SIAM J. OPTIM.
Vol. 29, No. 1, pp. 31–54

OPTIMALITY CONDITIONS AND CONSTRAINT QUALIFICATIONS FOR GENERALIZED NASH EQUILIBRIUM PROBLEMS AND THEIR PRACTICAL IMPLICATIONS*

LUÍS FELIPE BUENO†, GABRIEL HAESER‡, AND FRANK NAVARRO ROJAS‡

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 (CQs) and optimality conditions tailored for GNEPs, and we discuss their relations and implications for global convergence of algorithms. We show the surprising fact that, in contrast to the case of nonlinear programming, in general the Karush–Kuhn–Tucker (KKT) residual cannot 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 an introduced weak CQ, we prove global convergence to a KKT point of an augmented Lagrangian algorithm for GNEPs, and under the quasi-normality (QN) CQ 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

AMS subject classifications. 65K05, 90C30, 90C46, 91A10

DOI. 10.1137/17M1162524

1. Introduction.

In this paper we consider the generalized Nash equilibrium problem (GNEP) where, given N players, each player $v = 1, \ldots, N$ aims at minimizing

$$P^v(x^{-v}): \quad \min_{x^v} f^v(x^v, x^{-v}) \quad \text{s.t.} \quad g^v(x^v, x^{-v}) \leq 0 \tag{1}$$

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 if $g^1 = \cdots = g^N$ is convex as a function of the entire vector x. In the case when 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

*Received by the editors December 26, 2017; accepted for publication (in revised form) October 8, 2018; published electronically January 3, 2019.
http://www.siam.org/journals/siopt/29-1/M116252.html
Funding: This work was supported by FAPESP (grants 2013/05475-7, 2015/02528-8, and 2017/18308-2), CNPq, and CAPES.
†Institute of Science and Technology, Federal University of São Paulo, São José dos Campos SP, Brazil (lfelipebueno@gmail.com).
‡Department of Applied Mathematics, Institute of Mathematics and Statistics, University of São Paulo, São Paulo SP, Brazil (ghaeser@ime.usp.br, franknr@ime.usp.br).

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/her objective function by unilaterally changing his/her 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 solutions 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 them 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) [12], Mangasarian–Fromovitz (MFCQ) [29], quasi-normality (QN) [25, 13], and constant positive linear dependence (CPLD) [33, 10]. Other weak CQs associated with global convergence of first-order algorithms have been introduced in 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 CQs 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 QN extension for GNEPs presented here is not weaker than the CPLD condition, as is the case in optimization [10]. We show that, even more surprising, 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]: 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 in 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 CQ, while the one in [28] is based on the (stronger) CPLD. It is well known that CCP is much weaker than CPLD; for instance, a problem that does not satisfy CCP can be made into a problem that does satisfy it by including appropriate well-behaved redundant constraints, while CPLD does not have this property (see [5]). We also present a convergence result based on the QN CQ, which extends [2] from optimization to GNEPs but proves 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, this kind of method excludes the possibility of finding some solutions of the GNEP. 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, and 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, using the standard approach of fixing $\bar{x}^{-v}$ when looking at player v's problem. 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 (GNEP-AKKT). We give an example where GNEP-AKKT 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 GNEP-AKKT sequences. We prove that for important classes of GNEPs, which include NEPs, the GNEP-AKKT is a necessary optimality condition. Moreover, we prove that the GNEP-CCP condition introduced in section 4 is the weakest condition that ensures that GNEP-AKKT 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 GNEP-CCP, which is stronger than the corresponding result from [28] under GNEP-CPLD. An independent global convergence result under the GNEP-QN CQ 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 ith component of $x \in \mathbb{R}^n$. We denote by $\mathbb{R}^n_+$ the set of vectors with nonnegative components and by $a_+ :=$

$\max\{0,a\}$ the nonnegative part of $a \in \mathbb{R}$. We use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product, and $\|\cdot\|$ denotes 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 $a_i, i = 1, \dots, m$, is positively linearly dependent if there are scalars $\lambda_i \geq 0, i = 1, \dots, m$, 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 CQs for optimization problems ensuring that a local minimizer satisfies the 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 and complexity analysis of many algorithms. The fact that the limit points of an algorithm satisfy 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 notation, consider only inequality constraints, defined by continuously differentiable functions $g_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \dots, m$:

$$\min_x f(x) \quad \text{s.t.} \quad g_i(x) \leq 0 \quad \forall i = 1, \dots, m. \tag{2}$$

Let $X \subset \mathbb{R}^n$ denote the feasible set of problem (2). Some 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) 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

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

It is well known that when $\bar{x} \in X$ is a local minimizer of (2), the following geometric necessary optimality condition holds:

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

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 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 $\bar{x} \in X$ read 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 (nonperturbed) 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 is a constraint qualification. We define a few of these as follows.

DEFINITION 2.2. *Let $\bar{x}$ be a feasible point of the nonlinear problem* (2). *Then we say that $\bar{x}$ satisfies the*

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

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

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

(d) *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) *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 for conditions (a)–(g) [6]:

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

Moreover, CPLD is strictly stronger than QN, and QN implies ACQ but 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 2.3. *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 $I \subset \{1, \ldots, m\}$ and scalars $\tilde{\lambda}_i \in \mathbb{R}$ $(i \in I)$ with $\lambda_i \tilde{\lambda}_i > 0$ for all $i \in I$, 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 CQ, in the sense that it guarantees that a local minimum of problem (2) 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 CQ, 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 arbitrarily small perturbations of $\bar{x}$. This gives rise to so-called sequential optimality conditions for constrained optimization problems (2), 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 CQ 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. In our analysis, a major role will be played by the most popular of these conditions, approximate-KKT (AKKT) [33, 3, 13].

DEFINITION 2.4 (AKKT). *We say that $\bar{x} \in X$ satisfies the AKKT condition (or is an AKKT point) if there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m_+$, 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, $\{x^k\}$ 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 2.5. *Let $\bar{x}$ be a local solution of problem (2); then $\bar{x}$ is an AKKT point.*

Proof. See [3, 15]. ∎

Theorem 2.5 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) that satisfies approximately the KKT conditions with tolerance ε. This fact is the main reason why CQs do not play a significant 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, then local minimizers that do not satisfy a CQ are ruled out completely from the analysis. Under this point of view, a CQ 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 CCP [6], the minimal CQ that ensures that an AKKT point is a KKT point, independently of the objective function. See the discussions 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 (as long as the limit point is feasible). This is 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 their 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. This is the standard well-known approach for defining optimality conditions for GNEPs. 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) and 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 $v = 1, \dots, N$.

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

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

Definition 3.1 (GNEP-KKT). *We say that $x \in X$ is a GNEP-KKT point for* (1) *when x^v is a KKT point for problem $P^v(x^{-v})$ for each $v = 1, \dots, N$; i.e., for each*

$v=1,\dots,N$, there are vectors $\lambda^v \in \mathbb{R}^{m_v}_+$ such that

$$\nabla_{x^v} f^v(x) + \sum_{i=1}^{m_v} \lambda_i^v \nabla_{x^v} g_i^v(x) = 0, \tag{3}$$

$$\min\{\lambda_i^v, -g_i^v(x)\} = 0 \quad \text{for each} \quad i = 1,\dots,m_v. \tag{4}$$

A useful definition is the set $A^v(x)$ of indices of active constraints for player $v=1,\dots,N$ at $x \in X$, that is, $A^v(x) := \{i \in \{1,\dots,m_v\} : g_i^v(x) = 0\}$. With this definition one may rewrite the GNEP-KKT condition as (3) 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 GNEP-KKT condition is not a first-order necessary optimality condition without some constraint qualification. Formally, a CQ for a GNEP is a property that when satisfied at a solution $\bar{x}$, ensures that the GNEP-KKT 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

GNEP-KKT or Not-CQ

is fulfilled at every solution of the GNEP (1).

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

THEOREM 3.2 (minimality of a Guignard-type CQ). *Let $\bar{x} \in X$. The following conditions are equivalent:*

(i) *The nonlinear programming problem of each player $v=1,\dots,N$, with fixed constraint set $X(\bar{x}^{-v})$, satisfies GCQ at $\bar{x}^v$.*
(ii) *For any objective functions $f^v, v=1,\dots,N$, in the definition of GNEP* (1) *such that $\bar{x}$ is a solution, the GNEP-KKT conditions hold.*

Proof. Assume (i), and let $f^v, v=1,\dots,N$, be objective functions in the definition of GNEP (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 GNEP-KKT, 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 contradicts (ii). ∎

From now on, we call the minimal condition ensuring GNEP-KKT given by Theorem 3.2 "partial-Guignard." More generally, it is a natural way to obtain a CQ for GNEPs by assuming that at the solution, the constraints of each player conform to a CQ for optimization. If "CQ" is a constraint qualification for optimization, we say that "partial-CQ" is the corresponding constraint qualification for GNEPs obtained in this fashion. That is, we have the following.

DEFINITION 3.3 (partial-CQ). *Given a feasible point $\bar{x} \in X$ for the GNEP* (1) *and a constraint qualification "CQ" for optimization, we say that $\bar{x}$ satisfies partial-CQ 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 3.4. It is easy to see that partial-CQ is a constraint qualification for GNEPs. Namely, if a solution $\bar{x}$ of the GNEP satisfies partial-CQ, then it satisfies the GNEP-KKT conditions. We note that in order to guarantee that an equilibrium solution satisfies the GNEP-KKT conditions, it is not necessary for the CQ for each player to be the same.

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

DEFINITION 3.5 (partial-MFCQ). *We say that* $\bar{x} \in X$ *satisfies partial-MFCQ with respect to the GNEP* (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 3.6 (partial-CPLD). *We say that* $\bar{x} \in X$ *satisfies partial-CPLD with respect to the GNEP* (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 3.7 (partial-CCP). *We say that* $\bar{x} \in X$ *satisfies partial-CCP with respect to the GNEP* (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 3.8. It is clear from the definition of partial-CQ that if CQ1 is weaker than CQ2, then partial-CQ1 is weaker than partial-CQ2.

In [32, 27] the partial-EMFCQ, which is an extension of partial-MFCQ 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 partial-CQ concept is not appropriate for proving global convergence results of algorithms to stationary points. That is why we call it "partial." In the next section we will show another way to extend an optimization CQ to the context of GNEPs, which will be adequate for proving global convergence results. The reason the partial-CQ 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 $v = 1, \ldots, N$, and hence it is not reasonable to fix $\bar{x}^{-v}$ when defining the neighborhood used in Definitions 3.6 and 3.7 for player v. Instead, one should 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, 7, 6]. Our goal is to provide a way of extending those CQs to GNEPs in such a way that global convergence of algorithms to a GNEP-KKT point can be proved. These definitions will highlight the joint structure of the GNEPs with respect to each player. Following the definition of GNEP-CPLD from [28] (which independently made such an extension of CPLD), given some constraint qualification "CQ" for optimization, we introduce a constraint qualification for GNEP that we call "GNEP-CQ."

Given $\bar{x} \in X$, the partial-CQ is a straightforward extension of the optimization CQ, obtained by looking individually at each player's 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 partial-CQ conditions are inherited directly from the corresponding relations among the optimization CQs, the relations among different GNEP-CQ conditions are not the same as the relations among the corresponding CQs for optimization. The main difference in the GNEP-CQs presented here, with respect to their partial-CQ 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 partial-CQ and GNEP-CQ 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 GNEP-CPLD concept, which was already independently considered in [28].

Definition 4.1 (GNEP-CPLD). *We say that $\bar{x} \in X$ satisfies GNEP-CPLD 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 partial-CPLD and GNEP-CPLD is that in partial-CPLD, one just has to consider x of the form $x = (x^v, \bar{x}^{-v})$ in the neighborhood of $\bar{x}$, for each v, and in GNEP-CPLD, 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 (GNEP-CCP). *We say that $\bar{x} \in X$ satisfies GNEP-CCP 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 GNEP-CCP 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 GNEP-CCP can be written as the outer semicontinuity of the single multifunction

$$K(x) := \prod_{v=1}^{N} K^v(x). \tag{5}$$

Now we will prove that GNEP-CPLD and GNEP-CCP are in fact CQs; i.e., if one of them is satisfied at a solution of the GNEP (1), then the GNEP-KKT conditions hold. To do this, it is sufficient to show that GNEP-CPLD and GNEP-CCP imply, respectively, partial-CPLD and partial-CCP, which are CQs for GNEPs as mentioned in Remark 3.4. In fact, let us prove the more general result that GNEP-CQ implies partial-CQ for a very broad class of CQs that includes all of the ones that we have presented.

Theorem 4.3. *Let CQ be a constraint qualification for optimization that is defined in terms of making a requirement for all points in a neighborhood of a point $\bar{x}$. Then GNEP-CQ implies partial-CQ.*

Proof. In order for a point $\bar{x} \in X$ to satisfy partial-CQ, for each player $v = 1, \dots, N$, there must be some restriction (defined by CQ) 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 $\bar{x}$. The corresponding GNEP-CQ is simply an extension of this requirement for general neighboring points x of $\bar{x}$. Hence GNEP-CQ is always stronger than its counterpart partial-CQ. ∎

In particular, GNEP-CQ 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 GNEP-CPLD as a CQ for GNEP, but they do not provide a proof. However, the fact that GNEP-CPLD and GNEP-CCP are CQs for GNEPs is a direct consequence of Theorem 4.3 and Remark 3.4.

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

THEOREM 4.4. *GNEP-CCP is weaker than GNEP-CPLD.*

Proof. Suppose that GNEP-CPLD occurs at $\bar{x} \in X$, and let $v \in \{1, \dots, N\}$ 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 2.3, 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

$$w^k = \sum_{i \in J_v} \tilde{\lambda}_i^{k,v} \nabla_{x^v} g_i^v(x^k), \quad \tilde{\lambda}_i^{k,v} \geq 0, \tag{6}$$

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 (6) by $\|\tilde{\lambda}^{k,v}\|$ and taking limit in k, we have that $0 = \sum_{i \in J_v} \bar{\lambda}_i^v \nabla_{x^v} g_i^v(\bar{x})$, and 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, by using the definition of GNEP-CPLD, we have that $\left\{\nabla_{x^v} g^v(x^k)\right\}_{i \in J_v}$ is positively linearly dependent. However, this is a contradiction to 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 (6), 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 GNEP-CCP condition holds. ∎

Since optimization problems are a special case of GNEPs, it is clear that GNEP-CCP is strictly weaker than GNEP-CPLD [6]. The following example shows that the implications in Theorem 4.3 are strict for CPLD and CCP.

Example 4.5 (partial-CPLD does not imply GNEP-CPLD, and partial-CCP does not imply GNEP-CCP). Consider a GNEP with $N = 2$, $n_1 = n_2 = m_1 = 1$, $m_2 = 0$, and $g^1(x^1, x^2) = x^1 x^2$. 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 partial-CPLD holds at $\bar{x}$. Partial-CCP, being weaker than partial-CPLD, 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 GNEP-CCP is not satisfied at $\bar{x}$. By Theorem 4.4, we also have that GNEP-CPLD does not hold at $\bar{x}$.

Finally, we present an extension of the QN CQ for GNEPs with the following new definition.

DEFINITION 4.6 (GNEP-QN). *We say that $\bar{x} \in X$ satisfies GNEP-QN if for any $v = 1, \ldots, N$, 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 whenever $\lambda_i^v > 0$ for some i.

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

Example 4.7 (GNEP-CPLD does not imply GNEP-QN). 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 GNEP-CPLD holds at every point of X. Let us show that GNEP-QN 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 pathological 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 GNEP-MFCQ (= partial-MFCQ) is strictly stronger than GNEP-QN.

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, 2] give different global convergence results, while also generalizing 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 GNEP-CCP (in Theorem 6.1) and under GNEP-QN (in Theorem 6.2). While the first result is stronger than the one under GNEP-CPLD from [28], the second is independent from [28] and also from Theorem 6.1. This will be shown 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 are AKKT points. A similar extension has been made 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 CQ, we further discuss some interesting practical implications of this fact.

5. Approximate-KKT conditions for GNEPs.

As we have discussed in Theorem 2.5, 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 implies the optimality condition "KKT or Not-CCP." In [28], the authors use a slight modification of what we call here GNEP-AKKT, with a similar purpose for GNEPs, by showing that an augmented Lagrangian generates sequences whose limit points satisfy GNEP-AKKT, and that it implies the optimality condition "GNEP-KKT or Not-GNEP-CPLD." We will show that GNEP-AKKT in fact implies the stronger optimality condition "GNEP-KKT or Not-GNEP-CCP," but that GNEP-AKKT itself is not an optimality condition. This result contrasts with what is known for optimization. In this section we also discuss some important practical implications of this fact, while also proving that the proposed concept GNEP-AKKT is in fact an optimality condition for an important class of problems, which includes NEPs.

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

$$\lim_{k\to\infty} \left\| \nabla_{x^v} f^v(x^k) + \sum_{i=1}^{m_v} \lambda_i^{k,v} \nabla_{x^v} g_i^v(x^k) \right\| = 0, \tag{7}$$

and

$$\lim_{k\to\infty} \min \left\{ \lambda_i^{k,v}, -g_i^v(x^k) \right\} = 0 \tag{8}$$

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

The sequence $\{x^k\}$ used in the definition of GNEP-AKKT is called a GNEP-AKKT sequence. This concept is the same as the one independently defined in [28], with the difference that they do not require the nonnegativity 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 nonnegative.

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

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

$$\left\| \nabla_{x^v} f^v(x^v, x^{-v}) + \sum_{i=1}^{m_v} \lambda_i^v \nabla_{x^v} g_i^v(x^v, x^{-v}) \right\| \leq \varepsilon, \tag{9}$$

$$\left\| \min \left\{ \lambda_i^v, -g_i^v(x^v, x^{-v}) \right\} \right\| \leq \varepsilon, \quad i = 1, \dots, m_v. \tag{10}$$

Note that for $\varepsilon = 0$, an ε-inexact GNEP-KKT point is a standard GNEP-KKT point. Moreover, a point $\bar{x}$ satisfies the GNEP-AKKT 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 GNEP-KKT point.

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

Example 5.3 (GNEP-AKKT 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 GNEP-AKKT 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 $\lambda^{k,2} z^k \to 1$. This clearly contradicts the fact that $\lambda^{k,1}$ and $\lambda^{k,2}$ are nonnegative, and therefore $\bar{x}$ is not a GNEP-AKKT point.

Example 5.3 shows that algorithms that provably generate sequences whose limit points are GNEP-AKKT points, can never converge to any solution of this jointly convex problem. In particular, either the iterand x^k is not defined, with a subproblem that cannot 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 18]. 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 and 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 GNEP-AKKT sequences can also be attracted by a point that does not satisfy a true optimality condition (hence is not a solution). To formalize this we need the following concept.

Definition 5.4 (partial-AKKT). *Let $\bar{x} \in X$ be a feasible point of the GNEP* (1). *We say that $\bar{x}$ satisfies the partial-AKKT condition if*

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

By the definition of the GNEP and by Theorem 2.5, it is easy to see that partial-AKKT 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 5.3 showed that partial-AKKT does not imply GNEP-AKKT. The next example shows that our two AKKT concepts are in fact independent. For this example, GNEP-AKKT holds, while partial-AKKT does not. Hence, given that partial-AKKT is an optimality condition, an algorithm can generate a GNEP-AKKT sequence that will necessarily fail to converge to a solution.

Example 5.5 (GNEP-AKKT does not imply partial-AKKT). 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 $\{x^k\} \subset \mathbb{R}^2$ be any sequence such that $x^k \to (-1, 0)$ and $\{\lambda^{k,1}\} \subset \mathbb{R}_+$. For k large enough we have $x^{k,1} + \lambda^{k,1} 0 \to -1,$ *and* therefore partial-AKKT does not occur in x.

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

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

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

THEOREM 5.6. *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 GNEP-AKKT condition holds at* $\bar{x}$.

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

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

By Theorem 2.5, there exist $\{x^{k,v}\} \subset \mathbb{R}^{n_v}$ and $\{\lambda^{k,v}\} \subset \mathbb{R}_+^{|A^v(\bar{x})|}$ such that $x^{k,v} \to \bar{x}^v$ and

$$\left\| \nabla_{x^v} f^v(x^{k,v}, \bar{x}^{-v}) + \sum_{i \in A^v(\bar{x})} \lambda_i^{k,v} g_i^{v,2}(\bar{x}^{-v}) \nabla_{x^v} g_i^{v,1}(x^{k,v}) \right\| \to 0. \tag{11}$$

Obviously, one could redefine, if necessary, $\lambda_i^{k,v} = 0$ if $g_i^{v,2}(\bar{x}^{-v}) = 0$, and (11) would still be valid.

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

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

Note that $x^k \to \bar{x}$ and that 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 (11) and the triangular inequality,

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

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

and so $\bar{x}$ is a GNEP-AKKT 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 cannot 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 $N = 2$, $n_1 = n_2 = m_1 = m_2 = 1$, $f^1(x^1) := x^1$, $f^2(x^2) := -x^2$, $g^1(x) := x^1x^2 + \frac{(x^1)^4}{4}$, and $g^2(x) := x^2\left(x^1\right)^3 + \frac{(x^2)^2}{2}$, then the origin is a solution of the GNEP, but the approximate multipliers $\lambda^{k,1}$ and $\lambda^{k,2}$ must have opposite signs, and hence GNEP-AKKT does not hold. Notice that this gives another example, similar to Example 5.3, where GNEP-AKKT does not hold at a solution.

Let us recall that jointly convex GNEPs constitute an important special class of GNEPs that 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

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

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 5.7. *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 5.3, we have a jointly convex GNEP; however, there is no solution of the associated $VI(X, F)$. When a solution of the GNEP is also a solution of the associated $VI(X, F)$, we say that it is a variational equilibrium.

DEFINITION 5.8. *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 the corresponding $VI(X, F)$ 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 equilibrium of a jointly convex GNEP, GNEP-AKKT holds.

THEOREM 5.9. *Any variational equilibrium of a jointly convex GNEP is a GNEP-AKKT point.*

Proof. Since $\bar{x}$ is a solution of $VI(X, F)$, we have that $\bar{x}$ is a solution of the following nonlinear programming problem:

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

By Theorem 2.5, there exist sequences $x^k \to \bar{x}$, $\{\lambda^k\} \subset \mathbb{R}_+^{|A(\bar{x})|}$ such that

$$F(\bar{x}) + \sum_{i \in A(\bar{x})} \lambda_i^k \nabla g_i(x^k) \to 0, \tag{12}$$

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 a GNEP-AKKT point. ∎

Even though variational equilibria 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 equilibria for the auxiliary problem, which are GNEP-AKKT points for the auxiliary GNEP, are trivially also GNEP-AKKT 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 GNEP-AKKT points.

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

THEOREM 5.10. *The GNEP-CCP condition is the weakest property under which GNEP-AKKT implies GNEP-KKT, independently of the objective functions f^v of each player $v = 1, \dots, N$.*

Proof. Let $\bar{x} \in X$ satisfy GNEP-CCP. Suppose that GNEP-AKKT 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 $v = 1, \dots, N$ 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 (5). Since $w^{k,v} \to -\nabla_{x^v} f^v(\bar{x})$, we have that $w^k \to -F(\bar{x}) := -\left(\nabla_{x^v} f^v(\bar{x})\right)_{v=1}^N$, and therefore $-F(\bar{x}) \in \limsup_{x \to \bar{x}} K(x)$. By the GNEP-CCP condition, we have that $-F(\bar{x}) \in K(\bar{x})$, and so $\bar{x}$ is a GNEP-KKT point.

Reciprocally, assume that $\bar{x}$ is such that GNEP-AKKT implies GNEP-KKT, 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 $v = 1, \dots, N$ sequences $\{\lambda^{k,v}\} \subset \mathbb{R}_+^{|A^v(\bar{x})|}$ such that

$$w^{k,v} = \sum_{i \in A^v(\bar{x})} \lambda_i^{k,v} \nabla_{x^v} g_i^v(x^k). \tag{13}$$

Define for each player $v = 1, \dots, N$ 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 (13),

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

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

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

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 assumption than in [28]. We discuss feasibility and optimality results under the GNEP-CCP condition; moreover, we give an alternative global convergence result using the QN condition for GNEPs (GNEP-QN) that includes boundedness of the Lagrange multipliers approximation.

Augmented Lagrangian methods are useful when there exist efficient algorithms for solving their 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 those in [27, 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 (GNEP-CPLD) than in [32], where GNEP-MFCQ was used.

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 that presented in [28]; however, our convergence analysis is done with a weaker assumption.

We consider the GNEP with upper- and lower-level constraints:

$$
(14)\qquad \begin{array}{llll} Q_v(x^{-v}): & \min_{x^v} & f^v(x^v, x^{-v}) & \\ & \text{s.t.} & g_i^v(x^v, x^{-v}) \le 0 & i = 1, \ldots, m_v, \\ & & h_j^v(x^v, x^{-v}) \le 0 & j = 1, \ldots, l_v. \end{array}
$$

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

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

for all x such that $h^v(x) \le 0$ and for each $v = 1, \ldots, N$. 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 $L_{\rho^v}^v(x, u^v)$ 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, which we call GNEP(u, ρ), where the problem for player v is given by

$$\begin{array}{ll} \min_{x^v} & L^v_{\rho^v}(x, u^v), \\ & h^v_j(x^v, x^{-v}) \leq 0 \quad j = 1, \ldots, l_v. \end{array} \tag{16}$$

The algorithm presented here aims at finding an ε-inexact GNEP-KKT point for (14), in the sense of Definition 5.2, through the calculation of ε_k-inexact GNEP-KKT 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 and is essentially Algorithm 11 of [28].

Algorithm 1 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 positive sequence $\varepsilon_k \to 0^+$, and set $k = 0$.

Step 1: If (x^k, λ^k, μ^k) is an ε-inexact GNEP-KKT point of the original GNEP (14): STOP.

Step 2: Compute an ε_{k+1}-inexact GNEP-KKT 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 $v = 1, \ldots, N$, 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 presents the main global convergence result. To present it, we define the *game of infeasibility* such as in [28]:

$$\min_{x^v} \left\|g^v_+(x)\right\|^2 \text{ s.t. } h^v_j(x^v, x^{-v}) \leq 0, \quad j = 1, \ldots, l_v. \tag{17}$$

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 $\lim_{k\in K} x^k = \bar{x}$. Then $\bar{x}$ satisfies GNEP-AKKT for the game of infeasibility* (17). *In addition, if $\bar{x}$ satisfies GNEP-CCP for the game of infeasibility, then $\bar{x}$ is a GNEP-KKT point of the game of infeasibility.*

In the case when the limit point $\bar{x}$ is feasible, it satisfies the GNEP-AKKT condition for GNEP (14), *and if it also fulfills the GNEP-CCP constraint qualification for* (14), *then the limit point $\bar{x}$ is a GNEP-KKT point of the GNEP* (14).

Proof. We omit the proof of the theorem because it is a direct adaptation of Theorem 6.2 of [15]. The fact that GNEP-AKKT sequences are generated is also given by Lemma 13 and Theorem 18 of [28]. The results under GNEP-CCP are a consequence of Theorem 5.10. ∎

We also point out that in [28] two results guaranteeing feasibility of the limit points of $\{x^k\}$ are presented, one of them relying on convexity of the problem and the other relying on EMFCQ, an extension of MFCQ for infeasible points.

To finish our contributions, we next prove the convergence of Algorithm AL-GNEP under the GNEP-QN constraint qualification when the limit point is feasible. This result is not directly associated with the concept of GNEP-AKKT, as GNEP-CCP is the minimal condition ensuring that a GNEP-AKKT point is a GNEP-KKT point, but GNEP-CCP is independent of GNEP-QN. This property seems to be inherent to the augmented Lagrangian method because its iterates are usually infeasible. A similar result in the optimization context has appeared recently in [2]. Under GNEP-QN and the feasibility of the limit point, 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 when the limit point is feasible. Since GNEP-CPLD does not imply GNEP-QN, this result does not hold for GNEPs under GNEP-CPLD. Note that the result holds, in particular, when the lower-level constraints are empty, namely, when all constraints of problem (14) 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 GNEP-QN, 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 iteration. We also point out that this assumption was also present in the general lower-level augmented Lagrangian from [11]. See also [14].

THEOREM 6.2. *Assume that the sequence* $\{x^k\}$ *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* $k \in K$. *Then, if* $\bar{x}$ *fulfills the GNEP-QN condition with respect to* (14), *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 GNEP-KKT point for problem* (14).

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

$$\left\| \nabla_{x^v} f^v(x^k) + \sum_{i=1}^{m_v} \lambda_i^{k,v} \nabla_{x^v} g_i^v(x^k) + \sum_{i=1}^{l_v} \mu_i^{k,v} \nabla_{x^v} h_i^v(x^k) \right\| \leq \varepsilon_k, \tag{18}$$

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, in order 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 $j = 1, \dots, l_v$, while $\bar{\lambda}_i^v \geq 0$ for all $j = 1, \dots, l_v$, with $\bar{\lambda}^v \neq 0$.

Dividing the expression (18) by $\delta^{k,v}$ and taking limits on K_2, we get

$$\sum_{i=1}^{m_v} \bar{\lambda}_i^v \nabla_{x^v} g_i^v(\bar{x}) = 0. \tag{19}$$

Moreover, since the GNEP-AKKT condition yields $\lambda_i^{k,v} \to 0$ if $g_i^v(\bar{x}) < 0$, (19) gives a nontrivial 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 GNEP-QN 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 GNEP-AKKT 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 jointly convex problems. In addition, we show that the GNEP-AKKT concept is independent of the partial-AKKT 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 GNEP-AKKT sequences. Since it is natural to construct methods that generate GNEP-AKKT sequences (and some of them have recently been published), we believe it is important to draw attention to this fact.

For specific classes of GNEPs, we have proved that GNEP-AKKT is an optimality condition, and for these cases, it is more reasonable to use methods that generate GNEP-AKKT sequences. In this paper we formalize and establish relations among several CQs for GNEPs. In particular, we have defined the GNEP-CCP condition and proved that it is the weakest condition that ensures that a GNEP-AKKT point satisfies the GNEP-KKT 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 GNEP-QN condition, which we surprisingly proved to be independent of the GNEP-CPLD, 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 GNEP-AKKT is not an optimality condition is that it may not be appropriate to use as an algorithm's stopping criterion the fact 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 GNEP-KKT 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 satisfy its KKT conditions with precision ε_k. This raises questions on conditions for the algorithm to be well defined and which constraints should be chosen 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, 8] the authors show that Newton's method can generate sequences that are not AKKT sequences. 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 CQ 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 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 GNEP-AKKT points would be very welcome 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*, Math. Program., 111 (2008), pp. 5–32, https://doi.org/10.1007/s10107-006-0077-1.

[2] R. Andreani, N. S. Fazzio, M. L. Schuverdt, and L. D. Secchin, *A Sequential Optimality Condition Related to the Quasinormality Constraint Qualification and Its Algorithmic Consequences*, Tech. report, Optimization Online, 2017, http://www.optimization-online.org/DB_HTML/2017/09/6194.html.

[3] R. Andreani, G. Haeser, and J. M. Martínez, *On sequential optimality conditions for smooth constrained optimization*, Optimization, 60 (2011), pp. 627–641, https://doi.org/10.1080/02331930903578700.

[4] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva, *A relaxed constant positive linear dependence constraint qualification and applications*, Math. Program., 135 (2012), pp. 255–273, https://doi.org/10.1007/s10107-011-0456-0.

[5] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva, *Two new weak constraint qualifications and applications*, SIAM J. Optim., 22 (2012), pp. 1109–1135, https://doi.org/10.1137/110843939.

[6] R. Andreani, J. M. Martínez, A. Ramos, and P. J. S. Silva, *A cone-continuity constraint qualification and algorithmic consequences*, SIAM J. Optim., 26 (2016), pp. 96–110, https://doi.org/10.1137/15M1008488.

[7] 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 (2018), pp. 693–717, https://doi.org/10.1287/moor.2017.0879.

[8] R. Andreani, J. M. Martínez, and L. T. Santos, *Newton's method may fail to recognize proximity to optimal points in constrained optimization*, Math. Program., 160 (2016), pp. 547–555, https://doi.org/10.1007/s10107-016-0994-6.

[9] R. Andreani, J. M. Martínez, L. T. Santos, and B. F. Svaiter, *On the behaviour of constrained optimization methods when Lagrange multipliers do not exist*, Optim. Methods Softw., 29 (2014), pp. 646–657, https://doi.org/10.1080/10556788.2013.841692.

[10] R. Andreani, J. M. Martinez, and M. L. Schuverdt, *On the relation between constant positive linear dependence condition and quasinormality constraint qualification*, J. Optim. Theory Appl., 125 (2005), pp. 473–483, https://doi.org/10.1007/s10957-004-1861-9.

[11] R. Andreani, J. M. Martínez, and B. F. Svaiter, *A new sequential optimality condition for constrained optimization and algorithmic consequences*, SIAM J. Optim., 20 (2010), pp. 3533–3554, https://doi.org/10.1137/090777189.

[12] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Practical Methods of Optimization: Theory and Algorithms*, John Wiley and Sons, 2006.

[13] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1999.

[14] E. G. Birgin, G. Haeser, and A. Ramos, *Augmented Lagrangians with constrained subproblems and convergence to second-order stationary points*, Comput. Optim. Appl., 69 (2018), pp. 51–75, https://doi.org/10.1007/s10589-017-9937-2.

[15] E. G. Birgin and J. M. Martínez, *Practical Augmented Lagrangian Methods for Constrained Optimization*, Fundam. Algorithms 10, SIAM, 2014, https://doi.org/10.1137/1.9781611973365.

[16] A. Dreves, C. Kanzow, and O. Stein, *Nonsmooth optimization reformulations of player convex generalized Nash equilibrium problems*, J. Global Optim., 53 (2012), pp. 587–614, https://doi.org/10.1007/s10898-011-9727-9.

[17] A. Dreves and N. Sudermann-Merx, *Solving linear generalized Nash equilibrium problems numerically*, Optim. Methods Softw., 31 (2016), pp. 1036–1063, https://doi.org/10.1080/10556788.2016.1165676.

[18] F. Facchinei, A. Fischer, and V. Piccialli, *On generalized Nash games and variational inequalities*, Oper. Res. Lett., 35 (2007), pp. 159–164, https://doi.org/10.1016/j.orl.2006.03.004.

[19] F. Facchinei and C. Kanzow, *Generalized Nash equilibrium problems*, Ann. Oper. Res., 175 (2010), pp. 177–211, https://doi.org/10.1007/s10479-009-0653-x.

[20] F. Facchinei and C. Kanzow, *Penalty methods for the solution of generalized Nash equilibrium problems*, SIAM J. Optim., 20 (2010), pp. 2228–2253, https://doi.org/10.1137/090749499.

[21] F. Facchinei and L. Lampariello, *Partial penalization for the solution of generalized Nash equilibrium problems*, J. Global Optim., 50 (2011), pp. 39–57, https://doi.org/10.1007/s10898-010-9579-8.

[22] A. Fischer, M. Herrich, and K. Schönefeld, *Generalized Nash equilibrium problems—recent advances and challenges*, Pesq. Oper., 34 (2014), pp. 521–558, https://doi.org/10.1590/0101-7438.2014.034.03.0521.

[23] M. Fukushima, *Restricted generalized Nash equilibria and controlled penalty algorithm*, Comput. Manag. Sci., 8 (2011), pp. 201–218, https://doi.org/10.1007/s10287-009-0097-4.

[24] F. J. Gould and J. W. Tolle, *A necessary and sufficient qualification for constrained optimization*, SIAM J. Appl. Math., 20 (1971), pp. 164–172, https://doi.org/10.1137/0120021.

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

[26] R. Janin, *Directional derivative of the marginal function in nonlinear programming*, in Sensitivity, Stability and Parametric Analysis, Math. Programming Stud. 21, A. V. Fiacco, ed., North–Holland, 1984, pp. 110–126, https://doi.org/10.1007/BFb0121214.

[27] C. Kanzow, *On the multiplier-penalty-approach for quasi-variational inequalities*, Math. Program., 160 (2016), pp. 33–63, https://doi.org/10.1007/s10107-015-0973-3.

[28] C. Kanzow and D. Steck, *Augmented Lagrangian methods for the solution of generalized Nash equilibrium problems*, SIAM J. Optim., 26 (2016), pp. 2034–2058, https://doi.org/10.1137/16M1068256.

[29] O. L. Mangasarian and S. Fromovitz, *The Fritz-John necessary conditions in the presence of equality and inequality constraints*, J. Math. Anal. Appl., 17 (1967), pp. 37–47, https://doi.org/10.1016/0022-247X(67)90163-1.

[30] L. C. Matioli, W. Sosa, and J. Yuan, *A numerical algorithm for finding solutions of a generalized Nash equilibrium problem*, Comput. Optim. Appl., 52 (2012), pp. 281–292, https://doi.org/10.1007/s10589-011-9407-1.

[31] L. Minchenko and S. Stakhovski, *On relaxed constant rank regularity condition in mathematical programming*, Optimization, 60 (2011), pp. 429–440, https://doi.org/10.1080/02331930902971377.

[32] J.-S. Pang and M. Fukushima, *Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games*, Comput. Manag. Sci., 2 (2005), pp. 21–56, https://doi.org/10.1007/s10287-004-0010-0.

[33] L. Qi and Z. Wei, *On the constant positive linear dependence condition and its application to SQP methods*, SIAM J. Optim., 10 (2000), pp. 963–981, https://doi.org/10.1137/S1052623497326629.

[34] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer-Verlag, 1998.

[35] F. N. Rojas, *Augmented Lagrangian Methods for Generalized Nash Equilibrium Problems*, qualifying examination for Ph.D. candidacy, Department of Applied Mathematics, University of São Paulo, 2016.

[36] A. von Heusinger and C. Kanzow, *Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions*, Comput. Optim. Appl., 43 (2009), pp. 353–377, https://doi.org/10.1007/s10589-007-9145-6.