Gap functions for quasi-equilibria

Giancarlo Bigi

Dipartimento di Informatica, Università di Pisa,
Largo B.Pontecorvo 3, 56127 Pisa, Italia
giancarlo.bigi@unipi.it

Mauro Passacantando

Dipartimento di Informatica, Università di Pisa,
Largo B.Pontecorvo 3, 56127 Pisa, Italia
mauro.passacantando@unipi.it

An approach for solving quasi-equilibrium problems (QEPs) is proposed relying on gap functions, which allow reformulating QEPs as global optimization problems. The (generalized) smoothness properties of a gap function are analysed and an upper estimates of its Clarke directional derivative is given. Monotonicity assumptions on both the equilibrium and constraining bifunctions are a key tool to guarantee that all the stationary points of a gap function actually solve QEP. A few classes of constraints satisfying such assumptions are identified covering a wide range of situations. Relying on these results, a descent method for solving QEP is devised and its convergence proved. Finally, error bounds are given in order to guarantee the boundedness of the sequence generated by the algorithm.

1. Introduction

Variational inequalities and Nash equilibrium problems (see, for instance, [26, 50]) are not well suited to model noncooperative games in which players share resources or more generally the feasible strategies of each player depend on the choices of the others (see, for instance, the seminal prototype of an abstract competitive economy by Arrow and Debreu [1]). More general mathematical models such as quasi-variational inequalities (see [4, 5, 13]) and generalized Nash equilibrium problems (see [23]) have to be considered.

In this paper we focus on the following (abstract) *quasi-equilibrium problem*

$$\text{find } x^* \in C(x^*) \quad \text{s.t.} \quad f(x^*, y) \geq 0 \quad \forall y \in C(x^*), \qquad (QEP)$$

where the bifunction $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfies the equilibrium condition $f(x, x) = 0$ for any $x \in \mathbb{R}^n$ and the constraints are given by a set-valued map $C : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ that describes how the feasible region changes together with the considered point.

Clearly, QEPs are modelled upon quasi-variational inequalities (shortly QVIs), that is the case

$$f(x, y) = \langle F(x), y - x \rangle \qquad (1)$$

for some $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Generalized Nash equilibrium problems (shortly GNEPs) can be reformulated through (QEP) as well. Indeed, consider the situation in which each player i aims at maximizing an utility function $f_i(\cdot, x_{-i})$ over a set of feasible strategies $C_i(x_{-i})$ for some $f_i : \mathbb{R}^{n_1} \times ... \times \mathbb{R}^{n_\ell} \longrightarrow \mathbb{R}$ and some set-valued map $C_i : \mathbb{R}^{n_1 + ... + n_\ell - n_i} \rightrightarrows \mathbb{R}^{n_i}$, which both depend upon the strategies $x_{-i} = (x_j)_{j \neq i}$ chosen by the other players. Finding a (generalized) Nash equilibrium amounts to solving (QEP) with the Nikaido-Isoda aggregate bifunction

$$f(x, y) = \sum_{i=1}^{\ell} [f_i(x) - f_i(x_{-i}, y_i)]$$

1

and $C(x) = C_1(x_{-1}) \times ... \times C_\ell(x_{-\ell})$. It is also worth noting that (QEP) is a natural generalization of the so-called abstract equilibrium problem (shortly EP), i.e., the case in which the set-valued map C is constant. As EP subsumes optimization, multiobjective optimization, variational inequalities, fixed point and complementarity problems, Nash equilibria in noncooperative games and inverse optimization in a unique mathematical model (see, for instance, [10, 7]), further "quasi" type models could be analysed through the QEP format beyond QVIs and GNEPs.

QVIs have been introduced in [4, 5] to study impulse control problems and subsequently exploited to model several finite and infinite-dimensional equilibrium problems (see, for instance, [3, 13, 47]). Different approaches have been considered to devise solution methods: characterizations based on fixed points and projections [13, 51, 52], penalization of coupling constraints [55, 57, 22], KKT systems [24], minimization of dual gap functions in the affine case [36]. Also Newton type methods, which guarantee only local convergence, have been developed [53, 54]. A way to study GNEPs is to formulate them as QVIs (see, for istance, [34]) and exploit the corresponding theories and algorithms: projection methods are expolited in [60] while penalty techniques and barrier methods in [46]. The simultaneous resolution of the KKT conditions of the optimization problems describing a GNEP has been carried out through locally convergent Newton type techniques [21, 42] and globally convergent interior-point type techniques [17]. Ad hoc algorithms for particular classes of GNEPs and equilibria have been developed as well: normalized equilibria of jointly convex GNEPs are computed through the reformulation as optimization problems [56, 37, 38] and (standard) variational inequalities [20], restricted equilibria through penalization techniques [31], and equilibria of potential GNEP through decomposition algorithms [27].

Unlikely QVI and GNEP, the QEP format did not receive much attention: existence results have been given in [11, 16] and an extragradient type algorithm has been developed in [58]. To the best of our knowledge, no other algorithm has been devised up to now. The goal of the paper is to reformulate (QEP) as an optimization problem through a suitable gap function and develop an ad-hoc descent algorithm, supposing that the set-valued map C can be described by constraining bifunctions (see Section 2.1).

Gap functions have been originally conceived for variational inequalities [15, 29] and later extended to EPs [48, 49], QVIs [32, 19, 30, 59, 2, 33, 35], jointly convex GNEPs via the Nikaido-Isoda binfunction [37, 38, 18, 56] and generic GNEPs via QVI reformulations [46]. Though descent type methods based on gap functions have been extensively developed for EPs (see, for instance, [49, 44, 12, 43, 45, 6, 8, 9] and Section 3.2 in the survey paper [7]), the analysis of gap functions for QVIs is focused on smoothness properties [19, 30, 59, 18, 35] and error bounds [2, 33] while no algorithm is developed. A descent method has been developed in [38] for jointly convex GNEPs; anyway, the choice of restricting to the computation of normalized equilibria makes the problem actually fall within the EP (and not the QEP) framework.

Indeed, the reformulation of (QEP) as an optimization problem brings some difficult issues in devising descent methods which are not met in the EP case: the gap function is not necessarily differentiable even though the equilibrium and the constraining bifunctions are differentiable; the feasible region is given by the fixed points of the set-valued constraining map C and is therefore more difficult to handle; the so-called stationarity property (see, for instance, [6]), which guarantees all the stationary points of the gap function to be actually global minimizers and therefore solutions of (QEP), requires monotonicity assumptions both on the equilibrium and constraining bifunctions. These issues are dealt with in Section 2. After the gap function has been introduced and the reformulation of (QEP) as an optimization problem shown, Section 2.1 analyses the smoothness properties of the gap function; in particular, an upper estimate of its Clarke directional derivative is given, which provides a key tool in devising the descent method. Section 2.2 deals with stationarity and descent directions; furthermore, classes of constraints which allow guaranteing the stationarity property are identified. Section 3 describes the descent method and its convergence is proved under

standard assumptions, while Section 4 provides error bounds, which guarantee that the sequence generated by the algorithm is bounded. Finally, Section 5 addresses possible improvements of the results and directions for further investigations.

2.   Gap functions

Throughout the paper we suppose that f is continuously differentiable and $f(x, \cdot)$ is convex for all $x \in \mathbb{R}^n$, while the values $C(x)$ of the constraining set-valued map $C : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ are supposed closed and convex (maybe empty) for any $x \in \mathbb{R}^n$. We also consider the sets $C(x)$ to be explicitly described by convex constraints, i.e.,

$$C(x) = \{ \ y \in \mathbb{R}^n \ : \ g_i(x, y) \le 0 \quad i = 1, \dots, m \ \}$$

for some continuously differentiable functions $g_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ such that $g_i(x, \cdot)$ is convex for all $x \in \mathbb{R}^n$. Let D denote the domain of the mapping C, i.e.,

$$D = \{ x \in \mathbb{R}^n : \ C(x) \ne \emptyset \}$$

and let D_S denote the set of all the points $x \in D$ such that $C(x)$ satisfies the Slater condition, i.e., $x \in D_S$ if and only if there exists $\hat{y} \in \mathbb{R}^n$ such that $g_i(x, \hat{y}) < 0$ for any $i = 1, \dots, m$. A simple continuity argument shows that D_S is an open set. Finally, let X denote the set of the fixed points of C, i.e.,

$$X = \{ x \in \mathbb{R}^n : \ x \in C(x) \} = \{ x \in \mathbb{R}^n : \ g_i(x, x) \le 0 \quad i = 1, \dots, m \}.$$

Clearly, $X \subset D$ and it can be considered the feasible region of (QEP) since any candidate solution of the problem has to belong to X. For the sake of simplicity, just inequality constraints are considered. Actually, linear equality constraints can be incorporated into the analysis as a pair of inequalities at no cost. It is enough to restate the Slater condition in the standard way, that is $\hat{y} \in \mathbb{R}^n$ satisfies $g_i(x, \hat{y}) < 0$ for the inequality constraints and $g_i(x, \hat{y}) = 0$ for the (implicitly given) linear equalities.

The minimum value of $f(x, \cdot)$ over $C(x)$ provides a measure of the quality of the candidate solution $x \in X$. Anyway, $f(x, \cdot)$ may be unbounded below on $C(x)$. Adding an auxiliary term to f allows overcoming this drawback. In fact, consider any continuously differentiable bifunction $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that

- $h(x, y) \ge 0$ for all $x, y \in \mathbb{R}^n$ and $h(z, z) = 0$ for all $z \in \mathbb{R}^n$,
- $h(x, \cdot)$ is strongly convex with modulus $\tau > 0$ for all $x \in \mathbb{R}^n$,
- $\nabla_y h(z, z) = 0$ for all $z \in \mathbb{R}^n$,
- $\langle \nabla_x h(x, y) + \nabla_y h(x, y), y - x \rangle \ge 0$ for all $x, y \in \mathbb{R}^n$.

The value function

$$\varphi(x) = - \min \{ \ f(x, y) + h(x, y) \ : \ y \in C(x) \ \} \tag{2}$$

is finite at any $x \in D$ since the inner optimization problem has a strongly convex objective function $f(x, \cdot) + h(x, \cdot)$ and the feasible region $C(x)$ is closed and convex. Moreover, it admits a unique minimizer $y(x)$: this feature is useful to prove continuity and smoothness properties of φ (see Section 2.1). The most straightforward choice for the auxiliary term is $h(x, y) = \|x - y\|^2 / 2$.

Just like in the particular case of QVIs (see, for instance, [30, 59]), φ is a gap function. Indeed, it allows turning (QEP) into an optimization problem through the following characterizations.

THEOREM 1.
a) $x^* \in X$ solves (QEP) if and only if $y(x^*) = x^*$.
b) $\varphi(x) \ge 0$ for all $x \in X$.
c) x^* solves (QEP) if and only if $x^* \in X$ and $\varphi(x^*) = 0$.

PROOF.

a) If $x^* \in X$ solves (QEP), then

$$f(x^*, y) + h(x^*, y) \geq f(x^*, y) \geq 0 = f(x^*, x^*) + h(x^*, x^*)$$

holds for any $y \in C(x^*)$. Since $y(x^*)$ is the unique minimizer of $f(x^*, \cdot) + h(x^*, \cdot)$ over $C(x^*)$, then clearly $y(x^*) = x^*$.

Vice versa suppose $y(x^*) = x^*$. Therefore, $x^* \in C(x^*)$ and thus also $x^* \in X$ holds. Moreover, the optimality conditions for x^* read

$$\langle \nabla_y f(x^*, x^*) + \nabla_y h(x^*, x^*), y - x^* \rangle \geq 0, \qquad \forall \; y \in C(x^*).$$

Since $\nabla_y h(x^*, x^*) = 0$, the convexity of $f(x^*, \cdot)$ implies

$$f(x^*, y) \geq f(x^*, x^*) + \langle \nabla_y f(x^*, x^*), y - x^* \rangle \geq 0, \qquad \forall \; y \in C(x^*),$$

hence x^* solves (QEP).

b) Any $x \in X$ satisfies $x \in C(x)$, therefore

$$\varphi(x) = -\min \{ \; f(x, y) + h(x, y) \; : \; y \in C(x) \; \} \geq -f(x, x) - h(x, x) = 0.$$

c) If x^* solves (QEP), then $x^* \in X$ and $x^* = y(x^*)$ by a) so that

$$\varphi(x^*) = -[f(x^*, x^*) + h(x^*, x^*)] = 0.$$

Vice versa, $\varphi(x^*) = 0$ reads

$$f(x^*, y) + h(x^*, y) \geq 0 = f(x^*, x^*) + h(x^*, x^*) \qquad \forall \; y \in C(x^*),$$

so that $x^* \in X$ implies $y(x^*) = x^*$. Therefore, x^* solves (QEP) by a). □

The theorem shows that (QEP) can be equivalently formulated as the global optimization problem

$$\min \{ \; \varphi(x) \; : \quad x \in X \; \}. \tag{3}$$

Anyway, (3) is generally a difficult problem: φ is neither convex nor differentiable and X may have a complex structure. Nonsmoothness is handled in Section 2.1, while monotonicity conditions that allow overcoming the lack of convexity are addressed in Section 2.2.

2.1. Continuity and differentiability

Once the continuity of the solution map $y(x)$ of the inner optimization problem in (2) is guaranteed, the continuity of the gap function φ is achieved as well since $\varphi(x) = -f(x, y(x)) - h(x, y(x))$.

THEOREM 2. *The mapping $x \longmapsto y(x)$ is continuous at any $x \in D_S$.*

PROOF. Since g_i are continuous, the set valued map C is closed on D. Moreover, C is open at any $x \in D_S$ since $g_i(x, \cdot)$ are continuous and convex and satisfy Slater condition (see [41, Theorem 12]). Thus, C is continuous at x. Since f is continuous and the mapping $x \longmapsto y(x)$ is single-valued, it is uniformly compact near any $x \in D$ by Corollary 9.1 in [41] and hence also continuous at x by Corollary 8.1 in [41]. □

Directional derivatives of φ can be computed or estimated through Danskin type formulas that exploit Lagrange multipliers associated to $y(x)$ in the inner optimization problem. To this aim, let $\Lambda(x)$ be the set of the vectors $\lambda \in \mathbb{R}_+^m$ such that the optimality conditions

$$\begin{cases} \nabla_y f(x, y(x)) + \nabla_y h(x, y(x)) + \displaystyle\sum_{i=1}^m \lambda_i \nabla_y g_i(x, y(x)) = 0, \\ \lambda_i \, g_i(x, y(x)) = 0, \quad i = 1, \ldots, m \end{cases} \tag{4}$$

for (2) hold at $x \in D_S$. Notice that the Slater condition guarantees the compactness of $\Lambda(x)$.

THEOREM 3.
a) φ *is directionally differentiable at any point* $x \in D_S$ *and any direction* $d \in \mathbb{R}^n$ *with*

$$\varphi'(x; d) = \min_{\lambda \in \Lambda(x)} \left[\left\langle -\nabla_x f(x, y(x)) - \nabla_x h(x, y(x)) - \sum_{i=1}^m \lambda_i \nabla_x g_i(x, y(x)), d \right\rangle \right].$$

b) φ *is locally Lipschitz continuous near any* $x \in D_S$ *and the Clarke directional derivative satisfies the inequality*

$$\varphi^\circ(x; d) \le \max_{\lambda \in \Lambda(x)} \left[\left\langle -\nabla_x f(x, y(x)) - \nabla_x h(x, y(x)) - \sum_{i=1}^m \lambda_i \nabla_x g_i(x, y(x)), d \right\rangle \right] \tag{5}$$

at x *in any direction* $d \in \mathbb{R}^n$.

PROOF.
a) It follows immediately from Theorem 2 in [40].
b) Considering any fixed $\bar{u} \in D_S$, the inner convex optimization problem in (2) can be written as

$$\min \left\{ \, f(\bar{u}, y) + h(\bar{u}, y) \; : \; g(\bar{u}, y) \le 0, \; y \in \mathbb{R}^n \, \right\}.$$

Since the Slater condition holds, the set of multipliers $\Lambda(\bar{u})$ is nonempty and $(y(\bar{u}), \lambda(\bar{u}))$ is a saddle point of the Lagrangian function

$$L(\bar{u}, y, \lambda) := f(\bar{u}, y) + h(\bar{u}, y) + \langle \lambda, g(\bar{u}, y) \rangle,$$

for any $\lambda(\bar{u}) \in \Lambda(\bar{u})$, i.e., the inequalities

$$L(\bar{u}, y(\bar{u}), \lambda) \le L(\bar{u}, y(\bar{u}), \lambda(\bar{u})) \le L(\bar{u}, y, \lambda(\bar{u}))$$

hold for any $y \in \mathbb{R}^n$ and any $\lambda \in \mathbb{R}_+^m$. Thanks to the complementary slackness conditions, $L(\bar{u}, y(\bar{u}), \lambda(\bar{u})) = f(\bar{u}, y(\bar{u})) + h(\bar{u}, y(\bar{u})) = -\varphi(\bar{u})$ holds and the above inequalities can be equivalently turned into

$$-L(\bar{u}, y, \lambda(\bar{u})) \le \varphi(\bar{u}) \le -L(\bar{u}, y(\bar{u}), \lambda).$$

Given any $u, v \in D_S$, the right inequality with $\bar{u} = u$ and $\lambda = \lambda(v)$ and the left inequality with $\bar{u} = v$ and $\lambda = \lambda(v)$ provide

$$\varphi(u) - \varphi(v) \le L(v, y(u), \lambda(v)) - L(u, y(u), \lambda(v))$$

$$= f(v, y(u)) + h(v, y(u)) - [f(u, y(u)) + h(u, y(u))]$$

$$+ \langle \lambda(v), g(v, y(u)) - g(u, y(u)) \rangle$$

$$\le f(v, y(u)) + h(v, y(u)) - [f(u, y(u)) + h(u, y(u))]$$

$$+ \|\lambda(v)\| \, \|g(v, y(u)) - g(u, y(u))\|.$$

Furthermore, the mean value theorem guarantees that

$$f(v, y(u)) + h(v, y(u)) - [f(u, y(u)) + h(u, y(u))]$$

$$= \langle \nabla_x f(v', y(u)) + \nabla_x h(v', y(u)), v - u \rangle$$

holds for some v' in the line segment between u and v. Given any $\bar{x} \in D_S$, then there exist $L_1 > 0$ and $\delta_1 > 0$ such that $B(\bar{x}, \delta_1) \subseteq D_S$ and

$$f(v, y(u)) + h(v, y(u)) - [f(u, y(u)) + h(u, y(u))] \leq L_1 \|v - u\|$$

holds for all $u, v \in B(\bar{x}, \delta_1)$ since the mappings $\nabla_x f$, and $\nabla_x h$ are continuous on $\mathbb{R}^n$ and $x \longmapsto y(x)$ is continuous at $\bar{x}$. On the other hand, the functions $g_i(\cdot, y(u))$ are continuously differentiable so that there exist $L_2 > 0$ and $\delta_2 > 0$ such that

$$\|g(v, y(u)) - g(u, y(u))\| \leq L_2 \|v - u\|$$

holds for all $u, v \in B(\bar{x}, \delta_2)$. Moreover, Lemma 2 in [40] guarantees the existence of $L_3 > 0$ and $\delta_3 > 0$ such that $\|\lambda(v)\| \leq L_3$ holds for all $v \in B(\bar{x}, \delta_3)$ and all $\lambda(v) \in \Lambda(v)$. Therefore, the last three inequalities imply that

$$\varphi_\alpha(u) - \varphi_\alpha(v) \leq (L_1 + L_2 L_3) \|v - u\|$$

holds for all $u, v \in B(\bar{x}, \delta)$, where $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Therefore φ is locally Lipschitz near $\bar{x}$.

In order to prove (5) consider a fixed $x \in D_S$ and a fixed direction $d \in \mathbb{R}^n$. Setting $z_t := z + t\, d$ for any $z \in B(x, \delta')$ with $\delta' > 0$ and $t > 0$, then both $z \in D_S$ and $z_t \in D_S$ hold if δ' and t are small enough. Therefore, considering any $\lambda(z_t) \in \Lambda(z_t)$, the inequality

$$\varphi(z_t) - \varphi(z) \leq f(z, y(z_t)) + h(z, y(z_t)) - [f(z_t, y(z_t)) + h(z_t, y(z_t))]$$
$$+ \langle \lambda(z), g(z, y(z_t)) - g(z_t, y(z_t)) \rangle$$

can be proved arguing as in the first part of the proof. The mean value theorem guarantees that

$$f(z, y(z_t)) + h(z, y(z_t)) - [f(z_t, y(z_t)) + h(z_t, y(z_t))]$$

$$= \langle \nabla_x f(\tilde{z}(z,t), y(z_t)) + \nabla_x h(\tilde{z}(z,t), y(z_t)), z - z_t \rangle$$

$$= t \langle -\nabla_x f(\tilde{z}(z,t), y(z_t)) - \nabla_x h(\tilde{z}(z,t), y(z_t)), d \rangle$$

and

$$g_i(z, y(z_t)) - g_i(z_t, y(z_t)) = \langle \nabla_x g_i(\tilde{z}_i'(z,t), y(z_t)), z - z_t \rangle$$

$$= -t \langle \nabla_x g_i(\tilde{z}_i'(z,t), y(z_t)), d \rangle$$

hold for some $\tilde{z}(z,t)$ in the line segment between z and z_t and some $\tilde{z}_i'(z,t)$ in the line segment between z and z_t.

Let $z^k \to x$ and $t_k \downarrow 0$ be two sequences such that $\varphi^\circ(x; d) = \lim_{k \to \infty} t_k^{-1} [\varphi(z_{t_k}^k) - \varphi(z^k)]$. The last three formulas above with $z = z^k$ and $t = t_k$ (and therefore $z_t = z_{t_k}^k$) lead to

$$[\varphi(z_{t_k}^k) - \varphi(z^k)]/t_k \leq \langle -\nabla_x f(\tilde{z}(z^k, t_k), y(z_{t_k}^k)) - \nabla_x h(\tilde{z}(z^k, t_k), y(z_{t_k}^k)), d \rangle$$

$$- \langle \lambda(z^k), w(x, z_{t_k}^k, t_k) \rangle,$$

where $w(x, z_{t_k}^k, t_k) \in \mathbb{R}^m$ is the vector of components $\langle \nabla_x g_i(\tilde{z}_i'(z^k, t_k), y(z_{t_k})), d \rangle$ $(i = 1, \ldots, m)$. Since $z^k \to x$ and $t_k \downarrow 0$, then $z_{t_k}^k \to x$, $\tilde{z}(z^k, t_k) \to x$, and $y(z_{t_k}^k) \to y(x)$ by Theorem 2. Hence, the limit

$$\lim_{k \to \infty} \langle -\nabla_x f(\tilde{z}(z^k, t_k), y(z_{t_k}^k)) - \nabla_x h(\tilde{z}(z^k, t_k), y(z_{t_k}^k)), d \rangle$$
$$= -\langle \nabla_x f(x, y(x)) + \nabla_x h(x, y(x)), d \rangle$$

holds since $\nabla_x f$ and $\nabla_x h$ are continuous. Furthermore, Lemma 2 in [40] guarantees that $\Lambda(z)$ is uniformly bounded on a neighborhood of x and closed at x. Hence, taking a subsequence if necessary, there exists $\hat{\lambda} \in \Lambda(x)$ such that $\lambda(z^k) \to \hat{\lambda}$. Moreover, $\tilde{z}_i'(z^k, t_k) \to x_i$ and the continuity of $\nabla_x g$ imply

$$\lim_{k \to \infty} -\langle \lambda(z^k), w(x, z_{t_k}^k, t_k) \rangle = -\langle \hat{\lambda}, \nabla_x g(x, y(x)) \, d \rangle.$$

As a consequence, the inequality

$$\varphi^{\circ}(x; d) \leq \langle -\nabla_x f(x, y(x)) - \nabla_x h(x, y(x)) - \sum_{i=1}^{m} \hat{\lambda}_i \nabla_x g_i(x, y(x)), d \rangle$$

holds, yielding (5). $\qquad \square$

When (QEP) takes the shape of a quasi-variational inequality, that is f is given by (1), Theorem 2.3 a) coincides with the results given for polyhedral constraints in [30] and for general convex constraints in [59].

If $\Lambda(x)$ is a singleton, then $\varphi'(x; \cdot)$ is actually linear and therefore the gap function φ is differentiable at x. This happen, for instance, in the case $y(x) \in \operatorname{int} C(x)$ since the complementarity slackness conditions in (4) imply $\Lambda(x) = \{0\}$.

Differentiability may hold also when the sets $C(x)$ have particular structures. For instance, φ is continuously differentiable if the set-valued mapping C describes a (generalized) moving set. Indeed, suppose

$$C(x) = Q(x) K + t(x) \tag{6}$$

holds at each $x \in \mathbb{R}^n$ for some closed convex $K \subseteq \mathbb{R}^p$ with $p \leq n$, $Q(x) = (Q_{ij}(x)) \in \mathbb{R}^{n \times p}$ with $Q_{ij} : \mathbb{R}^n \to \mathbb{R}$ and $t : \mathbb{R}^n \to \mathbb{R}^n$ being continuously differentiable.

As a consequence, φ can be equivalently written as

$$\varphi(x) = -\min\{ \ f(x, Q(x)z + t(x)) + h(x, Q(x)z + t(x)) \ : z \in K \ \}.$$

If $Q(x)$ has full column rank, then the inner optimization problem above has a unique solution $z(x)$ and moreover $y(x) = Q(x)z(x) + t(x)$. Since K does not depend upon x, the classical Danskin's theorem can be exploited and provides the following result.

THEOREM 4. *If (6) holds at each $x \in \mathbb{R}^n$, all the functions Q_{ij} and t are continuously differentiable on $\mathbb{R}^n$, $Q(x)$ has full column rank for all $x \in \mathbb{R}^n$ and K is closed and convex, then φ is continuously differentiable on $\mathbb{R}^n$ and*

$$\nabla \varphi(x) = -\Big[\nabla_x f(x, y(x)) + \nabla_x h(x, y(x)) +$$
$$\Big(\nabla t(x) + \sum_{i=1}^{p} z_i(x) \nabla Q_i(x) \Big) (\nabla_y f(x, y(x)) + \nabla_y h(x, y(x))) \Big], \tag{7}$$

where $\nabla Q_i(x)$ denotes the matrix whose j-row is $\nabla Q_{ij}(x)$.

Notice that (7) turns into the simpler form

$$\nabla \varphi(x) = -[\nabla_x f(x, y(x)) + \nabla_x h(x, y(x)) + \nabla t(x)(\nabla_y f(x, y(x)) + \nabla_y h(x, y(x)))]$$

whenever $p = n$ and $Q = I$. If furthermore $t \equiv 0$, then $C(x)$ is always K and (QEP) reduces to the classical EP format while Theorem 4 collapses into Theorem 2.1 of [49].

2.2. Stationary points

The reformulation of (QEP) as the minimization problem (3) allows exploiting optimality conditions: the stationarity condition for $x \in D_S$ reads

$$\varphi'(x;d) \geq 0 \qquad \forall\, d \in T(X,x), \tag{8}$$

where $T(X,x)$ denotes the Bouligand tangent cone of X at x. Anyway, global minima are sought: stationary points may be not even local minima since φ and X are not necessarily convex. Adding suitable monotonicity conditions is a way to overcome this issue. When EPs are considered, i.e., $C(x) = C$ for some convex set C and any x, the sets of stationary points and global minima coincide provided that f is *strictly ∇-monotone* on C, i.e., f satisfies

$$\langle \nabla_x f(x,y) + \nabla_y f(x,y), y - x \rangle > 0$$

for any $x, y \in C$ (see, for instance, [6, 49]). In the general case some assumptions on the constraints are required as well. Since only active constraints are involved in the directional derivative, it is reasonable to restrict monotonicity assumptions accordingly: the constraining function g_i is said to satisfy the *active ∇-monotonicity condition* at $x \in D$ if any $y \in C(x)$ satisfies

$$\langle \nabla_x g_i(x,y) + \nabla_y g_i(x,y), y - x \rangle \geq 0 \qquad \text{if } g_i(x,y) = 0. \tag{9}$$

The equivalence between stationarity and global optimality holds under these monotonicity assumptions provided that the solution $y(x)$ of the inner optimization problem (2) lies inside the first order approximation $x + T(X,x)$ of X around x.

THEOREM 5. *Let $x \in X \cap D_S$ such that $y(x) \in x + T(X,x)$ be given. Suppose f is strictly ∇-monotone on $C(x)$ and the constraint functions g_i satisfy the active ∇-monotonicity condition at x. Then, the following statements are equivalent:*
  a) x solves (QEP);
  b) $\varphi'(x;d) \geq 0$ for any $d \in T(X,x)$;
  c) $\varphi^\circ(x;d) \geq 0$ for any $d \in T(X,x)$.

PROOF. The necessary optimality conditions for (3) guarantee that *a)* implies *b)*, while *b)* implies *c)* since $\varphi^\circ(x;d) \geq \varphi'(x;d)$ holds for any $d \in \mathbb{R}^n$. Hence, it is enough to prove that *c)* implies *a)*.

By contradiction, suppose x does not solve (QEP). Hence, $y(x) \neq x$ and the following chain of equalities and inequalities hold

$$
\begin{aligned}
\varphi^\circ(x;y(x)-x) &\leq \langle -\nabla_x f(x,y(x)) - \nabla_x h(x,y(x)) - \sum_{i=1}^m \lambda_i \nabla_x g_i(x,y(x)), y(x)-x \rangle \\
&\leq \langle -\nabla_x f(x,y(x)) + \nabla_y h(x,y(x)) - \sum_{i=1}^m \lambda_i \nabla_x g_i(x,y(x)), y(x)-x \rangle \\
&= \langle -\nabla_x f(x,y(x)) - \nabla_y f(x,y(x)), y(x)-x \rangle + \\
&\quad - \sum_{i=1}^m \lambda_i \langle \nabla_x g_i(x,y(x)) + \nabla_y g_i(x,y(x)), y(x)-x \rangle \\
&\leq \langle -\nabla_x f(x,y(x)) - \nabla_y f(x,y(x)), y(x)-x \rangle \\
&< 0,
\end{aligned}
$$

where $\lambda \in \Lambda(x)$ provides an upper bound on the value of the directional derivative according to Theorem 3 b). The other inequalities are due to the ∇-monotonicity of h, f and g_i paired with the

complementarity slackness conditions in (4), while the equality is due to the multiplier rule in (4). Therefore, c) does not hold since $y(x) - x \in T(X, x)$. $\square$

The following descent property is achieved as a by-product of the above proof, and it provides the basic tool to design the solution algorithm of Section 3.

COROLLARY 1. *Let $x \in X \cap D_S$ such that $y(x) \in x + T(X, x)$ be given. Suppose f is strictly ∇-monotone on $C(x)$ and the constraint functions g_i satisfy the active ∇-monotonicity condition at x. If x is not a solution of (QEP), then $y(x) - x$ is a descent direction for φ at x, i.e., $\varphi^\circ(x; y(x) - x) < 0$.*

In [59] the equivalence between *a)* and *b)* of Theorem 5 is proved for QVIs under a slightly stronger assumption than (9), which involves only the gradients $\nabla_x g_i(x, y)$. Furthermore, the same equivalence has been exploited in [46] in a very particular framework: the QVI reformulation of a generalized Nash equilibrium problem with only linear equality shared constraints. Indeed, linear equality shared constraints satisfy both the active ∇-monotonicity condition (9) and the additional assumption on $y(x)$ (see the last paragraph of this section).

In the case of EPs Theorem 5 and Corollary 1 collapse to well-known results (see, for instance, [6, 49]). Notice that active ∇-monotonicity and the additional requirement on $y(x)$ are not explicitly given in [6, 49] since they are always met. Indeed, if $g_i(x, y) = c_i(y)$ for suitable continuously differentiable convex functions c_i for all i's, that is $C(x) = C = \{y : c_i(y) \leq 0\}$ for any x, then $X = C$ and hence $y(x) \in X$ always holds true. As a consequence, the convexity of X guarantees $y(x) \in x + T(X, x)$. Moreover, the convexity of the constraning functions c_i's guarantees the active ∇-monotonicity condition:

$$\begin{aligned} \langle \nabla_x g_i(x, y(x)) + \nabla_y g_i(x, y(x)), y(x) - x \rangle &= \langle \nabla c_i(y(x)), y(x) - x \rangle \\ &\geq c_i(y(x)) - c_i(x) \\ &= -c_i(x) \\ &\geq 0. \end{aligned}$$

Actually, the active ∇-monotonicity condition is satisfied by a few classes of constraints, which cover a wide range of situations, as summarized below.

PROPOSITION 1. *Let c_i be convex and (twice) continuously differentiable, $F_i : \mathbb{R}^n \to \mathbb{R}^n$ be monotone and differentiable, $P_i, R_i \in \mathbb{R}^{n \times n}$ be such that R_i and $P_i + R_i$ are positive semidefinite, $Q_i \in \mathbb{R}^{n \times p}$, $a, r_i \in \mathbb{R}^n$ and $\beta_i \in \mathbb{R}$, $\alpha, \alpha_i \leq 1$, $v_i \leq 0$. Then, any of the following bifunctions*
  a) $g_i(x, y) = c_i(y) - c_i(x)$
  b) $g_i(x, y) = c_i(Q_i[y - (\alpha x + a)])$
  c) $g_i(x, y) = c_i(x) + \langle \nabla c_i(x), y - x \rangle$
  d) $g_i(x, y) = \langle F_i(x), y - x \rangle + v_i$
  e) $g_i(x, y) = \langle P_i x + R_i y + r_i, y - x \rangle + v_i$
  f) $g_i(x, y) = y_i - \alpha_i x_i - \beta_i$
  g) $g_i(x, y) = \alpha_i x_i + \beta_i - y_i$
satisfies the active ∇-monotonicity condition (9) at any $x \in X$.

PROOF. Given any $x \in D$, let $y \in C(x)$ satisfy $g_i(x, y) = 0$.
a) The convexity of c_i implies that ∇c_i is monotone and therefore

$$\langle \nabla_x g_i(x, y) + \nabla_y g_i(x, y), y - x \rangle = \langle \nabla c_i(y) - \nabla c_i(x), y - x \rangle \geq 0.$$

b) The assumptions guarantee the convexity of $g_i(x, \cdot)$ that can be exploited to obtain (9) in the following way

$$
\begin{aligned}
\langle \nabla_x g_i(x,y) + \nabla_y g_i(x,y), y - x \rangle &= (1-\alpha) \langle Q_i \nabla c_i(Q_i[y - (\alpha\, x + a))], y - x \rangle \\
&= (1-\alpha) \langle \nabla_y g_i(x,y), y - x \rangle \\
&\geq (1-\alpha)[g_i(x,y) - g_i(x,x)] \\
&= -(1-\alpha)\, g_i(x,x) \\
&\geq 0,
\end{aligned}
$$

where the last inequality holds since $x \in X$ requires $g_i(x,x) \leq 0$.

c) The convexity of c_i implies that $\nabla^2 c_i(x)$ is positive semidefinite and therefore

$$
\begin{aligned}
\langle \nabla_x g_i(x,y) + \nabla_y g_i(x,y), y - x \rangle &= \langle y - x, \nabla^2 c_i(x)[y - x] \rangle + \langle \nabla c_i(x), y - x \rangle \\
&\geq \langle \nabla c_i(x), y - x \rangle \\
&= -c_i(x) \\
&\geq 0,
\end{aligned}
$$

where the last inequality holds since $x \in X$ implies $c_i(x) = g_i(x,x) \leq 0$.

d) The monotonicity of F_i implies that $\nabla F_i(x)$ is positive semidefinite and therefore

$$
\langle \nabla_x g_i(x,y) + \nabla_y g_i(x,y), y - x \rangle = \langle y - x, \nabla F_i(x)(y - x) \rangle \geq 0.
$$

e) The positive semidefiniteness of $P_i + R_i$ guarantees

$$
\langle \nabla_x g_i(x,y) + \nabla_y g_i(x,y), y - x \rangle = \langle y - x, (P_i + R_i)(y - x) \rangle \geq 0.
$$

f) The assumptions $g_i(x,y) = 0$ and $x \in X$ read $x_i \leq \alpha_i x_i + \beta_i = y_i$. Therefore, $\alpha_i \leq 1$ guarantees

$$
\langle \nabla_x g_i(x,y) + \nabla_y g_i(x,y), y - x \rangle = (1-\alpha_i)(y_i - x_i) \geq 0.
$$

g) Analogous to *f)*. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

If all the constraints g_i fall within the kinds *a)*, *d)* and *e)* of the above proposition, then $D = X = \mathbb{R}^n$. This is no longer true if some of them are of the kinds *b)*, *c)*, *f)* or *g)*.

Suppose all the constraints are of the kind *b)*: it turns out that $C(x) = K + t(x)$ is the moving set provided by the convex set $K = \bigcap_i \{y \in \mathbb{R}^n \ : \ c_i(Q_i y) \leq 0\}$ and $t(x) = \alpha\, x + a$. If $K \neq \emptyset$, then $D = \mathbb{R}^n$ and

$$
X = \begin{cases} (K + a)/(1 - \alpha) & \text{if } \alpha < 1, \\ \emptyset & \text{if } \alpha = 1 \text{ and } -a \notin K, \\ \mathbb{R}^n & \text{if } \alpha = 1 \text{ and } -a \in K. \end{cases} \tag{10}
$$

In this framework $C(x) \subseteq X$ and therefore also $y(x) \in X$ hold for any $x \in X$ whenever X is not empty. Indeed, if $\alpha < 1$, then

$$
\begin{aligned}
(1-\alpha)C(x) &= (1-\alpha)K + \alpha(1-\alpha)x + (1-\alpha)a \\
&\subseteq (1-\alpha)K + \alpha(K + a) + (1-\alpha)a \\
&= K + a.
\end{aligned}
$$

Suppose all the constraints are of the kind *c)*: it turns out that $X = \bigcap_i \{x \in \mathbb{R}^n \ : \ c_i(x) \leq 0\}$. If X satisfies the Slater condition, then $C(x) \subseteq x + T(X, x)$ (see, for instance, [8, Theorem 3]) and the requirement of Theorem 5 is fulfilled. Anyway, $y(x) \in X$ does not necessarily hold: the convexity of the c_i's guarantees just $X \subseteq C(x)$ and not the opposite inclusion. Penalization techniques might

be exploited in algorithmic frameworks as done in [8] while linearizing the constraints in descent type methods for EPs.

If all the variables y_i's are bounded by above through box constraints of the kind f) with $\alpha_i < 1$ and there are no other constraints, then $X = \prod_i(-\infty, \beta_i/(1-\alpha_i)]$ is a box unbounded by below. Similarly, X is a box unbounded by above if only the constraints of the kind g) exist while it is a bounded box if both kinds of constraints describe the feasible region. In all these situations $x \in X$ guarantees also $y(x) \in X$ if moreover $\alpha_i > 0$.

Beyond box constraints, more general linear constraints with variable right-hand side, i.e., $g_i(x, y) = \langle d, y \rangle - c_i(x)$, satisfy the active ∇-monotonicity condition at any $x \in X$ provided that c_i is a convex function such that $C(x)$ includes the sublevel set constraint $c_i(y) \le c_i(x)$ as well:

$$
\begin{aligned}
\langle \nabla_x g_i(x, y) + \nabla_y g_i(x, y), y - x \rangle &= \langle d - \nabla c_i(x), y - x \rangle \\
&= -\langle \nabla c_i(x), y - x \rangle + (c_i(x) - \langle d, x \rangle) \\
&\ge -\langle \nabla c_i(x), y - x \rangle \\
&\ge c_i(x) - c_i(y) \\
&\ge 0.
\end{aligned}
$$

Shared constraints are a frequent feature in noncooperative games. A linear equality shared constraint is given by

$$
\langle a_1, x_1 \rangle + \cdots \cdots + \langle a_\ell, x_\ell \rangle = b \tag{11}
$$

for some $a_i \in \mathbb{R}^{n_i}$, $b \in \mathbb{R}$. In the reformulation as (QEP) through the Nikaido-Isoda bifunction it is described by the 2ℓ bifunctions

$$
g_i(x, y) = \langle a_i, y_i \rangle + \sum_{j \neq i} \langle a_j, x_j \rangle - b, \qquad i = 1, \ldots \ell,
$$

and $-g_i(x, y)$. All the above constraints satisfy the active ∇-monotonicity condition (9) at any $x \in X$. In fact, $x \in X$ guarantees that (11) holds and therefore any $y \in C(x)$ satisfies

$$
\langle a_i, y_i \rangle - \langle a_i, x_i \rangle = b - \sum_{j \neq i} \langle a_j, x_j \rangle - \langle a_i, x_i \rangle = b - \sum_{j=i}^{\ell} \langle a_i, x_i \rangle = 0.
$$

Therefore,

$$
\langle \nabla_x g_i(x, y) + \nabla_y g_i(x, y), y - x \rangle = \sum_{j=1}^{\ell} (\langle a_j, y_j \rangle - \langle a_j, x_j \rangle) = 0
$$

holds for any $y \in C(x)$ as $g_i(x, y) = 0$ for all i's. Moreover, the condition $y(x) \in X$ is guaranteed as well if there are no other constraints since any $y \in C(x)$ satisfies

$$
\sum_{j=1}^{\ell} \langle a_j, y_j \rangle = \sum_{j=1}^{\ell} \langle a_j, x_j \rangle = b.
$$

REMARK 1. In [24] some classes of constraints are explicitly covered within KKT type methods for QVIs: moving set, box constraints, linear constraints with variable right-hand side, binary and bilinear constraints. The first three are covered in the framework of this paper too but the last two are not. On the other side, Proposition 1 provides additional classes of constraints that have not been considered in [24].

3.   Descent algorithm

Theorem 5 and Corollary 1 provide the basic tools to design a solution method in the footsteps of descent algorithms for EPs (see, for instance, [49, 12, 9]): if the current iterate x^k is not a stationary point of the gap function φ, a step along the descent direction $y(x^k) - x^k$ is taken exploiting some inexact line search.

Algorithm

(0) Choose $\beta, \gamma \in (0,1)$, $x^0 \in X$ and set $k = 0$.
(1) Compute $y(x^k) = \arg\min\{f(x^k, y) + h(x^k, y) \ : \ y \in C(x^k)\}$.
(2) If $d^k := y(x^k) - x^k = 0$, then STOP.
(3) Compute the smallest non-negative integer s such that

$$\varphi(x^k + \gamma^s d^k) - \varphi(x^k) \le -\beta\,\gamma^{2s}\,\|d^k\|.$$

4) Set $t_k = \gamma^s$, $x^{k+1} = x^k + t_k\,d^k$, $k = k+1$ and goto Step 1.

Anyway, some meaningful differences hold with the EP case. Beyond the nonsmoothness of the gap function which can be anyhow handled well as shown in Section 2, the sequence $\{x^k\}$ may be unfeasible: $x^{k+1} \notin X$ may occur even if $x^k \in X$, unlikely the case of EPs. In fact, $y(x^k) \in C(x^k)$ does not necessarily yield $y(x^k) \in X$ and the convexity of X is not guaranteed unless the mappings $x \longmapsto g_i(x, x)$ are quasiconvex. If both conditions are met, which clearly happens if $X = \mathbb{R}^n$, then the algorithm generates a feasible sequence and convergence is achieved.

THEOREM 6. *Suppose $X \subseteq D_S$, X is convex and $C(x) \subseteq X$ holds for any $x \in X$. If f is strictly ∇-monotone on X and the constraining functions g_i satisfy the active ∇-monotonicity condition (9) on X, then either Algorithm stops at a solution of (QEP) after a finite number of iterations or produces a sequence $\{x^k\}$ such that any of its cluster points solves (QEP).*

PROOF. The line search procedure is finite. Otherwise, some iteration k satisfies

$$\varphi(x^k + \gamma^s d^k) - \varphi(x^k) > -\beta\,\gamma^{2s}\,\|d^k\|$$

for all $s \in \mathbb{N}$. Therefore, taking the maximum limit as $s \to +\infty$ leads to

$$\varphi^\circ(x^k; d^k) \ge \limsup_{s \to \infty}\ \gamma^{-s}\left[\varphi(x^k + \gamma^s d^k) - \varphi(x^k)\right] \ge 0,$$

in contradiction with Theorem 5 b) since x^k does not satisfy the stopping criterion of Step 2 and therefore does not solve (QEP). Indeed, the algorithm stops at x^k after a finite number of iterations at Step 2 if and only if x^k solves (QEP) thanks to Theorem 1.

Suppose the algorithm generates an infinite sequence $\{x^k\}$ and let x^* be any cluster point of the sequence. Since the map C is closed, then X is a closed set and thus $x^* \in X$. Moreover, $\hat{x}^\ell \to x^*$ for some subsequence $\{\hat{x}^\ell\}$, i.e., $\hat{x}^\ell = x^{k_\ell}$ for some $k_\ell \uparrow +\infty$ as $\ell \uparrow +\infty$. Moreover, Theorem 2 guarantees $\hat{d}^\ell = y(\hat{x}^\ell) - \hat{x}^\ell \to d^* = y(x^*) - x^*$. Therefore, $d^* = 0$ guarantees that x^* solves (EP) by Theorem 1. Ab absurdo, suppose $d^* \ne 0$. Since the sequence $\{\varphi(x^k)\}$ is monotone decreasing and bounded by below, it has a limit and thus

$$\lim_{\ell \to \infty}\left[\varphi(\hat{x}^\ell) - \varphi(\hat{x}^{\ell+1})\right] = 0$$

holds as well. Moreover, the stepsize rule guarantees

$$\varphi(\hat{x}^\ell) - \varphi(\hat{x}^{\ell+1}) \ge \varphi(x^{k_\ell}) - \varphi(x^{k_\ell+1}) \ge \beta\,\hat{t}_\ell^2\,\|\hat{d}^\ell\| > 0$$

for $\hat{t}_\ell = t_{k_\ell}$, Therefore, $\hat{t}_\ell \to 0$ as $\ell \to +\infty$ since $d^* \ne 0$. Moreover, the inequality

$$\varphi\left(\hat{x}^\ell + \hat{t}_\ell\,\gamma^{-1}\,\hat{d}^\ell\right) - \varphi(\hat{x}^\ell) > -\beta\,(\hat{t}_\ell\,\gamma^{-1})^2\,\|\hat{d}^\ell\|$$

holds for all $\ell \in \mathbb{N}$ by the line search procedure, while the mean value theorem (see [14, Theorem 2.3.7]) guarantees the existence of some $\theta_\ell \in (0,1)$ and some generalized gradient $\hat{\xi}^\ell$ of φ at $\hat{x}^\ell + \theta_\ell \hat{t}_\ell \gamma^{-1} \hat{d}^\ell$ such that

$$\varphi\left(\hat{x}^\ell + \hat{t}_\ell \gamma^{-1} \hat{d}^\ell\right) - \varphi(\hat{x}^\ell) = \langle \xi^\ell, \hat{t}_\ell \gamma^{-1} \hat{d}^\ell \rangle.$$

Hence, the two inequalities together provide

$$\langle \xi^\ell, \hat{d}^\ell \rangle > -\beta \hat{t}_\ell \gamma^{-1} \|\hat{d}^\ell\|.$$

and thus

$$\varphi^\circ\left(\hat{x}^\ell + \theta_\ell \hat{t}_\ell \gamma^{-1} \hat{d}^\ell; \hat{d}^\ell\right) > -\beta \hat{t}_\ell \gamma^{-1} \|\hat{d}^\ell\|$$

follows immediately from the definition of the generalized gradient. The upper semicontinuity of φ° as a function of (x,d) (see [14, Proposition 2.1.1]) guarantees

$$\varphi^\circ(x^*; d^*) \geq \limsup_{\ell \to \infty} \varphi^\circ\left(\hat{x}^\ell + \theta_\ell \hat{t}_\ell \gamma^{-1} \hat{d}^\ell; \hat{d}^\ell\right) \geq 0$$

since $\hat{x}^\ell \to x^*$, $\hat{d}^\ell \to d^*$, and $\hat{t}_\ell \to 0$ imply $\hat{x}^\ell + \theta_\ell \hat{t}_\ell \gamma^{-1} \hat{d}^\ell \to x^*$. Therefore, x^* solves (QEP), otherwise it should be $\varphi^\circ(x^*; d^*) < 0$ by Theorem 5. $\qquad\square$

Notice that the existence of cluster points is not guaranteed by the assumption of the above theorem. Obviously, they do exist whenever X is compact, which may hold true, for instance, in the case of a compact moving set (see (10) for $\alpha < 1$). Anyway, X may be unbounded or even the whole space $\mathbb{R}^n$ in meaningful cases as the examples of Section 2 show. Therefore, the existence of cluster points is a relevant issue that the error bound given in the next section helps addressing.

4. Error bounds

Error bounds can be developed to estimate the distance of a given point from some solution of (QEP). In this section an error bound is given relying on the value of the gap function φ at the considered point. Since the algorithm of Section 3 is a descent method for the minimization of φ, this error bound provides conditions that guarantee the boundedness of the sequence generated by the algorithm.

Clearly, the error bound requires the existence of a solution of (QEP) (see, for instance, [11, 16] for existence results). Also, beyond suitable monotonicity and continuity conditions on the data, the result relies on the possibility to control the optimal solution

$$y(x,z) := \arg\min\{f(x,y) + h(x,y) \ : \ y \in C(z)\} \tag{12}$$

of the inner optimization problem in the gap function subject to some perturbation $C(z)$ of the feasible region $C(x)$.

THEOREM 7. *Let $x^* \in X$ be a solution of (QEP). Suppose there exist $\mu, L_1, L_2, L_3 > 0$ such that*
 a) $f(x,x^*) + f(x^*,x) \leq -\mu\|x - x^*\|^2$,
 b) $\|\nabla_y f(x,z)\| \leq L_1$,
 c) $\|\nabla_y f(x,z) - \nabla_y f(x^*,z)\| \leq L_2 \|x - x^*\|$,
 d) $\|y(x,x) - y(x,x^*)\| \leq L_3 \|x - x^*\|$
hold for any $x \in X$ and $z \in C(x)$. Then, any $x \in X$ satisfies the inequality

$$\|x - x^*\| \leq (L_2\sqrt{\varphi(x)} + \sqrt{\tau}(L_1 + L_1 L_3))/\mu\sqrt{\tau}. \tag{13}$$

PROOF. Given any $x \in X$, the optimality conditions for $y(x) = y(x, x)$ imply

$$\langle \nabla_y f(x, y(x)), x - y(x) \rangle \geq \langle \nabla_y h(x, y(x)), y(x) - x \rangle$$

while the strong convexity of $h(x, \cdot)$ guarantees

$$0 = h(x, x) \geq h(x, y(x)) + \langle \nabla_y h(x, y(x)), x - y(x) \rangle + \tau \|y(x) - x\|^2.$$

Therefore, the inequalities

$$\langle \nabla_y f(x, y(x)), x - y(x) \rangle \geq h(x, y(x)) + \tau \|y(x) - x\|^2 \geq 0 \tag{14}$$

follow taking into account that h is non-negative.

Setting $y^*(x) := y(x, x^*)$, then the following inequalities hold

$$
\begin{aligned}
0 &\leq \langle \nabla_y f(x, y(x)), x - y(x) \rangle = \langle \nabla_y f(x, y(x)) - \nabla_y f(x^*, y(x)) + \nabla_y f(x^*, y(x)), x - y(x) \rangle \\
&\leq \|\nabla_y f(x, y(x)) - \nabla_y f(x^*, y(x))\| \, \|x - y(x)\| + f(x^*, x) - f(x^*, y(x)) \\
&\leq L_2 \|x - x^*\| \, \|x - y(x)\| + f(x^*, x) + f(x^*, y^*(x)) - f(x^*, y(x)) \\
&\leq L_2 \|x - x^*\| \, \|x - y(x)\| - f(x, x^*) - \mu \|x - x^*\|^2 + \langle \nabla_y f(x^*, y^*(x)), y^*(x) - y(x) \rangle \\
&\leq L_2 \|x - x^*\| \, \|x - y(x)\| - f(x, x^*) - \mu \|x - x^*\|^2 + \|\nabla_y f(x^*, y^*(x))\| \, \|y^*(x) - y(x)\| \\
&\leq L_2 \|x - x^*\| \, \|x - y(x)\| + \langle \nabla_y f(x, x), x - x^* \rangle - \mu \|x - x^*\|^2 + L_1 L_3 \|x - x^*\| \\
&\leq L_2 \|x - x^*\| \, \|x - y(x)\| + L_1 \|x - x^*\| - \mu \|x - x^*\|^2 + L_1 L_3 \|x - x^*\|.
\end{aligned}
$$

The second inequality follows from the convexity of $f(x^*, \cdot)$ coupled with the Schwarz inequality, the third from assumption c) and x^* solving (QEP), the forth from assumption a) and the convexity of $f(x^*, \cdot)$, the fifth again from the Schwarz inequality while the last two from the convexity of $f(x, \cdot)$ and assumptions b) and d). Therefore, the inequality

$$\|x - x^*\| \leq (L_2 \|x - y(x)\| + L_1 + L_1 L_3)/\mu$$

holds as well. The error bound (13) follows since $\|x - y(x)\| \leq \sqrt{\varphi(x)/\tau}$. Indeed, the gap function φ satisfies

$$
\begin{aligned}
\varphi(x) &= -f(x, y(x)) - h(x, y(x)) \\
&\geq \langle \nabla_y f(x, y(x)), x - y(x) \rangle - h(x, y(x)) \\
&\geq \tau \|y(x) - x\|^2,
\end{aligned}
$$

where the first inequality is due to the convexity of $f(x, \cdot)$ and the second is actually the left inequality in (14). □

Condition a) amounts to the strong monotonicity of f somehow restricted to the solution x^*. It is worth stressing that strong monotonicity does not guarantee the uniqueness of the solution differently from EPs.

EXAMPLE 1. Consider (QEP) with $n = 1$, the strongly monotone bifunction $f(x, y) = x(y - x)$ and the moving set $C(x) = [x - 1, x] = [-1, 0] + x$. It turns ou that $X = \mathbb{R}$ while the solution set is $S = (-\infty, 0]$.

Condition d) addressed the Lipschitz behaviour of the optimal solution (12) with respect to perturbations of the feasible region. Since the objective function $f(x, \cdot) + h(x, \cdot)$ is strongly convex, a classical sensitivity result (see, for instance, [28, Theorem 5.2]) leads to the following sufficient conditions for d) to hold.

PROPOSITION 2. *Let $x^* \in X$ be any solution of (QEP). If the vectors $\nabla_y g_i(x^*, y(x^*))$ for all i's such that $g_i(x^*, y(x^*)) = 0$ are linearly independent, then there exists $\delta > 0$ such that condition d) of Theorem 7 holds at any $x \in D \cap B(x^*, \delta)$.*

Notice that the above result holds true just in a neighborhood of a solution x^* of (QEP). As a consequence, it yields the error bound as well as the convergence of the sequence generated by the algorithm locally.

A global result for condition $d)$ can be proved if the constraining mapping C describes either a moving set

$$C(x) = K + t(x) \tag{15}$$

for some translation function $t : \mathbb{R}^n \to \mathbb{R}^n$, or a so-called "expanding set"

$$C(x) = s(x)\, K, \tag{16}$$

where $s : \mathbb{R}^n \to \mathbb{R}_{++}$ provides the magnitude of the expansion at each point.

PROPOSITION 3. *Let $x^* \in X$ be any solution of (QEP). Suppose the set $K \subseteq \mathbb{R}^n$ is convex, closed and bounded, the functions t, s and $\nabla_y f(x, \cdot) + \nabla_y h(x, \cdot)$ are Lipschitz continuous for any $x \in \mathbb{R}^n$. If the constraining multifunction C is given by (15) or it is given by (16), then condition d) of Theorem 7 holds at any $x \in D$.*

PROOF. Given any $x \in D$, consider $p(z) = f(x, z) + h(x, z)$, $y = y(x)$ and denote the unique minimum point $y(x, x^*)$ of p over $C(x^*)$ by y^*. Therefore, the optimality conditions for y and y^* read

$$\langle \nabla p(y), z - y \rangle \geq 0 \qquad \forall\, z \in C(x),$$

$$\langle \nabla p(y^*), z - y^* \rangle \geq 0 \qquad \forall\, z \in C(x^*).$$

If C is given by (15), then $y^* - t(x^*) + t(x) \in C(x)$ and $y - t(x) + t(x^*) \in C(x^*)$ so that summing the optimality conditions with these choices of z gives

$$\langle \nabla p(y) - \nabla p(y^*), y^* - t(x^*) + t(x) - y \rangle \geq 0.$$

As $h(x, \cdot)$ is strongly convex with modulus τ, then also p is strongly monotone and therefore its gradient $\nabla p = \nabla_y f(x, \cdot) + \nabla_y h(x, \cdot)$ is strongly monotone with the same modulus (see, for instance, [39, Chap IV - Theorem 4.1.4]). Hence, the above inequality implies

$$\langle \nabla p(y) - \nabla p(y^*), t(x) - t(x^*) \rangle \geq \langle \nabla p(y) - \nabla p(y^*), y - y^* \rangle \geq \tau \|y - y^*\|^2.$$

Since t and ∇p are Lipschitz, the inequalities

$$\langle \nabla p(y) - \nabla p(y^*), t(x) - t(x^*) \rangle \leq \|\nabla p(y) - \nabla p(y^*)\| \|t(x) - t(y)\|$$

$$\leq L\, \|y - y^*\| \|x - x^*\|$$

hold for some $L > 0$. As a consequence, $\|y - y^*\| \leq \tau^{-1} L \|x - x^*\|$, i.e., condition d) of Theorem 7 holds with $L_4 = \tau^{-1} L$.

If C is given by (16), then $s(x)y^*/s(x^*) \in C(x)$ and $s(x^*)y/s(x) \in C(x^*)$ so that summing the optimality conditions for y and y^* with these choices of z after being multiplied respectively by $s(x^*)$ and $s(x)$ gives

$$\langle \nabla p(y) - \nabla p(y^*), s(x)y^* - s(x^*)y \rangle \geq 0.$$

Therefore, the strong monotonicity of ∇p implies

$$[s(x) - s(x^*)]\langle \nabla p(y) - \nabla p(y^*), y^* \rangle \geq s(x^*)\langle \nabla p(y) - \nabla p(y^*), y - y^* \rangle \geq \tau s(x^*)\|y - y^*\|^2.$$

Since s and ∇p are Lipschitz, $y^*/s(x^*) \in K$ and K is bounded, the inequalities

$$
\begin{aligned}
[s(x) - s(x^*)]\langle \nabla p(y) - \nabla p(y^*), y^* \rangle &\leq |s(x) - s(x^*)| \|\nabla p(y) - \nabla p(y^*)\| \|y^*\| \\
&\leq L \|x - x^*\| \|y - y^*\| \|y^*\| \\
&\leq LM s(x^*) \|x - x^*\| \|y - y^*\|
\end{aligned}
$$

hold for some $L, M > 0$. As a consequence, $\|y - y^*\| \leq \tau^{-1} LM \|x - x^*\|$, i.e., condition d) of Theorem 7 holds with $L_4 = \tau^{-1} LM$. $\square$

REMARK 2. If C is given by (16) and $0 \in \operatorname{int} K$, then $x \in X$ if and only if $s(x) \geq \gamma_K(x)$, where $\gamma_K(x) := \inf\{\lambda > 0 : x \in \lambda K\}$ denotes the gauge of K.

REMARK 3. The result of Proposition 3 holds also if the multifunction C is given by $C(x) = s(x)K + t(x)$ (just put together the tricks of both proofs). Furthermore, notice that this is the particular case of the moving set (6) with $p = n$ and $Q(x) = s(x)I$.

A further error bound can be achieved when the squared distance is chosen as h in the footsteps of a known result for quasi-variational inequalities [2].

THEOREM 8. *Let x^* be a solution of (QEP). Suppose*
a) there exists $\mu > 0$ such that $f(x,y) + f(y,x) \leq -\mu\|x - y\|^2$ holds for any $x, y \in X$
b) $x^* \in C(x)$ for any $x \in C(x^*)$,
c) $h(x,y) = \alpha \|y - x\|^2$ for some $\alpha \in (0, \mu)$.
Then any $x \in C(x^*)$ satisfies

$$
\|x - x^*\| \leq \sqrt{\varphi(x)/(\mu - \alpha)}. \tag{17}
$$

PROOF. Given any $x \in C(x^*)$, the following inequalities hold

$$
\begin{aligned}
\varphi(x) &\geq -f(x, x^*) - \alpha \|x - x^*\|^2 \\
&\geq f(x^*, x) + \mu \|x - x^*\|^2 - \alpha \|x - x^*\|^2 \\
&\geq (\mu - \alpha) \|x - x^*\|^2,
\end{aligned}
$$

where the first follows from the definition of φ coupled with assumption c) as assumption b) guarantees $x^* \in C(x)$, the second from assumption a) while the last from x^* solving (QEP). $\square$

Notice that the assumptions of Theorem 8 are simpler than those of Theorem 7. However, the error bound (17) holds just locally, that is within $C(x^*)$, while the error bound (13) given by Theorem 7 is global.

5. Conclusions

The paper analyses gap functions for QEPs and provides a first basic descent algorithm for quasi-equilibria together with some error bounds. Further work can be carried out trying to improve the results in a few directions.

First, the algorithm of this paper converges under monotonicity assumptions that guarantee the equivalence between stationary points of the gap function and quasi-equilibria. It would be worthwhile to develop solution methods under assumptions not implying such an equivalence (see, for instance, [6, 8] for the case of EPs).

Next, the algorithm relies on the condition $y(x) \in X$ to guarantee the feasibility of the generated sequence. Though this assumption is satisfied in a number of cases, it appears a bit restrictive. Hence, new solution algorithms not requiring it are worth to be investigated. These algorithms could be devised either by combining descent and penalization techniques or exploiting D-gap functions to avoid feasibility issues.

Finally, a new class of gap functions based on the linearization of the constraining functions g_i's could be developed in order to simplify the computation of the gap function at a given point. Penalization techniques properly applied to this kind of gap function could drive to further solution algorithms (see [8] for the case of EPs).

Acknowledgments.

References

[1] K.J. Arrow , G. Debreu, Existence of an equilibrium for a competitive economy, Econometrica 22 (1954), 265-290.

[2] D. Aussel, R. Correa, M. Marechal, Gap functions for quasivariational inequalities and generalized Nash equilibrium problems, J. Optim. Theory Appl. 151 (2011), 474-488.

[3] C. Baiocchi, A. Capelo, Variational and quasivariational inequalities: applications to free boundary problems, Wiley, New York, 1984.

[4] A. Bensoussan, M.Goursat, J.-L. Lions, Contrôle impulsionnel et inéquations quasi-variationnelles stationnaires. C. R. Acad. Sci. Paris Sér A 276 (1973), 1279-1284.

[5] A. Bensoussan, J.-L. Lions, Nouvelle formulation de problèmes de contrôle impulsionnel et applications. C. R. Acad. Sci. Paris Sér A 276 (1973), 1189-1192.

[6] G. Bigi, M. Castellani, M. Pappalardo, A new solution method for equilibrium problems, Optim. Methods Softw. 24 (2009), 895-911.

[7] G. Bigi, M. Castellani, M. Pappalardo, M. Passacantando, Existence and solution methods for equilibria, European J. Oper. Res. 227 (2013), 1-11.

[8] G. Bigi, M. Passacantando, Gap functions and penalization for solving equilibrium problems with nonlinear constraints, Comput. Optim. Appl. 53 (2012), 323-346.

[9] G. Bigi, M. Passacantando, Descent and penalization techniques for equilibrium problems with nonlinear constraints, J. Optim. Theory Appl. 164 (2015), 804-818.

[10] E. Blum, W. Oettli, From optimitazion and variational inequalities to equilibrium problems, Math. Student 63 (1994), 123-145.

[11] M. Castellani, M. Giuli, An existence result for quasiequilibrium problems in separable Banach spaces, J. Math. Anal. Appl. 425 (2015), 85-95.

[12] O. Chadli, I.V. Konnov, J.C. Yao, Descent methods for equilibrium problems in a Banach space, Comput. Math. Appl. 48 (2004), 609-616.

[13] D. Chan, J.-S. Pang, The generalized quasi-variational inequality problem, Math. Oper. Res. 7 (1982), 211-222.

[14] F.H. Clarke, Optimization and nonsmooth analysis, John Wiley and Sons, New York, 1983.

[15] G. Cohen, Auxiliary problem principle extended to variational inequalities, J. Optim. Theory Appl. 59 (1988), 325-333.

[16] P. Cubiotti, Existence of solutions for lower semicontinuous quasi-equilibrium problems, Comput. Math. Appl. 30 (1995), 11-22.

[17] A. Dreves, F. Facchinei, C. Kanzow, S. Sagratella, On the solution of the KKT conditions of generalized Nash equilibrium problems, SIAM J. Optim. 21(2011), 1082-1108.

[18] A. Dreves, C. Kanzow, Nonsmooth optimization reformulations characterizing all solutions of jointly convex generalized Nash equilibrium problems, Comput. Optim. Appl. 50 (2011), 23-48.

[19] H. Dietrich, Optimal control problems for certain quasivariational inequalities, Optimization 49 (2001), 67-93.

[20] F. Facchinei, A. Fischer, V. Piccialli, On generalized Nash games and variational inequalities, Oper. Res. Lett. 35 (2007), 159-164.

[21] F. Facchinei, A. Fischer, V. Piccialli, Generalized Nash equilibrium problems and Newton methods, Math. Program. 117 (2009), 163-194.

[22] F. Facchinei, C. Kanzow, Penalty methods for the solution of generalized Nash equilibrium problems, SIAM J. Optim. 20 (2010), 2228-2253.

[23] F. Facchinei, C. Kanzow, Generalized Nash equilibrium problems, Ann. Oper. Res. 175 (2010), 177-211.

[24] F. Facchinei, C. Kanzow, S. Sagratella, Solving quasi-variational inequalities via their KKT conditions, Math. Program. 144 (2014), 369-412.

[25] F. Facchinei, C. Kanzow, S. Sagratella, QVILIB: A library of quasi-variational inequality test problems, Pac. J. Optim. 9 (2013), 225-250.

[26] F. Facchinei, J.S. Pang, Finite-dimensional variational inequalities and complementarity problems, Springer, New York, 2003.

[27] F. Facchinei, V. Piccialli, M. Sciandrone, Decomposition algorithms for generalized potential games, Comput. Optim. Appl. 50 (2011), 237-262.

[28] A.V. Fiacco, Y. Ishizuka, Sensitivity and stability analysis fornonlinear programming, Ann. Oper. Res. 27 (1990), 215-236.

[29] M. Fukushima, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, Math. Program. 53 (1992), 99-110.

[30] M. Fukushima, A class of gap functions for quasi-variational inequality problems, J. Ind. Manag. Optim. 3 (2007), 165-171.

[31] M. Fukushima, Restricted generalized Nash equilibria and controlled penalty algorithm, Comput. Manag. Sci. 8 (2011), 201-218.

[32] F. Giannessi, Separation of sets and gap functions for quasi-variational inequalities, Variational inequalities and network equilibrium problems (F. Giannessi and A. Maugeri, ed.s), Plenum:, New York, 2005, pp. 101-121.

[33] R. Gupta, A. Mehra, Gap functions and error bounds for quasivariational inequalities. J. Glob. Optim. 53 (2012), 737-748.

[34] P.T. Harker, Generalized Nash games and quasi-variational inequalities, European J. Oper. Res. 54 (1991), 81-94.

[35] N. Harms, C. Kanzow, O. Stein, Smoothness properties of a regularized gap function for quasi-variational inequalities, Optim. Methods Softw. 29 (2014), 720-750.

[36] N. Harms, T. Hoheisel, C. Kanzow, On a smooth dual gap function for a class of quasi-variational inequalities, J. Optim. Theory Appl. 163 (2014), 413-438.

[37] A. von Heusinger, C. Kanzow, Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions, Comput. Optim. Appl. 43 (2009), 353-377.

[38] A. von Heusinger, C. Kanzow, Relaxation methods for generalized Nash equilibrium problems with inexact line search, J. Optim. Theory Appl. 143 (2009), 159-183.

[39] J.-B. Hiriart-Urruty, C. Lemarèchal, Convex analysis and minimization algorithms I, Springer, Berlin,1993.

[40] W. Hogan, Directional derivatives for extremal-value functions with applications to the completely convex case, Oper. Res. 21 (1973), 188-209.

[41] W. Hogan, Point-to-set maps in mathematical programming, SIAM Rev. 15 (1973), 591-603.

[42] A.F. Izmailov, M.V. Solodov, On error bounds and Newton-type methods for generalized Nash equilibrium problems, Comput. Optim. Appl. 59 (2014), 201-218.

[43] I.V. Konnov, M.S.S. Ali, Descent methods for monotone equilibrium problems in Banach spaces, J. Comput. Appl. Math. 188 (2006), 165-179.

[44] I.V. Konnov, O.V. Pinyagina, Descent method with respect to the gap function for nonsmooth equilibrium problems, Russian Math. 47 (2003), 67-73.

[45] I.V. Konnov, O.V. Pinyagina, A descent method with inexact linear search for nonsmooth equilibrium problems, Comput. Math. Math. Phys. 48 (2008), 1777-1783.

[46] K. Kubota, M. Fukushima, Gap function approach to the generalized Nash equilibrium problem, J. Optim. Theory Appl. 144 (2010), 511-531.

[47] U. Mosco, Implicit variational problems and quasi variational inequalities, Nonlinear operators and the calculus of variations (J.P. Gossez, E.J. Lami Dozo, J. Mawhin and L. Waelbroeck ed.s), Lecture Notes in Mathematics 543, 1979, pp. 83-156.

[48] G. Mastroeni, On auxiliary principle for equilibrium problems, Equilibrium problems and variational models (P. Daniele, F. Giannessi, A. Maugeri ed.s), Kluwer , Dordrecht, 2003, pp. 289-298.

[49] G. Mastroeni, Gap functions for equilibrium problems. J. Glob. Optim. 27 (2003), 411-426.

[50] J.F. Nash, Equilibrium points in n-person games, Proc. Natl. Acad. Sci. USA 36 (1950), 48-49.

[51] Y. Nesterov, L. Scrimali, Solving strongly monotone variational and quasi-variational inequalities, Discrete Cont. Dyn. Syst. 31 (2011), 1383-1396.

[52] T.T.V. Nguyen, T.P.D. Nguyen, J.J. Strodiot, V.H. Nguyen, A class of hybrid methods for quasi-variational inequalities, Optim. Lett. 8 (2014), 2211-2226.

[53] J. Outrata, M. Kocvara, On a class of quasi-variational inequalities, Optim. Methods Softw. 5 (1995), 275-295.

[54] J. Outrata, J. Zowe, A Newton method for a class of quasi-variational inequalities, Comput. Optim. Appl. 4 (1995), 5-21.

[55] J.-S. Pang, M. Fukushima, Quasi-variational inequalities, generalized Nash equilibria, and multileader-follower games, Comput. Manag. Sci. 2 (2005), 21-56.

[56] B. Panicucci, M. Pappalardo, M. Passacantando, On solving generalized Nash equilibrium problems via optimization, Optim. Lett. 3 (2009), 419-435.

[57] I.P. Ryazantseva, First-order methods for certain quasi-variational inequalities in Hilbert space, Comput.l Math. Math. Phys. 47 (2007), 183-190.

[58] J.J. Strodiot, T.T.V. Nguyen, V.H. Nguyen, A new class of hybrid extragradient algorithms for solving quasi-equilibrium problems, J. Glob. Optim. 56 (2013), 373-397.

[59] K. Taji, On gap functions for quasi-variational inequalities, Abstr. Appl. Anal. 2008 (2008), Art. Numb. 531361.

[60] J. Zhang, B. Qu, N. Xiu, Some projection-like methods for the generalized Nash equilibria, Comput. Optim. Appl. 45 (2010), 89-109.