THE SEMISMOOTH NEWTON METHOD FOR THE SOLUTION OF QUASI-VARIATIONAL INEQUALITIES

Francisco Facchinei[1], Christian Kanzow[2],
Sebastian Karl[2], and Simone Sagratella[1]

Preprint 319 October 2013

[2]University of Rome La Sapienza
Department of Computer, Control,
and Management Engineering
Via Ariosto 25
00185 Roma
Italy
e-mail: facchinei@dis.uniroma1.it
sagratella@dis.uniroma1.it

[2]University of Würzburg
Institute of Mathematics
Emil-Fischer-Str. 30
97074 Würzburg
Germany
e-mail: kanzow@mathematik.uni-wuerzburg.de
sebastian.karl@mathematik.uni-wuerzburg.de

October 16, 2013

Abstract. We consider the application of the globalized semismooth Newton method to the solution of (the KKT conditions of) quasi variational inequalities. We show that the method is globally and locally superlinearly convergent for some important classes of quasi variational inequality problems. We report numerical results to illustrate the practical behavior of the method.

Key Words: Quasi-variational inequality; KKT conditions; Semismooth method; Global convergence; Superlinear convergence.

1 Introduction

We consider the following *quasi-variational inequality* QVI(K, F): Find a vector $x^* \in K(x^*)$ such that

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

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a (point-to-point) mapping and $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a point-to-set mapping with closed and convex images. We will always assume that F is C^1 and that the feasible set mapping K is given by a parametric set of inequality constraints:

$$K(x) \triangleq \{y \in \mathbb{R}^n \mid g(y, x) \leq 0\}, \tag{2}$$

where $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ is twice continuously differentiable and, for each $i = 1, \ldots, m$, $g_i(\cdot, x)$ is convex on $\mathbb{R}^n$, for each $x \in \mathbb{R}^n$. We could easily consider also linear equality constraints, but for simplicity of notation we drop them from our subsequent analysis.

QVIs were introduced by Bensoussan and Lions [5, 6, 7] and they are a powerful tool that can be used to model complex equilibrium situations that appear in many different fields such as generalized Nash games, mechanics, physics, economics, statistics, transportation, and biology, see e.g. [3, 4, 8, 9, 14, 26, 27, 30, 31, 33, 34, 36, 41, 42, 46, 49, 51, 54] and references therein. The interested reader can consult the Mosco [37] and Baiocchi and Capelo [2] monographs for a more comprehensive description of the properties of QVIs and their applicability.

Due to the difficulty of the topic, developments of algorithms for the solution of QVIs are still scarce and lag far behind those of methods for Variational Inequalities (VIs). Motivated by earlier research on the implicit complementarity problem [37, 44, 45], Chan and Pang proposed in [10] what is probably the first globally convergent algorithm for a QVI. In this seminal paper the authors use a fixed point argument to show convergence of a projection-type algorithm in the case in which $K(x) = c(x) + Q$, where Q is a closed convex set and $c : \mathbb{R}^n \to \mathbb{R}^n$ a mapping satisfying certain conditions. Most of the subsequent papers, where globally convergent algorithms are analyzed, consider variants or extensions of the basic setting proposed in [10] and also follow a fixed point approach, see e.g. [38, 40, 39, 50, 52] and references therein.

In a departure from this setting, Pang and Fukushima [46] introduced a sequential penalty approach to the solution of general QVIs. The method in [46] reduces the solution of a QVI to the solution of a sequence of VIs; however, in spite of its many merits, the global convergence properties of

this method are in jeopardy since they ultimately hinge of the capability of solving a sequence of possibly very challenging VIs.

Another approach for the solution of QVIs is based on a reformulation as a constrained or unconstrained global optimization problem with the help of so-called gap functions, see [1, 15, 25, 28, 29, 35, 53] for some contributions in this direction. In general, these gap functions are nonsmooth which makes it difficult to solve the corresponding optimization problem. In some cases, the gap function is smooth, cf. [15, 28, 29], but this is true only for certain classes of QVIs, and global convergence is still difficult to guarantee.

More recently an interior point approach in [20] was shown to enjoy very favorable global convergence properties, in that convergence can be established not only for the case $K(x) = c(x) + Q$, but, for the first time, for several other classes of interesting QVIs.

If we turn our attention to local (Newton-type) methods, the pioneering work was done by Outrata and co-workers who studied some fast local methods using the semismooth method or other nonsmooth algorithms to the solution of suitable reformulations of a QVI [41, 42, 43]. While local fast convergence can be established under suitable assumptions, the globalization of these semismooth Newton methods is difficult and not discussed in these papers.

Very recently a so called LP-Newton method has been developed which can successfully be applied to nonsmooth systems of equations with non-isolated solutions [17, 18]. These features of the LP-Newton method seem very well-suited to the solution of (the KKT conditions of) QVIs, where nonisolatedness and nonsmoothness arise naturally. Indeed, the application of the LP-Newton method to some particular QVIs arising from generalized Nash equilibrium problems has been described in [16]. Although the LP-Newton method leads to fast local methods in this case, it seems that it is suitable only for problems which are not too large, since otherwise the computational cost of each iteration can become excessive.

The aim of this paper is to investigate the *global and local* convergence behavior of a semismooth Newton method for the solution of the the KKT system of a QVI. Although the fast local convergence rate of the semismooth Newton methods hinges on assumptions that are stronger than those required by the LP-Newton method, it essentially calls for the (possibily inexact) solution of a linear system at each iteration, thus making it a practically attractive alternative to the computationally more intensive LP-Newton method. Furthermore, its globalization is rather simple, thus overcoming open issues related to the approaches in [41, 42, 43]. The globalized Newton method is a well-established technique for the solution of KKT-like systems and its use in the solution of the KKT system arising from a QVI is rather natu-

ral and, in principle, simple, its applicability depending in essence on the non-singularity of certain matrices. Unfortunately, some of the results in [19] show that in some significant cases arising from game theory, these matrices are inherently singular, thus questioning the use of the semismooth approach in the solution of QVIs.

The main focus of this paper is then an in-depth analysis of these nonsingularity issues. Building on the recent results in [20], we are able to show that nonsingularity can be guaranteed under very reasonable assumptions for important classes of QVIs, thus providing for the first time a firm theoretical basis to the use of semismooth methods in the solution of the KKT conditions of QVIs. We also report numerical experiences on the same set of test problems used in [20]. The semismooth Newton method turns out to be less robust than the interior-point method in [20]. However, when started from a point near a solution, convergence is very fast and a high-precision in the computation of a solution can be achieved, a precision which is usually impossible to get with the interior-point method. Furthermore, when global convergence occurs, the semismooth Newton method is usually much faster than the interior-point method. This suggests that the interior-point method of [20] and the globalized semismooth Newton method could be suitably combined to obtain a robust and efficient solution method. Preliminary numerical results in this direction are also discussed.

The paper is organized as follows. In the next section we describe the globalized semismooth Newton method and recall its main features. Section 3 contains the main theoretical results, all related to some nonsingularity issues key to the globalized semismooth Newton method. In Section 4 we report the results of our numerical tests and finally, in Section 5 we draw our conclusions.

Notation: All vector and matrix norms are l_2 norms. Positive definite matrices are not necessarily symmetric. The symbol $\mu_m^s(A)$ denotes the minimum eigenvalue of the symmetric part $\frac{1}{2}(A + A^T)$ of a matrix $A \in \mathbb{R}^{n\times n}$. For a continuously differentiable function $F : \mathbb{R}^n \to \mathbb{R}^m$, we write $JF(x)$ for the Jacobian of F at a point $x \in \mathbb{R}^n$, whereas $\nabla F(x)$ denotes the transposed Jacobian. If F is only locally Lipschitz continuous, then $\partial F(x)$ is the generalized Jacobian of F at x in the sense of Clarke [11]. Similarly, given a smooth mapping $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m, (y, x) \mapsto g(y, x)$, then $\nabla_y g(y, x)$ denotes the transpose of the partial Jacobian of g with respect to the y-variables. For a vector $x \in \mathbb{R}^n$ and a subset $I \subseteq \{1, \ldots, n\}$, we write x_I for the subvector consisting of the elements $x_i, i \in I$. Furthermore, for a matrix $A \in \mathbb{R}^{n\times n}$ and two subsets $I, J \subseteq \{1, \ldots, n\}$, the symbol A_{IJ} stands for the submatrix with entries a_{ij} for $i \in I, j \in J$. On the other hand, $A_{I\bullet}$ contains all rows of the matrix A that belong to the index set I, whereas $A_{\bullet J}$ consists of all columns of A corresponding to the index set J. The matrix $A \in \mathbb{R}^{n\times n}$ is then called

a P-matrix if $\det(A_{II}) > 0$ holds for all index sets $I \subseteq \{1, \ldots, n\}$. Finally, $x \circ y$ is the Hadamard (pointwise) product of two vectors $x, y \in \mathbb{R}^n$, and x^{-1} denotes the pointwise inverse of a vector whose components are nonzero.

2 The semismooth Newton method

In this section we briefly review the globalized semismooth Newton method as applied to the solution of the KKT system of a QVI and taking for granted familiarity with semismoothness and related issues, see [22, 47, 48]. Although, as far as we are aware of, this specific application has never been explicitly considered previously, the resulting method can certainly be considered standard. Consequently, all results in this section are given without proof, since they can be derived easily from known results, see for example [12, 13, 22], with very minor modifications.

We say that a point $x \in \mathbb{R}^n$ satisfies the KKT conditions if multipliers $\lambda \in \mathbb{R}^m$ exist such that

$$\begin{array}{l} F(x) + \nabla_y g(x,x)\lambda = 0, \\ \lambda_i \geq 0,\ g_i(x,x) \leq 0, \lambda_i g_i(x,x) = 0 \quad \forall i = 1, \ldots, m. \end{array} \tag{3}$$

Note that $g(x,x) \leq 0$ means that $x \in K(x)$ and that by $\nabla_y g(x,x)$ we indicate the partial transposed 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 [22], and it is quite easy to show that if a point x, together with a suitable vector $\lambda \in \mathbb{R}^m$ 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 $\lambda \in \mathbb{R}^m$ exist such that the pair (x, λ) satisfies the KKT conditions (3). 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.

As made clear above, there is a very close relationship between a QVI and its corresponding KKT conditions. Our aim is then to find a solution x^* of the QVI by solving the KKT conditions (3). To this end, let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ denote the Fischer-Burmeister function

$$\varphi(a,b) := \sqrt{a^2 + b^2} - a - b,$$

cf. [24]. Its most interesting property is the fact that its zeros can be char-

acterized in the following way:

$$\varphi(a,b)=0 \Longleftrightarrow a\geq 0, b\geq 0, ab=0.$$

It therefore follows that (x^*,λ^*) is a KKT point of the QVI if and only if (x^*,λ^*,w^*) with $w^*:=-h(x^*)$ is a solution of the nonlinear (and nonsmooth) system of equations

$$H(x,\lambda,w)=0, \quad \text{with} \quad H(x,\lambda,w):=\begin{pmatrix} L(x,\lambda) \\ h(x)+w \\ \Phi(\lambda,w) \end{pmatrix}, \tag{4}$$

where, for simplicity of notation, we set

$$\begin{aligned} L(x,\lambda) &:= F(x)+\nabla_y g(x,x)\lambda, \\ h(x) &:= g(x,x), \\ \Phi_i(\lambda,w) &:= \varphi(\lambda_i,w_i) \quad \forall i=1,\ldots,m; \end{aligned}$$

the mapping L defined in this way is called the *Lagrangian* of the QVI. Associated with the system $H(x,\lambda,w)=0$, we consider also its natural merit function:

$$\Psi(z):=\frac{1}{2}\|H(z)\|^2, \tag{5}$$

where we set $z:=(x,\lambda,w)$.

The following proposition collects some differentiability properties that are central to the semismooth approach.

Proposition 2.1 *Let the mapping H be defined by* (4). *Then the following statements hold:*

(a) If F is continuously differentiable and g is twice continuously differentiable, then H is semismooth and

$$\partial H(x,\lambda,w)\subseteq \left\{ \begin{pmatrix} J_x L(x,\lambda) & \nabla_y g(x,x) & 0 \\ Jh(x) & 0 & I \\ 0 & diag(a) & diag(b) \end{pmatrix} \right\},$$

where $diag(a):=diag(a_1,\ldots,a_m)$, $diag(b):=diag(b_1,\ldots,b_m)$, $(a_i,b_i)\in \partial\varphi(\lambda_i,w_i)$ $\forall i=1,\ldots,m$ and

$$\partial\varphi(\lambda_i,w_i)=\begin{cases} \left(\frac{\lambda_i}{\|(\lambda_i,w_i)\|}-1, \frac{w_i}{\|(\lambda_i,w_i)\|}-1\right), & \text{if } (\lambda_i,w_i)\neq(0,0), \\ \{(\xi_i-1,\zeta_i-1) \mid \|(\xi_i,\zeta_i)\|\leq 1\}, & \text{if } (\lambda_i,w_i)=(0,0) \end{cases}$$

is the usual subdifferential of the convex function φ.

(b) If, in addition, JF and $\nabla^2 g_i$ $(i = 1, \ldots, m)$ are locally Lipschitz, then H is strongly semismooth.

(c) Let the merit function Ψ be defined by (5). If F is continuously differentiable and g is twice continuously differentiable, Ψ is continuously differentiable, and its gradient is given by

$$\nabla\Psi(z) = V^T H(z)$$

for an arbitrary element $V \in \partial H(z)$.

Essentially, the semismooth Newton method is nothing else but a standard globalization of the semismooth Newton method from [47, 48] for the system of equations $H(z) = 0$, which uses Ψ as a merit function. The algorithm below follows closely that first proposed in [12].

Algorithm 2.2 *Globalized Semismooth Method* `(S.0)`: *Choose $z^0 = (x^0, \lambda^0, w^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, $\rho > 0, \beta \in (0,1), \sigma \in \left(0, \frac{1}{2}\right), p > 2, \varepsilon \geq 0$, and set $k := 0$.*

`(S.1)`: *If $\|\nabla\Psi(z^k)\| \leq \varepsilon$: STOP.*

`(S.2)`: *Choose an arbitrary element $V_k \in \partial H(z^k)$, and compute d^k as a solution of the linear system of equations*

$$V_k d = -H(z^k). \tag{6}$$

If either this system is not solvable or the sufficient decrease condition

$$\nabla\Psi(z^k)^T d^k \leq -\rho\|d^k\|^p$$

is not satisfied, then take $d^k := -\nabla\Psi(z^k)$.

`(S.3)`: *Compute a stepsize t_k as the maximum of the numbers $\beta^\ell, \ell = 0, 1, 2, \ldots$, such that the following Armijo condition holds:*

$$\Psi(z^k + t_k d^k) \leq \Psi(z^k) + \sigma t_k \nabla\Psi(z^k)^T d^k.$$

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

The main global convergence result for Algorithm 2.2 is stated in the following theorem. It assumes implicitly that $\varepsilon = 0$ and that Algorithm 2.2 does not terminate after finitely many iterations with an exact stationary point.

Theorem 2.3 *Every accumulation point of a sequence* $\{z^k\} = \{(x^k, \lambda^k, w^k)\}$ *is a stationary point of the merit function* Ψ*. Furthermore, let* $z^* = (x^*, \lambda^*, w^*)$ *be an accumulation point of* $\{z^k\}$ *such that all elements* $V \in \partial H(z^*)$ *are nonsingular. Then the following statements hold:*

(a) The entire sequence $\{z^k\}$ *converges to* z^**.*

(b) The search direction d^k *is eventually given by the solution of the linear system* (6)*.*

(c) The full stepsize $t_k = 1$ *is accepted for all* k *sufficiently large.*

(d) The rate of convergence is superlinear.

(e) If, in addition, JF *and* $\nabla^2 g_i$ $(i = 1, \ldots, m)$ *are locally Lipschitz, then the rate of convergence is quadratic.*

The above theorem raises two central questions: (a) when is a stationary point of Ψ a solution of the QVI? (b) under what conditions can we guarantee superlinear convergence of the method? It turns out that both questions are related to the nonsingularity of certain matrices. The study of these nonsingularity conditions turns out to be by no means trivial and is the subject of the following section.

3 Nonsingularity Conditions

Theorem 2.3 guarantees convergence of Algorithm 2.2 to stationary points of the merit function Ψ. As we already mentioned, the first key question we want to investigate is when such a stationary point is a solution of the QVI. The following result unlocks the possibility of applying in a direct and simple way the results in [20] to this issue.

Proposition 3.1 *Let* $z = (x, \lambda, w) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ *be an arbitrary point. Then the gradient of* Ψ *at* z *can be represented as*

$$\nabla \Psi(z) = \tilde{V}^T H(z),$$

where $\tilde{V}$ *is defined by*

$$\tilde{V} := \begin{pmatrix} J_x L(x, \lambda) & \nabla_y g(x, x) & 0 \\ Jh(x) & 0 & I \\ 0 & diag(\tilde{a}) & diag(\tilde{b}) \end{pmatrix}$$

with $diag(\tilde{a}) := diag(\tilde{a}_1, \ldots, \tilde{a}_m)$, $diag(\tilde{b}) := diag(\tilde{b}_1, \ldots, \tilde{b}_m)$ *and*

$$\tilde{a}_i := \begin{cases} \frac{\partial \varphi}{\partial a}(\lambda_i, w_i) & \text{if } \varphi(\lambda_i, w_i) \neq 0, \\ -1 & \text{if } \varphi(\lambda_i, w_i) = 0, \end{cases}$$

$$\tilde{b}_i := \begin{cases} \frac{\partial \varphi}{\partial b}(\lambda_i, w_i) & \text{if } \varphi(\lambda_i, w_i) \neq 0, \\ -1 & \text{if } \varphi(\lambda_i, w_i) = 0. \end{cases}$$

Proof. By Proposition 2.1 (c) we already know that $\partial\Psi(z) = V^T\Psi(z)$, for any $V \in \partial H(z)$. Note that the only possible difference between an element $V \in \partial H(z)$ and the matrix $\tilde{V}$ is in the diagaonl entries a_i, b_i of V and $\tilde{a}_i, \tilde{b}_i$ of $\tilde{V}$, all other entries obviously coincide. However, since we have $\nabla\Psi(z) = V^T H(z)$, it follows that, in this representation, the diagonal elements a_i and b_i of V^T get post-multiplied by the component $\varphi(\lambda_i, w_i)$. Hence, if this element is zero, we can replace the corresponding diagonal entry a_i and b_i by any number (in our case, we take -1) without changing the gradient of Ψ. □

Suppose now that $z^* = (x^*, \lambda^*, w^*)$ is an accumulation point of a sequence generated by Algorithm 2.2. Then Theorem 2.3 guarantees that z^* is a stationary point of Ψ. Taking into account Proposition 3.1, we therefore have $0 = \nabla\Psi(z^*) = \tilde{V}_*^T H(z^*)$ for a suitable matrix $\tilde{V}_*$ as defined in that result. Now, if this matrix $\tilde{V}_*$ is nonsingular, we would immediately obtain that $H(z^*) = 0$ holds, i.e., that $z^* = (x^*, \lambda^*, w^*)$ is a KKT point and therefore that x^* itself is a solution of the QVI. Of course, in general, one cannot expect the matrix $\tilde{V}_*$ to be nonsingular. However, the definition of $\tilde{V}_*$ implies that all elements of the two diagonal matrices with entries $\tilde{a}_i$ and $\tilde{b}_i$ are strictly negative (note that a partial derivative of φ can be zero only if the function value of φ is equal to zero). Hence $\tilde{V}_*$ has exactly the same structure as the standard Jacobian matrices that arise within the potential-reduction interior-point method for QVIs as discussed in the recent paper [20]. Consequently, all the nonsingularity conditions given in [20] in that framework can now be applied to our case in order to guarantee that a stationary point is a solution of the underlying QVI. Proposition 3.1, together with Theorem 2.3, therefore shows that global convergence of Algorithm 2.2 can be guaranteed under conditions that are structurally very similar to those needed by the interior point method in [20].

From the above there would seem to be no clear theoretical advantage to use Algorithm 2.2 instead of the interior point method in [20]. The reason of our interest in the semismooth method stems from the fact that, as indicated in Theorem 2.3, this simple algorithm can achieve a superlinear, or even quadratic, local convergence rate. Neither the interior point method in [20]

nor any other method we are aware of has the property that it is both globally and locally fast convergent, with the only exception of [16] where a hybrid method for the solution of generalized Nash equilibrium problems, that can be reformulated as QVI with a particular structure, is developed. However, as mentioned in the introduction, the method [16] is rather complex and not suitable for large scale problems.

We are therefore naturally led to consider the second question we asked at the end of the previous section: under what conditions can we guarantee the superlinear convergence of the method or, more concretely, taking into account Theorem 2.3, under which assumptions can we guarantee that all elements from the generalized Jacobian $\partial H(z^*)$ are nonsingular? Answering this question is not at all easy, and the results in the remaining part of this section constitute the main theoretical contribution of this paper. The leading idea is to follow the arguments given in [20], but the situation here is different because in [20], only points $z = (x, \lambda, w)$ with $(\lambda, w) > 0$ had to be considered, whereas here we are interested in nonsingularity results at a KKT point $z = z^*$ whose components λ^* and w^* might have zero entries. Hence the assumptions have to be different and, furthermore, one has to expect stronger assumptions than in [20] since, in our semismooth Newton setting, the nonsingularity of $\partial H(z^*)$ implies a local superlinear/quadratic rate of convergence, whereas the corresponding nonsingularity statements in the paper [20] only guarantee global convergence.

The first main nonsingularity result is given in the following theorem. This theorem is still rather abstract and difficult to interpret; therefore we will successively consider several concrete settings and show how this main theorem can be used by exploting specific problem structures to obtain more concrete results. These specific cases we will consider are the same ones considered in [20] and, as it will be clear, they are of great practical interest.

Theorem 3.2 *Let $z^* = (x^*, \lambda^*, w^*)$ be a KKT point and let*

$$\delta := \{i \mid w_i^* = 0\} = \{i \mid h_i(x^*) = 0\} = \{i \mid g_i(x^*, x^*) = 0\} \tag{7}$$

be the index set of active inequality constraints. Suppose that $J_x L(x^*, \lambda^*)$ is nonsingular and that the submatrix $M(x^*, \lambda^*)_{\delta\delta}$ is a P-matrix, where

$$M(x^*, \lambda^*) := Jh(x^*) J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*). \tag{8}$$

Then all elements from $\partial H(z^*)$ are nonsingular.

Proof. Choose an arbitrary $V \in \partial H(z^*)$. According to Proposition 2.1, we have

$$V = \begin{pmatrix} J_x L(x^*, \lambda^*) & \nabla_y g(x^*, x^*) & 0 \\ Jh(x^*) & 0 & I \\ 0 & \mathrm{diag}(a) & \mathrm{diag}(b) \end{pmatrix}$$

with $(a_i, b_i) \in \partial\varphi(\lambda_i^*, w_i^*)$ for all $i = 1, \ldots, m$. We then define the index sets

$$I := \{i \mid b_i < 0\} \quad \text{and} \quad J := \{1, \ldots, m\} \setminus I = \{i \mid b_i = 0\}$$

(recall from Proposition 2.1 that $b_i \leq 0$ holds for all $i = 1, \ldots, m$). Moreover, note that $i \notin \delta$ implies $w_i^* > 0$, hence $\lambda_i^* = 0$ in view of the KKT conditions which, in turn, gives $b_i = 0$, so that $i \notin I$. It therefore follows that $I \subseteq \delta$. Furthermore, from the expression of the subdifferential for the Fischer-Burmeister function in Proposition 2.1, we immediately obtain

$$b_i < 0, a_i \leq 0 \ \ \forall i \in I, \qquad b_i = 0, a_i < 0 \ \ \forall i \in J. \tag{9}$$

Now consider the homogeneous linear system $Vq = 0$ with a suitably partitioned vector $q = \left(q^{(1)}, q^{(2)}, q^{(3)}\right)$; we get

$$\begin{aligned} J_x L(x^*, \lambda^*) q^{(1)} + \nabla_y g(x^*, x^*) q^{(2)} &= 0, & (10) \\ Jh(x^*) q^{(1)} + q^{(3)} &= 0, & (11) \\ \mathrm{diag}(a) q^{(2)} + \mathrm{diag}(b) q^{(3)} &= 0. & (12) \end{aligned}$$

In view of (9), we obtain from (12) that

$$q_J^{(2)} = 0 \tag{13}$$

and

$$q_I^{(3)} = -\mathrm{diag}(b_I)^{-1} \mathrm{diag}(a_I) q_I^{(2)}. \tag{14}$$

Since $J_x L(x^*, \lambda^*)$ is nonsingular, we obtain

$$q^{(1)} = -J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*) q^{(2)} \tag{15}$$

from (10). Inserting this into (11) yields

$$-Jh(x^*) J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*) q^{(2)} + q^{(3)} = 0. \tag{16}$$

Looking only at the components in the index set I, we obtain from (16)

$$\begin{aligned} 0 &= -\left[Jh(x^*) J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*) q^{(2)}\right]_I + q_I^{(3)} \\ &= -\left[Jh(x^*) J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*)\right]_{II} q_I^{(2)} \\ &\quad -\left[Jh(x^*) J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*)\right]_{IJ} q_J^{(2)} + q_I^{(3)} \\ &= -\left[Jh(x^*) J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*)\right]_{II} q_I^{(2)} + q_I^{(3)}, \end{aligned}$$

where the last equality follows from (13). Together with (14), we get

$$\begin{aligned} 0 &= \Big(\big[Jh(x^*) J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*) \big]_{II} + \mathrm{diag}(b_I)^{-1} \mathrm{diag}(a_I) \Big) q_I^{(2)} \\ &= \big(M(x^*, \lambda^*)_{II} + \mathrm{diag}(b_I)^{-1} \mathrm{diag}(a_I) \big) q_I^{(2)}. \end{aligned} \tag{17}$$

Since $M(x^*, \lambda^*)_{\delta\delta}$ is a P-matrix and $I \subseteq \delta$, it follows that $M(x^*, \lambda^*)_{II}$ is also a P-matrix. Since $\mathrm{diag}(b_I)^{-1}\mathrm{diag}(a_I)$ is a positive semidefinite diagonal matrix, it therefore follows from a known characterization of P-matrices, see, e.g., [32], that $M(x^*, \lambda^*)_{II} + \mathrm{diag}(b_I)^{-1}\mathrm{diag}(a_I)$ is nonsingular. Hence (17) yields $q_I^{(2)} = 0$ which, together with (13), immediately gives $q^{(1)} = 0$ in view of (15). We then also have $q^{(3)} = 0$ because of (11). Hence $q = 0$, so that V is nonsingular. □

Note that the matrix (or suitable submatrices of) $M(x^*, \lambda^*)$ also plays a central role in the related semismooth Newton works by Outrata and co-workers [41, 43]. These papers use a completely different reformulation of the QVI, and the authors therefore need some assumptions which are different from ours, but a central condition also exploited in that paper (though written down using a different notation) is the nonsingularity of certain submatrices of $M(x^*, \lambda^*)$.

3.1 Moving Set Case

We now consider the most popular class of QVIs, where the set $K(x)$ is defined by

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

for some function $c : \mathbb{R}^n \to \mathbb{R}^n$ and a fixed set $Q \subseteq \mathbb{R}^n$ which we assume to have a representation of the form

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

with another mapping $q : \mathbb{R}^n \to \mathbb{R}^m$ such that each component function q_i is convex. Throughout this subsection, we assume that c is at least continuously differentiable and q is at least twice continuously differentiable. It is easy to see that we can rewrite the set $K(x)$ as

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

Putting this into our general framework, this means that we are dealing with QVIs whose constraints are given by the function

$$g(y, x) := q(y - c(x)), \tag{18}$$

so that the mapping h is given by

$$h(x) := q(x - c(x)). \tag{19}$$

We now exploit this particular structure in order to derive a nonsingularity result in the moving set case from the general condition given in Theorem 3.2.

The proof of the subsequent result uses the fact that the inverse of a positive definite matrix is again positive definite. In principle, this statement is well-known. However, we recall that, in our case, we do not require symmetry when speaking of a positive definite matrix. That the above statement also holds in the possibly nonsymmetric case can be seen as follows: Let A be positive definite. Since $d^T Ad = d^T A^T d$ for all $d \in \mathbb{R}^n$, this implies that A^T is also positive definite. It therefore follows from

$$d^T A^{-1} d = d^T A^{-1} A^T A^{-T} d = (A^{-T} d)^T A^T (A^{-T} d) > 0 \quad \forall d \neq 0$$

that A^{-1} is indeed positive definite. This observation allows us to state the following result.

Theorem 3.3 *Consider the above setting, and let $z^* = (x^*, \lambda^*, w^*)$ be a KKT point. Suppose that $JF(x^*)$ is positive definite, the gradients $\nabla q_i(x^* - c(x^*))$ $(i \in \delta)$ are linearly independent, with δ being the index set from (7), and that*

$$\|Jc(x^*)\| < \frac{\mu_m^s(JF(x^*)^{-1})}{\|JF(x^*)^{-1}\|} \tag{20}$$

holds. Then all elements $V \in \partial H(z^*)$ are nonsingular.

Proof. First note that (20) implies that $\|Jc(x^*)\| < 1$, hence $I - Jc(x^*)$ is nonsingular. We claim that

$$(I - Jc(x^*))JF(x^*)^{-1} \text{ is positive definite.} \tag{21}$$

In fact, we have for all $v \in \mathbb{R}^n \setminus \{0\}$ that

$$\begin{aligned} v^T Jc(x^*)JF(x^*)^{-1} v &\leq \|Jc(x^*)\| \, \|JF(x^*)^{-1}\| \, \|v\|^2 \\ &< \mu_m^s(JF(x^*)^{-1})\|v\|^2 \quad \text{from (20)} \\ &\leq v^T JF(x^*)^{-1} v, \end{aligned}$$

which immediately gives (21). Hence the corresponding inverse matrix

$$JF(x^*)(I - Jc(x^*))^{-1}$$

is also positive definite. Since we are in a KKT point, we also have $\lambda_i^* \geq 0$, and the convexity of q_i yields the positive semidefiniteness of the Hessians $\nabla^2 q_i(x^* - c(x^*))$. We therefore obtain that the matrix

$$J_x L(x^*, \lambda^*)(I - Jc(x^*))^{-1} = JF(x^*)(I - Jc(x^*))^{-1} + \sum_{i=1}^{m} \lambda_i^* \nabla^2 q_i(x^* - c(x^*))$$

is positive definite, too. In particular, this matrix is nonsingular, which in turn implies that $J_x L(x^*, \lambda^*)$ is nonsingular. In view of Theorem 3.2, it therefore remains to show that the submatrix $M(x^*, \lambda^*)_{\delta\delta}$ is a P-matrix, where $M(x^*, \lambda^*)$ is defined by (8) and δ denotes the index set of active inequalities from (7).

To investigate the properties of the matrix $M(x^*, \lambda^*)$, recall that we have

$$M(x^*, \lambda^*) = Jh(x^*) J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*),$$

with

$$\nabla_y g(x^*, x^*) = \nabla q(x^* - c(x^*)) \quad \text{and} \quad Jh(x^*) = Jq(x^* - c(x^*))(I - Jc(x^*))$$

by (18), (19). We can therefore rewrite the matrix $M(x^*, \lambda^*)$ as

$$M(x^*, \lambda^*) = Jq(x^* - c(x^*))(I - Jc(x^*)) J_x L(x^*, \lambda^*)^{-1} \nabla q(x^* - c(x^*)).$$

This implies

$$[M(x^*, \lambda^*)]_{\delta\delta} = Jq(x^* - c(x^*))_{\delta\bullet}(I - Jc(x^*)) J_x L(x^*, \lambda^*)^{-1} \nabla q(x^* - c(x^*))_{\bullet\delta}.$$

Since $(I - Jc(x^*)) J_x L(x^*, \lambda^*)^{-1}$ is the inverse of a matrix that was already shown to be positive definite, we obtain that the submatrix $M(x^*, \lambda^*)_{\delta\delta}$ is positive definite, hence a P-matrix, from the assumed linear independence constraint qualification. □

Note that in condition (20) neither the function q nor the variables λ and w are involved. The fact that q is not involved indicates that the nonsingularity of ∂H is not related to the shape of the set Q, but only to the trajectory the moving set follows. More precisely, condition (20) stipulates that the trajectory described by c be not "too steep", where the exact meaning of "too steep" is given by (20).

We also note that the other assumptions, namely the positive definiteness of $JF(x^*)$ and the linear independence condition, are also used in the related framework discussed in [41, 43]. Hence, although there are no well-accepted assumptions for QVIs that guarantee local fast convergence, simply because there are not many methods with this property, it seems that these two conditions are somewhat natural. In fact, both assumptions will also be used in the subsequent results (at least implicitly), whereas the condition on the mapping c (or g, in the general setting) and therefore on the shape of the feasible set will be different since it depends very much on the particular structure of the parameterized set $K(x)$.

3.2 Constraints with Variable Right-hand Side

Here we consider the case where the set $K(x)$ is defined by

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

for some twice continuously differentiable mappings $c, q : \mathbb{R}^n \to \mathbb{R}^m$ such that each component q_i is convex. Hence the set $K(x)$ is described by the system of inequalities q with a variable right-hand-side $c(x)$ depending on x.

For constraints given in this form, we have the following nonsingularity result.

Theorem 3.4 *Consider the above setting, and let $z^* = (x^*, \lambda^*, w^*)$ be a KKT point. Suppose that $JF(x^*)$ is positive definite, and that*

$$\|Jc(x^*)_{\delta\bullet}\| < \frac{\mu_m^s([Jq(x^*)J_xL(x^*,\lambda^*)^{-1}\nabla q(x^*)]_{\delta\delta})}{\|J_xL(x^*,\lambda^*)^{-1}\|\,\|Jq(x^*)_{\delta\bullet}\|} \tag{22}$$

holds, where δ denotes the index set from (7). Then all elements $V \in \partial H(z^*)$ are nonsingular.

Proof. We have

$$\nabla_y g(x^*,x^*) = \nabla q(x^*), \;\; Jh(x^*) = Jq(x^*) - Jc(x^*),$$
$$J_xL(x^*,\lambda^*) = JF(x^*) + \textstyle\sum_{i=1}^m \lambda_i^* \nabla^2 q_i(x^*).$$

Since all Hessian matrices $\nabla^2 q_i(x^*)$ are positive semidefinite due to the assumed convexity of each q_i, and since $\lambda_i^* \geq 0$ for all $i = 1, \ldots, m$ since we are in a KKT point, it follows immediately from the positive definiteness of $JF(x^*)$ that the Jacobian $J_xL(x^*,\lambda^*)$ is also positive definite, hence nonsingular. In view of Theorem 3.2, it therefore suffices to show that the submatrix

$[M(x^*,\lambda^*)]_{\delta\delta}$ is a P-matrix, where $M(x^*,\lambda^*)$ denotes the matrix defined in (8) which, in our current setting, is given by

$$M(x^*,\lambda^*) = \big(Jq(x^*) - Jc(x^*)\big) J_x L(x^*,\lambda^*)^{-1} \nabla q(x^*),$$

hence we have

$$M(x^*,\lambda^*)_{\delta\delta} = \big(Jq(x^*)_{\delta\bullet} - Jc(x^*)_{\delta\bullet}\big) J_x L(x^*,\lambda^*)^{-1} Jq(x^*)^T_{\delta\bullet}.$$

For an arbitrary nonzero vector $v \in \mathbb{R}^{|\delta|}$, we obtain from (22)

$$\begin{aligned} v^T Jc(x^*)_{\delta\bullet} J_x L(x^*,\lambda^*)^{-1} Jq(x^*)^T_{\delta\bullet} v \\ &\le \|Jc(x^*)_{\delta\bullet}\| \cdot \|J_x L(x^*,\lambda^*)^{-1}\| \bullet \|Jq(x^*)_{\delta\bullet}\| \cdot \|v\|^2 \\ &< \mu^s_m\big(Jq(x^*)_{\delta\bullet} J_x L(x^*,\lambda^*)^{-1} Jq(x^*)^T_{\delta\bullet}\big) \cdot \|v\|^2 \\ &\le v^T Jq(x^*)_{\delta\bullet} J_x L(x^*,\lambda^*)^{-1} Jq(x^*)^T_{\delta\bullet} v. \end{aligned}$$

This implies $v^T[M(x^*,\lambda^*)]_{\delta\delta} v > 0$ for all nonzero v, so that, in particular, $[M(x^*,\lambda^*)]_{\delta\delta}$ is a P-matrix. □

In the linear case where $q(y) = Ey - b$ for some matrix $E \in \mathbb{R}^{m\times n}$ and a fixed vector $b \in \mathbb{R}^m$, we have $J_x L(x,\lambda) = JF(x)$ and $Jq(y) = E$. Hence we obtain the following special case as an immediate corollary of Theorem 3.4.

Corollary 3.5 *Consider the linear case with $q(y) = Ey - b$, and let $z^* = (x^*,\lambda^*,w^*)$ be a KKT point. Suppose that $JF(x^*)$ is positive definite, and that*

$$\|Jc(x^*)_{\delta\bullet}\| < \frac{\mu^s_m([EJF(x^*)^{-1}E^T]_{\delta\delta})}{\|JF(x^*)^{-1}\|\,\|E_{\delta\bullet}\|} \tag{23}$$

holds, where δ denotes the index set from (7). Then all elements $V \in \partial H(z^*)$ are nonsingular.

The meaning of the conditions (22) and (23) is that the right-hand-side $c(x)$ should not vary "too quickly". In contrast to the moving set case, considered in the previous subsection though, this does not necessarily imply that $K(x)$ changes "slowly" with x, since a polyhedron, for example, can have abrupt changes when the right-hand side changes even slightly.

Let us close this section with a few of comments on the condition (23) (similar comments hold for the nonlinear case (22)).

Given a matrix $A \in \mathbb{R}^{n\times n}$ and arbitrary index sets $I, J \subseteq \{1,\dots,n\}$, it is easy to see that $\|A_{IJ}\| \le \|A\|$ and $\mu^s_m(A_{II}) \ge \mu^s_m(A)$. Consequently, the condition

$$\|Jc(x^*)\| < \frac{\mu^s_m(EJF(x^*)^{-1}E^T)}{\|JF(x^*)^{-1}\| \cdot \|E\|} \tag{24}$$

implies that the central assumption (23) from the previous result holds. Note, however, that the condition from (24) is typically much stronger than the requirement from (23). We further note that assumption (23) can hold only if the matrix $E_{\delta\bullet}$ has full rank.

3.3 Bilinear Constraints

In this subsection, we consider the case where the feasible set $K(x)$ is defined by the mapping

$$g(y,x) := \begin{pmatrix} q_1(y) \\ \vdots \\ q_p(y) \\ x^TQ_1y - c_1 \\ \vdots \\ x^TQ_by - c_b \end{pmatrix},$$

where each mapping $q_i : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable and convex, $Q_i \in \mathbb{R}^{n\times n}$ are symmetric positive semidefinite matrices for all $i = 1,\ldots,b$, and $c = (c_1,\ldots,c_b)^T \in \mathbb{R}^b$ denotes a fixed vector. This structure of the feasible set mapping 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.

For this class of QVIs, we have the following central nonsingularity result.

Theorem 3.6 *Consider the above setting, and let $z^* = (x^*,\lambda^*,w^*)$ be a KKT point. Suppose that $JF(x^*)$ is positive definite, and let the active gradients $\nabla_y g_i(x^*,x^*)$ $(i \in \delta)$ be linearly independent, where δ denotes the index set from (7). Then all elements $V \in \partial H(z^*)$ are nonsingular.*

Proof. Note that, in the current setting, it follows that

$$\nabla h(x) = \nabla_y g(x,x)D \tag{25}$$

with the diagonal matrix

$$D := \operatorname{diag}\big(\underbrace{1,\ldots,1}_{p-\text{times}},\underbrace{2,\ldots,2}_{b-\text{times}}\big). \tag{26}$$

This observation will be crucial within our proof.

We further note that, in the bilinear case, we have

$$J_xL(x^*,\lambda^*) = JF(x^*) + \sum_{i=1}^{p}\lambda_i^*\nabla^2 q_i(x^*) + \sum_{i=1}^{b}\lambda_{p+i}^*Q_i.$$

Since $JF(x^*)$ is positive definite by assumption, $\lambda_i^* \geq 0$ because $z^* = (x^*, \lambda^*, w^*)$ is a KKT point, Q_i is positive semidefinite, and since each $\nabla^2 q_i(x^*)$ is also positive semidefinite due to the convexity of q_i, it follows that $J_x L(x^*, \lambda^*)$ is positive definite and, therefore, also nonsingular.

In order to apply once again Theorem 3.2, it remains to show that the submatrix $[M(x^*, \lambda^*)]_{\delta\delta}$ is a P-matrix, where $M(x^*, \lambda^*)$ is defined in (8). In our particular setting, exploiting (25), we have

$$M(x^*, \lambda^*) = D\nabla_y g(x^*, x^*)^T J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*),$$

hence

$$[M(x^*, \lambda^*)]_{\delta\delta} = D_{\delta\delta} \nabla_y g(x^*, x^*)_{\bullet\delta}^T J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*)_{\bullet\delta}.$$

Since $J_x L(x^*, \lambda^*)$ was already shown to be positive definite, also the inverse $J_x L(x^*, \lambda^*)^{-1}$ is positive definite. The linear independence assumption implies that $\nabla_y g(x^*, x^*)_{\bullet\delta}$ has full column rank. Hence

$$\nabla_y g(x^*, x^*)_{\bullet\delta}^T J_x L(x^*, \lambda^*)^{-1} \nabla_y g(x^*, x^*)_{\bullet\delta}$$

is positive definite, too. In particular, it is a P-matrix. But then it follows immediately from the positive definiteness of the diagonal matrix D that $[M(x^*, \lambda^*)]_{\delta\delta}$ is also a P-matrix, and this completes the proof. □

We stress that, in the previous proof, it is quite important that the submatrix $[M(x^*, \lambda^*)]_{\delta\delta}$ has to be a P-matrix only and not positive definite (which would have been sufficient for the other proofs). This has to do with the fact that, for a nonsymmetric matrix A, the positive definiteness of A does, in general, not imply that DA is also positive definite for a positive diagonal matrix D. To see this, consider the matrices

$$A := \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix} \quad \text{and} \quad D := \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

(note that D has precisely the structure of the diagonal matrix from (26)). Then it is easy to see that A is indeed positive definite, whereas DA is obviously a P-matrix, but not positive definite since, e.g., we have $x^T DAx = 0$ for the nonzero vector $x := (1, 1)^T$.

4 Numerical Results

In this section we report the results obtained by a Matlab implementation of the semismooth Newton method on QVILIB, a varied set of QVI test

problems [21]. In order to show the practical behavior of the semismooth Newton method on QVI problems and to compare it with the state-of-the-art interior-point method in [20], we report numerical experiences on the same set of test problems from QVILIB that is used in [20]. All the computations in this paper 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.

4.1 Implementation Details

The algorithm we implemented corresponds to the scheme given in Algorithm 2.2, in what follows we give implementation details.

At step `(S.2)`, Algorithm 2.2 calls for the solution of an $n + 2m$ square linear system in order to determine the search direction d^k. This system is very structured and some simple manipulations permit to reduce its solution to a linear system of lower dimension. More precisely, we must find a solution $(\bar{d}_1, \bar{d}_2, \bar{d}_3)$ of the following linear system of dimension $n + 2m$

$$\begin{pmatrix} J_x L(x,\lambda) & \nabla_y g(x,x) & 0 \\ Jh(x) & 0 & I \\ 0 & \operatorname{diag}(a) & \operatorname{diag}(b) \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} -L(x,\lambda) \\ -h(x) - w \\ -\Phi(\lambda, w) \end{pmatrix}, \qquad (27)$$

where all the quantities involved are defined in detail in Section 2. Let us consider the following index sets:

$$S := \{s \in \{1,\dots,m\} \,:\, a_s = 0\}, \quad J := \{j \in \{1,\dots,m\} \,:\, b_j = 0\},$$
$$K := \{k \in \{1,\dots,m\} \,:\, a_k \neq 0,\, b_k \neq 0\}.$$

Then we can rewrite system (27) in the following way

$$\begin{pmatrix} J_x L(x,\lambda) & \nabla_y g(x,x)_{\bullet K} & 0 & \nabla_y g(x,x)_{\bullet S} & 0 \\ Jh(x)_{K\bullet} & 0 & I & 0 & 0 \\ 0 & \operatorname{diag}(a_K) & \operatorname{diag}(b_K) & 0 & 0 \\ Jh(x)_{S\bullet} & 0 & 0 & 0 & 0 \\ Jh(x)_{J\bullet} & 0 & 0 & 0 & I \end{pmatrix} \begin{pmatrix} d_1 \\ (d_2)_K \\ (d_3)_K \\ (d_2)_S \\ (d_3)_J \end{pmatrix} =$$
$$= \begin{pmatrix} -L(x,\lambda) - \nabla_y g(x,x)_{\bullet J}(\bar{d}_2)_J \\ -h(x)_K - w_K \\ -\Phi(\lambda, w)_K \\ -h(x)_S - w_S - (\bar{d}_3)_S \\ -h(x)_J - w_J \end{pmatrix},$$

where

$$(\bar{d}_2)_J = -\operatorname{diag}(a_J)^{-1}\Phi(\lambda, w)_J \quad \text{and} \quad (\bar{d}_3)_S = -\operatorname{diag}(b_S)^{-1}\Phi(\lambda, w)_S.$$

It is easy to verify, by substitution and by the fact that all components of a_K and b_K are different from zero, that if we compute $(\bar{d}_1, (\bar{d}_2)_S, (\bar{d}_3)_J)$ as a solution of

$$\begin{pmatrix} A & \nabla_y g(x,x)_{\bullet S} & 0 \\ Jh(x)_{S\bullet} & 0 & 0 \\ Jh(x)_{J\bullet} & 0 & I \end{pmatrix} \begin{pmatrix} d_1 \\ (d_2)_S \\ (d_3)_J \end{pmatrix} = \begin{pmatrix} B \\ -h(x)_S - w_S - (\bar{d}_3)_S \\ -h(x)_J - w_J \end{pmatrix},$$

where

$$\begin{aligned} A &= J_x L(x,\lambda) + \nabla_y g(x,x)_{\bullet K} \mathrm{diag}(a_K^{-1} \circ b_K) Jh(x)_{K\bullet}, \\ B &= -L(x,\lambda) - \nabla_y g(x,x)_{\bullet J} (\bar{d}_2)_J + \nabla_y g(x,x)_{\bullet K} \mathrm{diag}(a_K)^{-1} \Phi(\lambda, w)_K - \\ &\quad \nabla_y g(x,x)_{\bullet K} \mathrm{diag}(a_K^{-1} \circ b_K)(h(x)_K + w_K), \end{aligned}$$

and $(\bar{d}_2)_K$, $(\bar{d}_3)_K$ by

$$\begin{aligned} (\bar{d}_3)_K &= -h(x)_K - w_K - J_x h(x)_{K\bullet} \bar{d}_1, \\ (\bar{d}_2)_K &= -\mathrm{diag}(a_K)^{-1} \Phi(\lambda, w)_K - \mathrm{diag}(a_K^{-1} \circ b_K)(\bar{d}_3)_K, \end{aligned}$$

this is indeed a solution of (27). This shows clearly that the main computational burden in solving the linear system (27) is actually the solution of an $(n + |S| + |J|) \times (n + |S| + |J|)$ square linear system. As can be easily expected, and as it is confirmed by the numerical experiments, usually in early iterations $|S| + |J| = 0$ and the core system has dimension n. However, as iterations progress and we get closer to a KKT point, $|S| + |J|$ increases and tends to approach m.

If any entry of the solution given by *mldivide* is a NaN or it is equal to $\pm\infty$ or the sufficient decrease condition is not satisfied then a scaled anti gradient direction is used:

$$d^k := -\tau^k \nabla \Psi(z^k), \quad \tau^k := \min\left\{1, \frac{2\max\{10^{-6}, \Psi(z^{k-1}) - \Psi(z^k)\}}{\|\nabla \Psi(z^k)\|_2^2}\right\}.$$

In order to obtain an element of $\partial H(z)$ we set $a_i = -1$ and $b_i = -1$ for all $i \in \{1, \ldots, m\}$ such that $\|(a_i, b_i)\|_2 \leq 10^{-30}$. When we compute the set of indices S, J, K used to reduce the linear system, we consider $a_i = 0$ (or $b_i = 0$) whenever $a_i \geq -10^{-30}$ ($b_i \geq -10^{-30}$) (recall again that these numbers are always nonpositive).

We further take $\rho = 10^{-10}$, $\beta = 0.5$, $\sigma = 10^{-2}$, and $p = 2.1$. The starting points are taken with $\lambda^0 = 0$ and $w^0 = 0$, while x^0 is defined separately for each run.

Unlike the scheme of Algorithm 2.2, the stopping criterion is based on an equation reformulation of the KKT conditions which uses the Fischer-

Burmeister function. The equation reformulation is then defined by

$$Y(x,\lambda) = \left\| \begin{array}{c} L(x,\lambda) \\ \Phi(\lambda, -h(x)) \end{array} \right\|_\infty .$$

The main termination criterion is $Y(x^k, \lambda^k) \leq 10^{-4}$. We chose this stopping criterion, which is actually the most natural one, being directly connected to the violation of the KKT conditions, also to be able to perform a fair comparison with the results in [20], where the same termination criterion is used. The iterations are also stopped if the number of iterations exceeds 1000, the running time exceeds one hour or the stepsize t_k computed at step `(S.3)` is less than 10^{-6}.

4.2 Numerical results

We solved several test problems whose detailed description can be found in [21]. In Table 1 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);
- the number of iterations, which is equal to the number of evaluations of $\nabla\Psi$;
- the number of evaluations of Ψ;
- the value of the KKT violation measure $Y(x,\lambda)$ at termination;
- elapsed CPU time in seconds.

We see that overall Algorithm 2.2 fails on 16 runs that is a third of the experiments. These failures are due to time limit (KunR32), iteration limit (KunR31 and RHS1A1 [with x^0=10]) or computed stepsize too small (all other failures). In cases where the termination is due to time limit or iteration limit, the algorithm uses massively the antigradient direction which proves to be ineffective. Therefore we tried to recover these failures by doing further tests with $\rho = 10^{-30}$, thus forcing the use of the semismooth Newton direction, however, the same time/iteration failures occured. In failures due to small stepsize, Algorithm 2.2 uses only the semismooth Newton direction (except for RHS1A1 and RHS2B1). We tried to recover these failures by setting $\rho = 10^{-4}$ in order to use earlier the antigradient direction when the

Problem	x^0	It	Ψ	$Y(x,\lambda)$	Time	Problem	x^0	It	Ψ	$Y(x,\lambda)$	Time
OutZ40	0	5	5	2.5270e-08	<0.1	Wal2	10	Failure			
OutZ41	0	5	5	2.8078e-08	<0.1	Wal3	0	Failure			
OutZ42	0	6	7	9.2896e-07	<0.1	Wal3	10	29	86	1.0008e-06	0.3
OutZ43	0	4	4	5.2459e-05	<0.1	Wal5	0	67	712	3.4386e-07	2.7
OutZ44	0	4	4	4.9362e-05	<0.1	Wal5	10	Failure			
OutKZ31	0	7	10	2.6128e-06	<0.1	MovSet3A1	0	Failure			
OutKZ31	10	7	7	2.6128e-06	<0.1	MovSet3A2	0	Failure			
OutKZ41	0	11	21	1.7394e-05	<0.1	MovSet3B1	0	Failure			
OutKZ41	10	17	86	1.6720e-05	0.1	MovSet3B2	0	Failure			
Scrim22	0	10	20	3.5249e-06	0.2	MovSet4A1	0	11	27	2.6602e-06	0.3
Scrim22	10	10	20	3.5249e-06	0.2	MovSet4A2	0	13	43	4.0769e-07	1.6
KunR11	0	14	26	7.1161e-05	25	MovSet4B1	0	11	28	8.0836e-07	0.4
KunR11	10	22	53	4.7736e-05	43.3	MovSet4B2	0	13	42	8.7163e-07	1.6
KunR12	0	20	53	5.1480e-05	144.7	Box2A	10	24	77	4.1771e-06	0.8
KunR12	10	19	38	3.4970e-05	120.4	Box2B	10	26	89	8.1684e-06	0.8
KunR21	0	5	5	2.8047e-05	8	Box3A	10	10	14	5.7049e-07	0.2
KunR21	10	14	28	1.1084e-05	25.8	Box3B	10	Failure			
KunR22	0	5	5	2.4567e-05	27.4	RHS1A1	0	Failure			
KunR22	10	13	22	3.6380e-06	78	RHS1A1	10	Failure			
KunR31	0	Failure				RHS2B1	0	Failure			
KunR31	10	Failure				RHS2B1	10	Failure			
KunR32	0	Failure				Ex3	10	5	5	7.4873e-09	<0.1
KunR32	10	Failure				Ex7	10	14	14	5.9159e-05	<0.1
Wal2	0	20	78	2.2062e-07	0.2	Ex8	10	5	5	4.6680e-08	<0.1

Table 1: Numerical results for Algorithm 2.2.

semismooth Newton direction starts to be almost orthogonal to $\nabla\Psi$, but the algorithm got into iteration limit failure in all cases. Difficulties were encountered by the linear system solver only when dealing with problems OutKZ41 and Wal5, in fact, as a thorough analysis showed, iterations of Algorithm 2.2, while solving these problems, produce linear systems with 1-norm condition number estimate of the matrix greater than 10^{35}. However, as Table 1 shows the antigradient direction works well in these cases by bringing iterations to better conditioned areas.

We can numerically compare Algorithm 2.2 with the interior-point method in [20] by confronting Table 1 in this text with Tables 2 and 3 in [20]. We can see that Algorithm 2.2 is significantly less robust than the interior-point method (16 failures versus 4), but when it works, it is more efficient both in terms of iterations and CPU time.

It is then quite natural to attempt to combine the robustness of the interior-point method in [20] and the fast convergence properties of the semismooth Newton method by defining a hybrid algorithm. This kind of hybrid techniques, in the field of variational inequalities, go back at least to [23] and have often given very good good results; the definition of such a hybrid method certainly requires further study. Below, to illustrate the potential of such a combination, we tried a very simple approach: run the interior-point method until iterations are near a solution and then leave to the semismooth method the task of fastly computing a solution with high accuracy. In what follows we are interested in computing solutions (x^*, λ^*, w^*) with high accuracy, that is $Y(x^*, \lambda^*) \leq 10^{-8}$. We report both results for the interior point method as it is described in [20] and then for the natural hybrid algorithm in

which the interior point method iterates until $Y(x,\lambda) \leq 10^{-3}$ and then the semismooth method goes on until termination, see Table 2.

		Standard $Y \leq 10^{-8}$			Hybrid $Y \leq 10^{-3}$ / $Y \leq 10^{-8}$					
		interior-point method			interior-point method			Semismooth method		
Problem	x^0	It	$Y(x,\lambda)$	Time	It	$Y(x,\lambda)$	Time	It	$Y(x,\lambda)$	Time
OutZ40	0	12	7.4975e-09	<0.1	7	6.8145e-04	<0.1	2	2.7242e-14	<0.1
OutZ41	0	Failure			14	9.9912e-04	<0.1	1	1.5950e-14	<0.1
OutZ42	0	12	1.4442e-09	<0.1	7	1.5321e-04	<0.1	2	7.0531e-09	<0.1
OutZ43	0	12	1.8886e-09	<0.1	7	2.1270e-04	<0.1	3	4.4409e-16	<0.1
OutZ44	0	12	2.9244e-09	<0.1	7	3.3333e-04	<0.1	2	9.0487e-09	<0.1
OutKZ31	0	Failure			17	3.6904e-04	<0.1	2	1.0387e-09	<0.1
OutKZ31	10	Failure			16	1.8527e-04	<0.1	2	7.3199e-10	<0.1
OutKZ41	0	Failure			18	8.6946e-04	<0.1	16	1.5657e-13	0.1
OutKZ41	10	Failure			18	7.9010e-04	<0.1	5	2.6164e-09	<0.1
Scrim22	0	Failure			16	2.1102e-04	0.2	4	1.6533e-12	0.1
Scrim22	10	Failure			19	8.9414e-05	0.2	4	1.6185e-11	0.1
KunR11	0	Failure			10	9.4101e-04	13.7	37	4.2274e-11	100.6
KunR11	10	Failure			14	9.0437e-04	21.2	30	9.2655e-09	79.7
KunR12	0	Failure			13	8.4709e-04	65.2	43	1.1759e-09	434.2
KunR12	10	Failure			17	9.2783e-04	92.3	34	1.1759e-09	334.4
KunR21	0	Failure			5	3.6678e-04	6.8	Failure		
KunR21	10	Failure			8	7.1537e-04	11	Failure		
KunR22	0	Failure			5	2.2015e-04	23.6	21	3.5844e-11	197.7
KunR22	10	Failure			6	5.9085e-04	28.5	Failure		
KunR31	0	Failure			154	4.6276e-05	441.5	Failure		
KunR31	10	Failure			Failure					
KunR32	0	Failure			138	9.3637e-04	1266.2	Failure		
KunR32	10	Failure			Failure					
Wal2	0	Failure			33	2.5828e-04	0.2	2	1.3107e-09	<0.1
Wal2	10	Failure			46	5.0478e-04	0.3	2	4.7743e-09	<0.1
Wal3	0	Failure			47	4.2073e-04	0.4	8	8.5265e-14	<0.1
Wal3	10	Failure			62	3.3684e-04	0.6	9	1.1369e-13	<0.1
Wal5	0	Failure			45	5.4037e-04	0.8	2	1.3587e-09	<0.1
Wal5	10	Failure			41	7.9511e-04	0.7	2	8.4838e-10	<0.1
MovSet3A1	0	15	2.7942e-09	1.4	10	2.7948e-04	0.9	3	6.6336e-10	0.4
MovSet3A2	0	15	5.6034e-09	7.2	10	5.6052e-04	4.8	5	1.0093e-07	3.8
MovSet3B1	0	15	1.7801e-09	1.4	10	1.9099e-04	0.9	3	1.7101e-09	0.3
MovSet3B2	0	15	3.5279e-09	7.2	10	3.8034e-04	5.7	Failure		
MovSet4A1	0	Failure			11	7.7953e-04	0.3	2	4.6264e-14	<0.1
MovSet4A2	0	Failure			11	7.7920e-04	1	2	4.6163e-14	0.3
MovSet4B1	0	Failure			11	4.7714e-04	0.3	2	1.1586e-14	<0.1
MovSet4B2	0	Failure			11	7.7654e-04	1	2	4.5521e-14	0.3
Box2A	10	Failure			166	5.4048e-04	2.8	3	3.9143e-12	<0.1
Box2B	10	Failure			193	8.7363e-04	3.2	4	8.1224e-13	0.1
Box3A	10	Failure			Failure					
Box3B	10	Failure			Failure					
RHS1A1	0	87	6.6880e-09	2.7	87	6.6880e-09	2.7	0	6.6880e-09	<0.1
RHS1A1	10	24	6.9907e-09	0.7	19	3.5596e-05	0.6	1	7.2433e-12	<0.1
RHS2B1	0	85	6.5264e-09	1.1	84	1.1662e-08	1	1	7.2110e-12	<0.1
RHS2B1	10	23	9.2802e-09	0.3	19	2.4006e-05	0.2	1	7.2727e-12	<0.1
Ex3	10	12	2.1650e-09	<0.1	7	2.2010e-04	<0.1	2	1.0086e-15	<0.1
Ex7	10	26	7.1530e-09	<0.1	14	1.4907e-04	<0.1	8	5.1223e-09	<0.1
Ex8	10	38	7.3202e-09	<0.1	31	2.4909e-04	<0.1	1	3.2791e-09	<0.1

Table 2: High accuracy results for the standard interior-point method and for the hybrid interior-point/semismooth method.

Looking at the Table 2, it is evident that the interior-point method [20] has severe difficulties in achieving a high accuracy, while the hybrid algorithm combines well properties of the two methods. In fact, the hybrid algorithm is not only faster than the state-of-the-art interior-point method described in [20], but it is also much more robust when computation of high-accuracy solutions is required. These are only very preliminary results, but they certainly point to the usefulness of such a hybrid strategy; further research on this topic, however, is beyond the scope of this paper.

5 Final Remarks

In this paper we have studied the globalized semismooth Newton method for the solution of (the KKT conditions of) QVIs. The theoretical properties of the method are good and compare well with existing methods. The situation from the numerical point of view is more complex though. On the one hand, the method is capable of computing quickly high precision approximations of a solution if started in a small neighborhood of the solution itself; furthermore when global convergence occurs, the algorithm behaves very well. However, the practical global convergence behavior of the method leaves much to be desired and (global) convergence occurred only on two thirds of the cases in our tests. This suggests that a good strategy would be to combine the semismooth Newton method with the interior-point algorithm [20] that is very robust but lacks a good local convergence rate. Our preliminary results show that this strategy is certainly of interest; it will be explored more in detail in future research.

References

[1] D. Aussel, R. Correa, and M. Marechal: *Gap functions for quasivariational inequalities and generalized Nash equilibrium problems.* Journal of Optimization Theory and Applications 151, 2011, pp. 474–488.

[2] C. Baiocchi and A. Capelo: *Variational and Quasivariational Inequalities: Applications to Free Boundary Problems.* Wiley, New York, 1984.

[3] J.W. Barrett and L. Prigozhin: *A quasi-variational inequality problem in superconductivity.* Math. Models Methods Appl. Sci. 20, 2010, pp. 679–706.

[4] A. Bensoussan: *Points de Nash dans le cas de fonctionnelles quadratiques et jeux differentiels lineaires a N personnes.* SIAM Journal on Control 12, 1974, pp. 460–499.

[5] A. Bensoussan, M. Goursat, and J.-L. Lions: *Contrôle impulsionnel et inéquations quasi-variationnelles stationnaires.* C. R. Acad. Sci. Paris Sér. A 276, 1973, pp. 1279–1284.

[6] A. Bensoussan and J.-L. Lions: *Nouvelle formulation de problèmes de contrôle impulsionnel et applications.* C. R. Acad. Sci. Paris Sér. A 276, 1973, pp. 1189–1192.

[7] A. Bensoussan and J.-L. Lions: *Nouvelles méthodes en contrôle impulsionnel.* Applied Mathematics and Optimization 1, 1975, pp. 289–312.

[8] P. Beremlijski, J. Haslinger, M. Kocvara, and J.V. Outrata: *Shape optimization in contact problems with Coulomb friction.* SIAM Journal on Optimization 13, 2002, pp. 561–587.

[9] M. Bliemer and P. Bovy: *Quasi-variational inequality formulation of the multiclass dynamic traffic assignment problem.* Transportation Research Part B: Methodological 37, 2003, pp. 501–519.

[10] D. Chan, and J.-S. Pang: *The generalized quasi-variational inequality problem.* Mathematics of Operations Research 7, 1982, pp. 211–222.

[11] F.H. Clarke: *Optimization and Nonsmooth Analysis.* John Wiley & Sons, New York, 1983 (reprinted by SIAM, Philadelphia, PA, 1990).

[12] T. De Luca, F. Facchinei, and C. Kanzow: *A semismooth equation approach to the solution of nonlinear complementarity problems.* Mathematical Programming 75, 1996, pp. 407–439.

[13] T. De Luca, F. Facchinei, and C. Kanzow: *A theoretical and numerical comparison of some semismooth algorithms for complementarity problems.* Computational Optimization and Applications 16, 2000, pp. 173–205.

[14] M. De Luca and A. Maugeri: *Discontinuous quasi-variational inequalities and applications to equilibrium problems.* In: F. Giannessi (ed.): *Nonsmooth Optimization: Methods and Applications.* Gordon and Breach, Amsterdam, 1992, pp. 70–75.

[15] H. Dietrich: *A smooth dual gap function solution to a class of quasi-variational inequalities.* Journal of Mathematical Analysis and Applications 235, 1999, pp. 380–393.

[16] A. Dreves, F. Facchinei, A. Fischer, and M. Herrich: *A new error bound result for Generalized Nash Equilibrium Problems and its algorithmic application.* Computational Optimization and Applications, in print. Published online, 2013, DOI 10.1007/s10589-013-9586-z

[17] F. Facchinei, A. Fischer, and M. Herrich: *An LP-Newton method: nonsmooth equations, KKT systems, and nonisolated solutions.* Mathematical Programming, in print. Published online, 2013, DOI 10.1007/s10107-013-0676-6

[18] F. Fachhinei, A. Fischer, and M. Herrich: *A family of Newton methods for nonsmooth constrained systems with nonisolated solutions.* Mathematical Methods of Operations Research 77, 2013, pp. 433-443.

[19] F. Facchinei, A. Fischer, and V. Piccialli: *Generalized Nash equilibrium problems and Newton methods.* Mathematical Programming 117, 2009, pp. 163–194.

[20] F. Facchinei, C. Kanzow, and S. Sagratella: *Solving quasi-variational inequalities via their KKT-conditions.* Mathematical Programming, in print. Published online DOI 10.1007/s10107-013-0637-0

[21] F. Facchinei, C. Kanzow, and S. Sagratella: *QVILIB: A library of quasi-variational inequality test problems.* Pacific Journal on Optimization 9, 2013, pp. 225–250

[22] F. Facchinei and J.-S. Pang: *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer-verlag, New York NY, 2003.

[23] F. Facchinei and J. Soares: *Testing a new class of algorithms for nonlinear complementarity problems.* In: F. Giannessi and A. Maugeri (eds.): *Variational Inequalities and Network Equilibrium Problems.* Plenum Press, New York, 1995, pp. 69–83.

[24] A. Fischer: *A special Newton-type optimization method.* Optimization 24, 1992, pp. 269–284.

[25] M. Fukushima: *A class of gap functions for quasi-variational inequality problems.* Journal of Industrial and Management Optimization 3, 2007, pp. 165–171.

[26] P. Hammerstein and R. Selten: *Game theory and evolutionary biology.* In: R.J. Aumann and S. Hart (eds.): *Handbook of Game Theory with Economic Applications, Vol. 2.* North-Holland, Amsterdam, 1994, pp. 929–993.

[27] P.T. Harker: *Generalized Nash games and quasi-variational inequalities.* European Journal of Operational Research 54, 1991, pp. 81–94.

[28] N. Harms, T. Hoheisel, and C. Kanzow: *On a smooth dual gap function for a class of quasi-variational inequalities.* Technical Report, Institute of Mathematics, University of Würzburg, Germany, October 2013.

[29] N. Harms, C. Kanzow, and O. Stein: *Smoothness properties of a regularized gap function for quasi-variational inequalities.* Optimization, to appear.

[30] J. Haslinger and P.D. Panagiotopoulos: *The reciprocal variational approach to the Signorini problem with friction. Approximation results.* Proc. Royal Soc. Edinburgh: Sect. A Math. 98, 1984, pp. 365–383.

[31] T. Ichiishi: *Game Theory for Economic Analysis.* Academic Press, New York, 1983.

[32] C. Kanzow and H. Kleinmichel: *A new class of semismooth Newton-type methods for nonlinear complementarity problems.* Computational Optimization and Applications 11, 1998, pp. 227–251.

[33] I. Kharroubi, J. Ma, H. Pham, and J. Zhang: *Backward SDEs with constrained jumps and quasi-variational inequalities.* Ann. Probab. 38, 2010, pp. 794–840.

[34] A.S. Kravchuk, and P.J. Neittaanmäki: *Variational and Quasi-Variational Inequalities in Mechanics.* Springer, Dordrecht, 2007.

[35] K. Kubota and M. Fukushima: *Gap function approach to the generalized Nash equilibrium problem.* Journal of Optimization Theory and Applications 144, 2010, pp. 511–531.

[36] M. Kunze and J.F. Rodrigues: *An elliptic quasi-variational inequality with gradient constraints and some of its applications.* Math. Methods Appl. Sci. 23, 2000, pp. 897–908.

[37] U. Mosco: *Implicit variational problems and quasi variational inequalities.* In: J. Gossez, E. Lami Dozo, J. Mawhin, and L. Waelbroeck (eds.): *Nonlinear Operators and the Calculus of Variations.* Lecture Notes in Mathematics Vol. 543, Springer, Berlin, 1976, pp. 83–156.

[38] Y. Nesterov and L. Scrimali: *Solving strongly monotone variational and quasi-variational inequalities.* CORE Discussion Paper

2006/107, Catholic University of Louvain, Center for Operations Research and Econometrics, 2006.

[39] M.A. Noor: *An iterative scheme for a class of quasi variational inequalities.* Journal of Mathematical Analysis and Applications 110, 1985, pp. 463–468.

[40] M.A. Noor: *On general quasi-variational inequalities.* J. King Saud University 24, 2012, pp. 81–88.

[41] J.V. Outrata and M. Kocvara: *On a class of quasi-variational inequalities.* Optimization Methods and Software 5, 1995, pp. 275–295.

[42] J.V. Outrata, M. Kocvara, and J. Zowe: *Nonsmooth approach to optimization problems with equilibrium constraints.* Kluwer Academic Publishers, Dordrecht and Boston, 1998.

[43] J.V. Outrata and J. Zowe: *A Newton method for a class of quasi-variational inequalities.* Computational Optimization and Applications 4, 1995, pp. 5–21.

[44] J.-S. Pang: *The implicit complementarity problem.* In: O.L. Mangasarian, R.R. Meyer, and S.M. Robinson (eds.): *Nonlinear Programming 4.* Academic Press, New York, 1981, pp. 487–518.

[45] J.-S. Pang: *On the convergence of a basic iterative method for the implicit complementarity problem.* Journal of Optimization Theory and Applications 37, 1982, pp. 149–162.

[46] J.-S. Pang and M. Fukushima: *Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games.* Computational Management Science 2, 2005, pp. 21–56 (erratum: ibid 6, 2006, pp. 373–375).

[47] L. Qi: *Convergence analysis of some algorithms for solving nonsmooth equations.* Mathematics of Operations Research 18, 1993, pp. 227–244.

[48] L. Qi and J. Sun: *A nonsmooth version of Newton's method.* Mathematical Programming 58, 1993, pp. 353–367.

[49] J.F. Rodrigues and L. Santos: *Quasivariational solutions for first order quasilinear equations with gradient constraint.* Archive for Rational Mechanics and Analysis 205, 2012, pp. 493–514.

[50] I.P. Ryazantseva: *First-order methods for certain quasi-variational inequalities in Hilbert space.* Computational Mathematics and Mathematical Physics 47, 2007, pp. 183–190.

[51] L. Scrimali: *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.

[52] A.H. Siddiqi and Q.H. Ansari: *Strongly nonlinear quasivariational inequalities.* Journal of Mathematical Analysis and Applications 149, 1990, pp. 444–450.

[53] K. Taji: *On gap functions for quasi-variational inequalities.* Abstract and Applied Analysis, 2008, Article ID 531361.

[54] J.Y. Wei and Y. Smeers: *Spatial oligopolistic electricity models with Cournot generators and regulated transmission prices.* Operations Research 47, 1999, pp. 102–112.