Optimality Conditions and Constraint Qualifications for Generalized Nash Equilibrium Problems and their Practical Implications

Luís Felipe Bueno* Gabriel Haeser† Frank Navarro Rojas‡

December 21, 2017

Abstract

Generalized Nash Equilibrium Problems (GNEPs) are a generalization of the classic Nash Equilibrium Problems (NEPs), where each player's strategy set depends on the choices of the other players. In this work we study constraint qualifications and optimality conditions tailored for GNEPs and we discuss their relations and implications for global convergence of algorithms. Surprisingly, differently from the case of nonlinear programming, we show that, in general, the KKT residual can not be made arbitrarily small near a solution of a GNEP. We then discuss some important practical consequences of this fact. We also prove that this phenomenon is not present in an important class of GNEPs, including NEPs. Finally, under a weak constraint qualification introduced, we prove global convergence to a KKT point of an Augmented Lagrangian algorithm for GNEPs and under the quasinormality constraint qualification for GNEPs, we prove boundedness of the dual sequence.

Key words: Generalized Nash Equilibrium Problems, Optimality conditions, Approximate-KKT conditions, Constraint qualifications, Augmented Lagrangian methods.

1 Introduction

In this paper we consider the Generalized Nash Equilibrium Problem (GNEP) where, given N players, each player $v = 1, \dots, N$ aims at minimizing

$$P^v(x^{-v}): \quad \min_{x^v} f^v(x^v, x^{-v}) \quad s.t. \quad g^v(x^v, x^{-v}) \le 0, \tag{1.1}$$

*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

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

by controlling his/her own variables $x^v \in \mathbb{R}^{n_v}$, given the choices of the remaining players denoted by x^{-v}. Formally, we define $n := n_1 + \cdots + n_N$, the total number of variables, and $m := m_1 + \cdots + m_N$, the total number of constraints. Here, $f^v : \mathbb{R}^n \to \mathbb{R}$ denotes the objective function of player v and $g^v : \mathbb{R}^n \to \mathbb{R}^{m_v}$ defines its constraints. As usual in the context of GNEPs, for every player $v = 1, \ldots, N$, we often write $(x^v, x^{-v}) \in \mathbb{R}^n$, instead of $x \in \mathbb{R}^n$, where the vector x^{-v} is defined by $x^{-v} := (x^u)_{u=1, u \neq v}^N$.

The GNEP is called player convex if all functions f^v and g^v are continuous and also convex as a function of x^v. The GNEP is called jointly convex if it is player convex and $g^1 = \cdots = g^N = g$ is convex as a function of the entire vector x. In the case that g^v depends only on x^v, the GNEP is reduced to the standard Nash Equilibrium Problem (NEP).

We will say that a point x is feasible for the GNEP if $g^v(x) \leq 0$ for each player $v = 1, \ldots, N$. A feasible point $\bar{x} = (\bar{x}^1, \ldots, \bar{x}^N)$ is a generalized Nash equilibrium, or simply a solution of the GNEP, if for each player $v = 1, \ldots, N$ it holds that

$$f^v(\bar{x}) \leq f^v(x^v, \bar{x}^{-v}) \quad \forall x^v : g^v(x^v, \bar{x}^{-v}) \leq 0.$$

The concept of solution of a GNEP means that no player v can improve his objective function by unilaterally changing his strategy $\bar{x}^v$.

For simplicity of notation, we assume throughout the paper that the feasible sets of the players are defined by inequalities only. We note that all results obtained can be easily extended when equality constraints are incorporated for each player. On the other hand, our approach does not require convexity of any kind, either of the objective functions or of the players' constraints. In contrast, most of the literature on GNEPs deals only with the player convex or jointly convex cases, see [16, 19, 20, 22, 36, 30] and references therein for more details.

For nonlinear programming problems, optimality conditions are conditions that are satisfied by all local solution of the problem. One of the main subjects in the theory of nonlinear optimization is the characterization of optimality. This is often done through conditions that use the derivatives of the objective function and of the constraints of the problem. Among such conditions, arguably the most important are the Karush-Kuhn-Tucker (KKT) conditions. The KKT conditions are extensively used in the development of algorithms for solving optimization problems and we say that a point that satisfies it is a stationary point. In order to ensure that the KKT conditions are necessary for optimality, a constraint qualification (CQ) is needed. A CQ is an assumption made about the functions that define the constraints of the problem that, when satisfied by a local minimizer, ensures that it is stationary. Among the most common CQs in the literature are the linear independence (LICQ, [13]), Mangasarian-Fromovitz (MFCQ, [29]), Quasinormality (QN, [25, 14]) and Constant Positive Linear Dependence (CPLD, [33, 10]). Other weak CQs associated with global convergence of first-order algorithms have been introduced in the recent years ([26, 31, 4, 5, 6, 7]). One of these CQs, of particular interest in this paper, is the Cone Continuity Property (CCP), defined in [6].

In this work we study optimality conditions and constraints qualification for GNEPs. Moreover, we are interested in optimality conditions and CQs that can be associated with the development of iterative algorithms for GNEPs. It turns out, however, that some results are different from those that are known for standard optimization problems. For example, the Quasinormality extension for GNEPs presented here is not weaker than the CPLD condition, as it is the case in optimization [10]. Even more surprisingly, we show that the Approximate-KKT (AKKT) condition [3] is not an optimality condition for a general GNEP.

Currently there is a considerable variety of methods for solving GNEPs. However, most of them are focused on the case of player or jointly convex GNEPs. We refer the interested reader to the survey papers [19, 22] and the references therein for a quite complete overview of the existing approaches.

In this work, we are interested in penalty-type methods for GNEPs, with proved global convergence. The first penalty method for GNEPs that we are aware of is due to Fukushima [23]. Other variants were studied in [20, 21], where a full penalty is considered in [20] and a partial penalty in [21].

Augmented Lagrangian methods for constrained optimization problems are known to be an alternative for penalty-type methods. Taking this into account, it is natural to apply an augmented Lagrangian-type approach in order to solve GNEPs. This idea was already studied by Pang and Fukushima [32] for quasi-variational inequalities (QVIs). An improved version of that method can be found in [27], also for QVIs. Under certain conditions, it is known that a GNEP can be reformulated as a QVI, so in both papers the authors briefly discuss the application of the method for this class of GNEPs.

Based on [27], Kanzow and Steck developed an augmented Lagrangian method specifically for GNEPs in [28]. Independently, we also proposed an augmented Lagrangian method for GNEPs in [35], which we briefly present here. The main differences of our approaches are that we focus on optimality conditions and CQs that are associated with the proposed method, in particular, our global convergence proof is based on the CCP constraint qualification, while the one in [28] is based on the (stronger) CPLD. We also present a convergence result based on the QN constraint qualification, which extends [2] from optimization to GNEPs, but proving in addition that the dual sequence is bounded. A main contribution of this paper is that AKKT is not an optimality condition for a general GNEP. Since the augmented Lagrangian method proposed in [28] (and also ours) generates AKKT sequences, it means that this kind of method excludes the possibility of finding some solutions of the GNEPs. Also, due to the general nature of the definition of AKKT, it is expected (as in the optimization case) that most algorithms would generate AKKT sequences, hence, the fact that AKKT is not an optimality condition for GNEPs is a fundamental aspect of the problem that should be taken into account when developing an algorithm. For a special class of GNEPs (including NEPs), we prove that this inherent difficulty is not present, as we prove that AKKT is indeed an optimality condition within this class.

This paper is organized as follows. In Section 2, we review some CQs and the AKKT concept for standard nonlinear programming. In Sections 3 and 4 we define and state a comprehensive study on several CQs for GNEPs. In Section 3 we discuss concepts where the analysis considers the constraints of the players somewhat independently. On the other hand, in Section 4 we explore concepts where the connection of the decisions of all the players in the constraints plays a fundamental role. The latter have a much greater impact on the analysis of numerical algorithms.

Section 5 is the most important of this paper. In this section we formally extend the concept of Approximate-KKT for GNEPs (AKKT-GNEP). We give an example where AKKT-GNEP does not hold even at a solution of a jointly convex GNEP. Inspired by this, we discuss some practical issues related to limit points of methods that generate AKKT-GNEP sequences. We prove that for important classes of GNEPs, which include the NEPs, the AKKT-GNEP is a necessary optimality condition. Moreover, we prove that the CCP-GNEP condition introduced in Section 4 is the weakest condition that ensures that AKKT-GNEP implies KKT for GNEPs.

In Section 6 we give a precise description of our Augmented Lagrangian method for GNEPs. Using the results of Sections 3, 4 and 5, we present a refined convergence analysis under CCP-GNEP, which is stronger than the corresponding result from [28] under CPLD-GNEP. An independent global convergence result under QN-GNEP constraint qualification is also presented, where we show, in addition, the boundedness of the dual variables, which is a new result even in the optimization case. Finally, conclusions and remarks are stated in the last section.

Our notation is standard in optimization and game theory. We denote by $\mathbb{R}^n$ the n-dimensional real Euclidean space, where x_i denotes the i-th component of $x \in \mathbb{R}^n$. We denote by $\mathbb{R}^n_+$ the set of vectors with non-negative components and by $a_+ := \max\{0, a\}$ the non-negative part of $a \in \mathbb{R}$. We use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product, and $\|\cdot\|$ the associated norm. For a real-valued function $f(x, y)$ involving variables $x \in \mathbb{R}^{n_1}$ and $y \in \mathbb{R}^{n_2}$, the partial derivatives with respect to x and y are denoted by $\nabla_x f(x, y) \in \mathbb{R}^{n_1}$ and $\nabla_y f(x, y) \in \mathbb{R}^{n_2}$, respectively.

We say that a finite set of vectors $a_i, i = 1, \ldots, m$, is positively linearly dependent if there are scalars $\lambda_i \geq 0, i = 1, \ldots, m$, not all zero, such that $\sum_{i=1}^m \lambda_i a_i = 0$; otherwise these vectors are called positively linearly independent.

2 Preliminaries

In this section we recall the definitions of several constraint qualifications for optimization problems ensuring that a local minimizer satisfies the Karush-Kuhn-Tucker (KKT) conditions. We also define the concept of an Approximate-KKT (AKKT) point for standard nonlinear programming, which is fundamental for the development of stopping criteria and for the global convergence analysis of many algorithms. The fact that the limit points of an algorithm satisfies the AKKT condition implies that they satisfy the KKT conditions under weak constraint qualifications. Let us begin by considering a nonlinear optimization problem defined by a continuously differentiable objective function $f : \mathbb{R}^n \to \mathbb{R}$ and, for simplicity of the notation, let us consider only inequality constraints, defined by continuously differentiable functions $g_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$:

$$\min_x f(x) \quad s.t. \quad g_i(x) \leq 0 \quad \forall i = 1, \ldots, m. \tag{2.1}$$

Let $X \subset \mathbb{R}^n$ denote the feasible set of problem (2.1). One of the most important conditions associated with the convergence analysis of many algorithms are the KKT conditions, which we state below:

Definition 2.1 (KKT). *We say that $\bar{x} \in X$ is a KKT (or stationary) point if there* *exists* $\lambda \in \mathbb{R}^m_+$ *such that*

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) = 0,$$

and

$$\min\{\lambda_i, -g_i(\bar{x})\} = 0$$

for all $i = 1, \dots, m$.

Unfortunately, it is not true that any local solution of problem (2.1) satisfies the KKT conditions. In order to ensure that a solution satisfies these conditions, we must assume a so-called constraint qualification (CQ). To present some of the most important CQs in the literature, we will need the following concepts.

The (Bouligand) tangent cone of X at a feasible point $\bar{x} \in X$ is defined by

$$T_X(\bar{x}) := \left\{ d \in \mathbb{R}^n : \exists \left\{x^k\right\} \subset X,\, \{t_k\} \downarrow 0, x^k \to \bar{x} \text{ and } \lim_{k\to\infty} \frac{x^k - \bar{x}}{t_k} = d \right\}.$$

It is well known that when $\bar{x} \in X$ is a local minimizer of (2.1), then the following geometric necessary optimality condition holds:

$$\langle \nabla f(\bar{x}), d \rangle \geq 0, \forall d \in T_X(\bar{x}).$$

Given $x \in X$, we define the set $A(x)$ of the indices of active constraints, that is, $A(x) := \{i \in \{1, \dots, m\} : g_i(x) = 0\}$. The corresponding linearized cone of X at $\bar{x} \in X$ is given by

$$L_X(\bar{x}) := \left\{ d \in \mathbb{R}^n : \nabla g_i(\bar{x})^T d \leq 0, i \in A(\bar{x}) \right\}.$$

Given $\bar{x} \in X$ and $x \in \mathbb{R}^n$, we define

$$K_X(x, \bar{x}) := \left\{ d \in \mathbb{R}^n : d = \sum_{i \in A(\bar{x})} \lambda_i \nabla g_i(x), \lambda_i \geq 0 \right\},$$

the perturbed cone generated by the gradients of the active constraints at $\bar{x}$. When $\bar{x}$ is fixed, and so there is no chance of misunderstanding, we will use $K_X(x)$ instead of $K_X(x, \bar{x})$, for simplicity of notation.

We recall that, given an arbitrary set $C \subset \mathbb{R}^n$, its polar cone is defined by

$$P(C) := \{v \in \mathbb{R}^n | \langle v, d \rangle \leq 0, \quad \forall d \in C\}.$$

Under the definition of the polar cone, the geometric optimality conditions at a local minimizer $\bar{x} \in X$ reads as $-\nabla f(\bar{x}) \in P(T_X(\bar{x}))$. In order to arrive at the KKT conditions, we note that Farkas' Lemma implies that the polar cone of $L_X(\bar{x})$ coincides with the cone generated by the (non-perturbed) gradients of active constraints at $\bar{x}$, namely, $P(L_X(\bar{x})) = K_X(\bar{x}, \bar{x})$, and it is easy to see that the KKT conditions hold at $\bar{x} \in X$ if, and only if, $-\nabla f(\bar{x}) \in P(L_X(\bar{x}))$. Hence, it is easy to see that any condition that implies the equality of the polars of the linearized and the tangent cones are constraint qualifications. We define below a few of them:

Definition 2.2. *Let $\bar{x}$ be a feasible point of the nonlinear problem* (2.1)*. Then we say that $\bar{x}$ satisfies the*

(a) Linear Independence Constraint Qualification (LICQ) if the gradients of the active constraints $\nabla g_i(\bar{x})$, $i \in A(\bar{x})$, are linearly independent;

(b) Mangasarian-Fromovitz Constraint Qualification (MFCQ) if the gradients of the active constraints $\nabla g_i(\bar{x})$, $i \in A(\bar{x})$, are positively linearly independent;

(c) Constant Positive Linear Dependence condition (CPLD) if for any subset $I \subset A(\bar{x})$ such that the gradient vectors $\nabla g_i(x)$, $i \in I$, are positively lineararly dependent at $x := \bar{x}$, they remain (positively) linearly dependent for all x in a neighborhood of $\bar{x}$;

(d) Quasinormality condition (QN) if for any $\lambda \in \mathbb{R}^m_+$ such that

$$\sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) = 0, \quad \lambda_i g_i(\bar{x}) = 0,$$

there is no sequence $x^k \to \bar{x}$ such that $g_i(x^k) > 0$ for all k whenever $\lambda_i > 0, i = 1, \ldots, m$;

(e) Cone Continuity Property (CCP) if the set-valued mapping (multifunction) $K_X : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, given by $x \mapsto K_X(x) := K_X(x, \bar{x})$, is outer semicontinuous at $\bar{x}$, that is, $\limsup_{x\to\bar{x}} K_X(x) \subset K_X(\bar{x})$, where

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

(f) Abadie constraint qualification (ACQ) if $T_X(\bar{x}) = L_X(\bar{x})$ holds;

(g) Guignard constraint qualification (GCQ) if $P(T_X(\bar{x})) = P(L_X(\bar{x}))$ holds.

The following implications hold between these conditions [6]:

$$LICQ \Rightarrow MFCQ \Rightarrow CPLD \Rightarrow CCP \Rightarrow ACQ \Rightarrow GCQ.$$

Moreover, CPLD is srtictly stronger than QN and QN is independent of CCP.

A useful tool associated with results relying on CPLD, which we also use in our analysis, is the following Carathéodory-type result, whose proof is a simple adaptation of [15, Lemma 3.1]:

Lemma 2.1. *Assume that a given vector $w \in \mathbb{R}^n$ has a representation of the form*

$$w = \sum_{i=1}^{m} \lambda_i a_i$$

with $a_i \in \mathbb{R}^n$ and $\lambda_i \in \mathbb{R}$ for all $i = 1, \ldots, m$. Then, there exist an index set $I \subset \{1, \ldots, m\}$ as well as scalars $\tilde{\lambda}_i \in \mathbb{R}$ $(i \in I)$ with $\lambda_i \tilde{\lambda}_i > 0, \forall i \in I$, such that

$$w = \sum_{i \in I} \tilde{\lambda}_i a_i$$

and such that the vectors a_i $(i \in I)$ are linearly independent.

It is possible to show that GCQ is the weakest possible constraint qualification, in the sense to guarantees that a local minimum of problem (2.1) is also a stationary point, independently of the objective function. This was originally proved in [24], as a consequence of the fact that the polar cone of $T_X(\bar{x})$ coincides with the set of vectors of the form $-\nabla \tilde{f}(\bar{x})$, with $\tilde{f}$ ranging over all continuously differentiable functions that assume a local minimum constrained to X at $\bar{x}$. See also [34].

Without assuming a constraint qualification, one can still rely on the KKT conditions to formulate a necessary optimality condition, but one must consider the validity of the condition not at the point $\bar{x}$ itself, but at arbitrarilly small perturbations of $\bar{x}$. This gives rise to so-called sequential optimality conditions for constrained optimization problems (2.1), which are necessarily satisfied by local minimizers, independently of the fulfillment of constraint qualifications. These conditions are used in practice to justify the stopping criteria for several important methods such as augmented Lagrangian methods. On the other hand, the separate analysis of the sequential optimality condition generated by the algorithm, together with the analysis of the minimal constraint qualification needed for a point satisfying the condition to be a KKT point, is a powerful tool for proving global convergence of algorithms to a KKT point under a weak constraint qualification. A major role in our analysis will be played by the most popular of these conditions, called Approximate-KKT (AKKT) [33, 3, 14]:

Definition 2.3 (AKKT)**.** *We say that $\bar{x} \in X$ satisfies the AKKT condition (or is an AKKT point) if there exist sequences $\left\{x^k\right\} \subset \mathbb{R}^n$, $\left\{\lambda^k\right\} \subset \mathbb{R}^m_+$, such that* $\lim_{k\to\infty} x^k = \bar{x}$,

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

and

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

for all $i = 1, \ldots, m$. In this case, $\left\{x^k\right\}$ is called an AKKT sequence.

The following theorem proves that the AKKT condition is a true optimality condition independently of the validity of any constraint qualification.

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

Proof. See [3, 15]. □

Theorem 2.1 says that regardless of the fixed tolerance $\varepsilon > 0$, it is always possible to find a point sufficiently close to a local solution of problem (2.1) that satisfies approximately the KKT conditions with tolerance ε. This fact is the main reason why constraint qualifications play little role in the development of practical algorithms, however, this result justifies the stopping criterion of most algorithms based on approximately satisfying the KKT conditions. Hence, it is reasonable that algorithms will aim at finding AKKT points, regardless of the validity of the KKT conditions or not at a solution. If one aims at finding KKT points, local minimizers that do not satisfy a constraint qualifications are ruled out completely from the analysis. Under this point of view, a constraint qualification is needed only in order to compare a global convergence result to an AKKT point, with a global convergence result to a KKT point under a constraint qualification. It turns out that the connection is made with the Cone Continuity Property (CCP) [6], the minimal constraint qualitication that ensures that an AKKT point is a KKT point, independently of the objective function. See the discussion in [7, 15].

Of course, under this generality, one must bear in mind that the sequence of approximate Lagrange multipliers, $\{\lambda^k\}$, can be unbounded, which introduces numerical difficulties. However, our analysis of an augmented Lagrangian for GNEPs will show that under an extension of QN to GNEPs, the sequence of Lagrange multipliers generated by the algorithm is necessarily bounded. A new result also in the optimization framework, which in particular implies boundedness of the dual sequence under CPLD. As we discuss later, the boundedness result under CPLD does not extend in general for GNEPs.

In fact, we will show that the situation is dramatically different for GNEPs, in the sense that there are problems where most algorithms are bound to disregard its solutions, even in the jointly convex case. Namely, we will extend the concept of AKKT for GNEPs and we will show that, unlike in nonlinear programming, we do not have that AKKT is an optimality condition in general. We will also discuss the role that AKKT plays in the study of global convergence of an augmented Lagrangian-type algorithm for GNEPs.

3 Partial Constraint Qualifications for GNEPs

Throughout this section we extend the definitions of several CQs from optimization to GNEPs, considering separately the feasible set of each player. We start by defining a Guignard-type CQ and proving that it is the minimal CQ to ensure a KKT-type condition at a solution of a GNEP. From now on, we consider the GNEP defined by (1.1) and we assume that the objective functions $f^v : \mathbb{R}^n \to \mathbb{R}$, as well as the constraint functions $g^v : \mathbb{R}^n \to \mathbb{R}^{m_v}$, are continuously differentiable for all $v = 1, \ldots, N$.

Given $x^{-v} \in \mathbb{R}^{n-n_v}$, the feasible set for player v will be denoted by

$$X^v(x^{-v}) := \left\{ x^v \in \mathbb{R}^{n_v} : g^v(x^v, x^{-v}) \leq 0 \right\}.$$

The feasible set for the GNEP will be given by

$$X := \{ x \in \mathbb{R}^n : g^v(x) \leq 0, \quad v = 1, \ldots, N \}.$$

Definition 3.1 (KKT-GNEP). *We say that $x \in X$ is a KKT-GNEP point for* (1.1) *when x^v is a KKT point for problem $P^v(x^{-v})$ for each $v = 1, \ldots, N$, i.e, for each $v = 1, \ldots, N$, there are vectors $\lambda^v \in \mathbb{R}^{m_v}_+$ such that*

$$\nabla_{x^v} f^v(x) + \sum_{i=1}^{m_v} \lambda_i^v \nabla_{x^v} g_i^v(x) = 0, \tag{3.1}$$

$$\min\{\lambda_i^v, -g_i^v(x)\} = 0 \quad \text{for each} \quad i = 1, \ldots, m_v. \tag{3.2}$$

A useful definition is the set $A^v(x)$ of indices of active constraints for player $v = 1, \ldots, N$ at $x \in X$, that is, $A^v(x) := \{i \in \{1, \ldots, m_v\} : g_i^v(x) = 0\}$. With this definition one may rewrite the KKT-GNEP condition simply as (3.1) with the sum ranging over $A^v(x)$ for all $v = 1, \dots, N$.

Since GNEPs are generalizations of optimization problems, it is straightforward that the KKT-GNEP condition is not a first-order necessary optimality condition without some constraint qualification. Formally, a constraint qualification CQ for a GNEP is a property that when satisfied at a solution $\bar{x}$, ensures that the KKT-GNEP conditions hold at $\bar{x}$ for appropriate Lagrange multipliers $\lambda^v \in \mathbb{R}^{m_v}_+$ for each $v = 1, \dots, N$. In other words, the property

$$\text{KKT-GNEP or Not-CQ}$$

is fulfilled at every solution of the GNEP (1.1).

Now we state the minimality of a Guignard-type condition with respect to guaranteeing the KKT-GNEP conditions:

Theorem 3.1 (Minimality of a Guignard-type CQ)**.** *Let $\bar{x} \in X$. The following conditions are equivalent:*

i) The constraints $X(\bar{x}^{-v})$ of each player $v = 1, \dots, N$ conforms to GCQ at $\bar{x}^v$;

ii) For any objective functions $f^v, v = 1, \dots, N$, in the definition of GNEP (1.1) *such that $\bar{x}$ is a solution, the KKT-GNEP conditions hold.*

Proof. Assume *i)* and let $f^v, v = 1, \dots, N$, be objective functions in the definition of GNEP (1.1) such that $\bar{x}$ is a solution. For each player v, $\bar{x}^v$ is a minimizer of $P^v(\bar{x}^{-v})$. Since GCQ holds at $\bar{x}^v$ for the constraints $X(\bar{x}^{-v})$, the KKT conditions hold for $P^v(\bar{x}^{-v})$. Combining the KKT conditions for each player v yields KKT-GNEP, hence *ii)*. Now suppose that *i)* does not hold, that is, for some player v, GCQ for $P(\bar{x}^{-v})$ does not hold. From the minimality of GCQ in optimization, there is an objective function f^v assuming a minimizer at $\bar{x}^v$ constrained to $X(\bar{x}^{-v})$ that does not conform to the KKT conditions. This constradicts *ii)*. □

From now on, we call the minimal condition ensuring KKT-GNEP given by Theorem 3.1 by "Guignard-Partial". More generally, it is a natural way to obtain a constraint qualification for GNEPs by assuming that at the solution, the constraints of each player conform to a constraint qualification for optimization. If "CQ" is a constraint qualification for optimization, we say that "CQ-Partial" is the corresponding constraint qualification for GNEPs obtained in this fashion. That is, we have the following:

Definition 3.2 (CQ-Partial)**.** *Given a feasible point $\bar{x} \in X$ for the GNEP* (1.1) *and a constraint qualification "CQ" for optimization, we say that $\bar{x}$ satisfies CQ-Partial for the GNEP if for all $v = 1, \dots, N$, $\bar{x}^v$ satisfies "CQ" with respect to the constraints $x \in X^v(\bar{x}^{-v})$.*

Remark 3.1. *It is easy to see that CQ-Partial is a constraint qualification for GNEPs. Namely, if a solution $\bar{x}$ of the GNEP satisfies CQ-Partial, then it satisfies the KKT-GNEP conditions. We note that in order to guarantee that an equilibrium solution satisfies the KKT-GNEP conditions, it is not necessary for the constraint qualification for each player to be the same.*

Let us now exemplify the concept of CQ-Partial by considering the well-known MFCQ, CPLD and CCP conditions for optimization.

Definition 3.3 (MFCQ-Partial)**.** *We say that $\bar{x} \in X$ satisfies MFCQ-Partial with respect to the GNEP* (1.1) *if for each $v = 1, \ldots, N$,*

$$\{\nabla_{x^v} g_i^v(\bar{x})\}_{i \in A^v(\bar{x})}$$

is positively linearly independent in $\mathbb{R}^{n_v}$.

Definition 3.4 (CPLD-Partial)**.** *We say that $\bar{x} \in X$ satisfies CPLD-Partial with respect to the GNEP* (1.1) *if for each $v = 1, \ldots, N$, there is a neighborhood $U^v \subset \mathbb{R}^{n_v}$ of $\bar{x}^v$ such that if $I^v \subset A^v(\bar{x})$ is such that $\{\nabla_{x^v} g_i^v(\bar{x})\}_{i \in I^v}$ is positively linearly dependent, then $\{\nabla_{x^v} g_i^v(x^v, \bar{x}^{-v})\}_{i \in I^v}$ is positively linearly dependent for each $x^v \in U^v$.*

Definition 3.5 (CCP-Partial)**.** *We say that $\bar{x} \in X$ satisfies CCP-Partial with respect to the GNEP* (1.1) *if for each $v = 1, \ldots, N$, $\limsup_{x^v \to \bar{x}^v} K_{X^v(\bar{x}^{-v})}(x^v) \subset K_{X^v(\bar{x}^{-v})}(\bar{x}^v)$, where*

$$K_{X^v(\bar{x}^{-v})}(x^v) = \left\{ w^v \in \mathbb{R}^{n_v} : w^v = \sum_{i \in A^v(\bar{x})} \lambda_i^v \nabla_{x^v} g_i^v(x^v, \bar{x}^{-v}), \quad \lambda_i^v \geq 0 \right\}.$$

Remark 3.2. *It is clear from the definition of CQ-Partial that if CQ1 is weaker than CQ2 then CQ1-Partial is weaker than CQ2-Partial.*

In [32, 27] the EMFCQ-Partial, which is an extension of MFCQ-Partial to infeasible points, was used to prove the feasibility of limit points of a penalty-type method for QVIs. The same result has been obtained for GNEPs in [28]. In the feasible case, a weaker CQ can be employed, which we discuss in the next Section.

It turns out that when dealing with CQs weaker than MFCQ, the CQ-Partial concept is not the appropriate one to prove global convergence results of algorithms to stationary points. That is the reason we give them the "partial" adjective. In the next section we will show another way to exend an optimization CQ to the context of GNEPs, which will be adequate for proving global convergence results. The reason the CQ-Partial concept is too weak for this purpose is because a sequence $\{x^k\} \subset \mathbb{R}^n$ generated by an algorithm, when converging to some $\bar{x} \in \mathbb{R}^n$, does so in such a way that, typically, $x^{k,v} \neq \bar{x}^v$ for all $v = 1, \ldots, N$, hence, it is not reasonable to fix $\bar{x}^{-v}$ when defining the neighborhood used in Definitions 3.4 and 3.5 for player v. One should, instead, consider jointly the full neighborhood in $\mathbb{R}^n$, even when stating the conditions to be satisfied for each player.

4 Joint Constraint Qualifications for GNEPs

Recently, many new CQs for optimization have been defined in order to prove global convergence results to KKT points under weaker assumptions. See [4, 5, 6, 7]. Our goal is to provide a way to extend those CQs to GNEPs in such a way that global convergence of algorithms to a KKT-GNEP point can be proved. These definitions will highlight the joint structure of the GNEPs with respect to each player. Following the definition of

CPLD-GNEP from [28], which independently did such extension of CPLD, given some constraint qualification CQ for optimization, we introduce a constraint qualification for GNEP that we call CQ-GNEP.

Given $\bar{x} \in X$, the CQ-Partial is a straightforward extension of the optimization CQ, obtained by looking individually at each player problem with constraint $x \in X^v(\bar{x}^{-v})$. The concept is adequate for NEPs, but for GNEPs, we must take into account how the feasible set of each player is perturbed at points nearby $\bar{x}^{-v}$ in order to obtain a meaningful definition. While the relations among all CQ-Partial conditions are inherited directly from the corresponding relations among the optimization CQs, the relations among different CQ-GNEP conditions are not the same as the relations among the corresponding CQs for optimization. The main difference of the CQ-GNEPs presented here, with respect to their CQ-Partial counterparts, is that the requirements for points close to the analyzed point should be considered taking into account simultaneous perturbations in the variables of all players.

For the case of the LICQ, MFCQ (and EMFCQ) conditions, the definitions of CQ-Partial and CQ-GNEP coincide, since there is no explicit use of neighborhoods in the definition of these CQs. This is not the case for the CPLD, for example. We next define the CPLD-GNEP concept, which was already independently considered in [28].

Definition 4.1 (CPLD-GNEP)**.** *We say that $\bar{x} \in X$ satisfies CPLD-GNEP if there exits a neighborhood $U \subset \mathbb{R}^n$ of $\bar{x}$ such that, for all $v = 1, \dots, N$, if $I^v \subset A^v(\bar{x})$ is such that $\{\nabla_{x^v} g_i^v(\bar{x})\}_{i \in I^v}$ is positively linearly dependent, then $\{\nabla_{x^v} g_i^v(x)\}_{i \in I^v}$ is positively linearly dependent for all $x \in U$.*

The difference between CPLD-Partial and CPLD-GNEP is that in CPLD-Partial, one just have to consider x of the form $x = (x^v, \bar{x}^{-v})$ in the neighborhood of $\bar{x}$, for each v, and in CPLD-GNEP, the positive linear dependency must occur for all $x \in \mathbb{R}^n$ close to $\bar{x}$. For the case of CCP we have the following new definition.

Definition 4.2 (CCP-GNEP)**.** *We say that $\bar{x} \in X$ satisfies CCP-GNEP if, for each $v = 1, \dots, N$, $\limsup_{x \to \bar{x}} K^v(x) \subset K^v(\bar{x})$, where*

$$K^v(x) := \left\{ w^v \in \mathbb{R}^{n_v} : w^v = \sum_{i \in A^v(\bar{x})} \lambda_i^v \nabla_{x^v} g_i^v(x) \quad \lambda_i^v \geq 0 \right\}.$$

In some sense, fixing the sums to be over the active constraints at $\bar{x}$, the definition of CCP-GNEP asks for the multifunctions $K_{X^v(x^{-v})}(x^v)$, as a function of the whole vector $x \in \mathbb{R}^n$, to be outer semicontinuous at $\bar{x}$, for each $v = 1, \dots, N$. It is easy to see that CCP-GNEP can be written as the outer semicontinuity of the single multifunction

$$K(x) := \prod_{v=1}^{N} K^v(x). \tag{4.1}$$

Now we will prove that CPLD-GNEP and CCP-GNEP are in fact CQs, i.e. if one of them is satisfied at a solution of the GNEP (1.1) then the KKT-GNEP conditions hold. To do this, it is sufficient to show that CPLD-Partial and CCP-Partial imply, respectively, CPLD-GNEP and CCP-GNEP, which are CQs for GNEPs as mentioned at Remark 3.1.

Theorem 4.1. *CPLD-Partial is weaker than CPLD-GNEP and CCP-Partial is weaker than CCP-GNEP.*

Proof. Suppose that CPLD-GNEP occurs at $\bar{x} \in X$, then there is an open set $U \subset \mathbb{R}^n$ with $\bar{x} \in U$ such that if $I^v \subset A^v(\bar{x})$ is such that $\{\nabla_{x^v} g_i^v(\bar{x})\}_{i \in I^v}$ is positively linearly dependent, then $\{\nabla_{x^v} g_i^v(x)\}_{i \in I^v}$ is positively linearly dependent for all $x \in U$. Since $U^v := \{x^v : (x^v, \bar{x}^{-v}) \in U\} \subset \mathbb{R}^{n_v}$ is an open neighborhood of $\bar{x}^v$ such that, for all $x^v \in U^v$, $\{\nabla_{x^v} g_i^v(x^v, \bar{x}^{-v})\}_{i \in I^v}$ is positively linearly dependent we have that CPLD-Partial holds at $\bar{x}$.

Suppose now that CCP-GNEP occurs at $\bar{x} \in X$, then $\limsup_{x \to \bar{x}} K^v(x) \subset K^v(\bar{x})$ for each $v = 1, \ldots, N$. It is straightforward that $K^v_{X^v(\bar{x}^{-v})}(x) \subset K^v(x)$ and so

$$\limsup_{x \to \bar{x}^v} K^v_{X^v(\bar{x}^{-v})}(x) \subset \limsup_{x \to \bar{x}} K^v(x) \subset K^v(\bar{x}).$$

Since $K^v(\bar{x}) = K^v_{X^v(\bar{x}^{-v})}(\bar{x})$ we have the desired result. □

In fact, at $\bar{x} \in X$, since each CQ-Partial, for each player $v = 1, \ldots, N$, makes a restriction on how the constraint function $g^v : \mathbb{R}^n \to \mathbb{R}^{m_v}$ should behave on neighboring points $(x^v, \bar{x}^{-v})$ of $\bar{x}$, while the corresponding CQ-GNEP extends the requirement for general neighboring points x of $\bar{x}$, CQ-GNEP is always stronger than their counterpart CQ-Partial. In particular, CQ-GNEP is a constraint qualification for GNEP. In optimization, it is correct to say that a property P such that AKKT + P implies KKT is a constraint qualification. This is true because AKKT is an optimality condition in optimization. For GNEPs, we will see that this does not hold. In [28], the authors consider CPLD-GNEP as a constraint qualification for GNEP, but they do not provide a proof. However, the fact that CPLD-GNEP and CCP-GNEP are CQs for GNEPs is a direct consequence of Theorem 4.1 and Remark 3.1.

Now we are going to prove that indeed, CCP-GNEP is weaker than CPLD-GNEP.

Theorem 4.2. *CCP-GNEP is weaker than CPLD-GNEP.*

Proof. Suppose that CPLD-GNEP occurs at $\bar{x} \in X$ and let $v \in \{1, \ldots, N\}$ be fixed and arbitrary. Let $w \in \limsup_{x \to \bar{x}} K^v(x)$, then there are sequences $\{x^k\}$ and $\{w^k\}$ such that $x^k \to \bar{x}$, $w^k \to w$ and $w^k \in K^v(x^k)$, that is, $w^k = \sum_{i \in A^v(\bar{x})} \lambda_i^{k,v} \nabla_{x^v} g_i^v(x^k)$ with $\lambda_i^{k,v} \geq 0$. By Carathéodory's Lemma 2.1, there is a set $J_v^k \subset A^v(\bar{x})$ and scalars $\tilde{\lambda}_i^{k,v} \geq 0$ with $i \in J_v^k$ such that $w^k = \sum_{i \in J_v^k} \tilde{\lambda}_i^{k,v} \nabla_{x^v} g_i^v(x^k)$ and $\left\{\nabla_{x^v} g_i^v(x^k)\right\}_{i \in J_v^k}$ is linearly independent. Since $A^v(\bar{x})$ is finite, we can take a subsequence such that $J_v^k = J_v$ and so

$$w^k = \sum_{i \in J_v} \tilde{\lambda}_i^{k,v} \nabla_{x^v} g_i^v(x^k), \quad \tilde{\lambda}_i^{k,v} \geq 0, \tag{4.2}$$

with $\left\{\nabla_{x^v} g_i^v(x^k)\right\}_{i \in J_v}$ linearly independent.

Now, suppose that the sequence $\{\tilde{\lambda}_i^{k,v}\}_{i \in J_v}$ is unbounded. Without loss of generality, we can assume that $\frac{\tilde{\lambda}^{k,v}}{\|\tilde{\lambda}^{k,v}\|} \to \bar{\lambda}^v \neq 0$. Dividing (4.2) by $\|\tilde{\lambda}^{k,v}\|$ and taking limit in k, we have that $0 = \sum_{i \in J_v} \bar{\lambda}_i^v \nabla_{x^v} g_i^v(\bar{x})$, then $\{\nabla_{x^v} g^v(\bar{x})\}_{i \in J_v}$ is positively linearly dependent. Since $x^k \to \bar{x}$, for k large enough, we have by using the definition of CPLD-GNEP that $\left\{\nabla_{x^v} g^v(x^k)\right\}_{i\in J_v}$ is positively linearly dependent. However, this is a contradiction with the way J_v was chosen, and therefore we have that $\{\tilde{\lambda}_i^{k,v}\}_{i\in J_v}$ is bounded. So we can assume that $\{\tilde{\lambda}_i^{k,v}\}_{i\in J_v}$ is convergent to $\tilde{\lambda}_i^v \geq 0$ and, taking limit in (4.2), we obtain that $w = \sum_{i\in J_v} \tilde{\lambda}_i^v \nabla_{x^v} g_i^v(\bar{x}) \in K^v(\bar{x})$ which means that the CCP-GNEP condition holds. □

Since optimization problems are a special case of GNEPs, it is clear that CCP-GNEP is strictly weaker than CPLD-GNEP [6]. The following example shows that the inclusions in Theorem 4.1 are strict.

Example 4.1 (CPLD-Partial does not imply CPLD-GNEP and CCP-Partial does not imply CCP-GNEP)**.** *Consider a GNEP with $N = 2$, $n_1 = n_2 = m_1 = 1$, $m_2 = 0$ and $g^1(x^1, x^2) = x^1x^2$. In this case, $\bar{x} := (1, 0) \in X$, and it is obvious that $\bar{x}^1 = 1$ satisfies CPLD on the set $X^1(0) = \mathbb{R}$, so CPLD-Partial holds at $\bar{x}$. CCP-Partial, being weaker than CPLD-Partial, also holds.*

On the other hand, $K^1(\bar{x}) = \{0\}$ and let us prove that $\limsup_{x\to\bar{x}} K^1(x) = \mathbb{R}$. Given $w \in \mathbb{R}$, consider the sequences $x^k := \left(1, \frac{w}{k}\right)$ and $w^k := w = k\frac{w}{k}$. It is straightforward that $w^k \in K^1(x^k)$, so $w \in \limsup_{x\to\bar{x}} K^1(x)$ and thus CCP-GNEP is not satisfied at $\bar{x}$. By Theorem 4.2 we also have that CPLD-GNEP does not hold at $\bar{x}$.

Finally, we present the extension of the quasinormality CQ for GNEPs.

Definition 4.3 (QN-GNEP)**.** *We say that $\bar{x} \in X$ satisfies QN-GNEP if for any $v = 1, \dots, N$, and any scalars $\lambda^v \in \mathbb{R}_+^{m_v}$ satisfying*

$$\sum_{i=1}^{m_v} \lambda_i^v \nabla_{x^v} g_i^v(\bar{x}) = 0, \quad \lambda_i^v g_i^v(\bar{x}) = 0,$$

there is no sequence $x^k \to \bar{x}$ such that $g_i^v(x^k) > 0$ for all k and $\lambda_i^v > 0$ for some i.

As a consequence of the results for optimization problems, we have that QN-GNEP does not imply CPLD-GNEP [10] and it is independent of CCP-GNEP [6]. However, differently from the optimization case, the following example shows that CPLD-GNEP does not imply QN-GNEP.

Example 4.2. *(CPLD-GNEP does not imply QN-GNEP) Consider a GNEP with two players where $n_1 = n_2 = 1$, $m_1 = 2$, $m_2 = 0$, $g_1^1(x) = x^1$ and $g_2^1(x) = -x^1 + x^2$. Since the gradients of $g_1^1(x)$ and $g_2^1(x)$ are constant for all x we have that CPLD-GNEP holds at every point of X. Let us show that QN-GNEP fails at $\bar{x} = (0, 0)$. Consider $\lambda_1^1 = \lambda_2^1 = 1 > 0$, then $\lambda_1^1 \nabla_{x^1} g_1^1(\bar{x}) + \lambda_2^1 \nabla_{x^1} g_2^1(\bar{x}) = 0$, however, the sequence $x^k = (1/k, 2/k)$ converges to $\bar{x}$ and $g_1^1(x^k) = g_2^1(x^k) = 1/k > 0$.*

Note that the failure occurs because of the presence of x^2 in the constraints but not in its gradients, which is particular to the situation of a GNEP. An even more pathologic example would be a two player game with $n_1 = n_2 = m_1 = 1$, $m_2 = 0$ and $g_1(x) = x^2$. Note however that MFCQ-GNEP (=MFCQ-Partial) is strictly stronger than QN-GNEP.

We are interested in global convergence results of algorithms to a KKT point under a constraint qualification. In the augmented Lagrangian literature of optimization, global convergence has been proved under CPLD in [1], with improvements in [4, 5], and more recently, under CCP in [6]. The recent paper [2] shows global convergence under QN. Since QN is independent of CCP, the results from [6] and from [2] give different global convergence results, while [2] and [6] generalize the original global convergence proof under CPLD in [1]. In a similar fashion, we will prove global convergence of an augmented Lagrangian method for GNEPs under CCP-GNEP (in Corollary 6.1) or QN-GNEP (in Theorem 6.3). While the first result is stronger than the one under CPLD-GNEP from [28], the second one is independent from [28] and also from Corollary 6.1. This will be done in Section 6. In pursuit of this goal, in the next section, we extend the notion of an AKKT point for GNEPs, where we later prove that all limit points of a sequence generated by an augmented Lagrangian method is an AKKT point. A similar extension has been done independently in [28]. Surprisingly, we discovered that the proposed extension of AKKT is not an optimality condition for a general GNEP. While this does not impact our results under a constraint qualification, we further discuss some interesting practical implications of this fact.

5 Approximate-KKT conditions for GNEPs

As we have discussed in Theorem 2.1, AKKT is an optimality condition for optimization problems, without constraint qualifications. Also, most algorithms for nonlinear programming generate sequences whose limit points satisfy AKKT [5]. Moreover, this optimality condition is strong, in the sense that it is essentially equivalent to the optimality condition "KKT or Not-CCP". In [28], the authors use a slight modification of what we call here AKKT-GNEP with a similar purpose for GNEPs, by showing that an augmented Lagrangian generates sequences whose limit points satisfies AKKT-GNEP, and that it implies the optimality condition "KKT-GNEP or Not-CPLD-GNEP". We will show that AKKT-GNEP in fact implies the stronger optimality condition "KKT-GNEP or Not-CCP-GNEP", but that itself, AKKT-GNEP is not an optimality condition. This result contrasts with what is known for optimization. In this section we discuss also some important practical implications of this fact, while also proving that the proposed concept AKKT-GNEP is in fact an optimality condition for an important class of problems, which includes NEPs.

Definition 5.1 (AKKT-GNEP)**.** *We say that $\bar{x} \in X$ satisfies AKKT-GNEP if there exist sequences $\left\{x^k\right\} \subset \mathbb{R}^n$ and $\left\{\lambda^{k,v}\right\} \subset \mathbb{R}^{m_v}_+$ for each $v = 1, \ldots, N$, such that $\lim_{k\to\infty} x^k = \bar{x}$,*

$$\lim_{k\to\infty} \left\| \nabla_{x^v} f^v(x^k) + \sum_{i=1}^{m_v} \lambda_i^{k,v} \nabla_{x^v} g_i^v(x^k) \right\| = 0, \tag{5.1}$$

and

$$\lim_{k\to\infty} \min\left\{\lambda_i^{k,v}, -g_i^v(x^k)\right\} = 0 \tag{5.2}$$

for all $i = 1, \ldots, m_v$, $v = 1, \ldots, N$.

The sequence $\{x^k\}$ used in the definition of AKKT-GNEP is called an AKKT-GNEP sequence. This concept is the same as the one independently defined in [28], with the difference that they do not require the non-negativity of the Lagrange multipliers sequence $\{\lambda^{k,v}\}$. However, the equivalence between the concepts is straightforward, replacing a possibly negative $\lambda^{k,v}$ by $\max\{\lambda^{k,v}, 0\}$ and using the continuity of the gradients. Thus, for simplicity, we will assume that the sequence of Lagrange multipliers is always non-negative.

To define a useful stopping criterion for an algorithm that generates AKKT-GNEP sequences, we need to define the ε-inexact KKT-GNEP concept, for an $\varepsilon \geq 0$ given.

Definition 5.2 (ε-inexact KKT-GNEP point)**.** *Consider the GNEP defined by* (1.1) *and let $\varepsilon \geq 0$. We call $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m$ an ε-inexact KKT-GNEP point if the following inequalities hold for each $v = 1, \ldots, N$:*

$$\left\| \nabla_{x^v} f^v(x^v, x^{-v}) + \sum_{i=1}^{m_v} \lambda_i^v \nabla_{x^v} g_i^v(x^v, x^{-v}) \right\| \leq \varepsilon, \tag{5.3}$$

$$\left\| \min\left\{\lambda_i^v, -g_i^v(x^v, x^{-v})\right\} \right\| \leq \varepsilon, \quad i = 1, ..., m_v. \tag{5.4}$$

Note that for $\varepsilon = 0$, an ε-inexact KKT-GNEP point is a standard KKT-GNEP point. Moreover, a point $\bar{x}$ satisfies the AKKT-GNEP condition if and only if there are sequences $\varepsilon^k \geq 0$, $x^k \in \mathbb{R}^n$, $\varepsilon_k \to 0$, $x^k \to \bar{x}$ such that x^k is an ε_k-inexact KKT-GNEP point.

The following example shows that even for a well-behaved GNEP, jointly convex, the condition AKKT-GNEP is not an optimality condition.

Example 5.1 (AKKT-GNEP is not an optimality condition)**.** *Consider the two-player GNEP, with $n_1 = n_2 = 1$ defined by the following problems, for $v = 1, 2$:*

$$\text{Player } v: \quad \min_{x^v} x^v \quad \text{s.t.} \quad \frac{(x^1)^2}{2} - x^1 x^2 + \frac{(x^2)^2}{2} \leq 0.$$

The solution set of this GNEP is $\{(x^1, x^2) \in \mathbb{R}^2 : x^1 = x^2\}$. Let $\bar{x}$ be a solution and suppose, by contradiction, that AKKT-GNEP occurs at $\bar{x}$. Therefore there are sequences $\{x^k\} \subset \mathbb{R}^2$ with $x^k \to \bar{x}$ and $\{\lambda^k\} \subset \mathbb{R}^2_+$ such that

$$\left|1 + \lambda^{k,1}(x^{k,1} - x^{k,2})\right| \to 0 \quad \text{and} \quad \left|1 + \lambda^{k,2}(x^{k,2} - x^{k,1})\right| \to 0.$$

Let $z^k = x^{k,1} - x^{k,2}$, then $z^k \to 0$, $\lambda^{k,1} z^k \to -1$ and $\lambda^{k,2} z^k \to 1$. For k large enough we have that $z^k \neq 0$, hence $\frac{\lambda^{k,1} z^k}{\lambda^{k,2} z^k} \to -1$, which implies $\frac{\lambda^{k,1}}{\lambda^{k,2}} \to -1$. This is a contradiction with the fact that $\lambda^{k,1}$ and $\lambda^{k,2}$ are nonnegative, therefore $\bar{x}$ is not an AKKT-GNEP point.

Example 5.1 shows that algorithms that provably generate sequences whose limit points are AKKT-GNEP points, can never converge to any solution of this well-behaved problem. In particular, either the iterand x^k is not defined, with a subproblem that can not be solved, or the sequence $\{x^k\}$ generated is unbounded, without a limit point. The possibility of converging to an infeasible point is ruled out by [28, Theorem 4.3].

The situation is strikingly different from the augmented Lagrangian for optimization, which can be defined with safeguarded boxes in such a way that a limit point of $\{x^k\}$ always exists, that when feasible, is an AKKT point. Also, this does not exclude any solution *a priori* such as in GNEPs.

On the other hand, the next example shows that algorithms that generate AKKT-GNEP sequences can also be attracted by a point that does not satisfy a true optimality condition (hence, not a solution). To formalize this we need the following concept.

Definition 5.3 (AKKT-Partial)**.** *Let $\bar{x} \in X$ be a feasible point of the GNEP* (1.1). *We say that $\bar{x}$ satisfies the AKKT-Partial condition if*

$$\bar{x}^v \quad \textit{satisfies AKKT for the optimization problem} \quad P^v(\bar{x}^{-v}), \quad \forall v = 1, \ldots, N.$$

By the definition of the GNEP and by Theorem 2.1, it is easy to see that AKKT-Partial is an optimality condition for GNEPs, without constraint qualifications. However, this condition is too strong for one to expect that an algorithm would generate this type of sequence, as $\{x^{k,v}\}$ would be an AKKT-type sequence for each $v = 1, \ldots, N$, but only when paired with the true limit, namely, $\{(x^{k,v}, \bar{x}^{-v})\}$. Example 5.1 showed that AKKT-Partial does not imply AKKT-GNEP. The next example shows that our two AKKT concepts are in fact independent. For this example, AKKT-GNEP holds while AKKT-Partial does not. Hence, given that AKKT-Partial is an optimality condition, an algorithm can generate an AKKT-GNEP sequence that will necessarily fail to converge to a solution.

Example 5.2 (AKKT-GNEP does not imply AKKT-Partial)**.** *Consider a GNEP with $N = 2$, $n_1 = n_2 = m_1 = 1$, $m_2 = 0$, $f_v(x) = \frac{(x^v)^2}{2}, v = 1, 2$ and $g^1(x^1, x^2) = x^1 x^2$.*

The unique solution of this GNEP is $(0, 0)$*. Now, consider the point $\bar{x} := (-1, 0) \in X$ and let $\left\{x^k\right\} \subset \mathbb{R}^2$ be any sequence such that $x^k \to (-1, 0)$ and $\left\{\lambda^{k,1}\right\} \subset \mathbb{R}_+$. For k large enough we have $x^{k,1} + \lambda^{k,1} 0 \to -1$ therefore AKKT-Partial does not occur in x.*

To see that AKKT-GNEP holds, consider the sequences $x^k = (-1, \frac{1}{k})$ and $\lambda^{k,1} = k$, then it is clear that:

$$\left|x^{k,1} + \lambda^{k,1} x^{k,2}\right| \to 0 \quad \textit{and} \quad \left|x^{k,2}\right| \to 0.$$

After seeing some bad properties of AKKT-GNEP, a natural question is if there are some class of GNEPs for which AKKT-GNEP is an optimality condition. The answer is affirmative for problems with somewhat separable constraints and for variational equilibrium of a jointly convex GNEP.

Theorem 5.1. *Consider a GNEP where the constraints of each player have the structure $g_i^v(x) := g_i^{v,1}(x^v) g_i^{v,2}(x^{-v}) + g_i^{v,3}(x^{-v})$ for each $v = 1, \ldots, N$ and $i \in \{1, \ldots, m_v\}$. If $\bar{x}$ is a solution of this GNEP then the AKKT-GNEP condition holds at $\bar{x}$.*

Proof. Let us show that an AKKT-Partial sequence is in fact an AKKT-GNEP sequence. By the definition of solution of a GNEP we have that $\bar{x}^v$ is the solution of the following optimization problem:

$$P^v(\bar{x}^{-v}) : \min_{x^v} f^v(x^v, \bar{x}^{-v}) \text{ s.t. } g_i^{v,1}(x^v) g_i^{v,2}(\bar{x}^{-v}) + g_i^{v,3}(\bar{x}^{-v}) \leq 0, \, i = 1, \ldots, m_v.$$

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

$$\|\nabla_{x^v} f^v(x^{k,v}, \bar{x}^{-v}) + \sum_{i \in A^v(\bar{x})} \lambda_i^{k,v} g_i^{v,2}(\bar{x}^{-v}) \nabla_{x^v} g_i^{v,1}(x^{k,v})\| \to 0. \tag{5.5}$$

Let $x^k := (x^{k,v})_{v=1}^N$ and define

$$\bar{\lambda}_i^{k,v} := \begin{cases} 0, \text{ if } g_i^{v,2}(\bar{x}^{-v}) g_i^{v,2}(x^{k,-v}) = 0 \\ \lambda_i^{k,v} \frac{g_i^{v,2}(\bar{x}^{-v})}{g_i^{v,2}(x^{k,-v})}, \text{otherwise.} \end{cases}$$

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

$$\nabla_{x^v} g_i^v(x) = g_i^{v,2}(x^{-v}) \nabla_{x^v} g_i^{v,1}(x^v)$$

we have that

$$\bar{\lambda}_i^{k,v} \nabla_{x^v} g_i^v(x^k) = \lambda_i^{k,v} g_i^{v,2}(\bar{x}^{-v}) \nabla_{x^v} g_i^{v,1}(x^{k,v}).$$

Therefore, by (5.5) and the triangular inequality,

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

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

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

□

This class of somewhat separable GNEPs contains important subclasses such as the linear GNEPs that were studied in [17] or even the NEPs.

Note that we can not combine the sum with the product as $g_i^v(x) := g_i^{v,1}(x^v) g_i^{v,2}(x^{-v}) + g_i^{v,3}(x^v)$. For example, if $N = 2$, $n_1 = n_2 = m_1 = m_2 = 1$, $f^1(x^1) := x^1$, $f^2(x^2) := -x^2$, $g^1(x) := x^1 x^2 + \frac{x^1}{4}$ and $g^2(x) := x^2 \left(x^1\right)^3 + \frac{x^2}{2}$ then the origin is a solution of the GNEP, but the approximate multipliers $\lambda^{k,1}$ and $\lambda^{k,2}$ must have opposite signs.

Let us recall that jointly convex GNEPs constitute an important special class of GNEPs, since it arises in some interesting applications, and for which a much more complete theory exists than for the arbitrary GNEP [22]. Consider the feasible set $X := \{x \in \mathbb{R}^n : g(x) \leq 0\}$ defined by a convex function $g : \mathbb{R}^n \to \mathbb{R}^m$ and a jointly convex GNEP associated with it. Related to this GNEP, the variational inequality problem $VI(X, F)$ is defined as:

$$\text{Find} \quad \bar{x} \in X \quad \text{such that} \quad \langle F(\bar{x}), x - \bar{x} \rangle \geq 0, \quad \text{for all} \quad x \in X,$$

where $F(x) := (\nabla_{x^v} f^v(x))_{v=1}^N$.

The following theorem establishes the relation between the solutions of $VI(X, F)$ and of the jointly convex GNEP; the proof may be found in [18].

Theorem 5.2. *Let a jointly convex GNEP be given with C^1-functions f^v. Then, every solution of the $VI(X, F)$ is also a solution of the GNEP.*

We remark that the reciprocal implication is not true in general. In Example 5.1, we have a jointly convex GNEP, however, there is no solution of the associated $VI(X, F)$. When the reciprocal implication holds, we say that the solution is a variational equilibrium:

Definition 5.4. *Let a jointly convex GNEP be given with C^1-functions f^v. We call a solution of the GNEP that is also a solution of $VI(X, F)$ a variational equilibrium.*

The alternative name normalized equilibrium is also frequently used in the literature instead of variational equilibrium. Let us show that at a variational solution of a jointly convex GNEP, AKKT-GNEP holds.

Theorem 5.3. *Any variational equilibrium of a jointly convex GNEP is an AKKT-GNEP point.*

Proof. A point $\bar{x}$ is a solution of the $VI(X, F)$ if, and only if $\bar{x}$ is a solution of the nonlinear programming problem:

$$\min \langle F(\bar{x}), x\rangle \quad s.t. \quad x \in X.$$

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

$$F(\bar{x}) + \sum_{i \in A(\bar{x})} \lambda_i^k \nabla g_i(x^k) \to 0 \tag{5.6}$$

where $A(\bar{x}) = \{i \in \{1, \ldots, m\} : g_i(\bar{x}) = 0\} = A^v(\bar{x})$ for all $v = 1, \ldots, N$. By the definition of F we have that

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

Since $\nabla_{x^v} f^v(x^k) \to \nabla_{x^v} f^v(\bar{x})$, we conclude that

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

and therefore $\bar{x}$ is an AKKT-GNEP point. $\square$

Even though variational solutions are, in general, a proper subset of the solution set of a GNEP, an auxiliary GNEP where each objective function is multiplied by a positive constant can be considered. In this case, the solution set is unchanged but the variational solutions for the auxiliary problem, which are AKKT-GNEP points for the auxiliary GNEP, are trivialy also AKKT-GNEP points for the original GNEP. In the case of the Example in [18, Section 5], even though only the point $(\frac{3}{4}, \frac{1}{4})$ is a variational equilibrium, while the whole set of solutions is the closed line segment bounded by $(\frac{1}{2}, \frac{1}{2})$ and $(1, 0)$, we can show in this fashion that all solutions in the relative interior of the line are AKKT-GNEP points.

In the following theorem we show that CCP-GNEP plays, with respect to AKKT-GNEP, the same role as the Guignard-Partial plays with respect to optimality. Namely, Guignard-Partial is the weakest constraint qualification that guarantees that any solution satisfies KKT-GNEP. In the same sense, CCP-GNEP is the weakest condition that guarantees that AKKT-GNEP implies KKT-GNEP.

Theorem 5.4. *The CCP-GNEP condition is the weakest property under which AKKT-GNEP implies KKT-GNEP, independently of the objective functions f^v of each player $v = 1, \ldots, N$.*

Proof. Let $\bar{x} \in X$ satisfy CCP-GNEP. Suppose that AKKT-GNEP occurs in $\bar{x}$, then there are sequences $\{x^k\} \subset \mathbb{R}^n$ with $x^k \to \bar{x}$ and $\{\lambda^{k,v}\} \subset \mathbb{R}_+^{|A^v(\bar{x})|}$ such that for each $v = 1, \ldots, N$ we have:

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

Let $w^{k,v} := \sum_{i \in A^v(\bar{x})} \lambda_i^{k,v} \nabla_{x^v} g_i^v(x^k)$ and $w^k := \left(w^{k,v}\right)_{v=1}^N$ then $w^k \in K(x^k)$ as defined in (4.1). Since $w^{k,v} \to -\nabla_{x^v} f^v(\bar{x})$, we have that $w^k \to -F(\bar{x}) := -\left(\nabla_{x^v} f^v(\bar{x})\right)_{v=1}^N$, therefore $-F(\bar{x}) \in \limsup_{x \to \bar{x}} K(x)$. By the CCP-GNEP condition, we have that $-F(\bar{x}) \in K(\bar{x})$ and so $\bar{x}$ is a KKT-GNEP point.

Reciprocally, assume that $\bar{x}$ is such that AKKT-GNEP implies KKT-GNEP, independently of the objective functions of each player. Let $w \in \limsup_{x \to \bar{x}} K(x)$, so there are sequences $x^k \to \bar{x}$ and $w^k \to w$ with $w^k \in K(x^k)$. Since $w^k \in K(x^k)$, there exists for each $v = 1, \ldots, N$ sequences $\{\lambda^{k,v}\} \subset \mathbb{R}_+^{|A^v(\bar{x})|}$ such that

$$w^{k,v} = \sum_{i \in A^v(\bar{x})} \lambda_i^{k,v} \nabla_{x^v} g_i^v(x^k). \tag{5.7}$$

Define for each player $v = 1, \ldots, N$ the objective function $\tilde{f}^v(x) := -\langle x, w \rangle$, then $\nabla_{x^v} \tilde{f}^v(x) = -w^v \in \mathbb{R}^{n_v}$. So, by (5.7),

$$\nabla_{x^v} \tilde{f}^v(x^k) + \sum_{i \in A^v(\bar{x})} \lambda_i^{k,v} \nabla_{x^v} g_i^v(x^k) = -w^v + w^{k,v} \to 0$$

for each $v = 1, \ldots, N$. Therefore $\bar{x}$ satisfies AKKT-GNEP and by the assumption, $\bar{x}$ is a KKT-GNEP point, which implies that $w \in K(\bar{x})$ and therefore $\bar{x}$ satisfies CCP-GNEP. □

As a final remark, it is an imediate consequence of [6] that CCP-Partial is the weakest condition such that AKKT-Partial implies KKT-GNEP.

6 Algorithm and Convergence

In this section we first derive our augmented Lagrangian algorithm for GNEPs, with lower level constraints. After that, we present its convergence analysis, which is done independently and with a weaker assumptions than in [28]. We discuss feasibility and optimality results under the CCP-GNEP condition; moreover, we give an alternative global convergence result using the quasinormality condition for GNEPs (QN-GNEP) that includes boundedness of the Lagrange multipliers approximation.

Augmented Lagrangian methods are useful when there exists efficient algorithms for solving its subproblems. These subproblems may have what is called lower-level constraints, which means that those constraints would not be penalized. If all the joint constraints are penalized then the subproblems are NEPs, for which the theory is richer than for GNEPs. The method we present here was inspired by the ones in [27] and [15]. In [27], the authors present an augmented Lagrangian method for the resolution of QVIs, which generalizes the method presented in [32]. They analyze the boundedness of the penalty parameter and the global convergence analysis is done with a weaker constraint qualification (CPLD-GNEP) than in [32], where it was used MFCQ-GNEP.

Similarly to [28], we specialize the algorithm of [27] for GNEPs with upper and lower level constraints. Our algorithm was developed independently but is essentially the same as presented in [28], however, our convergence analysis is done with weaker assumptions.

We consider the GNEP with upper and lower level constraints:

$$
\begin{array}{lll}
Q_v(x^{-v}): & \min_{x^v} & f^v(x^v, x^{-v}) \\
& s.t. & g_i^v(x^v, x^{-v}) \le 0 \quad i = 1, ..., m_v, \\
& & h_j^v(x^v, x^{-v}) \le 0 \quad j = 1, ..., l_v.
\end{array}
\tag{6.1}
$$

Analogously to the penalty methods for classical optimization problems, we penalize the upper-level constraints and maintain the lower-level constraints in the subproblems. In this work we use the classic Powell-Hestenes-Rockafellar augmented Lagrangian. Given $u^v \in \mathbb{R}_+^{m_v}$ and $\rho^v > 0$, the augmented Lagrangian function for player v is given by

$$
L_{\rho^v}^v(x, u^v) := f^v(x^v, x^{-v}) + \frac{\rho^v}{2} \sum_{i=1}^{m_v} \max\left\{0, g_i^v(x^v, x^{-v}) + \frac{u_i^v}{\rho^v}\right\}^2 \tag{6.2}
$$

for all x such that $h^v(x) \le 0$ and for each $v = 1, \dots, N$. We define the vectors $u := (u^v)_{v=1}^N$, which combines the safeguarded estimates for the Lagrange multipliers associated with the functions g^v, and $\rho := (\rho^v)_{v=1}^N$, which encompasses the penalty parameters.

Note that the functions $L_{\rho^v}(x, u^v)$ are continuously differentiable with respect to x and the partial gradient with respect to x^v is given by:

$$
\nabla_{x^v} L_{\rho^v}^v(x, u^v) = \nabla_{x^v} f^v(x^v, x^{-v}) + \sum_{i=1}^{m_v} \max\left\{0, u_i^v + \rho^v g_i^v(x^v, x^{-v})\right\} \nabla_{x^v} g_i^v(x^v, x^{-v}).
$$

The augmented Lagrangian functions define a new GNEP, a GNEP(u, ρ), where the problem for player v is given by:

$$
\begin{array}{ll}
\min_{x^v} & L_{\rho^v}^v(x, u^v) \\
& h_j^v(x^v, x^{-v}) \le 0 \quad j = 1, ..., l_v.
\end{array}
\tag{6.3}
$$

The algorithm presented here aims at findind an ε-inexact KKT-GNEP point for (6.1), in the sence of Definition 5.2, through the calculation of ε_k-inexact KKT-GNEP points of a sequence of GNEPs given by GNEP(u^k, ρ^k). The following algorithm is a direct extension of Algorithms 4.1 of [15] and 3.1 of [27] for GNEPs.

Algorithm AL-GNEP: (Augmented Lagrangian for GNEPs)

Step 0) Let $u^{max} \in \mathbb{R}^m_+$, $0 < \tau < 1$, $\gamma > 1$ and $\varepsilon \geq 0$. Choose $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l$, $u^0 \in [0, u^{max}]$, $\rho^0 \in \mathbb{R}^N_+$, a bounded positive sequence $\{\varepsilon_k\}$ and set $k = 0$.

Step 1) If (x^k, λ^k, μ^k) is an ε-inexact KKT-GNEP point of the original GNEP (6.1): STOP.

Step 2) Compute an ε_{k+1}-inexact KKT-GNEP point (x^{k+1}, μ^{k+1}) of GNEP(u^k, ρ^k).

Step 3) Define $\lambda^{k+1,v} := \max\left\{0, u^{k,v} + \rho^{k,v} g^v(x^{k+1})\right\}$ for all $v = 1, \ldots, N$.

Step 4) For each $v = 1, ..., N$, if

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

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

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

By construction, the sequence $\left\{u^k\right\}$ generated by Algorithm AL-GNEP is nonnegative and bounded. A natural choice at Step 5 is $u^{k+1} := \min\left\{\lambda^{k+1}, u^{max}\right\}$. We could also consider different parameters γ^v, τ^v, but for simplicity we will take the same γ, τ for each player.

In our convergence analysis we consider the asymptotic behavior of the sequences generated by Algorithm AL-GNEP, for this we consider $\varepsilon = 0$ as the stopping criterion used in Step 2 of the algorithm. The following theorem proves that when the algorithm admits a feasible limit point, this point satisfies the AKKT-GNEP condition.

Theorem 6.1. *Assume that the sequence $\left\{x^k\right\}$ is generated by algorithm AL-GNEP and $K \subset \mathbb{N}$ is such that $\lim_{k \in K} x^k = \bar{x}$ and $\bar{x}$ is feasible. Moreover, assume that the bounded sequence $\{\varepsilon_k\}$ is such that $\lim_{k \in K} \varepsilon_k = 0$. Then, $\bar{x}$ satisfies the AKKT-GNEP conditions for GNEP* (6.1).

Proof. We omit the proof of the theorem because it is a direct adaptation of Theorem 6.2 of [15] and a very similar proof has also been presented in Theorem 4.6 of [28]. □

Corollary 6.1. *Under the asumptions of Theorem 6.1, if $\bar{x}$ fulfills the CCP-GNEP constraint qualification, then $\bar{x}$ is a KKT-GNEP point of GNEP* (6.1).

Proof. The proof is a consequence of Theorems 6.1 and 5.4. □

In order to deal with the feasibility property of the limit point, we define the *game of infeasibility*:

$$\min_{x^v} \left\|g^v_+(x)\right\|^2 \text{ subject to } h^v_j(x^v, x^{-v}) \leq 0, \quad j = 1, \ldots, l_v. \tag{6.4}$$

The following theorem states that, by means of Algorithm AL-GNEP, we necessarily find stationary points of the game of infeasibility. It is interesting that we do not need that ε_k converges to zero to prove this important property.

Theorem 6.2. *Assume that the sequence* $\left\{x^k\right\}$ *is generated by Algorithm AL-GNEP. Let* $\bar{x}$ *be a limit point of* $\left\{x^k\right\}$*. Then* $\bar{x}$ *satisfies AKKT-GNEP for the game of infeasibility* (6.4)*.*

Proof. Once again, we omit the proof because it is very similar of the one presented in Theorem 6.2 of [15] and Lemma 4.1 of [28]. □

This immediately yields the following corollary.

Corollary 6.2. *Every limit point of a sequence* $\{x^k\}$ *generated by Algorithm AL-GNEP that satisfy CCP-GNEP for the game of infeasibility* (6.4) *is a KKT-GNEP point of the game of infeasibility.*

Proof. The proof is a direct consequence of Theorems 6.2 and 5.4. □

To finish our contributions, we next prove the convergence of Algorithm AL-GNEP under the QN-GNEP constraint qualification. This result is not directly associated with the concept of AKKT-GNEP, as CCP-GNEP is the minimal condition ensuring that an AKKT-GNEP point is a KKT-GNEP point, but CCP-GNEP is independent of QN-GNEP. This property is inherent to the Augmented Lagrangian method, by taking into account that its iterates are usually infeasible. A similar result in the optimization context has appeared recently in [2].

Under QN-GNEP, we also prove that the dual sequence $\{\lambda^k\}$, associated with the penalized constraints, is bounded. This result is new and surprising, even in the optimization case ($N = 1$). A consequence of the result for optimization is that, since CPLD implies QN [10], the dual sequence generated by the augmented Lagrangian [1] under CPLD also generates bounded dual sequences. Since CPLD-GNEP does not imply QN-GNEP, this result does not hold for GNEPs under CPLD-GNEP. Note that the result holds, in particular, when the lower-level constraints are empty, namely, when all constraints of problem (6.1) are penalized. In this sense, there is no loss of information in passing from the external penalty method to the augmented Lagrangian method in terms of conditions ensuring boundedness of the Lagrange multiplier estimate.

To prove our global convergence result under QN-GNEP, we need an algorithmic assumption on the Lagrange multipliers sequence $\{\mu^k\}$, which are associated with the lower level constraints and are computed at Step 2. We will assume that $\{\mu^k\}$ is a bounded sequence, which can be ensured, for instance, if there are no lower level constraints. Also, this assumption is reasonable given that the lower level constraints should be easy to handle, as one needs to approximately solve an equilibrium problem subject to the lower level constraints at each iterations. We also point out that this assumption was also present in the general lower level augmented Lagrangian from [11]. See also [12].

Theorem 6.3. *Assume that the sequence* $\left\{x^k\right\}$ *is generated by algorithm AL-GNEP and* $K \subset \mathbb{N}$ *is such that* $\lim_{k\in K} x^k = \bar{x}$ *and* $\bar{x}$ *is feasible. Moreover, assume that there exists* $\bar{\mu} \geq 0$ *such that* $\|\mu^k\| \leq \bar{\mu}$ *for all* $k \in K$ *and that* $\lim_{k\in K} \varepsilon_k = 0$*. Then, if* $\bar{x}$ *fulfills the QN-GNEP condition with respect to* (6.1)*, we have that the dual sequence* $\{\lambda^k\}$ *is bounded on* K*, and its limit points are Lagrange multipliers associated with* $\bar{x}$*. In particular,* $\bar{x}$ *is a KKT-GNEP point for problem* (6.1)*.*

Proof. By Step 2 of the algorithm and Definition 5.2, we have for each $v = 1, \ldots, N$ that

$$\left\| \nabla_{x^v} f^v(x^k) + \sum_{i=1}^{m_v} \lambda_i^{k,v} \nabla_{x^{k,v}} g_i^v(x^k) + \sum_{i=1}^{l_v} \mu_i^{k,v} \nabla_{x^{k,v}} h_i^v(x^k) \right\| \leq \varepsilon_k \tag{6.5}$$

where $\lambda_i^{k,v} = \max\left\{0, u_i^{k-1,v} + \rho^{k-1,v} g_i^v(x^k)\right\}$. Let

$$\delta^{k,v} := \sqrt{1 + \sum_{i=1}^{m_v} \left(\lambda_i^{k,v}\right)^2 + \sum_{i=1}^{l_v} \left(\mu_i^{k,v}\right)^2},$$

and let us assume to obtain a contradiction that there is a subsequence $K_1 \subset K$ such that $\delta^{k,v} \to +\infty$ for $k \in K_1$. Since $\left\|\left(\frac{1}{\delta^{k,v}}, \frac{\lambda^{k,v}}{\delta^{k,v}}, \frac{\mu^{k,v}}{\delta^{k,v}}\right)\right\| = 1$ for every k, there exists $K_2 \subset K_1$ such that

$$\lim_{k \in K_2} \left(\frac{1}{\delta^{k,v}}, \frac{\lambda^{k,v}}{\delta^{k,v}}, \frac{\mu^{k,v}}{\delta^{k,v}}\right) = \left(\bar{\nu}^v, \bar{\lambda}^v, \bar{\mu}^v\right) \neq 0$$

with $\bar{\nu}^v = 0$ and $\bar{\mu}_j^v = 0$ for every $j = 1, \ldots, l_v$, while $\bar{\lambda}_i^v \geq 0$ for all $j = 1, \ldots, l_v$, with $\bar{\lambda}^v \neq 0$.

Dividing the expression (6.5) by $\delta^{k,v}$ and taking limits on K_2 we get

$$\sum_{i=1}^{m_v} \bar{\lambda}_i^v \nabla_{x^v} g_i^v(\bar{x}) = 0. \tag{6.6}$$

Moreover, since the AKKT-GNEP condition yields $\lambda_i^{k,v} \to 0$ if $g_i^v(\bar{x}) < 0$, equation (6.6) gives a non-trivial positive linear combination of the gradients of the active constraints $g^v(x) \leq 0$ at $\bar{x}$.

If $\bar{\lambda}_i^v > 0$, then there exists $a > 0$ and k_0 such that

$$\frac{\max\{0, u_i^{k-1,v} + \rho^{k-1,v} g_i^v(x^k)\}}{\delta^{k,v}} \geq a, \quad \forall k \geq k_0, \quad k \in K_2.$$

Since $\{u^{k-1,v}\}$ is bounded and $\bar{x}$ is feasible, we must have that $\rho^{k-1,v} g_i^v(x^k)$ goes to infinity and so $g_i^v(x^k) > 0$ for $k \in K_2$ large enough. However, this contradicts the QN-GNEP condition.

□

We note that an extension of the above result also holds when equality constraints are present.

7 Conclusions

The major contribution of this work is to show that the AKKT-GNEP condition is not an optimality condition for GNEPs. This observation has important practical implications because it shows that methods that generate sequences of this type may be neglecting solutions even in well-behaved problems. In addition, we show that the AKKT-GNEP concept is independent of the AKKT-Partial concept (which is an optimality condition) and therefore even points that do not satisfy a true optimality condition (hence not a solution) can be found by algorithms generating AKKT-GNEP sequences. Since it is natural to construct methods that generate AKKT-GNEP sequences (and some of them have recently been published), we believe it is important to draw the attention to this fact.

For specific classes of GNEPs, we have proved that AKKT-GNEP is an optimality condition, and for these cases, it is more reasonable to use methods that generate AKKT-GNEP sequences. In this paper we formalize and establish relation among several CQs for GNEPs. In particular, we have defined the CCP-GNEP condition and proved that it is the weakest condition that ensures that an AKKT-GNEP point satisfies the KKT-GNEP condition. Therefore, we obtain a more refined convergence result than the one presented in [28] for an Augmented Lagrangian type method for GNEPs. In addition, we have also shown a convergence result using the QN-GNEP condition, which we surprisingly proved to be independent of the CPLD-GNEP, in contrast with what is known in optimization. Boundedness of the dual sequence was also proved in this case.

Another important point about the fact that AKKT-GNEP is not an optimality condition is that it may not be appropriate to use as an algorithm's stopping criterion that the KKT conditions are satisfied within certain precision. This is due to the fact that it is not the case that solutions of a GNEP are approximated by perturbed KKT-GNEP points. An important research topic would be to define a true sequential optimality condition for GNEPs, and to adapt the augmented Lagrangian algorithm in such a way that the new type of sequences is always generated. Even in the algorithm presented here, we asked in Step 1 that the approximate solution of the subproblem satisfies its KKT conditions with precision ε_k. This raises questions on conditions for the algorithm to be well defined and which should be the choice of the constraints to be used at the upper and lower levels. We believe that ensuring that the subproblem has an approximate solution is closely related to the fact that the original problem has an approximate solution and this is an interesting point for future research.

A curious fact is that stopping criteria not based on KKT conditions would have implications even in traditional methods for standard nonlinear optimization. In [9] and [8] the authors show that Newton's method can generate sequences that are not AKKT. The authors present this fact with a pessimistic tone, emphasizing that the method may not recognize a solution of an optimization problem. Since Newton-type methods are also widely used for equilibrium problems and variational inequalities, we might assume that this would also be a deficiency of these methods. However, if a good alternate stopping criterion is developed, this feature can become a virtue, since the method would not have the drawback of avoiding non-AKKT solutions.

Finally, it is worth mentioning that global convergence results relying on a constraint qualification are mathematically correct, by establishing conditions on the limit points of the sequence generated by the algorithm, but they are incomplete. With simple examples in mind where the algorithm may fail, there is a greater concern about sufficient conditions to ensure that this type of algorithm does indeed generate a sequence and that this sequence has limit points. A study on the computational behavior of this type of method when applied to problems whose solutions are not AKKT-GNEP points would be very welcomed and certainly is within our future research perspectives.

References

[1] R. Andreani, E. G. Birgin, J. M. Martínez and M. L. Schuverdt. Augmented Lagrangian methods under the Constant Positive Linear Dependence constraint qualification. *Mathematical Programming* 111:5–32, 2008.

[2] R. Andreani, N. S. Fazzio, M. L. Schuverdt, L. D. Secchin. A sequential optimality condition related to the quasinormality constraint qualification and its algorithmic consequences. *Optimization Online*, 2017.

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

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

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

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

[7] R. Andreani, J. M. Martínez, A. Ramos, and P. J. S. Silva. Strict constraint qualifications and sequential optimality conditions for constrained optimization. *to appear in Mathematics of Operations Research*, 2017.

[8] R. Andreani, J. M. Martínez, L. T. Santos. Newton's method may fail to recognize proximity to optimal points in constrained optimization. *Mathematical Programming* 160, pp. 547–555, 2016.

[9] R. Andreani, J. M. Martínez, L. T. Santos, and B. F. Svaiter. On the behavior of constrained optimization methods when Lagrage multipliers do not exist. *Optimization Methods and Software* 29, pp. 646–657, 2014.

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

[11] R. Andreani, J. M. Martínez, and B. F. Svaiter. A new sequential optimality condition for constrained optimization and algorithmic consequences. *SIAM Journal on Optimization*, 20(6):3533–3554, 2010.

[12] E. G. Birgin, G. Haeser, and A. Ramos. Augmented Lagrangians with constrained subproblems and convergence to second-order stationary points. *Computational Optimization and Applications*, 2017. DOI: 10.1007/s10589-017-9937-2.

[13] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Practical methods of Optimization: theory and algorithms*. John Wiley and Sons, 2006.

[14] D. P. Bertsekas. *Nonlinear programming.* Athenas Scientific, 1999.

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

[16] 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.

[17] A. Dreves, N. Sudermann-Merx. Solving linear generalized Nash equilibrium problems numerically. *Optimization Methods & Software*, 31(5):1036–1063, 2016.

[18] F. Facchinei, A. Fischer, and V. Piccialli. On generalized nash games and variational inequalities. *Oper. Res. Lett.*, 35(2):159–164, 2007.

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

[20] F. Facchinei and C. Kanzow. Penalty methods for the solution of generalized nash equilibrium problems. *SIAM J. Optim.*, 20(5):2228–2253, 2010.

[21] F. Facchinei and L. Lampariello. Partial penalization for the solution of generalized nash equilibrium problems. *J. Global Optim.*, 50(1):39–57, 2011.

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

[23] M. Fukushima. Restricted generalized nash equilibria and controlled penalty algorithm. *Comput. Manag. Sci.*, 8(3):201–218, 2011.

[24] F. J. Gould and J. W. Tolle. A necessary and sufficient qualification for constrained optimization. *SIAM Journal of Applied Mathematics*, 20:164–172, 1971.

[25] M. R. Hestenes. *Optimization Theory: The Finite Dimensional Case.* John Wiley, 1975.

[26] R. Janin. Direction derivative of the marginal function in nonlinear programming. *Mathematical Programming Studies*, 21:127–138, 1984.

[27] C. Kanzow. On the multiplier-penalty-approach for quasi-variational inequalities. *Math. Program.*, 2016.

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

[29] O. L. Mangasarian and S. Fromovitz. The Fritz-John necessary conditions in presence of equality and inequality constraints. *Journal of Mathematical Analysis and applications*, 17:37–47, 1967.

[30] L. C. Matioli, W. Sosa, and J. Yuan. A numerical algorithm for finding solutions of a generalized nash equilibrium problem. *Computational Optimization and Applications*, 2017. DOI: 10.1007/s10589-011-9407-1.

[31] L. Minchenko and S. Stakhovski. On relaxed constant rank regularity condition in mathematical programming. *Optimization*, 60:429–440, 2011.

[32] 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.

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

[34] R. T. Rockafellar and R. Wets. *Variational Analysis*. Series: Grundlehren der mathematischen Wissenschaften, Vol. 317, 2009.

[35] F. N. Rojas. Augmented Lagrangian methods for generalized Nash equilibrium problems. *Department of Applied Mathematics, University of São Paulo*, June 15th, 2016. Qualifying examination for PhD candidacy.

[36] A. von Heusinger and C. Kanzow. Optimization reformulations of the generalized nash equilibrium problem using nikaido-isoda-type functions. *Comput. Optim. Appl.*, 43(3):353–377, 2009.