An Augmented Lagrangian Method for Quasi-Equilibrium Problems

Luís Felipe Bueno* Gabriel Haeser† Felipe Lara‡ Frank Navarro Rojas§

January 18, 2019

Abstract

In this paper, we propose an Augmented Lagrangian algorithm for solving a general class of possible non-convex problems called quasi-equilibrium problems (QEPs). We define an Augmented Lagrangian bifunction associated with QEPs, introduce a secondary QEP as a measure of infeasibility and we discuss several special classes of QEPs within our theoretical framework. For obtaining global convergence under a new weak constraint qualification, we extend the notion of an Approximate Karush-Kuhn-Tucker (AKKT) point for QEPs (AKKT-QEP), showing that in general it is not necessarily satisfied at a solution, differently from its counterpart in optimization. We study some particular cases where AKKT-QEP does hold at a solution, while discussing the solvability of the subproblems of the algorithm.

Keywords: Augmented Lagrangian methods; Quasi-equilibrium problems; Equilibrium problems; Constraint qualifications; Approximate-KKT conditions.

1 Introduction

Given a nonempty set K from $\mathbb{R}^n$ and an equilibrium bifunction f on K, i.e., a bifunction $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with $f(x, x) = 0$ for all $x \in K$, the equilibrium problem is defined by

$$\text{find } x \in K \text{ such that } \; f(x, y) \geq 0, \; \forall \, y \in K. \tag{EP}$$

As was noted in [13], equilibrium problems encompass several problems found in fixed point theory, continuous optimization and nonlinear analysis, e.g. minimization

*Institute of Science and Technology, Federal University of São Paulo, São José dos Campos SP, Brazil. E-mail: lfelipebueno@gmail.com

†Department of Applied Mathematics, Institute of Mathematics and Statistics, University of São Paulo, São Paulo SP, Brazil. E-mail: ghaeser@ime.usp.br

‡Escola de Matemática Aplicada, Fundação Getulio Vargas EMAp/FGV, Praia de Botafogo, Rio de Janeiro, Brazil. E-mail: felipelaraobreque@gmail.com

§Department of Applied Mathematics, Institute of Mathematics and Statistics, University of São Paulo, São Paulo SP, Brazil. E-mail: franknr@ime.usp.br

problems, linear complementary problems, variational inequalities (VIs from now on) and vector optimization problems, among others.

On the other hand, and mainly motivated by real life problems, quasi-variational inequalities (QVIs from now on) have been introduced and studied deeply in the recent years. Recall that, given a point-to-set operator $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and a point-to-point operator $F : \mathbb{R}^n \to \mathbb{R}^n$, the QVI problem consists of

$$\text{find } x \in K(x) \text{ such that } \langle F(x), x - y\rangle \geq 0, \ \forall\, y \in K(x). \tag{QVI}$$

We say that a point $x \in \mathbb{R}^n$ is feasible if $x \in K(x)$. If $K(x) := K$, then (QVI) reduces to the usual variational inequality problem, which is also a particular case the equilibrium problem (EP).

In order to unify both approaches, the quasi-equilibrium problem (QEP from now on) have been introduced and studied. Here the problem is defined by a point-to-set operator K and an equilibrium bifunction f, where the QEP consists of

$$\text{find } x \in K(x) \text{ such that } f(x, y) \geq 0, \ \forall\, y \in K(x). \tag{QEP}$$

Therefore, QEPs encompass both EPs and QVIs simultaneously, i.e., by extension, minimization problems, linear complementary problems, generalized Nash equilibrium problems (GNEPs from now on), and many others related to economics, management and mechanics among others (see [9, 26, 36]). Moreover, the tools used for providing existence results and optimality conditions goes from convex analysis and operator theory to generalized convexity, generalized monotonicity, fixed point theory and variational analysis, i.e., such problems provide a rich area for applying theoretical results and new developments from nonlinear applied analysis (see [8, 9, 15, 19] for instance).

With respect to algorithms for solving QEPs, several developments have been made in the past 10 years. We mention here different approaches for QEPs as Extragradient methods (see [40, 41]) and the gap function approach (see [10]). The case of the Augmented Lagrangian method, which is also the main topic of this paper, have been developed in [39] for the usual minimization problem, and in [29] for the variational inequality problems. Different variants of the Augmented Lagrangian method for QVIs may be found in [32, 33, 34, 37], extending the method from VIs to QVIs.

In this paper, we propose an Augmented Lagrangian algorithm for QEPs. The main difference of our algorithm is given by its global convergence properties under weak constraint qualifications. To do this, and after an study of optimality conditions and constraints qualification for QEPs, we adapt the so-called sequential optimality conditions from nonlinear programming to QEPs (see [1]). Furthermore, it turns out that the generalization of an Approximate-KKT (AKKT) point for QEPs, which is a natural sequential optimality condition in optimization, is not necessarily satisfied at the solutions of a general QEP. So, special classes of QEPs need to be studied.

In our algorithm, we divide the constraint set $K(x)$ in two parts and we penalize only one of these parts within our (partial) Augmented Lagrangian approach. Hence, we consider a whole class of methods which are quite flexible and that can take into account the special structure of the underlying QEP in a favourable way. Since Augmented Lagrangian methods are not expected to find feasible points without strong assumptions, we provide a tendency for finding feasible points by introducing a secondary QEP as a measure of infeasibility. Hence, our global convergence theory is split into a result concerning feasibility and another one concerning optimality, as motivated by similar results in optimization (see, e.g., [12]). Finally, we provide special classes of QEPs for which the resulting EP-subproblems are easy to solve, for instance, under a monotone or pseudomonotone assumption on the Lagrangian bifunction or when the AKKT conditions are necessarily satisfied at its solutions.

The paper is organized as follows. In Section 2, we set up notation, basic definitions and preliminaries on constraint qualifications and generalized monotonicity. In Section 3, we deal with QEP-tailored constraint qualifications (CQ-QEP) and we introduce the concept of Approximate Karush-Kuhn-Tucker (AKKT) condition for QEPs (AKKT-QEP). We present classes of QEPs for which AKKT-QEP is satisfied at a solution. In Section 4, we present our Augmented Lagrangian method. We provide a compact global convergence analysis considering both feasibility and optimality of a limit point generated by our algorithm. Finally, in Section 5, we deal with properties of the feasibility of QEPs and consider some special classes of QEPs via the study of monotonicity properties of its associated Lagrangian. Finally, an example for pseudomonotone equilibrium problems is also provided.

2 Preliminaries

Given $a \in \mathbb{R}$, we define $a_+ := \max\{0, a\}$. Similarly, for a real vector x, we write x_+ for the vector where the plus-operator is applied to each component. A vector-valued function $\psi : \mathbb{R}^n \to \mathbb{R}^m$ is called convex if all component functions are convex. Finally, for a continuously differentiable bifunction $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$, we denote the partial (with respect to the second argument y) transposed Jacobian by $\nabla_y g(x, y)$. Hence, for the i-th component, $\nabla_y g_i(x, y)$ is the gradient, viewed as a column vector. A collection $a_1, \dots, a_m$ of vectors is called positively linearly dependent (p.l.d. from now on) if $\sum_{i=1}^m t_i a_i = 0$ for some $t_1 \geq 0, \dots, t_m \geq 0$ not all zero. Otherwise the collection is called positively linearly independent (p.l.i. from now on).

Consider a nonlinear programming problem with inequality constraints (for simplicity),

$$\min_x u(x) \quad s.t.\ c_i(x) \leq 0,\ \forall\, i \in \{1, \dots, m\}, \tag{2.1}$$

where $u : \mathbb{R}^n \to \mathbb{R}$ and $c_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \dots, m$ are assumed to be continuously differentiable. Let X denote the feasible set of problem (2.1) and $A(\bar{x}) = \{i \mid c_i(\bar{x}) = 0\}$ the index set of active constraints at a point $\bar{x} \in X$.

Definition 2.1. *Let $\bar{x} \in X$ be a feasible point. We say that $\bar{x}$ satisfies the:*

- (a) *Linear Independence Constraint Qualification (LICQ) if the gradient vectors $\nabla c_i(\bar{x})$ for $i \in A(\bar{x})$ are linearly independent.*
- (b) *Mangasarian-Fromovitz Constraint Qualification (MFCQ) if the gradients $\nabla c_i(\bar{x})$ for $i \in A(\bar{x})$ are p.l.i.*
- (c) *Constant Positive Linear Dependence (CPLD) constraint qualification if for any subset $I \subseteq A(\bar{x})$ such that the gradient vectors $\nabla c_i(\bar{x})$ for $i \in I$ are p.l.d., they remain p.l.d. for all x in a neighborhood of $\bar{x}$.*

(*d*) *Cone Continuity Property (CCP) if the set-valued mapping $C : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous at $\bar{x}$, i.e., $\limsup_{x \to \bar{x}} C(x) \subseteq C(\bar{x})$, where*

$$C(x) := \left\{ w \in \mathbb{R}^n : \; w = \sum_{i \in A(\bar{x})} \lambda_i \nabla c_i(x), \; \lambda_i \geq 0 \right\}, \text{ and}$$
$$\limsup_{x \to \bar{x}} C(x) := \left\{ w \in \mathbb{R}^n : \; \exists \; x^k \to \bar{x}, \; \exists \; w^k \to w, w^k \in C(x^k) \right\}.$$

It is know that CCP is the weakest of the conditions presented (among others), while still being a constraint qualification, implying, e.g., Abadie's CQ (see [6]). That is, when CCP holds at a local minimizer, the KKT conditions are satisfied. On the other hand, sequential optimality conditions for constrained optimization are necessarily satisfied by local minimizers, independently of the fulfillment of constraint qualifications. These conditions are used for developing stopping criteria for several important methods such as the Augmented Lagrangian method and others and for proving global convergence results to a KKT point under a weak constraint qualification (CCP, for instance). The most popular of these sequential optimality conditions is the Approximate-KKT (AKKT) [1, 38] defined below:

Definition 2.2. *(AKKT) We say that $\bar{x} \in X$ satisfies AKKT if there exist sequences $\{x^k\} \subset \mathbb{R}^n$ and $\{\lambda^k\} \subset \mathbb{R}^m_+$ such that $\lim_{k \to \infty} x^k = \bar{x}$,*

$$\lim_{k \to \infty} \left\| \nabla u(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla c_i(x^k) \right\| = 0, \text{ and } \lim_{k \to \infty} \min \left\{ -c_i(x^k), \lambda_i^k \right\} = 0,$$

for all $i = 1, \ldots, m$. Sequences $\{x^k\}$ and $\{\lambda^k\}$ are called primal AKKT and dual AKKT sequences respectively.

The following theorem states that AKKT is a true necessary optimality condition independently of the validity of any constraint qualification (see [1, 12]).

Theorem 2.1. *Let $\bar{x}$ be a local solution of problem* (2.1)*, then $\bar{x}$ satisfies AKKT.*

When an AKKT point is such that the corresponding dual sequence is bounded, it is clear that the point is a true KKT point. However, even in the unbounded case, one may prove that the KKT conditions hold under different assumptions. The weakest of such assumptions, independently of the objective function, is CCP. Theorem 2.1 is also relevant without assuming constraint qualifications, as it shows that it is possible to find a point arbitrarily close to a local solution of problem (2.1) that satisfies the KKT conditions up to a given tolerance $\epsilon > 0$. This result suggests the use of perturbed KKT conditions as stopping criterion of numerical algorithms.

In our analysis, we consider the (QEP) with a continuously differentiable bifunction f, together with the multifunction K defined as

$$K(x) = \{y \in \mathbb{R}^n : g(x, y) \leq 0\}, \tag{2.2}$$

where $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable and denotes the parameterized constraints.

Note that equality constraints can also be included, but to keep the notation simple, we consider only inequality constraints. If g depends only on y, by abuse of notation, we replace $g(x, y)$ by $g(y)$. Thus, $K(x) = \{y \in \mathbb{R}^n : \; g(y) \leq 0\} = K$ for all $x \in \mathbb{R}^n$, and (QEP) reduces to (EP).

Let x^* be a solution of the QEP with K given as in equation (2.2). Then $x^* \in K(x^*)$ and $f(x^*, y) \geq 0$ for all $y \in K(x^*)$, or equivalently,

$$f(x^*, y) \geq 0, \quad \forall \; y : \; g(x^*, y) \leq 0.$$

As $f(x^*, x^*) = 0$, it follows that x^* is a solution of the problem

$$\min_y f(x^*, y) \text{ s.t } g_i(x^*, y) \leq 0, \; \forall \; i \in \{1, \dots, m\}. \tag{2.3}$$

Assuming that a suitable constraint qualification holds at the solution x^* with respect to the set $K(x^*) \subseteq \mathbb{R}^n$, it follows that there exists some Lagrange multiplier $\lambda^* \in \mathbb{R}^m_+$ such that (x^*, λ^*) satisfies the following KKT conditions:

$$\begin{aligned}
&\nabla_y f(x^*, x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla_y g_i(x^*, x^*) = 0, \\
&\lambda_i^* \geq 0, \; g_i(x^*, x^*) \leq 0, \; \lambda_i^* g_i(x^*, x^*) = 0, \; \forall \; i \in \{1, \dots, m\}.
\end{aligned}$$

This motivates the following definition of the KKT system for a QEP:

Definition 2.3 (KKT-QEP)**.** *Consider the* (QEP) *with K given by* (2.2)*. Then the system*

$$\begin{aligned}
&\nabla_y f(x, x) + \sum_{i=1}^{m} \lambda_i \nabla_y g_i(x, x) = 0, \\
&\lambda_i \geq 0, \; g_i(x, x) \leq 0, \; \lambda_i g_i(x, x) = 0, \; \forall \; i \in \{1, \dots, m\},
\end{aligned}$$

is called the KKT conditions of the underlying (QEP)*. Every (x, λ) satisfying these conditions is called a KKT-QEP pair.*

A QEP is said to be convex if $f(x, \cdot)$ and $g(x, \cdot)$ are convex for each x (a usual assumption for QEPs – which we do not assume). Then, the KKT-QEP conditions are sufficient for optimality.

Our aim is to compute a KKT-QEP point by solving a related sequence of KKT-QEP systems from (simpler) quasi-equilibrium subproblems. In fact, in our analysis we allow for inexact solutions of the underlying subproblems. The following definition, motivated by the similar concept for optimization suggested by Definition 2.2, introduces our notion of an ϵ-stationary point of this QEP.

Definition 2.4. *Consider the* (QEP) *with K defined by* (2.2)*, and let $\epsilon \geq 0$. We call (x, λ), with $\lambda \geq 0$, an ϵ-inexact KKT-QEP pair of the* (QEP) *if the following inequalities hold:*

$$\|\nabla_y f(x, x) + \sum_{i=1}^{m} \lambda_i \nabla_y g_i(x, x)\| \leq \epsilon, \tag{2.4}$$

$$|\min\{-g_i(x, x), \lambda_i\}| \leq \epsilon, \; \forall \; i \in \{1, \dots, m\}. \tag{2.5}$$

Note that for $\epsilon = 0$ an ϵ-inexact KKT-QEP point is a standard KKT-QEP point. A limit $\bar{x}$ of ϵ-inexact KKT-QEP points $\{x_\epsilon\}_{\epsilon \to 0^+}$ (with suitable multipliers $\{\lambda_\epsilon\}_{\epsilon \to 0^+}$ that may not be convergent) will be called an AKKT-QEP point. See Definition 3.1.

We finish this section with the following monotonicity notions which will be relevant in our forthcoming analysis.

Definition 2.5. *Let S be a nonempty set from $\mathbb{R}^n$. Then an equilibrium bifunction $f : S \times S \to \mathbb{R}$ is said to be*

(a) *strongly monotone on S, if there exists a constant $\gamma > 0$ such that*

$$f(x, y) + f(y, x) \leq -\gamma \|x - y\|^2, \ \forall \ x, y \in S; \tag{2.6}$$

(b) *mononote on S, if*

$$f(x, y) + f(y, x) \leq 0, \ \forall \ x, y \in S, \tag{2.7}$$

while f is strictly monotone on S if the previous inequality is strict whenever $y \neq x$;

(c) *pseudomonotone on S if for every $x, y \in S$,*

$$f(x, y) \geq 0 \implies f(y, x) \leq 0; \tag{2.8}$$

(d) *∇_{xy}-monotone on S, if the mapping $\nabla_x f(x, \cdot)$ is monotone for any $x \in S$, that is*

$$\langle \nabla_x f(x, y) - \nabla_x f(x, z), y - z \rangle \geq 0, \ \forall \ x, y, z \in S,$$

while f is strictly ∇_{xy}-monotone on S if the previous inequality is strict whenever $y \neq z$.

Clearly, every strongly monotone bifunction is strictly monotone and every monotone bifunction is pseudomonotone. If f is (strictly) $\nabla_{xy}-$monotone on S, then f is (strictly) monotone on S by [11, Theorem 3.1].

For a further study on generalized monotonicity we refer to [11, 25].

3 Approximate-Karush-Kuhn-Tucker condition and Constraint Qualifications for QEPs

In many problems, it is natural to stop the execution of an algorithm when a stationarity measure is approximately satisfied. In this section we will show that this procedure may avoid solutions a priori. This is in contrast with what is known in nonlinear programming.

We begin by extending the concept of AKKT for QEPs (AKKT-QEP), followed by the study of some important cases of QEPs where this condition is necessarily satisfied at a solution, whereas we show that this does not happen in general. Then we will see the relationship that exists between AKKT-QEP and constraint qualifications for QEPs, together with the Augmented Lagrangian method that will be presented in the next section. This analysis yields a global convergence proof to a KKT-QEP point under a weak constraint qualification.

Definition 3.1. *(AKKT-QEP) Consider the* (QEP) *with K defined by* (2.2)*. We say that a feasible $\bar{x} \in \mathbb{R}^n$ satisfies AKKT-QEP if there exist sequences $\{x^k\} \subset \mathbb{R}^n$ and $\{\lambda^k\} \subset \mathbb{R}^m_+$ such that $\lim_{k\to\infty} x^k = \bar{x}$,*

$$\lim_{k\to\infty} \left\| \nabla_y f(x^k, x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla_y g_i(x^k, x^k) \right\| = 0, \tag{3.1}$$

and

$$\lim_{k\to\infty} \min\left\{-g_i(x^k, x^k), \lambda_i^k\right\} = 0, \tag{3.2}$$

for all $i = 1, \ldots, m$. Sequences $\{x^k\}$ and $\{\lambda^k\}$ are called, respectively, primal AKKT-QEP sequence and dual AKKT-QEP sequence.

Example 3.1. *AKKT-QEP is not necessarily satisfied at a solution. Indeed, set $f, g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $f(x, y) = -x + y$, $g(x, y) = \frac{1}{2}(x - y)^2$, and $K(x) = \{y \in \mathbb{R} : \ g(x, y) \le 0\} = \{x\}$. Clearly, the solution set of* (QEP) *is the whole real line. Set any solution $x^* \in \mathbb{R}$. If x^* is an AKKT-QEP point, then we should find sequences $\{x^k\} \subset \mathbb{R}$ and $\{\lambda^k\} \subset \mathbb{R}_+$ such that $|1 + \lambda^k(x^k - x^k)| \to 0$, which is impossible. Hence x^* is not an AKKT-QEP point.*

In [22, 34, 32], some important classes of QVIs were analyzed in the study of the Augmented Lagrangian method and of a method based on a potential reduction approach for solving the KKT system of a QVI. Let us show that for some of these classes and some others, which are extensions to general QEPs, we have the necessity of AKKT-QEP at a solution.

Theorem 3.1. *Consider* (QEP) *where the constraints have the structure*

$$g_i(x, y) = g_i^1(x) g_i^2(y) + g_i^3(x), \ \forall\, i \in \{1, \ldots, m\}, \tag{3.3}$$

with continuously differentiable functions and $x, y \in \mathbb{R}^n$. If $\bar{x}$ is a solution then the AKKT-QEP condition holds at $\bar{x}$.

Proof. We have that $\bar{x}$ is a solution of the following optimization problem:

$$\min_y f(\bar{x}, y) \ \text{ s.t. } g_i^1(\bar{x}) g_i^2(y) + g_i^3(\bar{x}) \le 0, \ \forall\, i \in \{1, \ldots, m\}.$$

By Theorem 2.1, there exist sequences $\{x^k\} \subset \mathbb{R}^n$ and $\{\lambda^k\} \subset \mathbb{R}_+^{|A(\bar{x})|}$ such that $x^k \to \bar{x}$ and

$$\|\nabla_y f(\bar{x}, x^k) + \sum_{i \in A(\bar{x})} \lambda_i^k g_i^1(\bar{x}) \nabla g_i^2(x^k)\| \to 0, \tag{3.4}$$

where $\lambda_i^k \to 0$ for $i \notin A(\bar{x})$ were equivalently replaced by a null sequence. Without loss of generality, one could also redefine, if necessary, $\lambda_i^k = 0$ if $g_i^1(\bar{x}) = 0$, and (3.4) would still hold. Let us define for k large enough and all $i \in A(\bar{x})$:

$$\bar{\lambda}_i^k := \begin{cases} 0, \ \text{if } g_i^1(x^k) = 0, \\ \lambda_i^k \frac{g_i^1(\bar{x})}{g_i^1(x^k)}, \text{otherwise.} \end{cases}$$

Note that $x^k \to \bar{x}$ and for k large enough $\bar{\lambda}_i^k$ has the same sign of λ_i^k. Moreover, since

$$\nabla_y g_i(x,y) = g_i^1(x)\nabla g_i^2(y),$$

we have that

$$\bar{\lambda}_i^k \nabla_y g_i(x^k, x^k) = \lambda_i^k g_i^1(\bar{x})\nabla g_i^2(x^k).$$

Therefore, by (3.4) and the triangular inequality,

$$\|\nabla_y f(x^k, x^k) + \sum_{i \in A(\bar{x})} \bar{\lambda}_i^k \nabla_y g_i(x^k, x^k)\| \leq$$

$$\|\nabla_y f(x^k, x^k) - \nabla_y f(\bar{x}, x^k)\| + \|\nabla_y f(\bar{x}, x^k) + \sum_{i \in A(\bar{x})} \lambda_i^k g_i^1(\bar{x})\nabla g_i^2(x^k)\| \to 0,$$

and so $\bar{x}$ is an AKKT-QEP point. $\square$

Note that setting $g_i^1(x) = 1$ and $g_i^3(x) = 0$ for all i we obtain the classical EP. Moreover, Theorem 3.1 also includes QEPs with Linear Constraints with variable rigth-hand side (see [22, 34]), that is,

$$g(x,y) = Ay - b(x), \tag{3.5}$$

where $A \in \mathbb{R}^{m\times n}$ and $b : \mathbb{R}^n \to \mathbb{R}^m$ is a given continuously differentiable function. A particularly important class of problems of type (3.5) are the QEPs with box constraints, that is,

$$g(x,y) = \begin{pmatrix} b_l(x) - y \\ y - b_u(x) \end{pmatrix}, \tag{3.6}$$

where $b_l, b_u : \mathbb{R}^n \to \mathbb{R}^n$ are given mappings which describe lower and upper bounds on the variable y, which may depend on x.

In fact, one may show the validity of AKKT for more general constraints than (3.5), but which are not contemplated by Theorem 3.1. For this, it is enough to have a constraint qualification holding at $K(\bar{x})$ for $\bar{x}$ fixed. For example, consider a problem where the constraints are of the form

$$g(x,y) = M(x)y - b(x), \tag{3.7}$$

where $M : \mathbb{R}^n \to \mathbb{R}^{m\times n}$ and $b : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable functions.

This class includes QEPs with bilinear constraints [22, 34, 32]. That is,

$$g(x,y) = \begin{pmatrix} x^T Q_1 y - b_1 \\ \vdots \\ x^T Q_m y - b_m \end{pmatrix}, \tag{3.8}$$

where each $Q_i \in \mathbb{R}^{n\times n}$ are symmetric matrices for all $i = 1, \dots, m$ and $b_i \in \mathbb{R}$ are given real numbers. To see this, set $M(x)$ as the matrix with the i-th row given by $x^T Q_i$.

For a fixed $\bar{x}$ we have that the constraints $g(\bar{x}, y) \leq 0$ are linear and so it satisfies a constraint qualification. So every solution $\bar{x}$ of (2.3) is a KKT point associated with some Lagrange multiplier $\bar{\lambda}$. Taking $x^k = \bar{x}$ and $\lambda^k = \bar{\lambda}$ for all k we have the desired result.

New, let us consider (QEP) with binary constraints [22, 34], that is, each continuously differentiable constraint $g_i(x, y)$ depends on a single pair (x_j, y_j) for some $j = j(i) \in \{1, \dots, n\}$. Then

$$K(x) = \left\{ y \in \mathbb{R}^n : g_i(x_{j(i)}, y_{j(i)}) \le 0, \ \forall\, i \in \{1, \dots, m\} \right\}. \tag{3.9}$$

This class of problems reduces to problems in which each constraint depends on one argument. In this case, let us see that AKKT-QEP is necessary at a solution.

Theorem 3.2. *Consider problem* (QEP) *with* $K(x)$ *as in* (3.9)*. Let* $\bar{x}$ *be a solution. Then* $\bar{x}$ *is an AKKT-QEP point.*

Proof. Since $\bar{x}$ is a solution of the optimization problem

$$\min_y f(\bar{x}, y) \ \text{ s.t. } g_i(\bar{x}_{j(i)}, y_{j(i)}) \le 0, \ \forall\, i \in \{1, \dots, m\},$$

by Theorem 2.1, there exist $\{x^k\} \subset \mathbb{R}^n$ and $\{\lambda^k\} \subset \mathbb{R}_+^{|A(\bar{x})|}$ such that $x^k \to \bar{x}$ and

$$\|\nabla_y f(\bar{x}, x^k) + \sum_{i \in A(\bar{x})} \lambda_i^k \nabla_y g_i(\bar{x}, x^k)\| \to 0, \tag{3.10}$$

where $\nabla_y g_i(x, y) = \big(0, \dots, 0, \partial_{y_{j(i)}} g_i(x_{j(i)}, y_{j(i)}), 0, \dots, 0\big)^\top$. Once again, note that we can redefine $\lambda_i^k = 0$ if $\nabla_y g_i(\bar{x}, x^k) = 0$. Now, define

$$\bar{\lambda}_i^k := \begin{cases} 0, \ \text{if } \partial_{y_{j(i)}} g_i(x^k_{j(i)}, x^k_{j(i)}) = 0, \\ \lambda_i^k \frac{\partial_{y_{j(i)}} g_i(\bar{x}_{j(i)}, x^k_{j(i)})}{\partial_{y_{j(i)}} g_i(x^k_{j(i)}, x^k_{j(i)})}, \ \text{otherwise.} \end{cases}$$

Note that $x^k \to \bar{x}$ and that $\bar{\lambda}_i^k$ has the same sign of λ_i^k for k large enough. Moreover,

$$\bar{\lambda}_i^k \nabla_y g_i(x^k, x^k) = \lambda_i^k \nabla_y g_i(\bar{x}^k, x^k).$$

Therefore, by (3.10), the continuity $\nabla_y f(x, y)$ and the triangular inequality,

$$\|\nabla_y f(x^k, x^k) + \sum_{i \in A(\bar{x})} \bar{\lambda}_i^k \nabla_y g_i(x^k, x^k)\| \le$$

$$\|\nabla_y f(x^k, x^k) - \nabla_y f(\bar{x}, x^k)\| + \|\nabla_y f(\bar{x}, x^k) + \sum_{i \in A(\bar{x})} \lambda_i^k \nabla_y g_i(\bar{x}, x^k)\| \to 0,$$

so $\bar{x}$ is an AKKT-QEP point. $\square$

In nonlinear optimization, a constraint qualification is needed for ensuring that a solution satisfies the KKT conditions. The same holds ture for a solution of an EP to satisfy KKT-QEP. Since AKKT is a necessary optimality condition, any property on the feasible set that guarantees that an AKKT point is KKT, is actually a constraint qualification. A constraint qualification with this property has been called a *strict* constraint qualification in [5].

On the other hand, algorithms for nonlinear optimization usually generate sequences whose limit points satisfy AKKT. From [2, 3, 4, 5, 6], it is well-known that a separate analysis of the sequences generated by the algorithm, together with the (strict) constraint qualification needed for this limit point to satisfy KKT, yields global convergence results to a KKT point under a weak constraint qualifications.

In the context of QEPs, the fact that AKKT-QEP is not necessarily satisfied at a solution has some drawbacks for algorithms that generate AKKT-QEP sequences (see [14] for a discussion around this issue in the context of GNEPs). Moreover, for an algorithm that generates AKKT-QEP sequences, conditions for ensuring that AKKT-QEP points are KKT-QEP are an important issue. These conditions are weaker than the usual MFCQ for QEPs. Therefore, this analysis provides new global convergence results for QEPs under weaker assumptions.

We list some relevant conditions below:

Definition 3.2. *Consider a continuously differentiable constraint bifunction* $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ *and a feasible point* $\bar{x} \in \mathbb{R}^n$*. We say that:*

(*a*) *LICQ-QEP holds at* $\bar{x}$ *if* $\{\nabla_y g_i(\bar{x}, \bar{x}) : i \in A(\bar{x})\}$ *is linearly independent.*

(*b*) *MFCQ-QEP holds at* $\bar{x}$ *if* $\{\nabla_y g_i(\bar{x}, \bar{x}) : i \in A(\bar{x})\}$ *is p.l.i.*

(*c*) *WCPLD-QEP holds at* $\bar{x}$ *if there exits a neighborhood* U *from* $\mathbb{R}^n$ *of* $\bar{x}$ *such that, if* $I \subseteq A(\bar{x})$ *is such that* $\{\nabla_y g_i(\bar{x}, \bar{x})\}_{i \in I}$ *is p.l.d., then* $\{\nabla_y g_i(x, x)\}_{i \in I}$ *is p.l.d. for all* $x \in U$*.*

(*d*) *WCCP-QEP holds at* $\bar{x}$ *if the set-valued mapping* $C : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *is outer semicontinuous at* $\bar{x}$*, that is,* $\limsup_{x \to \bar{x}} C(x) \subseteq C(\bar{x})$*, where*

$$
\begin{aligned}
C(x) &= \left\{ w \in \mathbb{R}^n : w = \sum_{i \in A(\bar{x})} \lambda_i \nabla_y g_i(x, x),\ \lambda_i \geq 0 \right\}, \text{ and} \\
\limsup_{x \to \bar{x}} C(x) &= \left\{ w \in \mathbb{R}^n :\ \exists\, x^k \to \bar{x},\ \exists\, w^k \to w,\, w^k \in C(x^k) \right\}.
\end{aligned}
$$

When $\bar{x} \in \mathbb{R}^n$ *is not necessarily feasible, we say that*

(*e*) *EMFCQ-QEP (Extended-MFCQ-QEP) holds at* $\bar{x}$ *if* $\{\nabla_y g_i(\bar{x}, \bar{x}) : i \in A_E(\bar{x})\}$ *is p.l.i., where* $A_E(\bar{x}) = \{i : g_i(\bar{x}, \bar{x}) \geq 0\}$*.*

To show that LICQ-QEP, MFCQ-QEP and EMFCQ-QEP are CQs for QEPs, it is enough to show that each property implies the corresponding optimization CQ for problem (2.3) at $y = \bar{x}$. However, Example 3.1 shows that WCPLD and WCCP are not CQs for QEPs, since the KKT conditions do not hold at any solution of the problem.

In order to obtain a CQ, we proceed as in [14], i.e., we require the validity on an arbitrary neighborhood of $(\bar{x}, \bar{x})$ in $\mathbb{R}^n \times \mathbb{R}^n$, and not only in points of the form (x, x). That is, we arrive at the following constraint qualifications.

Definition 3.3. *Consider a continuously differentiable constraint bifunction* $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ *and a feasible point* $\bar{x} \in \mathbb{R}^n$*. We say that:*

(*a*) *CPLD-QEP holds at $\bar{x}$ if there exists a neighborhood U from $\mathbb{R}^n \times \mathbb{R}^n$ of $(\bar{x}, \bar{x})$ such that, if $I \subseteq A(\bar{x})$ is such that $\{\nabla_y g_i(\bar{x}, \bar{x})\}_{i \in I}$ is p.l.d., then $\{\nabla_y g_i(x, y)\}_{i \in I}$ is p.l.d. for all $(x, y) \in U$.*

(*b*) *CCP-QEP holds at $\bar{x}$ if the set-valued mapping $\bar{C} : \mathbb{R}^n \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous at $(\bar{x}, \bar{x})$, that is, $\limsup_{(x,y)\to(\bar{x},\bar{x})} \bar{C}(x, y) \subseteq \bar{C}(\bar{x}, \bar{x})$, where*

$$\bar{C}(x,y) = \left\{ w \in \mathbb{R}^n : w = \sum_{i \in A(\bar{x})} \lambda_i \nabla_y g_i(x,y),\ \lambda_i \geq 0 \right\}, \text{ and}$$

$$\limsup_{(x,y)\to(\bar{x},\bar{x})} \bar{C}(x,y) = \left\{ w \in \mathbb{R}^n :\ \exists\ (x^k, y^k) \to (\bar{x}, \bar{x}),\ \exists\ w^k \to w, w^k \in \bar{C}(x^k, y^k) \right\}.$$

Clearly CPLD-QEP (CCP-QEP) implies both WCPLD-QEP (WCCP-QEP) and the traditional CPLD (CCP) in the context of optimization for the constraints $g(\bar{x}, y) \leq 0$, which means that CPLD-QEP and CCP-QEP are CQs for QEPs. Using the same argments presented in [14], we have the following strict implications:

$$\text{LICQ-QEP} \implies \text{MFCQ-QEP} \implies \text{CPLD-QEP} \implies \text{CCP-QEP}.$$

In the next theorem we show that in order to arrive at a KKT-QEP point from an AKKT-QEP point, the WCCP-QEP condition is the weakest property that ensures this for every bifunction f.

Theorem 3.3. *The WCCP-QEP condition is equivalent to the fact that for any bifunction f, AKKT-QEP implies KKT-QEP.*

Proof. Let $\bar{x} \in \mathbb{R}^n$ be a feasible point satisfying WCCP-QEP. Let f be a bifunction such that AKKT-QEP occurs in $\bar{x}$. Then, there are sequences $\{x^k\} \subset \mathbb{R}^n$ with $x^k \to \bar{x}$ and $\{\lambda^k\} \subset \mathbb{R}_+^{|A(\bar{x})|}$ such that

$$\left\| \nabla_y f(x^k, x^k) + \sum_{i \in A(\bar{x})} \lambda_i^k \nabla_y g_i(x^k, x^k) \right\| \to 0.$$

Let $w^k = \sum_{i \in A(\bar{x})} \lambda_i^k \nabla_y g_i(x^k, x^k) \in C(x^k)$. Then $w^k \to -\nabla_y f(\bar{x}, \bar{x})$, i.e., $-\nabla_y f(\bar{x}, \bar{x}) \in \limsup_{x \to \bar{x}} C(x)$. From the WCCP-QEP condition, it follows that $-\nabla_y f(\bar{x}, \bar{x}) \in C(\bar{x})$, so $\bar{x}$ is a KKT-QEP point.

Reciprocally, assume that AKKT-QEP implies KKT-QEP for any bifunction. We will prove that WCCP-QEP holds. Indeed, let $w \in \limsup_{x \to \bar{x}} C(x)$. Then there are sequences $x^k \to \bar{x}$ and $w^k \to w$ such that $w^k \in C(x^k)$. Define $f(x, y) := \langle x - y, w \rangle$ with $\nabla_y f(x, y) = -w \in \mathbb{R}^n$. As $w^k \in C(x^k)$, there exists a sequence $\{\lambda^k\} \subset \mathbb{R}_+^{|A(\bar{x})|}$ such that

$$w^k = \sum_{i \in A(\bar{x})} \lambda_i^k \nabla_y g_i(x^k, x^k).$$

Since $\nabla_y f(x^k, x^k) = -w$ and $w^k \to w$, we have

$$\nabla_y f(x^k, x^k) + \sum_{i \in A(\bar{x})} \lambda_i^k \nabla_y g_i(x^k, x^k) \to 0.$$

Thus $\bar{x}$ satisfies AKKT-QEP, i.e., KKT-QEP holds. Hence $-\nabla_y f(\bar{x}, \bar{x}) = w \in C(\bar{x})$ and WCCP-QEP holds. □

4 An Augmented Lagrangian Method

In this section we propose an Augmented Lagrangian method for QEPs. From now on we will consider (QEP) where K is defined as follows:

$$K(x) = \{y \in \mathbb{R}^n : \, g(x, y) \leq 0, \, h(x, y) \leq 0\}, \tag{4.1}$$

where $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$ are continuously differentiable. In order to solve the QEP, we follow the approach from [32] where the authors compute a solution of a QVI by solving a sequence of suitable QVIs.

Similarly to the minimization problem, we separate the set of constraints (4.1) in two parts. The part described by g, with the difficult constraints, will be penalized, while the part described by h will define the constraints of the subproblems at each iteration. Therefore, our analysis includes the case when all the constraints are penalized, where the subproblems are unconstrained, and also when the subproblems are EPs.

Formally, at each iteration of the algorithm, the new mapping that defines the constraints of the subproblems will be defined as $K_h : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with

$$K_h(x) := \{y \in \mathbb{R}^n : \; h(x, y) \leq 0\}. \tag{4.2}$$

Given $u \in \mathbb{R}^m$ and $\rho > 0$, we define the Augmentend Lagrangian bifunction, with respect to the constraints bifunction g, as

$$L(x, y; u, \rho) = f(x, y) + \frac{\rho}{2} \sum_{i=1}^{m} \left[\max\{0, g_i(x, y) + \frac{u_i}{\rho}\} \right]^2 - \frac{\rho}{2} \sum_{i=1}^{m} \left[\max\{0, g_i(x, x) + \frac{u_i}{\rho}\} \right]^2,$$

where ρ is a suitable penalty parameter and u_i denotes a safeguarded estimate for the Lagrange multipliers λ_i associated with g_i. The Augmented Lagrangian bifunction, together with the mapping K_h, define in each iteration of the algorithm a new QEP denoted by QEP(u, ρ).

Remark 4.1. *If $g_i(x, y) = c_i(y)$ for all $i = 1, \ldots, m$, then the Augmented Lagrangian* (5.1) *reduces to the Augmented Lagrangian for EPs (see [29, Equation (2.4)]) by taking $\gamma = \frac{1}{\rho}$. Furthermore, if $f(x, y) = u(y) - u(x)$ and $g_i(x, y) = c_i(y)$ for all $i = 1, \ldots, m$, then the Augmented Lagrangian* (5.1) *reduces to the usual Augmented Lagrangian for the minimization problem* (2.1) *(see [39] for instance).*

Our algorithm, in each iteration, computes an ϵ-inexact KKT-QEP point for a tolerance $\epsilon \to 0^+$ of QEP(u, ρ) (for values of u and ρ that will be updated in each iteration) to find a KKT-QEP point of (QEP) under a weak constraint qualification.

The precise statement of our Augmented Lagrangian method is given in Algorithm 1.

Algorithm 1 Augmented Lagrangian (AL-QEP)

Step 0. Let $u^{max} \in \mathbb{R}^m_+$, $\tau \in (0,1)$, $\gamma > 1$, $\epsilon > 0$ and a sequence $\{\epsilon_k\} \subset \mathbb{R}_+, \epsilon_k \to 0$. Choose $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l$, $u^1 \in [0, u^{max}]$, $\rho^1 \geq 1$, and set $k := 1$.

Step 1. If $(x^{k-1}, (\lambda^{k-1}, \mu^{k-1}))$ is an ϵ-inexact KKT-QEP pair of (QEP): STOP.

Step 2. Compute an ϵ_k-inexact KKT-QEP pair (x^k, μ^k) of QEP(u^k, ρ^k) below:

$$\text{QEP}(u^k, \rho^k): \text{ find } x \in K_h(x) \text{ such that } L(x, y; u^k, \rho^k) \geq 0, \ \forall\, y \in K_h(x). \tag{4.3}$$

Step 3. Define $\lambda^k = \max\left\{0, u^k + \rho^k g(x^k, x^k)\right\}$.

Step 4. If

$$\left\|\max\left\{g(x^k, x^k), -\lambda^k\right\}\right\| \leq \tau \left\|\max\left\{g(x^{k-1}, x^{k-1}), -\lambda^{k-1}\right\}\right\|,$$

then set $\rho^{k+1} = \rho^k$, else set $\rho^{k+1} = \gamma\rho^k$.

Step 5. Choose $u^{k+1} \in [0, u^{max}]$, set $k := k+1$ and go to Step 1.

A natural choice of the sequence $\{u^k\}$ is $u^{k+1} = \min\{\lambda^k, u^{max}\}$. Recall that, from Definition 2.4, the pair (x^k, μ^k) computed in Step 2 must be such that:

$$\|\nabla_y L(x^k, x^k, u^k, \rho^k) + \nabla_y h(x^k, x^k)\mu^k\| \leq \epsilon_k, \tag{4.4}$$

$$\|\min\left\{-h(x^k, x^k), \mu^k\right\}\| \leq \epsilon_k. \tag{4.5}$$

Similarly to [32], our main result with respect to feasibility could be obtained requiring only that the expression in (4.4) is bounded, not necessarily converging to zero. We adopt the current presentation for clarity of exposition.

We proceed by considering the convergence properties of Algorithm 1. The analysis of the algorithm is divided into the study of feasibility and optimality. Regarding the former, note that (4.5) already implies that every limit point of $\{x^k\}$ satisfies the h-constraints.

For the discussion of feasibility with respect to the g-constraints, we introduce an auxiliary QEP, which consists of finding $x \in K_h(x)$ such that:

$$\Psi(x, y) \geq 0, \ \forall\, y \in K_h(x), \tag{4.6}$$

where

$$\Psi(x, y) = \frac{\|g_+(x, y)\|^2 - \|g_+(x, x)\|^2}{2}.$$

Note that its associated KKT-QEP system is given by:

$$\nabla_y g(x, x) g_+(x, x) + \textstyle\sum_{j=1}^l \mu_j \nabla_y h_j(x, x) = 0,$$
$$\mu_j \geq 0, \ h_j(x, x) \leq 0, \ \mu_j h_j(x, x) = 0, \ \forall\, j = 1, \ldots, l.$$

Clearly, a solution $\bar{x}$ of (QEP) related to (Ψ, K_h), denoted by QEP(Ψ, K_h), is such that $\|g_+(\bar{x}, y)\|$ is globally minimized for $y \in K_h(\bar{x})$ at $y = \bar{x}$. Hence, if the feasible region of (QEP) is non-empty, $\bar{x}$ is feasible for (QEP). Since we can only prove feasibility under strong assumptions, the following result shows that any limit point of a sequence generated by Algorithm 1 at least tends to be feasible, in the sense that it satisfies AKKT-QEP for QEP(Ψ, K_h).

Theorem 4.1. *Let $\{x^k\}$ be a sequence generated by Algorithm 1. Any limit point of $\{x^k\}$ satisfies the AKKT-QEP condition for* QEP(Ψ, K_h).

Proof. Let us assume that $x^k \to x^*$ in a subsequence. Since h is continuous and (4.5) holds, we have $h(x^*, x^*) \leq 0$ and hence x^* is feasible for QEP(Ψ, K_h).

If the sequence $\{\rho^k\}$ is bounded, we have by Step 4 that $\lim_{k\to\infty} \left\|\max\left\{g(x^k, x^k), -\lambda^k\right\}\right\| = 0$, which implies that $g(x^*, x^*) \leq 0$ and hence x^* is feasible for (QEP). This clearly gives an AKKT-QEP sequence with zero dual sequence for QEP(Ψ, K_h).

Let us now suppose that $\{\rho^k\}$ tends to infinity. From (4.4), we have that $\|\delta^k\| \leq \epsilon_k$ where

$$\delta^k = \nabla_y f(x^k, x^k) + \sum_{i=1}^{m} \max\{0, u_i^k + \rho^k g_i(x^k, x^k)\}\nabla_y g_i(x^k, x^k) + \sum_{j=1}^{l} \mu_j^k \nabla_y h_j(x^k, x^k).$$

Dividing by ρ^k, we obtain

$$\frac{\delta^k}{\rho^k} = \frac{\nabla_y f(x^k, x^k)}{\rho^k} + \sum_{i=1}^{m} \max\{0, \frac{u_i^k}{\rho^k} + g_i(x^k, x^k)\}\nabla_y g_i(x^k, x^k) + \sum_{j=1}^{l} \frac{\mu_j^k}{\rho^k} \nabla_y h_j(x^k, x^k),$$

where $\dfrac{\delta^k}{\rho^k} \to 0$, from the boundedness of $\{\epsilon_k\}$.

If $g_i(x^*, x^*) < 0$, from the boundedness of $\{u^k\}$, we have that $\max\{0, \dfrac{u_i^k}{\rho^k} + g_i(x^k, x^k)\} = 0$ for large enough k. Therefore, taking the limit we can see that

$$\lim_{k\in K} \| \sum_{i:g_i(x^*,x^*)\geq 0} g_i(x^k, x^k)\nabla_y g_i(x^k, x^k) + \sum_{j=1}^{l} \frac{\mu_j^k}{\rho^k} \nabla_y h_j(x^k, x^k)\| = 0. \tag{4.7}$$

Since $\mu_j^k \geq \dfrac{\mu_j^k}{\rho^k}$, from (4.5) and the fact that

$$\nabla_y \Psi(x^k, x^k) = \sum_{i:g_i(x^*,x^*)\geq 0} g_i(x^k, x^k)\nabla_y g_i(x^k, x^k),$$

we have that x^* satisfies the AKKT-QEP condition for QEP(Ψ, K_h). □

Corollary 4.1. *Under the assumptions of Theorem 4.1, if x^* fulfills WCCP-QEP with respect to the h-constraints describing K_h, then x^* is a KKT-QEP point of QEP(Ψ, K_h).*

Proof. It is a consequence of Theorems 3.3 and 4.1. □

Let us now state some particular cases of QEPs or additional conditions that ensure that a limit point of Algorithm 1, that is, an AKKT-QEP point for QEP(Ψ, K_h), is indeed feasible for QEP(f, K). Note that, from the proof of Theorem 4.1, this is the case when $\{\rho^k\}$ is bounded. The proofs are omitted since they are small adaptations of the ones from [34]. The first one uses the traditional argument that, under EMFCQ, a certain null linear combination of the constraint gradients can only occur with null coefficients.

Theorem 4.2. *Let x^* be a limit point of a sequence generated by Algorithm 1, and suppose that x^* satisfies EMFCQ-QEP regarding the constraints defined by g and h. Then x^* is feasible for* (QEP)*.*

In the next results, the main argument is that, under certain conditions, KKT points are indeed solutions to the problem.

Consider problem (QEP) with $K(x) = \{c(x) + S(x)w : w \in Q_1 \cap Q_2\}$, where $S(x) \in \mathbb{R}^{n\times n}$ is nonsingular for all x, $Q_i := \{x \in \mathbb{R}^n : q^i(x) \le 0\}$ for $i = 1, 2$, and $q^1 : \mathbb{R}^n \to \mathbb{R}^m$, $q^2 : \mathbb{R}^n \to \mathbb{R}^l$ are convex. Here $K(x)$ has the form of (4.1) with

$$g(x, y) = q^1(S^{-1}(x)[y - c(x)]) \quad \text{and} \quad h(x, y) = q^2(S^{-1}(x)[y - c(x)]). \tag{4.8}$$

Then we have the following

Theorem 4.3. *Let x^* be a limit point of a sequence generated by Algorithm 1 applied to a QEP of the form* (4.8) *with $Q_1 \cap Q_2 \neq \emptyset$. If x^* satisfies WCCP-QEP regarding the h-constraints, then x^* is feasible for* (QEP)*.*

The following results are also proved in a similar way

Theorem 4.4. *Consider a QEP with bilinear constraints, where g is given by* (3.8)*, where each $Q_i \in \mathbb{R}^{n\times n}$ is symmetric positive semidefinite for $i = 1, \dots, m$ and $h(x, y) = (h_1(y), \dots, h_l(y))^T$ has convex components. Let x^* be a limit point of a sequence generated by Algorithm 1. Suppose that x^* satisfies WCCP-QEP regarding the h-restrictions. Then $\bar{x}$ is feasible.*

Theorem 4.5. *Consider a QEPs with linear constraints with variable rigth-hand side given by* (3.5)*. Suppose that $rank(A) = m$. Let x^* be a limit point of a sequence generated by Algorithm 1. Suppose that x^* satisfies WCCP-QEP regarding the h-restrictions. Then $\bar{x}$ is feasible.*

Theorem 4.6. *Consider a QEPs with box constraints given by* (3.6)*. Suppose that $b_l(\bar{x}) \le b_u(\bar{x})$. Let x^* be a limit point of a sequence generated by Algorithm 1. Suppose that x^* satisfies WCCP-QEP regarding the h-restrictions. Then $\bar{x}$ is feasible.*

Let us now discuss the optimality properties of the limit points of sequences generated by the Algorithm. The following result says that when the Algorithm generates a sequence that has a feasible accumulation point, this is an AKKT-QEP point for the original problem (QEP).

Theorem 4.7. *Assume that the sequence $\{x^k\}$ generated by Algorithm 1 has a feasible limit point x^*. Then, x^* satisfies the AKKT-QEP condition for* (QEP)*.*

Proof. Let $\lim_{k\in K} x^k = x^*$. By Steps 2 and 3, we have that

$$\lim_{k\in K} \|\nabla_y f(x^k, x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla_y g_i(x^k, x^k) + \sum_{j=1}^{l} \mu_j^k \nabla_y h_j(x^k, x^k)\| = 0,$$
$$\text{and } \lim_{k\in K} \| \min\left\{-h(x^k, x^k), \mu^k\right\} \| = 0.$$

It remains to prove that $\lim_{k\in K} \min\{-g_i(x^k, x^k), \lambda_i^k\} = 0$ for all $i = 1, \ldots, m$. If $g_i(x^*, x^*) = 0$, the result follows from the continuity of g_i. Otherwise, for k large enough we have $g_i(x^k, x^k) < c < 0$ for some constant c. If $\{\rho^k\}$ is bounded, Step 4 of the algorithm implies that $\lambda_i^k \to 0$. On the other hand, the same result follows from the updating scheme of Step 3 and the boundedness of $\{u^k\}$. This concludes the proof. □

Corollary 4.2. *Under the assumptions of Theorem 4.7, if x^* fulfills WCCP-QEP with respect to the constraints g and h describing K, then x^* is a KKT-QEP point of* (QEP).

Proof. It is a consequence of Theorems 3.3 and 4.7. □

The fact that the usual assumptions (LICQ, MFCQ, CPLD, CCP, etc) used in global convergence theorems of nonlinear programming algorithms are constraint qualifications is related to the fact that the algorithm does not discard solutions a priori. If a property P is not a CQ, than there would be a problem whose solution satisfies P but not the KKT conditions. Thus, if a theorem says that under P, a limit point of a sequence generated by an algorithm satisfies KKT, such solution would never be found. As we discussed earlier, several algorithms generate AKKT sequences which, as a genuine necessary optimality condition in the context of nonlinear programming, also do not rule out solutions a priori. In the case of QEPs, the situation is different. The fact that AKKT-QEP is not an optimality condition already implies that algorithms discard solutions that do not satisfy it. Therefore, to study the impact of using an assumption that is not a CQ in such an algorithm, we must turn our attention to solutions that are AKKT-QEP. Among these, all that satisfy WCCP-QEP are KKT-QEP and therefore no additional solution would be discarded by an algorithm that generates AKKT-QEP sequences. In this way, there is no reason to worry that WCCP-QEP is not a constraint qualification. Thus, since WCCP-QEP is weaker than CCP-QEP, Corollary 4.2 is stronger than the analog one under CCP-QEP, which is indeed a constraint qualification.

5 Solution of EP-Subproblems

The previous convergence theory for the Augmented Lagrangian method works for general QEPs provided that we are able to find an approximate KKT-QEP point of the resulting QEP(u, ρ), that arises at each iteration. In this section, we consider

$$K(x) = \{y \in \mathbb{R}^n : \ g(x, y) \le 0, h(y) \le 0\},$$

because there is a larger amount of algorithms available for EPs than for QEPs to solve the subproblems in each iteration.

Some especial classes of QEPs where the EP-subproblems are not difficult to solve are studied below. We emphasize on the study of the Lagrangian

$$L(x,y) = f(x,y)+\frac{\rho}{2}\sum_{i=1}^{m}(\max\{0,g_i(x,y)+\frac{u_i}{\rho}\})^2-\frac{\rho}{2}\sum_{i=1}^{m}(\max\{0,g_i(x,x)+\frac{u_i}{\rho}\})^2, \quad (5.1)$$

to be monotone or pseudomonotone.

5.1 Monotone Bifunctions

In this subsection, we provide necessary and sufficient conditions for the Lagrangian L to be monotone, i.e., those conditions also ensure that our EP-subproblems admit a solution.

By definition, L is monotone on K_h if and only if $L(x,y) + L(y,x) \leq 0$ for all $x, y \in K_h$, or equivalently,

$$f(x,y) + \frac{\rho}{2}\left(\sum_{i=1}^{m}(\max\{0,g_i(x,y)+\frac{u_i}{\rho}\})^2 - (\max\{0,g_i(x,x)+\frac{u_i}{\rho}\})^2\right) + f(y,x)$$
$$+\frac{\rho}{2}\left(\sum_{i=1}^{m}(\max\{0,g_i(y,x)+\frac{u_i}{\rho}\})^2 - (\max\{0,g_i(y,y)+\frac{u_i}{\rho}\})^2\right) \leq 0, \;\forall\, x,y \in K_h.$$

Therefore, L is monotone if and only if

$$f(x,y)+f(y,x) \leq \frac{\rho}{2}\left(\alpha(x,x)+\alpha(y,y)-\alpha(x,y)-\alpha(y,x)\right), \;\forall\, x,y\in K_h, \quad (5.2)$$

where $\alpha(x,y) := \sum_{i=1}^{m}\max\{0,g_i(x,y)+\frac{u_i}{\rho}\}^2$.

As a consequence, we have the following general result.

Proposition 5.1. *The Lagrangian L is monotone if and only if equation* (5.2) *holds. In particular, if $g_i(x,x) = 0$ for all $x \in K_h$ and all $i = 1,\ldots,m$, and*

$$f(x,y)+f(y,x) \leq -\frac{\rho}{2}\left(\sum_{i=1}^{m}(\max\{0,g_i(x,y)+\frac{u_i}{\rho}\})^2+(\max\{0,g_i(y,x)+\frac{u_i}{\rho}\})^2\right)$$
$$+\rho\sum_{i=1}^{m}(\max\{0,\frac{u_i}{\rho}\})^2, \;\forall\, x,y\in K_h, \quad (5.3)$$

then L is monotone.

Since the sum of monotone bifunctions is also monotone, in the case when f is monotone, the monotonicity of L is ensured by the monotonicity of

$$\phi(x,y) := \frac{\rho}{2}\sum_{i=1}^{m}(\max\{0,g_i(x,y)+\frac{u_i}{\rho}\})^2-\frac{\rho}{2}\sum_{i=1}^{m}(\max\{0,g_i(x,x)+\frac{u_i}{\rho}\})^2. \quad (5.4)$$

Therefore, the following results follows easily:

Proposition 5.2. *The bifunction ϕ is monotone if and only if*

$$\alpha(x,y)+\alpha(y,x)\leq \alpha(x,x)+\alpha(y,y),\ \forall\ x,y\in K_h. \tag{5.5}$$

In particular, if $g_i(x,x)=0$ for all $x\in K_h$ and all $i=1,\ldots,m$, and

$$\alpha(x,y)+\alpha(y,x)\leq 2\sum_{i=1}^{m}(\max\{0,\frac{u_i}{\rho}\})^2,\ \forall\ x,y\in K_h, \tag{5.6}$$

then ϕ is monotone.

In order to provide more concrete sufficient conditions, we consider the following assumptions:

Assumption 5.1. *The functions f,g and h are such that:*

(a) *$f:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ is continuously differentiable on its second argument.*

(b) *$g:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^m$ is twice continuously differentiable on $\mathbb{R}^n\times\mathbb{R}^n$.*

(c) *$h:\mathbb{R}^n\to\mathbb{R}^l$ is a function for which K_h is a convex set.*

If h is $\mathbb{R}^l_+$-convex, i.e., each h_i is convex, then K_h is convex. The reverse statement does not hold, i.e., there are classes of vector-valued functions for which K_h is convex without the $\mathbb{R}^l_+$-convexity assumption on h, for instance, the class of $*$-quasiconvex functions (see [31, Definition 2.2]).

If ϕ is $\nabla_{xy}-$monotone, then ϕ is monotone (by [11, Theorem 3.1(f)]), i.e., a sufficient condition for ϕ to be monotone is that

$$\begin{aligned}\psi(y)=\nabla_x\phi(x,y)=\sum_{i=1}^{m}\max\{0,\rho g_i(x,y)+u_i\}\nabla_x g_i(x,y)&\\ -\sum_{i=1}^{m}\max\{0,\rho g_i(x,x)+u_i\}Jg_i(x,x)^{\top},&\end{aligned} \tag{5.7}$$

be monotone on K_h, i.e., $\langle\psi(x)-\psi(y),x-y\rangle\geq 0$ for all $x,y\in K_h$.

Clearly, $\psi:\mathbb{R}^n\to\mathbb{R}^n$ is locally Lipschitz without being continuously differentiable. By [30, Proposition 2.3(a)], we known that ψ is monotone on an open set D if and only if all generalized Jacobians (in the sense of Clarke [17]) from $\partial\psi(y)$ are positive semidefinite for all $y\in D$.

We estimate the generalized Jacobian of ψ below.

Proposition 5.3. *Suppose that Assumption 5.1 holds. Then the generalized Jacobian of ψ at $y\in\mathbb{R}^n$ satisfies $\partial\psi(y)\subseteq M(y)$ with*

$$M(y)=\left\{\sum_{i=1}^{m}\max\{0,\rho g_i(x,y)+u_i\}\nabla^2_{yx}g_i(x,y)+\rho\sum_{i=1}^{m}s_i\nabla_x g_i(x,y)\left[\nabla_y g_i(x,y)\right]^{\top}\right\},$$

where $s_i=1$ if $\rho g_i(x,y)+u_i>0$, $s_i=0$ if $\rho g_i(x,y)+u_i<0$, and $s_i\in[0,1]$ if $\rho g_i(x,y)+u_i=0$.

Proof. Note that ψ is nonsmooth only on its max-terms which are compositions of a smooth and a convex function, i.e., a regular mapping in the sense of Clarke [17]. $\square$

If the elements of $M(y)$ are positive semidefinite, then ψ is monotone, i.e., the monotonicity of L holds whenever f is monotone, a usual assumption for EPs. Hence, there are a large number of methods for solving the resulting EP-subproblem.

A sufficient condition for the elements from $M(y)$ to be positive semidefinite is given below.

Proposition 5.4. *Suppose that Assumption 5.1 holds. If the matrices*

$$\nabla^2_{xy} g_i(x,y),\ \forall\ i:\ u_i + \rho g_i(x,y) > 0,$$

$$\nabla_x g_i(x,y) \left[\nabla_y g_i(x,y)\right]^\top,\ \forall\ i:\ u_i + \rho g_i(x,y) \geq 0,$$

are positive semidefinite, then all elements in $M(y)$ are positive semidefinite.

Proof. By the representation of $M(y)$ and our assumptions, it follows that each element of $M(y)$ is a nonnegative sum of positive semidefinite matrices, i.e., $M(y)$ is positive semidefinite. $\square$

5.1.1 Example 1: The moving set case

An interesting special case of problem (QEP) is the moving set case [10, 22, 32, 34]. This is the case when $K(x) = c(x) + Q$ for some vector-valued function $c : \mathbb{R}^n \to \mathbb{R}^n$ and a closed and convex set Q from $\mathbb{R}^n$. Usually, Q is given by

$$Q := \{x \in \mathbb{R}^n :\ q(x) \leq 0\}, \tag{5.8}$$

where $q : \mathbb{R}^n \to \mathbb{R}$ is a function such that Q is closed and convex. As we noted in the previous subsection, function q may not be convex.

If q is convex, then we have the following sufficient condition for ensuring monotonicity.

Proposition 5.5. *Assume that $c : \mathbb{R}^n \to \mathbb{R}^n$ is given by $c(x) = (c_1(x_1), \ldots, c_n(x_n))^\top$ with $c_i'(x_i) < 0$ for all $i = 1, \ldots, m$. Then the elements of $M(y)$ are positive semidefinite.*

Proof. Since $g_i(x,y) = q_i(y - c(x))$, it follows that $\nabla_x g_i(x,y) = -Jc(x)^\top \nabla q_i(y - c(x))$ and $\nabla_y g_i(x,y) = \nabla q_i(y - c(x))$. Thus

$$\nabla_x g_i(x,y)(\nabla_y g_i(x,y))^\top = -Jc(x)^\top \nabla q_i(y - c(x)) \nabla q_i(y - c(x))^\top. \tag{5.9}$$

By assumption, $S = -Jc(x)$ is a positive definite diagonal matrix, so $S = DD$ with D a positive definite diagonal matrix. Then

$$\begin{aligned} v^\top \nabla_x g_i(x,y)(\nabla_y g_i(x,y))^\top v &= v^\top DD \nabla q_i(y - c(x)) \nabla q_i(y - c(x))^\top v \\ &= v^\top DD \nabla q_i(y - c(x)) \nabla q_i(y - c(x))^\top D^{-1} D v \\ &= w^\top D \nabla q_i(y - c(x)) \nabla q_i(y - c(x))^\top D^{-1} w, \end{aligned} \tag{5.10}$$

where $w = Dv$ for all $v \in \mathbb{R}^n$. Since $\nabla q_i(y-c(x))\nabla q_i(y-c(x))^\top$ is positive semidefinite, we have that the similar matrix $D\nabla q_i(y-c(x))\nabla q_i(y-c(x))^\top D^{-1}$ is also positive semidefinite.

On the other hand, a direct computation shows that

$$\nabla^2_{xy} g_i(x,y) = -Jc(x)\nabla^2 q_i(y-c(x)).$$

Since q_i is convex, its Hessian is symmetric and positive semidefinite. Hence, similarly to the previous computation, $\nabla^2_{xy} g_i(x,y)$ is positive semidefinite and the result follows from Proposition 5.4. $\square$

5.1.2 Example 2: The Binary Constraints case

We consider problem (QEP) with g given as in equation (3.9), i.e.,

$$K(x) = \left\{y \in \mathbb{R}^n : \; g_i(x_{j(i)}, y_{j(i)}) \le 0, \; \forall\, i \in \{1,\dots,m\}\right\}.$$

A sufficient condition for ensuring the monotonicity of the corresponding subproblems is given below.

Proposition 5.6. *Assume that for all $i = 1,\dots,m$*

$$\nabla_{x_{j(i)}} g_i(x_{j(i)}, y_{j(i)})\nabla_{y_{j(i)}} g_i(x_{j(i)}, y_{j(i)}) \ge 0, \;\; \text{and} \; \nabla^2_{y_{j(i)}x_{j(i)}} g_i(x_{j(i)}, y_{j(i)}) \ge 0. \quad (5.11)$$

Then all elements in $M(y)$ are positive semidefinite.

Proof. Clearly, $\nabla_y g_i(x,y) = \left(0,\dots,0,\nabla_{y_{j(i)}} g_i(x_{j(i)}, y_{j(i)}),0,\dots,0\right)^\top$, i.e., only position $j(i)$ could be nonzero. Then the Jacobian $\nabla^2_{xy} g_i(x,y)$ is $\nabla^2_{y_{j(i)}x_{j(i)}} g_i(x_{j(i)}, y_{j(i)})$ at the diagonal position $(j(i), j(i))$, and zero elsewhere. Therefore, it is positive semidefinite by assumption (5.11).

On the other hand, $\nabla_x g_i(x,y) = (0,\dots,0,\nabla_{x_{j(i)}} g_i(x_{j(i)}, y_{j(i)}),0,\dots,0)$. So, only position $j(i)$ could be nonzero. Then, the matrix $\nabla_x g_i(x,y)(\nabla_y g_i(x,y))^\top$ is diagonal with value $\nabla_{x_{j(i)}} g_i(x_{j(i)}, y_{j(i)})\nabla_{y_{j(i)}} g_i(x_{j(i)}, y_{j(i)})$ at position $(j(i), j(i))$, and zero elsewhere.

Therefore, the result follows from assumptions (5.11) and Proposition 5.4. $\square$

A special case of constraints with variable right-hand side which are also binary constraints is defined below

$$K(x) = \{y \in \mathbb{R} : \; g_i(x,y) = c_i(x_{j(i)}) + d_i(y_{j(i)}) \le 0, \; \forall\, i \in \{1,\dots,m\}\}, \quad (5.12)$$

where c_i, d_i are twice continuously differentiable functions for each i. Here

$$\nabla_{x_{j(i)}} g_i(x,y) = c_i'(x_{j(i)}), \; \nabla_{y_{j(i)}} g_i(x,y) = d_i'(y_{j(i)}), \; \text{and} \; \nabla^2_{y_{j(i)}x_{j(i)}} g_i(x,y) = 0.$$

The following result follows easily from the previous proposition.

Corollary 5.1. *If $c_i'(x_{j(i)})d_i'(y_{j(i)}) \ge 0$ for all $i = 1,\dots,m$, then all elements in $M(y)$ are positive semidefinite.*

Another example is the class of problems with box constraints (with variable right-hand side). Recall that

$$K(x) = \{y \in \mathbb{R}^n : y_i - \alpha_i x_i - \gamma_i \leq 0, \ \forall \ i \in \{1, \dots, n\}\}. \tag{5.13}$$

Clearly, if $\alpha_i \leq 0$ for all i, then all elements in $M(y)$ are positive semidefinite.

5.2 Pseudomonotone Equilibrium Problems

In this subsection, we provide an example of an interesting and usual variational inequality problem for which its associated Augmented Lagrangian is not a monotone bifunction, but for which there exists a positive answer for finding the solution of the related EP-subproblem.

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix and $a \in \mathbb{R}^n$. We consider the following variational inequality problem:

$$\text{find } \overline{x} \in K_h : \quad \langle A\overline{x} + a, y - \overline{x}\rangle \geq 0, \ \forall \ y \in K_h, \tag{5.14}$$

where $h := (h_1, h_2, \dots, h_m)$ is such that K_h is a convex set and for each $i = 1, 2, \dots, m$, the function $h_i : \mathbb{R}^n \to \mathbb{R}$ is continuous.

Given $u \in \mathbb{R}^n$ and $\rho > 0$. The Augmented Lagrangian associated to (5.14) is given by:

$$L^A_{u,\rho}(x, y) = \langle Ax + a, y - x\rangle + \frac{\rho}{2}\sum_{i=1}^{m}\left((\max\{0, h_i(y) + \frac{u_i}{\rho}\})^2 - (\max\{0, h_i(x) + \frac{u_i}{\rho}\})^2\right).$$

Note that solving Step 2 of Algorithm 1 is equivalent to finding a solution of the *mixed variational inequality problem* on $\mathbb{R}^n$:

$$\text{find } \overline{x} \in \mathbb{R}^n : \quad \langle A\overline{x} + a, y - \overline{x}\rangle + g(y) - g(\overline{x}) \geq 0, \ \forall \ y \in \mathbb{R}^n, \tag{5.15}$$

where $g(\cdot) := \frac{\rho}{2}\sum_{i=1}^{m}(\max\{0, h_i(\cdot) + \frac{u_i}{\rho}\})^2$. This class of problems is of special interest due to its applications in economics, mechanics and electronics (see [24, 35]).

Set $f^g_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$f^g_A(x, y) := \langle Ax + a, y - x\rangle + g(y) - g(x). \tag{5.16}$$

Clearly, $\overline{x} \in \mathbb{R}^n$ is a solution of problem (5.15) if and only if it is a solution of the following equilibrium problem:

$$\text{find } \overline{x} \in \mathbb{R}^n : \quad f^g_A(\overline{x}, y) \geq 0, \ \forall \ y \in \mathbb{R}^n. \tag{5.17}$$

Note that $f^g_A(x, x) = 0$ for all $x \in \mathbb{R}^n$, and that if g is continuous, then $f^g_A(x, \cdot)$ and $f^g_A(\cdot, y)$ are continuous for all $x, y \in \mathbb{R}^n$. If A is positive semidefinite, then f^g_A may be not monotone. However, we have a positive answer with pseudomonotonicity as the following proposition shows (see [28]). We emphasize that there is no convexity assumption on g.

Proposition 5.7. *If A is positive semidefinite, then f^g_A is pseudomonotone.*

Proof. Set $x, y \in \mathbb{R}^n$ such that $f_A^g(x,y) \geq 0$, that is,

$$\begin{aligned}
&\langle Ax + a, y - x\rangle + g(y) - g(x) \geq 0 \Longleftrightarrow \langle Ax + a, x - y\rangle - g(y) + g(x) \leq 0 \\
&\Longleftrightarrow \langle Ax + a, x - y\rangle + (\langle Ay, x - y\rangle - \langle Ay, x - y\rangle) - g(y) + g(x) \leq 0 \\
&\Longleftrightarrow \langle Ay + a, x - y\rangle - g(y) + g(x) + \langle A(x - y), x - y\rangle \leq 0.
\end{aligned}$$

Since A is positive semidefinite, $\langle A(x - y), x - y\rangle \geq 0$ for all $x, y \in \mathbb{R}^n$. Thus

$$\begin{aligned}
f_A^g(y,x) &= \langle Ay + a, x - y\rangle - g(y) + g(x) \\
&\leq \langle Ay + a, x - y\rangle - g(y) + g(x) + \langle A(x - y), x - y\rangle \leq 0.
\end{aligned}$$

Therefore, f_A^g is pseudomonotone. □

Existence results for pseudomonotone equilibrium problems for classes of nonconvex functions may be found in [18, 27]. In particular, existence results for problem (5.15) may be found in [24, 28]. An algorithm for a class of pseudomonotone equilibrium problems may be found in [7].

6 Conclusion

In this paper we described an Augmented Lagrangian method for Quasi-Equilibrium Problems, where we proved that it tends to find feasible limit points in the sense that an Approximate-KKT-QEP point is found for an auxiliary feasibility QEP. When a limit point is feasible, an Approximate-KKT-QEP point for the original problem is found. We also discuss in some details the notion of an approximate stationary point in the context of QEPs, where we showed that, differently from the case of nonlinear programming, the KKT-QEP residual can not be made arbitrarily small near any solution of a general QEP. Nonetheless, we were able to prove that feasible limit points of the sequence generated by the Augmented Lagrangian method are true KKT-QEP points under a new weak condition that we call Weak Cone Continuity Property (WCCP), which, surprisingly, is not even a constraint qualification.

The difficulties underlying the possibility of dealing with non-convex problems is somewhat subsumed in the assumption that the Augmented Lagrangian subproblems can be solved, at least approximately. Hence, we also provided a detailed discussion on several classes of problems where these subproblems can be properly solved, in the sense that the they yield monotone or pseudomonotone equilibrium problems.

References

[1] R. Andreani, G. Haeser, and J. M. Martínez. On sequential optimality conditions for smooth constrained optimization. *Optimization*, 60(5):627–641, 2011.

[2] R. Andreani, G. Haeser, M.L. Schuverdt, and P.J.S. Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Math. Programm.*, 135:255–273, 2012.

[3] R. Andreani, G. Haeser, M.L. Schuverdt, and P.J.S. Silva. Two new weak constraint qualifications and applications. *SIAM J. on Optim.*, 22:1109–1135, 2012.

[4] R. Andreani, J.M. Martínez, and M.L. Schuverdt. On the relation between constant positive linear dependence condition and quasinormality constraint qualification. *J. Optim. Theory Appl.*, 125(2):473–483, 2005.

[5] R. Andreani, J.M. Martínez, A. Ramos, and P.J.S. Silva. Strict constraint qualifications and sequential optimality conditions for constrained optimization. *Math. Oper. Res.*, 43(3):693-717, 2018.

[6] R. Andreani, J.M. Martínez, A. Ramos, and P.J.S. Silva. A cone-continuity constraint qualification and algorithmic consequences. *SIAM J. on Optim.*, 26(1):96–110, 2016.

[7] P. K. Anh, T. N. Hai. Novel self-adaptive algorithm for non-Lipschitz equilibrium problems with applications *J. of Global Optim.*, DOI: 10.1007/s10898-018-0722-2, 2018.

[8] D. Aussel, J. Cotrina, and A. Iusem, An existence result for quasi-equilibrium problems, *J. of Convex Anal.* 24:55–66, 2017.

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

[10] G. Bigi, M. Passacantando, Gap functions for quasi-equilibria. *J. Glob. Optim.* 66(4): 791–810, 2016.

[11] G. Bigi and M. Passacantando, Twelve monotonicity conditions arising from algorithms for equilibrium problems. Optim. Meth. and Soft., 30(2), 323–337, 2015.

[12] E. G. Birgin and J. M. Martínez. *Practical Augmented Lagrangian Methods for Constrained Optimization.* SIAM Publications, 2014.

[13] E. Blum, W. Oettli. From optimization and variational inequalities to equilibrium problems. *Math. Stud*, 63:1–23, 1993.

[14] L. F. Bueno, G. Haeser, and F. N. Rojas. *Optimality Conditions and Constraint Qualifications for Generalized Nash Equilibrium Problems and their Practical Implications.* SIAM J. Optim. **29**(1), 31–54 (2019)

[15] M. Castellani and M. Giuli, An existence result for quasiequilibrium problems in separable Banach spaces. *J. Math. Anal. Appl.* 425: 85–95, 2015.

[16] D. Chan and J.-S. Pang, The generalized quasi-variational inequality problem. *Math. Oper. Res.* 7: 211–222, 1982.

[17] F.H. Clarke, *Optimization and nonsmooth analysis.* SIAM Publications, 1990.

[18] J. Cotrina and Y. Garcia, Equilibrium problems: existence results and applications, *Set-Valued and Var. Anal.*, **26**, 159–177, (2018).

[19] P. Cubiotti, Existence of solutions for lower semicontinuous quasi-equilibrium problems. *Comput. Math.Appl.* 30: 11–22, 1995.

[20] A. Dreves, C. Kanzow, and O. Stein. Nonsmooth optimization reformulations of player convex generalized nash equilibrium problems. *J. Global Optim*, 53(4):587–614, 2012.

[21] F. Facchinei and C. Kanzow. Generalized nash equilibrium problems. *Ann. Oper. Res.*, 175:177–211, 2010.

[22] F. Facchinei, C. Kanzow, and S. Sagratella, Solving quasi-variational inequalities via their KKT conditions. *Math. Programm.*, 50(1):369–412, 2014.

[23] A. Fischer, M. Herrich, and K. Schönefeld. Generalized nash equilibrium problems - recent advances and challenges. *Pesq. Oper.*, 34:521–558, 2014.

[24] D. Goeleven. "Complementarity and Variational Inequalities in Electronics". Academic Press, London, (2017).

[25] N. Hadjisavvas, S. Komlosi and S. Schaible. Handbook of Generalized Convexity and Generalized Monotonicity. Springer, Boston (2005)

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

[27] A. Iusem and F. Lara, Optimality conditions for vector equilibrium problems with applications. *J. Optim. Theory Appl.*, DOI: 10.1007/s10957-018-1321-6, (2018).

[28] A. Iusem and F. Lara, Existence results for noncoercive mixed variational inequalities, *Submitted.*

[29] A. Iusem and M. Nasri, Augmented Lagrangian methods for variational inequality problems, *RAIRO, Oper. Res.*, **44**, 5–24 (2010).

[30] H.Y. Jiang and L. Qi, Local uniqueness and convergence of iterative methods for nonsmooth variational inequalities. *J. Math. Anal. Appl.*, 196: 314–331, 1995.

[31] V. Jeyakumar, W. Oettli, M. Natividad, A solvability theorem for a class of quasiconvex mappings with applications to optimization. *J. Math. Anal. Appl.*, 179, 537–546 (1993).

[32] C. Kanzow. On the multiplier-penalty-approach for quasi-variational inequalities. *Math. Program.*, 160 (1): 353–377, (2016).

[33] C. Kanzow and D. Steck. Augmented lagrangian methods for the solution of generalized nash equilibrium problems. *SIAM J. on Optim.*, 26(4): 2034–2058, (2016).

[34] C. Kanzow and D. Steck. Augmented Lagrangian and exact penalty methods for quasi-variational inequalities. *Comput. Optim. Appl.*, 69(3):801–824, (2018).

[35] D. Kinderlehrer and G. Stampacchia. "An Introduction to Variational Inequalities and Their Applications". Academic Press, New York (1980).

[36] A. S. Kravchuk, P. J. Neittaanmäki, Variational and Quasi-Variational Inequalities in Mechanics, **147**, Solid Mechanics and Its Applications. Springer, Dordrecht (2007).

[37] J.-S. Pang and M. Fukushima. Quasi-variational inequalities, generalized nash equilibria, and multi-leader-follower games. *Comput. Manag. Sci.*, 2(1):21–56, 2005.

[38] L. Qi and Z. Wei. On the constant positive linear dependence conditions and its application to SQP methods. *SIAM J. on Optim.*, 10:963–981, 2000.

[39] R. T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Math. Oper. Res.*, **1**, 97–116 (1976).

[40] J.J. Strodiot, T.T.V. Nguyen, and V.H. Nguyen, A new class of hybrid extragradient algorithms for solving quasi-equilibrium problems. *J. Glob. Optim.* 56(2): 373–397, 2013.

[41] N.T.T. Van, J.J. Strodiot, V.H. Nguyen, and P.T. Vuong. An extragradient-type method for solving nonmonotone quasi-equilibrium problems. *Optimization.*, 67(5):651–664, 2018.