotone with modulus $\frac{1}{L}$.*

Proof. The result can easily be derived from the Baillon-Haddad Theorem, see [1], when $D = \mathbb{R}^n$. We give here a direct proof which is valid also when $D \neq \mathbb{R}^n$. In view of Proposition A.4, we only have to verify the statement that F^{-1} is strongly monotone with constant $\frac{1}{L}$.

To this end, first consider a symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$, let $A^{1/2}$ be the corresponding (unique) symmetric positive definite square root of A so that $A^{1/2} A^{1/2} = A$, and let $A^{-1/2}$ be the inverse of $A^{1/2}$. Then the symmetry of $A^{1/2}$ together with the Cauchy-Schwarz inequality implies

$$\|d\|^2 = d^T d = d^T A^{1/2} A^{-1/2} d \le \|A^{1/2} d\| \cdot \|A^{-1/2} d\|, \quad \forall d \in \mathbb{R}^n.$$

Squaring both sides shows that

$$\|d\|^4 \le \big(d^T A d\big)\big(d^T A^{-1} d\big), \quad \forall d \in \mathbb{R}^n \tag{71}$$

holds. Since F is strongly monotone, the Jacobian $JF(x)$ is positive definite for all $x \in D$; furthermore, since F is a gradient mapping, this Jacobian is also symmetric. Hence we can apply inequality (71) to the matrix $A := JF(x)$ and obtain

$$\begin{aligned} \|d\|^4 &\le \big(d^T JF(x) d\big)\big(d^T JF(x)^{-1} d\big) \\ &\le \big(d^T JF(x)^{-1} d\big) \|d\|^2 \|JF(x)\| \\ &\le \big(d^T JF(x)^{-1} d\big) L \|d\|^2, \qquad \forall d \in \mathbb{R}^n, \end{aligned}$$

where the second inequality uses the Cauchy-Schwarz inequality once again, and the third inequality takes into account Proposition A.2. This implies

$$\frac{1}{L} d^T d = \frac{1}{L} \|d\|^2 \leq \big(d^T JF(x)^{-1} d\big), \quad \forall d \in \mathbb{R}^n \ \forall x \in D.$$

Since $JF(x)^{-1} = JF^{-1}(y)$ for $y = F(x)$ by the Inverse Function Theorem, this gives

$$\frac{1}{L} d^T d \leq d^T JF^{-1}(y) d, \quad \forall d \in \mathbb{R}^n \ \forall y \in F(D).$$

By a well-known result, see [50, Theorem 5.4.3] this is equivalent to saying that F^{-1} is strongly monotone on $F(D)$ with constant $1/L$. □ □

The sharper result from Proposition A.5 regarding the modulus of strong monotonicity does, in general, not hold for non-gradient mappings, see the corresponding discussion and (counter-) example at the end of Subsection 3.2.

B Appendix on Nonsymmetric Matrices

Positiveness and Beyond

In this appendix we recall some definitions and discuss some related results on square matrices not necessarily symmetric. See [10] for further readings.

Definition B.1 *Let $A \in \mathbb{R}^{n\times n}$ be a given square matrix. Then*

(a) A is positive definite *($A \succ 0$) if*

$$v^T Av > 0, \quad \forall\, v \in \mathbb{R}^n,\, v \neq 0$$

(b) A is positive semidefinite *($A \succeq 0$) if*

$$v^T Av \geq 0, \quad \forall\, v \in \mathbb{R}^n$$

(c) $A \in P$ if

$$\exists\, i \in \{1, \ldots, n\} : \quad v_i(Av)_i > 0, \quad \forall\, v \in \mathbb{R}^n,\, v \neq 0$$

(d) $A \in P_0$ if

$$\exists\, i \in \{1, \ldots, n\} : \quad v_i \neq 0, \quad v_i(Av)_i \geq 0, \quad \forall\, v \in \mathbb{R}^n,\, v \neq 0$$

Proposition B.2 *Let $A \in \mathbb{R}^{n\times n}$ be a given square matrix. Then*

(a) $A \succ 0$ if and only if all principal minors of $(A + A^T)$ are positive

(b) $A \succ 0$ if and only if all eigenvalues of $(A + A^T)$ are positive

(c) $A \succeq 0$ if and only if all principal minors of $(A + A^T)$ are nonnegative

(d) $A \succeq 0$ if and only if all eigenvalues of $(A + A^T)$ are nonnegative

(e) $A \succeq 0$ if and only if $A + \epsilon I \succ 0$, $\forall\, \epsilon > 0$

(f) $A \in P$ if and only if all its principal minors are positive

(g) $A \in P$ if and only if all real eigenvalues of A and of all principal submatrices of A are positive

(h) $A \in P_0$ if and only if all its principal minors are nonnegative

(i) $A \in P_0$ if and only if all real eigenvalues of A and of all principal submatrices of A are nonnegative

(j) $A \in P_0$ if and only if $A + \epsilon I \in P$, $\forall\, \epsilon > 0$

(k) if $A \succ 0$ *then* A *is nonsingular*

(l) if $A \in P$ *then* A *is nonsingular*

(m) if $A \succ 0$ *then* $A \succeq 0$, $A \in P$, $A \in P_0$

(n) if $A \succeq 0$ *then* $A \in P_0$

(o) if $A \in P$ *then* $A \in P_0$

Now we give some further propositions in order to prove results in Table 9.

Proposition B.3 *Let* $A \succ 0$ *and* $B \succeq 0$ *then* $A + B \succ 0$.

Proof. $\forall v \in \mathbb{R}^n$, $v \neq 0$ we have $v^T Av > 0$ and $v^T Bv \geq 0$, and then $v^T(A + B)v > 0$. □

Proposition B.4 *Let* $A \succeq 0$ *and* $B \succeq 0$ *then* $A + B \succeq 0$.

Proof. $\forall v \in \mathbb{R}^n$ we have $v^T Av \geq 0$ and $v^T Bv \geq 0$, and then $v^T(A + B)v \geq 0$. □

Proposition B.5 *There exist* $A \succeq 0$ *and* $B \succeq 0$ *such that* $A + B \notin P$.

Proof. By choosing A and B such that a vector $v \neq 0$ exists such that $Av = 0$ and $Bv = 0$ (for example if $A = B$), we obtain $(A + B)v = 0$. Therefore $A + B$ is singular and then it cannot be P. □

Proposition B.6 *There exist* $A \succ 0$ *and* $B \in P$ *such that* $A + B \notin P_0$.

Proof. By choosing

$$A := \begin{pmatrix} 1 & -2 \\ 0 & 2 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}$$

we obtain

$$A + B = \begin{pmatrix} 2 & -2 \\ -4 & 3 \end{pmatrix} \notin P_0,$$

and then we have the proof. Another example with A symmetric follows

$$A := \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}$$

then

$$A + B = \begin{pmatrix} 3 & -3 \\ -7 & 6 \end{pmatrix} \notin P_0.$$

□

Proposition B.7 *There exist $A \succ 0$ and $B \in P$ such that $A + B$ is singular.*

Proof. By choosing

$$A := \begin{pmatrix} 1 & -2 \\ 0 & 2 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$$

we obtain

$$A + B = \begin{pmatrix} 2 & -2 \\ -3 & 3 \end{pmatrix}$$

which is singular and then we have the proof. Another example with A symmetric follows

$$A := \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$$

then

$$A + B = \begin{pmatrix} 3 & -3 \\ -6 & 6 \end{pmatrix}$$

which is singular. □

Proposition B.8 *There exist $A \succ 0$ and $B \succ 0$ such that $AB \notin P_0$.*

Proof. By choosing

$$A := \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 3 \\ 1 & 5 \end{pmatrix}$$

we obtain

$$AB = \begin{pmatrix} 2 & 6 \\ -1 & -1 \end{pmatrix} \notin P_0,$$

then we have the proof. Another example with A and B symmetric follows

$$A := \begin{pmatrix} \frac{3}{2} & 4 \\ 4 & 11 \end{pmatrix}, \quad B := \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}$$

then

$$AB = \begin{pmatrix} -\frac{9}{2} & \frac{7}{2} \\ -13 & 10 \end{pmatrix} \notin P_0.$$

□

Proposition B.9 *Let A and B be two square nonsingular matrices, then AB is nonsingular.*

Proof. For all $v \neq 0$ we obtain $Bv = u \neq 0$ being B nonsingular, then $ABv = Au \neq 0$ being A nonsingular. □

Proposition B.10 *Let A and B be two square matrices, with A nonsingular and B singular, then AB is singular.*

Proof. A vector v exists such that $Bv = 0$, then we obtain $ABv = A0 = 0$, that is AB is singular. □

Proposition B.11 *Let $A \succ 0$ and D be a square diagonal matrix with positive diagonal entries, then $DA \in P$.*

Proof. Choosing an arbitrary subset of indices of rows (or columns) I, and exploiting the diagonal structure of D, we can write:

$$(DA)_{II} = D_{II}A_{II}.$$

Then it is not difficult to see that all real eigenvalues of $D_{II}A_{II}$ are positive, since they are the same of the matrix $D_{II}^{\frac{1}{2}}A_{II}D_{II}^{\frac{1}{2}} \succ 0$, where $D_{II}^{\frac{1}{2}}$ is the square root of D_{II} (which is diagonal). This is equivalent to say that $DA \in P$. □

Proposition B.12 *Let $A \succeq 0$ and D be a square diagonal matrix with nonnegative diagonal entries, then $DA \in P_0$.*

Proof. It is similar to that of Proposition B.11. □

	$A \succ 0$	$A \succeq 0$	$A \in P$	$A \in P_0$
$B \succ 0$	$A + B \succ 0$ $AB \notin P_0$ but nonsing.			
$B \succeq 0$	$A + B \succ 0$ $AB \notin P_0$ and sing.	$A + B \succeq 0$ but not P $AB \notin P_0$ and sing.		
$B \in P$	$A + B \notin P_0$ and can be sing. $AB \notin P_0$ but nonsing.	$A + B \notin P_0$ and can be sing. $AB \notin P_0$ and sing.	$A + B \notin P_0$ and can be sing. $AB \notin P_0$ but nonsing.	
$B \in P_0$	$A + B \notin P_0$ and can be sing. $AB \notin P_0$ and sing.	$A + B \notin P_0$ and can be sing. $AB \notin P_0$ and sing.	$A + B \notin P_0$ and can be sing. $AB \notin P_0$ and sing.	$A + B \notin P_0$ and can be sing. $AB \notin P_0$ and sing.

Table 9: Sum and product general properties.

1-Rank Matrices

Given two non null vectors $u, v \in \mathbb{R}^n$, we define the square matrix

$$A := uv^T. \tag{72}$$

The following considerations hold (pag. 61 in [35]):

- A has rank equal to 1;
- A has $n-1$ eigenvalues equal to 0 and one eigenvalue equal to $\lambda := u^T v$;
- the right eigenvector associated with λ is u and the left one is v;
- any vector orthogonal to v is a right eigenvector with eigenvalue equal to zero, while if it is orthogonal to u then it is a left one.

It is important to recall that any matrix with rank 1 can be expressed as in (72).

Proposition B.13 *Consider the matrix A in (72), the following holds:*

(a) $A \succeq 0$ if and only if $u = \alpha v$, with $\alpha \geq 0$;

(b) $A \in P_0$ if and only if $u_i v_i \geq 0$ for all $i = 1, \dots, n$.

Proof. (a) Suppose that $u \neq \alpha v$ or that $u = \alpha v$ with $\alpha < 0$, then it is not difficult to find a vector $z \in \mathbb{R}^n$ such that $z^T u < 0$ and $z^T v > 0$, but this is equivalent to say that $z^T uv^T z < 0$ that is $A \not\succeq 0$.

If $u = \alpha v$ with $\alpha \geq 0$ then for all $z \in \mathbb{R}^n$: $z^T u = \alpha z^T v$, but this is equivalent to say that $z^T uv^T z = \alpha z^T vv^T z \geq 0$ that is $A \succeq 0$.
(b) The determinant of any principal submatrix of A with dimension greater than 1 is equal to 0. The determinants of the principal submatrices of A with dimension equal to 1 are equal to $u_i v_i$ for $i = 1, \dots, n$. Then by the definition of matrix P_0 the proof follows. □

The corollary below follows directly from Proposition B.13 and from the fact that any matrix with rank equal to 1 can be rewritten as the product of two vectors.

Corollary B.14 *Let A be a matrix whose rank is equal to 1. Then the followings hold:*

(a) $A \succeq 0$ if and only if A is symmetric with all diagonal entries nonnegative;

(b) $A \in P_0$ if and only if all the diagonal entries of A are nonnegative.

Proposition B.15 *Consider $u, v \in \mathbb{R}^n$ and assume that $u^T v \neq 0$. Then a basis $\{b_1, \dots, b_n\}$ of non null right [left] unit eigenvectors of the matrix $A = uv^T$ exists such that $b_1 = \frac{u}{\|u\|}$ $\left[b_1 = \frac{v}{\|v\|}\right]$, $b_n = \frac{1}{\sqrt{\|u\|^2\|v\|^2 - (u^T v)^2}}\left(\|v\|u - \frac{u^T v}{\|v\|}v\right)$ $\left[b_n = \frac{1}{\sqrt{\|u\|^2\|v\|^2 - (u^T v)^2}}\left(\|u\|v - \frac{u^T v}{\|u\|}u\right)\right]$, $b_1 \perp b_i$ and $b_n \perp b_i$ for all $i = 2, \dots, n-1$ and $b_j \perp b_i$ for all $i, j = 2, \dots, n-1$ and $i \neq j$.*

Proof. It is well known that b_1 is an eigenvector of A with eigenvalue equal to $u^T v$ and it not difficult to see that b_n is an eigenvector of A with eigenvalue equal to 0. Given the two vectors $u, v \in \mathbb{R}^n$ it is always possible to define $n-2$ non null unit vectors mutually orthogonal and orthogonal to both u and v. Then, since b_n is a linear combination of u and v, we can conclude that $b_2, \ldots, b_{n-1}$ exist. $\square$

Proposition B.16 *Consider $u, v \in \mathbb{R}^n$ and assume that $u^T v = 0$. Consider the matrix $A = uv^T$, then either it is equal to the null matrix or it is not diagonalizable.*

Proof. It is known that A has only one eigenvalue equal to 0 with algebraic multiplicity equal to n. Unless A is equal to the null matrix (that is u or v are null vectors) then the rank of A is equal to 1 and then the geometric multiplicity of the eigenvalue 0 is equal to $n-1$, and then A is not diagonalizable. $\square$

Remark B.17 It is well known that the roots of a polynomial vary continuously with the coefficients. This means that the coefficients of the characteristic polynomial of a matrix vary continuously with the entries of the matrix, since they can be expressed in terms of sums of principal minors. Consequently, for any matrix the eigenvalues must vary continuously with the entries (see Example 7.1.3 in [35]), and then the same can be said for the determinant of the matrix (since it is equal to the product of its eigenvalues).

Theorem B.18 *Let $B \in \mathbb{M}_{n\times n}$ be a positive definite matrix and let $u, v \in \mathbb{R}^n$ be two vectors such that $u^T v \geq 0$. If any of u and v is null or if the following condition holds:*

$$\frac{\mu_m^s(B)}{\|B\|} \geq \sqrt{1 - \left(\frac{u^T v}{\|u\|\|v\|}\right)^2}, \tag{73}$$

then matrix $M(\alpha) := B + \alpha\, uv^T$ is nonsingular (in particular its determinant is positive) for all $\alpha \geq 0$.

Proof. First of all we can say that $M(0) = B$ is non singular by the positive definiteness of B, and the same can be said if any of u and v is null. Moreover we can also say that if $u = \rho v$ with $\rho \geq 0$ then uv^T is positive semidefinite and then, by the positive definiteness of B, matrix $M(\alpha)$ is positive definite (and then nonsingular) for all $\alpha \geq 0$. Then we must consider only the case in which $\alpha > 0$, $0 \leq \frac{u^T v}{\|u\|\|v\|} < 1$ and both u and v are non null.

We first consider the case in which $u^T v > 0$. Let $\{b_1, \ldots, b_n\}$ be the basis of non null right unit eigenvectors of matrix uv^T defined in Proposition B.15. Let us suppose that the thesis does not hold, that is a non null vector $z \in \mathbb{R}^n$ and a scalar $\bar{\alpha} > 0$ exist such that

$$M(\bar{\alpha})z = Bz + \bar{\alpha}\, uv^T z = 0_n. \tag{74}$$

Vector z can be expressed by using the vectors $\{b_1, \ldots, b_n\}$ and the scalars $\{\beta_1, \ldots, \beta_n\}$:

$$z := \sum_{i=1}^{n} \beta_i b_i.$$

We can write the following:

$$\begin{aligned}\bar{\alpha}\, uv^T z &= \bar{\alpha}\, uv^T \left(\sum_{i=1}^{n} \beta_i b_i\right) \\ &= \bar{\alpha}\, uv^T(\beta_1 b_1) \qquad (75) \\ &= \bar{\alpha}\beta_1 \frac{v^T u}{\|u\|}\, u,\end{aligned}$$

where (75) follows from the fact that $v \perp b_i$ for all $i = 2, \ldots, n$. Then, by (74), we can conclude that

$$Bz = \gamma u, \tag{76}$$

where $\gamma := -\bar{\alpha}\beta_1 \frac{v^T u}{\|u\|}$. Moreover by the positive defineteness of B, we can say that $\gamma \neq 0$ and then $\beta_1 \neq 0$.

The following chain of equalities holds:

$$\begin{aligned}\gamma z^T u &= \gamma \left(\beta_1 b_1 + \beta_n b_n + \sum_{i=2}^{n-1} \beta_i b_i\right)^T u \\ &= \gamma \left(\beta_1 \frac{u}{\|u\|} + \beta_n \frac{1}{\|u\|\|v\|\sqrt{1 - \frac{(u^T v)^2}{\|u\|^2\|v\|^2}}} \left(\|v\| u - \frac{u^T v}{\|v\|} v\right) + \sum_{i=2}^{n-1} \beta_i b_i\right)^T u \\ &= \gamma \left(\beta_1 \|u\| + \beta_n \frac{\|u\|}{\sqrt{1 - \frac{(u^T v)^2}{\|u\|^2\|v\|^2}}} - \beta_n \frac{(u^T v)^2}{\|u\|\|v\|^2\sqrt{1 - \frac{(u^T v)^2}{\|u\|^2\|v\|^2}}}\right) \\ &= \gamma \|u\| \left(\beta_1 + \beta_n \frac{1}{\sqrt{1 - \frac{(u^T v)^2}{\|u\|^2\|v\|^2}}} - \beta_n \frac{(u^T v)^2}{\|u\|^2\|v\|^2\sqrt{1 - \frac{(u^T v)^2}{\|u\|^2\|v\|^2}}}\right) \\ &= \gamma \|u\| \left(\beta_1 + \beta_n \sqrt{1 - \frac{(u^T v)^2}{\|u\|^2\|v\|^2}}\right).\end{aligned}$$

Then by (76) we obtain

$$\gamma \|u\| \left(\beta_1 + \beta_n \sqrt{1 - \frac{(u^T v)^2}{\|u\|^2\|v\|^2}}\right) = z^T B z \geq \mu_m^s(B)\|z\|^2,$$

which is equivalent to

$$\left(-\bar{\alpha}\beta_1^2 v^T u =\right)\ \gamma\|u\|\beta_1 \geq \mu_m^s(B)\|z\|^2 - \gamma\|u\|\beta_n\sqrt{1-\frac{(u^Tv)^2}{\|u\|^2\|v\|^2}}. \qquad (77)$$

Remembering that $\bar{\alpha} > 0$, $\beta_1 \neq 0$ and $u^Tv > 0$ we obtain $-\bar{\alpha}\beta_1^2 v^T u < 0$. Then, by the fact that $\mu_m^s(B)\|z\|^2 \geq 0$ and $\sqrt{1-\frac{(u^Tv)^2}{\|u\|^2\|v\|^2}} \geq 0$, we obtain that, in order to satisfy (77), the following must hold

$$\gamma\|u\|\beta_n > 0. \qquad (78)$$

Since by the triangular inequality

$$\begin{aligned}|\beta_n| = \|\beta_n b_n\| &\leq \|\beta_1 b_1\| + \|\beta_1 b_1 + \beta_n b_n\| = |\beta_1| + \sqrt{\beta_1^2+\beta_n^2+2\beta_1\beta_n b_1^T b_n}\\ &\leq |\beta_1| + \sqrt{\beta_1^2+\beta_n^2+\sum_{i=2}^{n-1}\beta_i^2+2\beta_1\beta_n b_1^T b_n} = |\beta_1| + \|z\|,\end{aligned}$$

then by (76) and (78)

$$\begin{aligned}\gamma\|u\|\beta_n &= |\gamma|\|u\||\beta_n| \leq |\gamma|\|u\||\beta_1| + |\gamma|\|u\|\|z\|\\ &= |\gamma|\|u\||\beta_1| + \|Bz\|\|z\| \leq |\gamma|\|u\||\beta_1| + \|B\|\|z\|^2,\end{aligned}$$

and then finally we can rewrite inequality (77):

$$\gamma\|u\|\beta_1 \geq \mu_m^s(B)\|z\|^2 - \left(|\gamma|\|u\||\beta_1| + \|B\|\|z\|^2\right)\sqrt{1-\frac{(u^Tv)^2}{\|u\|^2\|v\|^2}},$$

that is

$$\gamma\|u\|\beta_1\left(1-\sqrt{1-\frac{(u^Tv)^2}{\|u\|^2\|v\|^2}}\right) \geq \mu_m^s(B)\|z\|^2 - \|B\|\|z\|^2\sqrt{1-\frac{(u^Tv)^2}{\|u\|^2\|v\|^2}}. \qquad (79)$$

Since the left-hand side of (79) is negative while, by assumption (73), the right-hand one is nonnegative, then we can conclude that z and $\bar{\alpha}$ can not exist and this part of the proof is complete.

Now we consider the case in which $u^Tv = 0$. Also in this part of the proof we suppose that a non null vector $z \in \mathbb{R}^n$ and a scalar $\bar{\alpha} > 0$ exist such that (74) holds. However we now express z by using a different basis that is exactly the same as in the previous part of this proof except for the fact that $b_1 := \frac{v}{\|v\|}$ and $b_n := \frac{u}{\|u\|}$, then in this case the basis $\{b_1, \ldots, b_n\}$ is orthonormal. It is not difficult to see that (76) holds but in this case $\gamma := -\bar{\alpha}\beta_1\|v\|$. And again we can say that $\gamma \neq 0$ and then $\beta_1 \neq 0$. Moreover it is not difficult to see that $\gamma z^T u = \gamma\|u\|\beta_n$ and then

$$\gamma\|u\|\beta_n = z^T Bz \geq \mu_m^s(B)\|z\|^2 \geq 0.$$

Since $(0 \leq)\, \gamma\|u\|\beta_n = |\gamma|\|u\||\beta_n|$, $|\beta_n| \leq \|z\|$ (being the basis orthonormal) and $|\gamma|\|u\| = \|B\|\|z\|$ then the following must hold

$$1 \geq \frac{|\gamma|\|u\||\beta_n|}{\|B\|\|z\|^2} \geq \frac{\mu^s_m(B)\|z\|^2}{\|B\|\|z\|^2} \geq 1,$$

where the last inequality follows from assumption (73). Then we can conclude that $|\beta_n| = \|z\|$ but this contradicts the fact that $\beta_1 \neq 0$ and finally we can say that $M(\alpha)$ is nonsingular for all $\alpha \geq 0$.

Since $M(\alpha)$ is nonsingular for all $\alpha \geq 0$, $\det(M(0)) > 0$ (being $B \succ 0$) and remembering that for any matrix its determinant varies continuously with its entries (see Remark B.17), we can conclude that $\det(M(\alpha)) > 0$ for all $\alpha \geq 0$. Finally the proof is complete. □

Note that Theorem B.18 states the matrix $M(\alpha) = B + \alpha uv^T$ to be nonsingular for all $\alpha \geq 0$ under some assumptions on B, u and v, but it does not say anything about the positiveness of that matrix. In fact let $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, then assumptions of Theorem B.18 are satisfied, but it is easy to see that $M(2) = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}$ is not P.

Theorem B.19 *Let $B \in \mathbb{M}_{n\times n}$, $u, v \in \mathbb{R}^n$ and let $M(\alpha) := B + \alpha uv^T$ be singular then the following statements hold:*

(a) if B is nonsingular, then $M(\alpha - \beta)$ is nonsingular for all $\beta \neq 0$;

(b) if $B \succ 0$ and $\alpha > 0$, then $\det(M(\alpha - \beta)) > 0$ *for all $\beta > 0$.*

Proof. (a) By assumption a non null vector $z \in \mathbb{R}^n$ exists such that $M(\alpha)z = 0_n$ that is $Bz = -\alpha(v^T z)u$ and then being B nonsingular we obtain

$$z = \gamma B^{-1} u, \tag{80}$$

where $\gamma := -\alpha(v^T z)$. Note that $\alpha \neq 0$, both u and v must be non null and $v^T z \neq 0$, otherwise we would have $Bz = 0_n$ contradicting the nonsingularity of B.

Now suppose that $M(\alpha - \beta)$ is singular that is a non null vector $y \in \mathbb{R}^n$ exists such that $By = -(\alpha - \beta)(v^T y)u$ and then

$$y = \rho B^{-1} u, \tag{81}$$

where $\rho := -(\alpha - \beta)(v^T y)$. Note that $(\alpha - \beta) \neq 0$ and $v^T y \neq 0$, otherwise again we would have $Bz = 0_n$ contradicting the nonsingularity of B. It is not difficult to see that

$$M(\alpha - \beta)y = \frac{\rho}{\gamma}\left(B + \alpha uv^T\right)\gamma B^{-1}u - \beta uv^T y = -\beta(v^T y)u \neq 0_n,$$

but this means that y does not exists and then $M(\alpha - \beta)$ is nonsingular for all $\beta \neq 0$.

(b) By using result (a) we know that $M(\alpha - \beta)$ is nonsingular for all $\beta > 0$. Since $det(M(0)) > 0$ (being $B \succ 0$) and remembering that for any matrix its determinant varies continuously with its entries (see Remark B.17), we can conclude that $\det(M(\alpha - \beta)) > 0$ for all $\beta > 0$. $\square$

Note that results of Theorem B.19 can be derived also by using the well known Sherman-Morrison formula that we report below for readers convenience.

Proposition B.20 *(Sherman-Morrison) Let $B \in \mathbb{M}_{n\times n}$ be nonsingular and $u, v \in \mathbb{R}^n$. If $1 + v^T B^{-1} u \neq 0$, then $B + uv^T$ is nonsingular and*

$$(B + uv^T)^{-1} = B^{-1} - \frac{B^{-1}uv^T B^{-1}}{1 + v^T B^{-1} u}.$$

References

[1] Baillon, J.-B., Haddad, G.: Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones. Isr. J. Math. 26, 137–150 (1977)

[2] Baiocchi, C., Capelo, A.: Variational and Quasivariational Inequalities: Applications to Free Boundary Problems. Wiley, New York (1984)

[3] Bensoussan, A.: Points de Nash dans le cas de fonctionnelles quadratiques et jeux differentiels lineaires a N personnes. SIAM J. Control 12, 460–499 (1974)

[4] Bensoussan, A., Goursat, M., Lions, J.-L.: Contrôle impulsionnel et inéquations quasi-variationnelles stationnaires. C. R. Acad. Sci. Paris Sér. A 276, 1279–1284 (1973)

[5] Bensoussan, A., Lions, J.-L.: Nouvelle formulation de problèmes de contrôle impulsionnel et applications. C. R. Acad. Sci. Paris Sér. A 276, 1189–1192 (1973)

[6] Bensoussan, A., Lions, J.-L.: Nouvelles méthodes en contrôle impulsionnel. Appl. Math. Optim. 1, 289–312 (1975)

[7] Beremlijski, P., Haslinger, J., Kocvara, M., Outrata, J.: Shape optimization in contact problems with Coulomb friction. SIAM J. Optim. 13, 561–587 (2002)

[8] Bliemer, M., Bovy, P.: Quasi-variational inequality formulation of the multiclass dynamic traffic assignment problem. Transp. Res. B: Methodol. 37, 501–519 (2003)

[9] Chan, D., Pang, J.-S.: The generalized quasi-variational inequality problem. Math. Oper. Res. 7, 211–222 (1982)

[10] Cottle, R.W., Pang, J.-S., Stone, R.E.: The Linear Complementarity Problem. Academic Press, New York (1992)

[11] De Luca, M., Maugeri, A.: Discontinuous quasi-variational inequalities and applications to equilibrium problems. In: Giannessi, F. (eds.) Nonsmooth Optimization: Methods and Applications, pp. 70–75. Gordon and Breach, Amsterdam (1992)

[12] Dirkse, S.P., Ferris, M.C.: The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems. Optim. Methods Softw. 5, 123–156 (1995)

[13] Dreves, A., Facchinei, F., Kanzow, C., Sagratella, S.: On the solution of the KKT conditions of generalized Nash equilibrium problems. SIAM J. Optim. 21, 1082–1108 (2011)

[14] Dreves, A., Kanzow, C., Stein, O.: Nonsmooth optimization reformulations for player convex generalized Nash equilibrium problems. Preprint 300, Institute of Mathematics, University of Würzburg, Germany, October (2010)

[15] Donato, M.B., Milasi, M., Vitanza, C.: An existence result of a quasi-variational inequality associated to an equilibrium problem. J. Glob. Optim. 40, 87–97 (2008)

[16] Facchinei, F., Fischer, A., Piccialli, V.: Generalized Nash equilibrium problems and Newton methods. Math. Program. 117, 163–194 (2009)

[17] Facchinei, F., Kanzow, C.: Generalized Nash equilibrium problems. 4OR 5, 173–210 (2007)

[18] Facchinei, F., Kanzow, C.: Penalty methods for the solution of generalized Nash equilibrium problems (with complete test problems). Technical Report, Institute of Mathematics, University of Würzburg, Germany, February (2009)

[19] Facchinei, F., Kanzow, C.: Penalty methods for the solution of generalized Nash equilibrium problems. SIAM J. Optim. 20, 2228–2253 (2010)

[20] Facchinei, F., Kanzow, C.: Generalized Nash equilibrium problems. Ann. Oper. Res. 175, 177–211 (2010)

[21] Facchinei, F., Kanzow, C., Sagratella, S.: QVILIB: A Library of Quasi-Variational Inequality Test Problems. Pac. J. Optim. (to appear)

[22] Facchinei, F., Lampariello, L.: Partial penalization for the solution of generalized Nash equilibrium problems. J. Global Optim. (2010)

[23] Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, Volumes I and II. Springer, New York (2003)

[24] Facchinei, F., Pang, J.-S.: Nash equilibria: the variational approach. In Convex Optimization in Signal Processing and Communications, Palomar, D.P., Eldar, Y.C., ed., Cambridge University Press, 443–493 (2010)

[25] Facchinei, F., Sagratella, S.: On the computation of all solutions of jointly convex generalized Nash equilibrium problems. in print in Optim. Lett. 5 (3), 531–547 (2011)

[26] Ferris, M.C., Munson, T.S.: Interfaces to PATH 3.0: Design, implementation and usage. Comput. Optim. Appl. 12, 207–227 (1999)

[27] Fukushima, M.: A class of gap functions for quasi-variational inequality problems. J. Ind. Manag. Optim. 3, 165–171 (2007)

[28] Fukushima, M.: Restricted generalized Nash equilibria and controlled penalty algorithm. Comput. Manag. Sci. 8, 201–218 (2011)

[29] Glowinski, R.: Numerical methods for variational inequalities. Springer, Berlin (2008, reprint of the 1984 edition)

[30] Glowinski, R., Lions, J.L., and Trémolières, R.: Numerical analysis of variational inequalities. North-Holland, Amsterdam (1981)

[31] Hammerstein, P., Selten, R.: Game theory and evolutionary biology. In: Aumann, R.J., Hart, S. (eds.) Handbook of Game Theory with Economic Applications, Vol. 2, pp. 929–993. North-Holland, Amsterdam (1994)

[32] Harker, P.T.: Generalized Nash games and quasi-variational inequalities. Eur. J. Oper. Res. 54, 81–94 (1991)

[33] Haslinger, J., Panagiotopoulos, P.D.: The reciprocal variational approach to the Signorini problem with friction. Approximation results. Proc. Royal Soc. Edinburgh: Sect. A Math. 98, 365–383 (1984)

[34] Hintermüller, M., Rautenberg, C.N.: A sequential minimization technique for elliptic quasi-variational inequalities with gradient constraints. IFP-Report No. 45, Institute of Mathematics and Scientific Computing, University of Graz, Austria, June (2011)

[35] Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1990)

[36] Ichiishi, T.: Game Theory for Economic Analysis. Academic Press, New York (1983)

[37] Kharroubi, I., Ma, J., Pham, H., Zhang, J.: Backward SDEs with constrained jumps and quasi-variational inequalities. Ann. Probab. 38, 794–840 (2010)

[38] Kravchuk, A.S., Neittaanmaki, P.J.: Variational and quasi-variational inequalities in mechanics. Springer, Dordrecht (2007)

[39] Kubota, K., Fukushima, M.: Gap function approach to the generalized Nash equilibrium problem. J. Optim. Theory Appl. 144, 511–531 (2010)

[40] Kunze, M., Rodrigues, J.F.: An elliptic quasi-variational inequality with gradient constraints and some of its applications. Math. Methods Appl. Sci. 23, 897–908 (2000)

[41] Luna, J.P., Sagastizábal, C., Solodov, M.V.: A class od Dantzig-Wolfe type decomposition methods for variational inequality problems. IMPA Preprint A706, Instituto de Matematica Pura e Aplicada, Rio de Janeiro, November (2011)

[42] Meyer, C.D.: Matrix Analysis and Applied Linear Algebra. Society for Industrial and Applied Mathematics, Philadelphia (2000)

[43] Monteiro, R.D.C., Pang, J.-S.: A potential reduction Newton method for constrained equations. SIAM J. Optim. 9, 729–754 (1999)

[44] Mosco, U.: Implicit variational problems and quasi variational inequalities. In: Gossez, J., Lami Dozo, E., Mawhin, J., Waelbroeck, L. (eds.) Nonlinear Operators and the Calculus of Variations, Lecture Notes in Mathematics Vol. 543, pp. 83–156. Springer, Berlin (1976)

[45] Nesterov, Y., Scrimali, L.: Solving strongly monotone variational and quasi-variational inequalities. CORE Discussion Paper 2006/107, Catholic University of Louvain, Center for Operations Research and Econometrics (2006)

[46] Noor, M.A.: On general quasi-variational inequalities. J. King Saud Univ. 24, 81–88 (2012)

[47] Noor, M.A.: Existence results for quasi variational inequalities. Banach J. Math. Anal. 1, 186–194 (2007)

[48] Noor, M.A.: An iterative scheme for a class of quasi variational inequalities. J. Math. Anal. Appl. 110, 463–468 (1985)

[49] Noor, M.A., Oettli, W.: On general nonlinear complementarity problems and quasi-equilibria. Le Matematiche XLIX, 313–331 (1994)

[50] Ortega, J.M., Rheinboldt, J.C.: Iterative solution of nonlinear equations in several variables. Academic Press, New York (1970) (reprinted by SIAM, Philadelphia (2000))

[51] Outrata, J., Kocvara, M.: On a class of quasi-variational inequalities. Optim. Methods Softw. 5, 275–295 (1995)

[52] Outrata, J., Kocvara, M., Zowe, J.: Nonsmooth approach to optimization problems with equilibrium constraints. Kluwer Academic Publishers, Dordrecht and Boston (1998)

[53] Outrata, J., Zowe, J.: A Newton method for a class of quasi-variational inequalities. Comput. Optim. Appl. 4, 5–21 (1995)

[54] Pang, J.-S.: The implicit complementarity problem. In: Mangasarian, O.L., Meyer, R.R., Robinson, S.M. (eds.) Nonlinear Programming 4, pp. 487–518. Academic Press, New York (1981)

[55] Pang, J.-S.: On the convergence of a basic iterative method for the implicit complementarity problem. J. Optim. Theory Appl. 37, 149–162 (1982)

[56] Pang, J.-S.: Error bounds in mathematical programming. Math. Program. 79, 299–332 (1997)

[57] Pang, J.-S.: Computing generalized Nash equilibria. Technical Report, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD, October (2002)

[58] Pang, J.-S., Fukushima, M.: Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. Computational Management Science 2, 21–56 (2005) (Erratum: ibid 6, 373–375 (2009))

[59] Ryazantseva, I.P.: First-order methods for certain quasi-variational inequalities in Hilbert space. Comput. Math. Math. Phys 47, 183–190 (2007)

[60] Scrimali, L.: Quasi-variational inequality formulation of the mixed equilibrium in mixed multiclass routing games. CORE Discussion Paper 2007/7, Catholic University of Louvain, Center for Operations Research and Econometrics (2007)

[61] Siddiqi, A.H., Ansari, Q.H.: Strongly nonlinear quasivariational inequalities. J. Math. Anal. Appl. 149, 444–450 (1990)

[62] Wei, J.Y., Smeers, Y.: Spatial oligopolistic electricity models with Cournot generators and regulated transmission prices. Oper. Res. 47, 102–112 (1999)