On Stationary Conditions and the Convergence of Augmented Lagrangian methods for Generalized Nash Equilibrium Problems

Lennin Mallma Ramirez * Frank Navarro Rojas † Alberto Ramos ‡

February 22, 2026

Abstract

In this work, we study stationarity conditions and constraint qualifications (CQs) tailored to Generalized Nash Equilibrium Problems (GNEPs) and analyze their relationships and implications for the global convergence of algorithms. We recall that GNEPs generalize Nash Equilibrium Problems (NEPs) in that the feasible strategy set of each player depends on the strategies chosen by the other players, thereby introducing additional difficulties in both theoretical analysis and algorithm design.

Our stationary concepts provide a theoretical framework for analyzing the global convergence of several numerical methods. In particular, we establish new convergence results for the safeguarded augmented Lagrangian method and propose a new adaptation of the Hyperbolic Augmented Lagrangian Algorithm (HALA) tailored to GNEPs. More specifically, we investigate the convergence properties of these methods with respect to both feasibility and optimality of the limit points. Furthermore, we prove global convergence to a Karush–Kuhn–Tucker (KKT) point under a weak CQ and establish boundedness of the associated multipliers under a strong quasinormality-type CQ adapted to GNEPs. Numerical experiments are presented to demonstrate the effectiveness and robustness of the proposed methods.

Keywords: Stationary conditions; convergence; Hyperbolic augmented Lagrangian; generalized Nash equilibrium; numerical experiments.

1 Introduction

Generalized Nash equilibrium problems (GNEP) arise in settings where multiple agents make decisions simultaneously. Unlike classical Nash equilibria, in GNEPs the feasible set of each player depends on the strategies chosen by the others, leading to coupling effects that complicate both the theoretical analysis and the design of numerical methods.

In this work, we introduce new stationarity conditions and constraint qualifications (CQs) for (GNEP). These conditions provide a framework for analyzing global convergence of algorithms and for ensuring boundedness of the associated Lagrange multipliers.

*Federal University of Rio de Janeiro & Systems Engineering and Computer Science Program (COPPE), Rio de Janeiro, Brazil. Email: lenninmr@cos.ufrj.br

†Universidad Nacional Mayor de San Marcos & Escuela de Matemáticas, Lima, Perú. Email: fnavarror@unmsm.edu.pe

‡Department of Mathematics, University of Tarapacá, Arica, Chile. Email: aramosf@academicos.uta.cl

Here, we consider (GNEP) that consists of N players, where each player $\nu = \{1, \ldots, N\}$ tries to solve the following optimization problem

$$P^\nu(x^{-\nu}): \quad \text{minimize}_{\,x^\nu}\ f^\nu(x) \quad \text{subject to} \quad g_i^\nu(x) \leq 0,\ \forall i = 1, \ldots, m_\nu \ \text{ and } \ x^\nu \in C_\nu \quad \text{(GNEP)}$$

by minimizing over his/her own variables $x^\nu \in \mathbb{R}^{n_\nu}$, given the choices of the remaining players, denoted by $x^{-\nu}$. The variable x^ν denotes the variable of the player ν, and we subsume the remaining variables into the subvector $x^{-\nu}$. We also write $x = (x^\nu, x^{-\nu})$ to indicate the importance of x^ν with respect to the total vector x. Here, $f^\nu : \mathbb{R}^n \to \mathbb{R}$ denotes the objective function of the player ν, the functions $g_i^\nu : \mathbb{R}^n \to \mathbb{R}$, $\forall i = 1, \ldots, m_\nu$ define the constraints, and $C_\nu \subset \mathbb{R}^{n_\nu}$ is an abstract set that we assume to be closed. Formally, $n = n_1 + \cdots + n_N$ is the total number of variables, and $m = m_1 + \cdots + m_N$ the total number of constraints.

We say that $\bar{x} = (\bar{x}^1, \ldots, \bar{x}^N)$ is a *generalized* Nash equilibrium (or solution of (GNEP)) if $\bar{x}$ satisfies all constraints and for each $\nu = 1, \ldots, N$, we have

$$f^\nu(\bar{x}) \leq f^\nu(x^\nu, \bar{x}^{-\nu}) \quad \forall\, x^\nu \in C_\nu \text{ such that } \ g_i^\nu(x^\nu, \bar{x}^{-\nu}) \leq 0,\ \forall i = 1, \ldots, m_\nu. \tag{1}$$

In other words, a point $\bar{x}$ is a solution if and only if no player ν can improve his or her objective value by unilaterally changing his or her strategy.

For notational convenience, equality constraints are not explicitly included in (GNEP). However, our framework can be readily extended to accommodate both equality and inequality constraints. Finally, our setting is very general in the sense that no convexity assumptions are imposed on the mappings f^ν, g_i^ν (for all i and ν), or on the sets C_ν (for all ν).

The contributions of this paper revolve around the following results:

1. We propose approximate stationarity conditions for (GNEP) that extend well-established approximate stationarity concepts from nonlinear programming (NLP) introduced in the literature, and we analyze the relationships among these conditions;

2. Under the above concepts, we extend and refine the convergence theory of the safeguarded augmented Lagrangian method for (GNEP) developed in [19]; see Theorems 7.1 and 7.2. In particular, we establish convergence to a KKT point under very weak constraint qualifications;

3. We introduce a new augmented Lagrangian method for solving (GNEP) based on a hyperbolic penalty function, which has shown promising computational performance for nonlinear programming; see [13]. We refer to this algorithm as the Hyperbolic Augmented Lagrangian (HALA) method for (GNEP). We establish convergence of the HALA method with respect to both feasibility and optimality of limit points. In the HALA method for (GNEP), the multiplier update is multiplicative in nature, in contrast to classical augmented Lagrangian methods, where the update is additive. Consequently, our convergence analysis does not follow as a direct consequence of existing augmented Lagrangian frameworks;

4. We present a comprehensive numerical study illustrating the performance of the HALA method.

This paper is organized as follows. In Section 2, we state our standard notation in optimization and variational analysis. The Section 3 deals with GNEP-tailored stationary concepts that can be used to improve the global convergence analysis of augmented Lagrangian-type algorithms, we prove some basic results and their relations. Section 4 presents a new GNEP-tailored CQ based on the recently introduced approximate stationary concepts. We also see their relations with the classic GNEP-CQs. We present a precise statement of our algorithm based on an augmented Lagrangian method that uses the hyperbolic penalty function (pre-Huber function) in its formulation. Section 6 is then dedicated to a thorough convergence analysis of the HALA method. We point out that our convergence analysis is based on the approximate stationary concepts introduced in Section 3. In particular, we introduce a new GNEP-CQ, called GNEP-strong quasinormality (see Definition 6.1), which can be used to prove the boundedness of the dual sequence of the HALA method. Furthermore, in Section 7 we extend the global convergence theory of the augmented Lagrangian method, studied in [19]. Finally, Section 8 presents some numerical results of our method. The final remarks and conclusions are given in Section 9.

Notation. Given a function $f = f(x)$ of suitable dimension, we denote by ∇f the transpose Jacobian of f. If x^ν is a given subvector of x, then $\nabla_{x^\nu} f$ denotes the submatrix of ∇f that corresponds to the components x^ν. We use $\|\cdot\|$ and $\|\cdot\|_\infty$ to denote the Euclidean and infinity norms, respectively. $\mathbb{R}_+ = \{t \in \mathbb{R} \mid t \geq 0\}$, v_i is the i th component of the vector v. In particular, $\lim_k y^k = \lim_{k\in\mathbb{N}} y^k$. If $\{\gamma_k\} \subset \mathbb{R}, \gamma_k > 0$, and $\gamma_k \to 0$, we write $\gamma \downarrow 0$. The *sign function* $\operatorname{sgn} a$ for $a \in \mathbb{R} \setminus \{0\}$ is defined as $\operatorname{sgn} a = 1$ if $a > 0$ and $\operatorname{sgn} a = -1$ if $a < 0$. We have $\operatorname{sgn}(a \cdot b) = \operatorname{sgn} a \cdot \operatorname{sgn} b$.

2 Basic results and preliminaries

We use standard notation in optimization and variational analysis; see [23, 20].

We use upper letters, as $\mathbb{E}$, to denote real Euclidean finite-dimensional spaces. We denote by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ the norm and inner product on $\mathbb{E}$. For a non-empty subset $S \subset \mathbb{R}$, the distance function associated with S under the corresponding norm will be denoted by $\operatorname{dist}(\cdot, S) : \mathbb{E} \to \mathbb{R}$. If $Q \subset \mathbb{E}$ is a closed convex set and $u \in \mathbb{E}$, then $\operatorname{proj}_Q(u)$ is the projection of u onto Q, that is, $\|u - \operatorname{proj}_Q(u)\| \leq \|u - \hat{u}\|$, $\forall \hat{u} \in Q$. The closed ball of center a and radius $r > 0$ is given by $\mathbb{B}_r(a)$.

For a given set $A \subset \mathbb{E}$, the interior of A is denoted by int A, and its closure by $\overline{A}$. Given $A \subset \mathbb{E}$, the *polar cone* of A is defined as

$$A^\circ := \{w \in \mathbb{E} : \langle w, a\rangle \leq 0, \forall a \in A\}.$$

Observe that A° is a closed convex cone. We denote the annihilator of A, as $A^\perp := A^\circ \cap (-A)^\circ$. If A is a linear subspace, $A^\perp = A^\circ$.

For a closed set $Q \subset \mathbb{E}$ and $\bar{x} \in Q$, the tangent cone *(Bouligand)* to Q at $\bar{x}$ is defined as the set

$$T_Q(\bar{x}) := \{d \in X : \exists\{d^k\} \subset X, \exists\{t_k\} \subset \mathbb{R}_+, d^k \to d;\ t_k \to 0;\ \bar{x} + t_k d^k \in Q, \forall k\}.$$

If Q is a closed convex set and $\bar{x} \in Q$, then $T_Q(\bar{x}) = \overline{\mathbb{R}_+(Q - \bar{x})}$.

The *regular normal cone* to Q at $\bar{x} \in Q$ is the set

$$\widehat{N}_Q(\bar{x}) := \{w \in \mathbb{E} : \langle w, d\rangle \leq 0,\ \text{for every } d \in T_Q(\bar{x})\}, \tag{2}$$

and the *limiting normal cone* to Q at $\bar{x} \in Q$ is

$$N_Q(\bar{x}) := \{w \in \mathbb{E} : \exists\{w^k\} \subset \mathbb{E}, \exists\{x^k\} \subset Q, w^k \to w;\ x^k \to \bar{x};\ w^k \in \widehat{N}_Q(x^k), \forall k\}.$$

We note that $N_Q(\bar{x}) = \widehat{N}_Q(\bar{x})$ if Q is a closed convex set. Additionally, if Q is a convex cone, then $N_Q(\bar{x}) = Q^\circ \cap \{\bar{x}\}^\perp$, for every $\bar{x} \in Q$.

Finally, for a given set-valued mapping $M : \mathbb{E} \times \mathbb{Y} \rightrightarrows \mathbb{E}$, the Painlevé–Kuratowski upper limit of $M(x,z)$ at $(\bar{x},\bar{z})$ is given by

$$\limsup_{x\to\bar{x},z\to\bar{z}} M(y,z) := \left\{ v \in \mathbb{E} : \begin{array}{l} \exists\{x^k\} \subset \mathbb{E}, \{z^k\} \subset \mathbb{Y}, \exists\{v^k\} \subset \mathbb{E} \\ x^k \to \bar{x}, z^k \to \bar{z}, v^k \to v, \\ v^k \in M(x^k,z^k),\ \forall k \in \mathbb{N} \end{array} \right\}.$$

3 New Stationary Concepts for GNEPs

In this section, we discuss the appropriate notions of stationary for (GNEP), and in the next section we present new GNEP-tailored constraint qualifications (CQs) that can be useful in the convergence analysis of several augmented Lagrangian methods.

We first recall the definition of the KKT conditions for (GNEP).

Definition 3.1. *Let $\bar{x}$ be a feasible point for* (GNEP). *We say that $\bar{x}$ is a Karush-Kuhn-Tucker (KKT) point if there exist vectors of multipliers $\lambda^\nu \in \mathbb{R}^{m_\nu}_+$, $\forall \nu = 1,\dots,N$ such that*

$$0 \in \nabla_{x^\nu} f^\nu(\bar{x}) + \sum_{i=1}^{m_\nu} \lambda_i^\nu \nabla_{x^\nu} g_i^\nu(\bar{x}) + N_{C_\nu}(\bar{x}^\nu) \quad \textit{and} \quad \min\{-g_i^\nu(\bar{x}), \lambda_i^\nu\} = 0, \forall i = 1,\dots,m_\nu \tag{3}$$

for every $\nu = 1,\dots,N$.

As (GNEP) is a generalization of classical optimization problems, it is known that KKT conditions are not a first-order necessary optimality condition without the fulfillment of some condition over the feasible set, usually called Constraint Qualification (CQ). Recently, to analyze the global convergence of some algorithms, the concept of sequential optimality conditions has been emerged. Sequential optimality conditions can be considered as asymptotic versions of KKT, usually related to the stopping criteria of algorithms, and they differ essentially from each other by how the complementary slackness condition is described. An important question is whether it is possible to extend these concepts to a broad framework that includes (GNEP). In this line of research, we mention [14] where the authors extended the concept of approximate KKT (AKKT) condition to (GNEP) problems. Here, we complement this work by extending other stationary concepts found in the literature to (GNEP). We continue by presenting new sequential stationary conditions for (GNEP) and analyzing their relations.

Definition 3.2. *Let $\bar{x}$ be a feasible point for* (GNEP). *Suppose that there are sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^{\nu,k}\} \subset \mathbb{R}^{m_\nu}_+$, $\{e^{\nu,k}\} \subset \mathbb{R}^n$ for each $\nu = 1,\dots,N$ such that $x^k \to \bar{x}$, $e^{\nu,k} \to 0$ as well as*

$$0 \in \nabla_{x^\nu} f^\nu(x^k) + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^\nu(x^k) + e^{\nu,k} + N_{C_\nu}(x^{\nu,k}),\ \forall k \tag{4}$$

for every $\nu = 1,\dots,N$. Furthermore, we say that $\{x^k\}$ is

1. *an Approximate KKT (AKKT-GNEP) sequence if, additionally to* (4), *we have*

$$\min\{-g_i^\nu(x^k), \lambda_i^{\nu,k}\} \to 0, \forall i = 1,\dots,m_\nu,\ \forall \nu = 1,\dots,N. \tag{5}$$

   In this case, the limit $\bar{x}$ is an AKKT-GNEP point;

2. *an Approximate gradient projection (AGP-GNEP) sequence if, additionally to* (4), *we have*

$$\lambda_i^{\nu,k} \min\{g_i^\nu(x^k), 0\} \to 0, \forall i = 1, \dots, m_\nu, \ \forall \nu = 1, \dots, N. \tag{6}$$

   In this case, the limit $\bar{x}$ *is an AGP-GNEP point;*

3. *a Complementary Approximate KKT (CAKKT-GNEP) sequence if, additionally to* (4), *we have*

$$\lambda_i^{\nu,k} g_i^\nu(x^k) \to 0, \forall i = 1, \dots, m_\nu, \ \forall \nu = 1, \dots, N. \tag{7}$$

   In this case, $\bar{x}$ *is a CAKKT-GNEP point;*

4. *a Positive Approximate KKT (PAKKT-GNEP) sequence if, additionally to* (4), *condition* (5) *holds and*

$$\lambda_i^{\nu,k} g_i^\nu(x^k) > 0 \text{ if } \lim_k \frac{\lambda_i^{\nu,k}}{\delta_k^\nu} > 0, \tag{8}$$

   where $\delta_k^\nu := \|(1, \lambda^{\nu,k})\|_\infty \to \infty$. *In this case,* $\bar{x}$ *is a PAKKT-GNEP point;*

5. *a Positive Complementary Approximate KKT (PCAKKT-GNEP) sequence if, additionally to* (4), *conditions* (8) *and* (7) *hold. In this case,* $\bar{x}$ *is a PCAKKT-GNEP point.*

As we observe the main difference between these concepts are related on how they treat the complementarity condition. Clearly, condition (7) implies (6), and the latter implies (5). Thus, CAKKT-GNEP implies AGP-GNEP, and AGP-GNEP implies AKKT-GNEP. By definition, PAKKT-GNEP implies AKKT-GNEP. As such conditions are reduced to the classical sequential optimality conditions for $N = 1$, we affirm that AKKT-GNEP does not imply PAKKT-GNEP, and that CAKKT-GNEP and PAKKT-GNEP are independent of each other, see [2]. Certainly, the most stringent condition is the PCAKKT-GNEP. All these implications are strict, and they are illustrated in Figure 3.

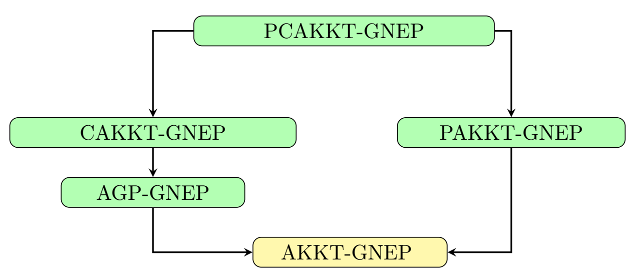

Figure 1: Relation between the sequential conditions for (GNEP). The arrows mean implications. AKKT-GNEP is the least stringent, while the more stringent condition is PCAKKT-GNEP. The conditions PAKKT-GNEP and CAKKT-GNEP are independent of each other.

Our stationary concepts are closely related to the KKT conditions for (GNEP). In fact, if the approximate multipliers $\{\lambda^{\nu,k} : \nu = 1, \dots, N\}$ are bounded sequences, it is not difficult to see, after taking an adequate subsequence of $\{\lambda^{\nu,k} : \nu = 1, \dots, N\}$, that $\bar{x}$ is a KKT point for (GNEP) and the corresponding multipliers are given by the limit point of the approximate multipliers. In particular, it holds when the GNEP-Mangasarian Fromovitz CQ (GNEP-MFCQ) is fulfilled at $\bar{x}$, see [19, Theorem 7]. We state this observation as a proposition.

Proposition 3.1. *If the sequence of multipliers $\{\lambda^{\nu,k}\}$ are bounded for all ν then $\bar{x}$ is a KKT point for* (GNEP)*. In this case, any limit point of $\{\lambda^{\nu,k} : \nu = 1, \dots, N\}$ is a Lagrange multiplier for $\bar{x}$.*

On the other hand, every KKT point is a PCAKKT-GNEP point. The proof follows from [2, Lemma 2.6]. In fact, suppose that $\bar{x}$ is a KKT point for (GNEP), where λ^ν is the corresponding multiplier. Following the proof [2, Lemma 2.6], it is possible to find a PAKKT-GNEP sequence $\{x^k\}$ with $x^{\nu,k} \in C_\nu$, $\forall \nu, \forall k$ having as approximate multipliers the vectors $\{\lambda^{\nu,k} := \lambda^\nu\}$. Certainly, $\{x^k\}$ is also a CAKKT-GNEP sequence, since $\lambda_i^{\nu,k} g_i^\nu(x^k) = \lambda_i^\nu g_i^\nu(x^k) \to \lambda_i^\nu g_i^\nu(\bar{x}) = 0$, $\forall i$, $\forall \nu$, since $\bar{x}$ is a KKT point. Therefore, we have the following result.

Proposition 3.2. *Every KKT point for* (GNEP) *is a PCAKKT-GNEP point.*

From [14, Example 5.1] it is known that AKKT-GNEP is not an optimality condition for (GNEP) with general inequality and equality constraints. As a consequence, our stationarity notions do not necessarily constitute optimality conditions. Therefore, a natural question is under what assumptions these stationarity notions guarantee optimality. We show that this is indeed the case for a well-behaved class of functions defining the constrained system. We have the following statement.

Theorem 3.1. *Suppose that for every $\nu = 1, \dots, N$, the constraint functions have the form:*

$$g_i^\nu(x) := a_i^\nu(x^{-\nu}) b_i^\nu(x^\nu) + c_i^\nu(x^{-\nu}), \quad \forall i = 1, \dots, m_\nu; \tag{9}$$

where all the above functions are C^1. Then, the following statements hold

(a) AKKT-GNEP is a necessary optimality condition;

(b) if additionally, $c_i^\nu(x^{-\nu}) = M_i^\nu a_i^\nu(x^{-\nu})$, $\forall i$ with $M_i^\nu \in \mathbb{R}$, then PCAKKT-GNEP is a necessary optimality condition.

Proof. The proof of item (a) follows from [14]. Now, we focus on item (b). Let $\bar{x}$ be a solution of (GNEP) where the constrained system is defined by functions of the form (9).

By definition of the solution of (GNEP), the component $\bar{x}^\nu$ is an optimal solution of the optimization problem:

$$\underset{x^\nu \in C_\nu}{\text{minimize}} \; f^\nu(x^\nu, \bar{x}^{-\nu}) \;\; \text{subject to} \;\; a_i^\nu(\bar{x}^{-\nu}) b_i^\nu(x^\nu) + c_i^\nu(\bar{x}^{-\nu}) \leq 0, \; \forall i = 1, \dots, m_\nu. \tag{10}$$

By [6, Theorem 2.4], we known that PCAKKT is an optimality condition, thus the component $\bar{x}^\nu$ is a PCAKKT point for the above optimization problem. Therefore, we can find sequences $\{\lambda^{\nu,k}\} \subset \mathbb{R}_+^{m_\nu}$, $\{x^{\nu,k}\}$ and $e^{\nu,k}$ with $x^{\nu,k} \to \bar{x}^\nu$ and $e^{\nu,k} \to 0$ satisfying

$$0 \in \nabla_{x^\nu} f^\nu(x^{\nu,k}, \bar{x}^{-\nu}) + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} a_i^\nu(\bar{x}^{-\nu}) \nabla_{x^\nu} b_i^\nu(x^{\nu,k}) + e^{\nu,k+1} + N_{C_\nu}(x^{\nu,k}), \; \forall k \tag{11}$$

$$\eta_i^{\nu,k} := \lambda_i^{\nu,k} [a_i^\nu(\bar{x}^{-\nu}) b_i^\nu(x^{\nu,k}) + c_i^\nu(\bar{x}^{-\nu})] \to 0, \; \forall i = 1, \dots, m_\nu, \; \forall \nu = 1, \dots, N. \tag{12}$$

and

$$\lambda_i^{\nu,k}[a_i^\nu(\bar{x}^{-\nu})b_i^\nu(x^{\nu,k}) + c_i^\nu(\bar{x}^{-\nu})] > 0 \text{ if } \lim_k \frac{\lambda_i^{\nu,k}}{\delta_k^\nu} > 0, \tag{13}$$

where $\delta_k^\nu := \|(1, \lambda^{\nu,k})\|_\infty \to \infty$. Now, we will show that $\bar{x}$ is a PCAKKT-GNEP point with an appropriate sequence of multipliers. So, set $x^k := (x^{\nu,k})_\nu$ and define the multipliers

$$\bar{\lambda}_i^{\nu,k} := \begin{cases} 0 & \text{if } a_i^\nu(\bar{x}^{-\nu}) = 0; \\ \lambda_i^{\nu,k} a_i^\nu(\bar{x}^{-\nu})(a_i^\nu(x^{-\nu,k}))^{-1} & \text{otherwise.} \end{cases}$$

Observe that $\bar{\lambda}_i^{\nu,k}$ is well-defined for k large enough, since $\{x^k\}$ converges to $\bar{x}$ and a_i^ν is continuous.

Now, since f^ν is a C^1 function, we get

$$\nabla_{x^\nu} f^\nu(x^{\nu,k}, x^{-\nu,k}) + \sum_{i=1}^{m_\nu} \bar{\lambda}_i^{\nu,k} a_i^\nu(x^{-\nu,k}) \nabla_{x^\nu} b_i^\nu(x^{\nu,k}) - [\nabla_{x^\nu} f^\nu(x^{\nu,k}, \bar{x}^{-\nu}) + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} a_i^\nu(\bar{x}^{-\nu}) \nabla_{x^\nu} b_i^\nu(x^{\nu,k})] \to 0. \tag{14}$$

So, from (11), the approximate stationary condition (4) holds for the approximate multipliers $\{\bar{\lambda}_i^{\nu,k}\}$. To verify the fulfillment of the approximate complementary condition (7) with the approximate multipliers $\{\bar{\lambda}_i^{\nu,k}\}$, we only analyze the case $a_i^\nu(\bar{x}^{-\nu}) \neq 0$. Thus, from (12), we see

$$\begin{aligned} \bar{\lambda}_i^{\nu,k}[a_i^\nu(x^{-\nu,k})b_i^\nu(x^{\nu,k}) + c_i^\nu(x^{-\nu,k})] &= \lambda_i^{\nu,k} a_i^\nu(\bar{x}^{-\nu})(a_i^\nu(x^{-\nu,k}))^{-1}[a_i^\nu(x^{-\nu,k})b_i^\nu(x^{\nu,k}) + c_i^\nu(x^{-\nu,k})] \\ &= \lambda_i^{\nu,k} a_i^\nu(\bar{x}^{-\nu}) b_i^\nu(x^{\nu,k}) + \lambda_i^{\nu,k} a_i^\nu(\bar{x}^{-\nu})(a_i^\nu(x^{-\nu,k}))^{-1} c_i^\nu(x^{-\nu,k}) \\ &= \eta_i^{\nu,k} - \lambda_i^{\nu,k} c_i^\nu(\bar{x}^{\nu,k}) + \lambda_i^{\nu,k} a_i^\nu(\bar{x}^{-\nu})(a_i^\nu(x^{-\nu,k}))^{-1} c_i^\nu(x^{-\nu,k}) \\ &= \eta_i^{\nu,k} + \lambda_i^{\nu,k} a_i^\nu(x^{-\nu,k})^{-1}[a_i^\nu(\bar{x}^{-\nu}) c_i^\nu(x^{-\nu,k}) - a_i^\nu(x^{-\nu,k}) c_i^\nu(\bar{x}^{-\nu})] \to 0, \end{aligned}$$

where in the last expression, we use $c_i^\nu(x^{-\nu}) = M_i^\nu a_i^\nu(x^{-\nu})$, $\forall i$. Finally, we need to analyze the fulfillment of (8). Set $\delta^\nu := \|(1, \lambda^{\nu,k})\|$ and $\bar{\delta}^\nu := \|(1, \bar{\lambda}^{\nu,k})\|$. From the definition of $\bar{\lambda}^{\nu,k}$, we get that if $\lim_k \bar{\lambda}_i^{\nu,k}/\bar{\delta}^\nu > 0$ then $\lim_k \lambda_i^{\nu,k}/\delta^\nu > 0$. So, by using $c_i^\nu(x^{-\nu}) = M_i^\nu a_i^\nu(x^{-\nu})$, $\forall i$, we get

$$\bar{\lambda}_i^{\nu,k}[a_i^\nu(x^{-\nu,k})b_i^\nu(x^{\nu,k}) + c_i^\nu(x^{-\nu,k})] = \lambda_i^{\nu,k}[a_i^\nu(\bar{x}^{-\nu})b_i^\nu(x^{\nu,k}) + c_i^\nu(\bar{x}^{-\nu})] > 0.$$

From above, we get that PCAKKT-GNEP holds at $\bar{x}$ with the sequence of multipliers $\{\bar{\lambda}^{\nu,k}\}$. □

Now, we turn our attention on the sufficiency of our conditions based on our new approximate stationary concepts. In [4], it was shown that CAKKT is equivalent to optimality under the convexity of the constraint functions. We extend this result for (GNEP) under an additional condition. We have the following result.

Theorem 3.2. *Suppose that CAKKT-GNEP holds at the feasible point $\bar{x}$ having $\{x^k\}$ as a CAKKT-GNEP sequence with approximate multipliers $\{\lambda^{\nu,k} : \nu = 1, \dots, N\}$. Let us suppose that the constraints and C_ν are convex and the following statement holds:*

For every $\nu = 1, \dots, N$ and every $i = 1, \dots, m_\nu$, we have

$$\lambda_i^{\nu,k}[g_i^\nu(x^{\nu,k}, x^{-\nu,k}) - g_i^\nu(x^{\nu,k}, \bar{x}^{-\nu})] \to 0, \text{ and } \lambda_i^{\nu,k}[\nabla_{x^\nu} g_i^\nu(x^{\nu,k}, x^{-\nu,k}) - \nabla_{x^\nu} g_i^\nu(x^{\nu,k}, \bar{x}^{-\nu})] \to 0. \tag{H}$$

Then $\bar{x}$ is a solution of (GNEP).

Proof. Let us show that $\bar{x}$ is a solution of (GNEP). Thus, take any z satisfying $g_i^\nu(z^\nu, \bar{x}^{-\nu}) \leq 0$, $\forall i = 1, \ldots, m_\nu$ and $\forall \nu = 1, \ldots, N$. For simplicity, we set

$$\begin{aligned} \gamma_i^{\nu,k} &:= \lambda_i^{\nu,k}[g_i^\nu(x^{\nu,k}, \bar{x}^{-\nu}) - g_i^\nu(x^{\nu,k}, x^{-\nu,k})] \text{ and} \\ \Gamma_i^{\nu,k} &:= \lambda_i^{\nu,k}[\nabla_{x^\nu} g_i^\nu(x^{\nu,k}, \bar{x}^{-\nu}) - \nabla_{x^\nu} g_i^\nu(x^{\nu,k}, x^{-\nu,k})] \end{aligned} \tag{15}$$

Clearly, $\gamma_i^{\nu,k} \to 0$ and $\Gamma_i^{\nu,k} \to 0$, $\forall i$, $\forall \nu$ due to Assumption H. By CAKKT-GNEP, we also have

$$0 = \nabla_{x^\nu} f^\nu(x^k) + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^\nu(x^k) + e^{\nu,k} + \eta^{\nu,k} \text{ with } \eta^{\nu,k} \in N_{C_\nu}(x^{\nu,k}). \tag{16}$$

Take $\nu = 1, \ldots, N$. Therefore, by convexity of $g_i^\nu(\cdot, \bar{x}^{-\nu})$, one has

$$0 \geq g_i^\nu(z^\nu, \bar{x}^{-\nu}) \geq g_i^\nu(x^{\nu,k}, \bar{x}^{-\nu}) + \nabla_{x^\nu} g_i^\nu(x^{\nu,k}, \bar{x}^{-\nu})[z^\nu - x^{\nu,k}],$$

and now multiplying by $\lambda_i^{\nu,k} \in \mathbb{R}_+$ the above expression and adding all $i = 1, \ldots, m_\nu$, one has

$$0 \geq \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} g_i^\nu(x^{\nu,k}, \bar{x}^{-\nu}) + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^\nu(x^{\nu,k}, \bar{x}^{-\nu})[z^\nu - x^{\nu,k}].$$

By using (15) and (16) in the above expression, we get

$$0 \geq \sum_{i=1}^{m_\nu} \gamma_i^{\nu,k} + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} g_i^\nu(x^k) + \sum_{i=1}^{m_\nu} \Gamma_i^{\nu,k} - \nabla_{x^\nu} f^\nu(x^k)[z^\nu - x^{\nu,k}] - \langle e^{\nu,k}, z^\nu - x^{\nu,k} \rangle - \langle \eta^{\nu,k}, z^\nu - x^{\nu,k} \rangle. \tag{17}$$

As $\nabla_{x^\nu} f^\nu(x^k)[z^\nu - x^{\nu,k}] \leq f^\nu(z^\nu, x^{-\nu,k}) - f^\nu(x^{\nu,k}, x^{-\nu,k})$ (due to the convexity of $f^\nu(\cdot, x^{-\nu,k})$) and $\langle \eta^{\nu,k}, z^\nu - x^{\nu,k} \rangle \leq 0$ (since $\eta^{\nu,k} \in N_{C_\nu}(x^{\nu,k})$), the expression (17) yields

$$f^\nu(z^\nu, x^{-\nu,k}) - f^\nu(x^{\nu,k}, x^{-\nu,k}) \geq \sum_{i=1}^{m_\nu} \gamma_i^{\nu,k} + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} g_i^\nu(x^k) + \sum_{i=1}^{m_\nu} \Gamma_i^{\nu,k} - \langle e^{\nu,k}, z^\nu - x^{\nu,k} \rangle.$$

Now, by taking the limit in the above expression and having in mind that $e^{\nu,k} \to 0$, $x^k \to \bar{x}$ and $\lambda_i^{\nu,k} g_i^\nu(x^k) \to 0$, $\forall i$ (since $\bar{x}$ is a CAKKT-GNEP); one has $f^\nu(z^\nu, \bar{x}^{-\nu}) - f^\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \geq 0$, that is, $\bar{x}$ is a solution of (GNEP). □

We continue with the following observation.

Remark 3.1. *We would like to mention some instances where Assumption H is fulfilled.*

- *One instance is when the GNEP is a Nash equilibrium problem (NEP), which arises when all constraint functions g_i^ν depend only on the strategy of player ν, that is, $g_i^\nu(x) = \hat{g}_i^\nu(x^\nu)$ for some C^1 function $\hat{g}_i^\nu$;*
- *Another important case is when the sequences $\{\lambda^{\nu,k}\} \subset \mathbb{R}^{m_\nu}$ are bounded, for every ν. This holds under very weak constraint qualification as GNEP-quasinormality (GNEP-QN), see Definition 4.3, which is weaker than GNEP-MFCQ, see Definition 4.2.*

4 New Constraint Qualifications for GNEP

In this section, we will introduce new CQs for GNEP based on the approximate stationary concept introduced in Section 3. As we will see in the next sections, several algorithms for (GNEP) are supported by one of our stationary sequential conditions, and thus converge to KKT points under well-known CQ, as GNEP-MFCQ, (see Theorem 4.2), but our approximate stationary concept says us more than that. Following [5], for each approximate stationary condition (ASC), one can define it is associated conditions, referred to strict constraint qualifications (SCQ), such that "ASC+SCQ $\implies$ KKT" and, among them, characterize the weakest one. Throughout this paper, we adopt the same terminology as in [5]. As example, the weakest SCQ associated with the AKKT-GNEP condition will be called of AKKT-GNEP regularity. Similar names are given for the other SCQs, associated to other approximate stationary conditions defined in Section 3.

We continue with the following definition.

Definition 4.1. *For every $\nu = 1, \ldots, N$; $\alpha^\nu, \beta^\nu \in \mathbb{R}_+$; and $r^\nu \geq 0$, we consider the following sets:*

$$\mathbb{K}^\nu_{PCAKKT}(x, \alpha^\nu, \beta^\nu, r^\nu) := \left\{ \sum_{i=1}^{m_\nu} \lambda^\nu_i \nabla_{x^\nu} g^\nu_i(x) + N_{C_\nu}(x^\nu) \;\middle|\; \begin{array}{l} \lambda^\nu_i g^\nu_i(x) \geq \alpha^\nu \;\; if \;\; \lambda^\nu_i > \beta^\nu \|(1, \lambda^\nu)\|_\infty, \\ \lambda^\nu \in \mathbb{R}^{m_\nu},\; \lambda^\nu_i \geq 0,\; \forall i = 1, \ldots, m_\nu \\ |\lambda^\nu_i g^\nu_i(x)| \leq r^\nu,\; \forall i = 1, \ldots, m_\nu \end{array} \right\} \tag{18}$$

$$\mathbb{K}^\nu_{PAKKT}(x, \alpha^\nu, \beta^\nu) := \left\{ \sum_{i=1}^{m_\nu} \lambda^\nu_i \nabla_{x^\nu} g^\nu_i(x) + N_{C_\nu}(x^\nu) \;\middle|\; \begin{array}{l} \lambda^\nu_i g^\nu_i(x) \geq \alpha^\nu \;\; if \;\; \lambda^\nu_i > \beta^\nu \|(1, \lambda^\nu)\|_\infty, \\ \lambda^\nu \in \mathbb{R}^{m_\nu}, \lambda^\nu_i \geq 0,\; \forall i = 1, \ldots, m_\nu \end{array} \right\}; \tag{19}$$

$$\mathbb{K}^\nu_{CAKKT}(x, r^\nu) := \left\{ \sum_{i=1}^{m_\nu} \lambda^\nu_i \nabla_{x^\nu} g^\nu_i(x) + N_{C_\nu}(x^\nu) \;\middle|\; \begin{array}{l} \lambda^\nu \in \mathbb{R}^{m_\nu}, \lambda^\nu_i \geq 0,\; \forall i = 1, \ldots, m_\nu \\ |\lambda^\nu_i g^\nu_i(x)| \leq r^\nu,\; \forall i = 1, \ldots, m_\nu \end{array} \right\}; \tag{20}$$

$$\mathbb{K}^\nu_{AGP}(x, r^\nu) := \left\{ \sum_{i=1}^{m_\nu} \lambda^\nu_i \nabla_{x^\nu} g^\nu_i(x) + N_{C_\nu}(x^\nu) \;\middle|\; \begin{array}{l} \lambda^\nu \in \mathbb{R}^{m_\nu}, \lambda^\nu_i \geq 0,\; \forall i = 1, \ldots, m_\nu, \\ |\lambda^\nu_i \min\{g^\nu_i(x), 0\}| \leq r^\nu,\; \forall i = 1, \ldots, m_\nu \end{array} \right\}; \tag{21}$$

$$\mathbb{K}^\nu_{AKKT}(x, r^\nu) := \left\{ \sum_{i=1}^{m_\nu} \lambda^\nu_i \nabla_{x^\nu} g^\nu_i(x) + N_{C_\nu}(x^\nu) \;\middle|\; \begin{array}{l} \lambda^\nu \in \mathbb{R}^{m_\nu}, \lambda^\nu_i \geq 0,\; \forall i = 1, \ldots, m_\nu \\ |\min\{\lambda^\nu_i, -g^\nu_i(x)\}| \leq r^\nu,\; \forall i = 1, \ldots, m_\nu \end{array} \right\}. \tag{22}$$

We say that the PCAKKT-GNEP regularity condition holds at $\bar{x}$, if the set-valued mapping $(x, \alpha^\nu, \beta^\nu, r^\nu) \rightrightarrows \mathbb{K}^\nu_{PCAKKT}(x, \alpha^\nu, \beta^\nu, r^\nu)$ is outer semi continuous at $(\bar{x}, 0, 0, 0)$, $\forall \nu$, that is,

$$\limsup_{x \to \bar{x}, \alpha^\nu \downarrow 0, \beta^\nu \downarrow 0, r^\nu \downarrow 0} \mathbb{K}^\nu_{PCAKKT}(x, \alpha^\nu, \beta^\nu, r^\nu) \subset \mathbb{K}^\nu_{PCAKKT}(\bar{x}, 0, 0, 0), \quad \forall \nu = 1, \ldots, N.$$

The constraint qualification conditions PAKKT-GNEP, CAKKT-GNEP, AGP-GNEP and AKKT-GNEP regularity have analogous definitions using the sets $\mathbb{K}^\nu_{PAKKT}(x, \alpha^\nu, \beta^\nu, r^\nu)$, $\mathbb{K}^\nu_{CAKKT}(x, r^\nu)$, $\mathbb{K}^\nu_{AGP}(x, r^\nu)$ and $\mathbb{K}^\nu_{AKKT}(x, r^\nu)$ respectively, for all $\nu = 1, \ldots, N$.

Remark 4.1. *If $\bar{x}$ is a feasible point, then for every $\nu = 1, \ldots, N$, we have* $\mathbb{K}^{\nu}_{PCAKKT}(\bar{x}, 0, 0, 0) = \mathbb{K}^{\nu}_{PAKKT}(\bar{x}, 0, 0) = \mathbb{K}^{\nu}_{CAKKT}(\bar{x}, 0) = \mathbb{K}^{\nu}_{AGP}(\bar{x}, 0) = \mathbb{K}^{\nu}_{AKKT}(\bar{x}, 0)$ *and all coincide with the KKT cone of the ν-player given by*

$$\mathbb{K}^{\nu}_{KKT}(\bar{x}) := \left\{ \sum_{i=1}^{m_\nu} \lambda^{\nu}_i \nabla_{x^\nu} g^{\nu}_i(\bar{x}) + N_{C_\nu}(\bar{x}^\nu) \,\Big|\, \lambda^\nu \in \mathbb{R}^{m_\nu}_+ \text{ and } \lambda^\nu_i g^\nu_i(\bar{x}) = 0, \ \forall i = 1, \ldots, m_\nu \right\}.$$

Clearly, a point $\bar{x}$ is a KKT point if and only if $-\nabla_{x^\nu} f^\nu(\bar{x}) \in \mathbb{K}^{\nu}_{KKT}(\bar{x})$, $\forall \nu = 1, \ldots, N$.

The next theorem states that each SCQ is, indeed, the weakest SCQ associated with each optimality condition.

Theorem 4.1. *A feasible point $\bar{x}$ of* (GNEP) *satisfies PCAKKT-GNEP regularity if, and only if, for every continuously differentiable collection of f^ν objective functions, the PCAKKT-GNEP condition at $\bar{x}$ implies the KKT conditions. Similar conclusions are valid for PAKKT-GNEP, CAKKT-GNEP, AGP-GNEP and AKKT-GNEP regularity.*

Proof. We the same techniques as [14, Theorem 5.4]. We just prove the statement for PCAKKT-GNEP regularity, since the other ones are analogous.

First, suppose that PCAKKT-GNEP regularity holds at the feasible point $\bar{x}$. Now, assume that PCAKKT-GNEP condition holds at $\bar{x}$ for ν-objective functions f^ν. Thus, there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^{\nu,k}\} \subset \mathbb{R}^{m_\nu}_+$, $\{\eta^{\nu,k}\} \subset \mathbb{R}^{n_\nu}$, $k \geq 1$, for each $\nu = 1, \ldots N$ such that $x^k \to \bar{x}$, $\eta^{\nu,k} \in N_{C_\nu}(x^{\nu,k})$, (4), (7) and (8) hold. Therefore, $\nabla_{x^\nu} f^\nu(x^k) + \omega^{\nu,k} \to 0$, where $\omega^{\nu,k} := \sum_{i=1}^{m_\nu} \lambda^{\nu,k}_i \nabla_{x^\nu} g^\nu_i(x^k) + \eta^{\nu,k}$.

For every $\nu = 1, \ldots, N$, we define the sequence $\{\delta^\nu_k := \|(1, \lambda^{\nu,k})\|_\infty\}$. We consider two cases:

Case 1: If δ^ν_k is bounded. Then, in this case, $\lambda^{\nu,k}$ is also bounded. So, after taking an adequate subsequence, we assume that $\lambda^{\nu,k} \to \bar{\lambda}^\nu$. It is not difficult to see that $\bar{x}$ satisfies (3) having multipliers $\bar{\lambda}^\nu$ as multiplier.

Case 2: If δ^ν_k is unbounded. Then, in this case, we suppose that $\delta^\nu_k \to \infty$. We define the sets

$$I^\nu_+ := \left\{ i \in A^\nu(\bar{x}) \,\Bigg|\, \lim_k \frac{\lambda^{\nu,k}_i}{\delta^\nu_k} > 0 \right\},$$

and for each $k \in \mathbb{N}$, we set $r^\nu_k := \max_{i=1,\ldots,m_\nu}\{|\lambda^{\nu,k}_i g^\nu_i(x^k)|\}$,

$$\alpha^\nu_k := \min\left\{ \frac{1}{k}, \ \min_{i \in I^\nu_+}\{\lambda^{\nu,k}_i g^\nu_i(x^k)\} \right\} \quad \text{and} \quad \beta^\nu_k := \max\left\{ \frac{1}{k}, \max_{i \notin I^\nu_+} \frac{\lambda^{\nu,k}_i}{\delta^\nu_k} \right\} + \frac{1}{k}.$$

We note that $\alpha^\nu_k \downarrow 0, \beta^\nu_k \downarrow 0$, $r^\nu_k \to 0$ and $\omega^{\nu,k} \in \mathbb{K}^\nu_{PCAKKT}(x^k, \alpha^\nu_k, \beta^\nu_k, r^\nu_k)$ for all k large enough. As $\bar{x}$ fulfills the GNEP-PCAKKT-regular condition, we have for each $\nu = 1, \ldots, N$

$$-\nabla_{x^\nu} f^\nu(\bar{x}) = \lim_k \omega^{\nu,k} \in \limsup_{x \to \bar{x}, \alpha \downarrow 0, \beta \downarrow 0} \mathbb{K}^\nu_{PCAKKT}(x, \alpha^\nu, \beta^\nu, r^\nu) \subset \mathbb{K}^\nu_{KKT}(\bar{x}), \tag{23}$$

that is, $\bar{x}$ is a KKT point. This proves the first statement.

Conversely, for each $\nu = 1, \ldots, N$, take $w^{*,\nu} \in \limsup_{x \to \bar{x}, \alpha^\nu \downarrow 0, \beta^\nu \downarrow 0, r^\nu \to 0} \mathbb{K}^\nu_{PCAKKT}(x, \alpha^\nu, \beta^\nu, r^\nu)$. Then there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\omega^k\} \subset \mathbb{R}^n$, $\{\alpha^\nu_k\} \subset \mathbb{R}$, $\{\beta^\nu_k\} \subset \mathbb{R}$ and $\{r^\nu_k\} \subset \mathbb{R}$ such that

$x^k \to \bar{x}$, $\omega^{\nu,k} \to \omega^{*,\nu}, \alpha_k^\nu \downarrow 0$, $\beta_k^\nu \downarrow 0$, $r_k^\nu \to 0$ and $\omega^{\nu,k} \in \mathbb{K}^\nu_{PCAKKT}(x^k, \alpha_k^\nu, \beta_k^\nu, r_k^\nu)$ for all $k \in \mathbb{N}$. For each k there are $\lambda^{\nu,k} \in \mathbb{R}^{m_\nu}_+$ such that

$$\omega^{\nu,k} = \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^v \left(x^k\right) + \eta^{\nu,k} \text{ where } \eta^{\nu,k} \in N_{C_\nu}(x^{\nu,k}).$$

For each $\nu = 1, ..., N$, we define $f^\nu(x) = -\langle \omega^{*,\nu}, x^\nu \rangle$, $\forall x$. Clearly, $\nabla_{x^\nu} f^\nu(x) = -w^{*,\nu}$. It is not difficult to see that $\bar{x}$ is a PCAKKT-GNEP point for f^ν, $\forall \nu$. By hypothesis, $\bar{x}$ is a KKT-GNEP point, and hence $\lim_k \omega^{\nu,k} = -\nabla_{x^\nu} f^\nu(\bar{x}) = w^{*,\nu} \in \mathbb{K}^\nu_{KKT}(\bar{x})$. This concludes the proof. □

In the next subsection, we discuss the relations of the strict CQ with other CQs for (GNEP) stated in the literature.

4.1 Relations with other Constraint Qualifications for GNEP

In this subsection, we discuss the relations with other CQ tailored for GNEP.

Definition 4.2. *Let ν be a given index and $\bar{x}$ be a feasible point. We say that g^ν satisfies MFCQ with respect to player ν or simply MFCQ$_\nu$ if there is a vector $d \in T_{C_\nu}(\bar{x})$ such that*

$$g_i^\nu(\bar{x}) \geq 0 \implies \nabla_{x_\nu} g_i^\nu(\bar{x}) d < 0, \ \forall i = 1, \dots, m_\nu$$

Furthermore, we say that GNEP-MFCQ holds at $\bar{x}$, if for every $\nu = 1, \dots, N$, the function g^ν satisfies MFCQ$_\nu$ in $\bar{x}$.

In some situation, it is important to known if the corresponding sequence of multipliers $\{\lambda^{\nu,k}\}$ associated to any approximate stationary concept is bounded. Thus, by following the proof of [19, Theorem 7], we have the following statement.

Theorem 4.2. *GNEP-MFCQ implies AKKT-GNEP regularity.*

Another important CQ for GNEP is the next CQ which generalizes the well-known quasinormality CQ, which implication to the exactness of penalty methods for NLP.

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

$$\sum_{i=1}^{m_\nu} \lambda_i^\nu \nabla_{x^\nu} g_i^\nu(\bar{x}) = 0 \ \text{and} \ \lambda_i^\nu g_i^\nu(\bar{x}) = 0, \ \forall i = 1, \dots, m_\nu$$

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

Theorem 4.3. *GNEP-QN implies PAKKT-GNEP regular.*

Proof. We assume that $\bar{x}$ is not PAKKT-GNEP regular. Then there exist $\nu \in \{1, \dots, N\}$ and $w^{*,\nu} \in \mathbb{R}^{n_\nu}$ such that $w^{*,\nu} \in \limsup_{x \to x^*, \alpha^\nu \downarrow 0, \beta^\nu \downarrow 0} \mathbb{K}^\nu_{PAKKT}(x, \alpha^\nu, \beta^\nu)$ and $w^{*,\nu} \notin \mathbb{K}^\nu_{KKT}(\bar{x})$. Hence, there exist sequences $\left\{x^k\right\} \subset \mathbb{R}^n, \left\{\omega^{\nu,k}\right\} \subset \mathbb{R}^{n_\nu}, \left\{\alpha^{\nu,k}\right\} \subset \mathbb{R}$, and $\left\{\beta^{\nu,k}\right\} \subset \mathbb{R}$ such that $x^k \to x^*$, $\omega^{\nu,k} \to \omega^{*,\nu}, \alpha_k^\nu \downarrow 0, \beta_k^\nu \downarrow 0$, and $\omega^{\nu,k} \in \mathbb{K}^\nu_{PAKKT}\left(x^k, \alpha_k^\nu, \beta_k^\nu\right)$ for all $k \geq 1$, where $\omega^{\nu,k} = \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^\nu(x^k) + \eta^{\nu,k}$.

Now, we define $\delta_k^\nu := \left\|(1,\lambda^{\nu,k})\right\|_\infty$. Certainly, the sequence $\{\delta_k^\nu\}$ is unbounded because, otherwise, $\{\lambda^{\nu,k}\}$ must be bounded, and after taking an adequate subsequence, we may suppose that $\lambda^{\nu,k} \to \bar{\lambda}^\nu$. The latter implies that $\omega^{*,\nu} \in \mathbb{K}^\nu_{KKT}(\bar{x})$ having $\bar{\lambda}^\nu$ as multiplier.

So, we suppose that $\delta_k^\nu \to \infty$. Now, dividing $\omega^{\nu,k}$ by δ_k^ν and after taking an adequate limit, we assume that $\lambda^{\nu,k}/\delta_k^\nu \to \lambda^{\nu,*}$ and

$$\sum_{i=1}^{m_\nu} \lambda_i^{\nu,*}\nabla_{x^\nu} g_i^\nu(\bar{x}) + \eta^\nu = 0, \text{ where } \lambda_i^{\nu,*} g_i^\nu(\bar{x}) = 0, \forall i \text{ and } \eta^\nu \in N_{C_\nu}(\bar{x}^\nu),$$

where $(\lambda^{\nu,*}) \neq 0$. Due to the continuity of the functions, $\operatorname{sgn}\left(\lambda_i^{\nu,*} g_i^\nu(x^k)\right) = \operatorname{sgn}\left(\lambda_i^{\nu,k} g_i^\nu(x^k)\right) = 1$ whenever $\lambda_i^{\nu,*} > 0$ (note that $\lim_k \lambda_i^{\nu,k}/\delta_k^\nu = \lambda_i^{\nu,*} > 0$ implies $\lambda_i^{\nu,k} > \beta_k^\nu \delta_k^\nu = \beta_k^\nu \left\|(1,\lambda^{\nu,k})\right\|_\infty$ for all k sufficiently large). Hence, $\bar{x}$ does not satisfy GNEP-QN, and it completes the proof. □

In Figure 4.1, we summarize the relations between our new CQs and some classical CQs for GNEP, as GNEP-MFCQ.

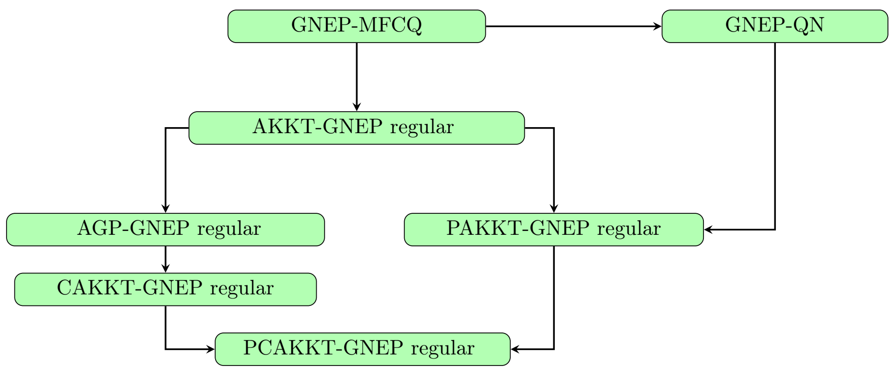

Figure 2: Relation between the CQ based on the sequential conditions for (GNEP) and other CQ for GNEP. The arrows mean implications.

5 A Hyperbolic Augmented Lagrangian method for GNEP

Here, we describe an augmented Lagrangian method based on the hyperbolic penalty function. We point out that employing augmented Lagrangian methods to find the solution of optimization problems is a very successful technique for solving finite-dimensional problems, and it is described in several standard textbooks on continuous optimization, for example, [10, 11, 12, 21, 3] to cite a few of them.

Augmented Lagrangian methods are useful when there exist efficient algorithms for solving their subproblems. 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 [13] for NLP. The hyperbolic function is a smoothing of the penalty function studied by Zangwill [26], and studied as an hyperbolic augmented Lagrangian function in [24]. This Lagrangian function showed interesting computational results in [25], [9] and [8].

We mention that for each player ν, the method penalizes the inequality constraints $g_i^\nu(x) \leq 0$, $\forall i$, but explicitly leaves the additional abstract constraint $x^\nu \in C_\nu$ in the constraints.

In this work, for each player $\nu = 1, ..., N$, and given $\lambda^\nu \in \mathbb{R}_+^{m_\nu}$ and $\rho_\nu > 0$, the *hyperbolic augmented Lagrangian* (HALA) function for the player ν is given by

$$L_H^\nu(x, \lambda^\nu; \rho_\nu) := f^\nu(x) + \sum_{i=1}^{m_\nu} \left(\lambda_i^\nu g_i^\nu(x) + \sqrt{(\lambda_i^\nu g_i^\nu(x))^2 + \frac{1}{\rho_\nu^2}} \right).$$

Formally, each player ν has to solve problems of the form (GNEP). Specifically, each associated NEP consisting of minimization problems

$$\text{minimize}_{x^\nu} \ L_H^\nu\left(x^\nu, x^{-\nu}, \lambda^\nu; \rho_\nu\right) \ \text{subject to} \ x^\nu \in C_\nu,$$

for some parameters λ^ν and ρ_ν that vary in each iteration under some criterion. Unlike other augmented Lagrangian functions, the HALA function is twice continuously differentiable, which allows us to use standard methods of optimization as the Newton method. We observe that the partial gradient with respect to x^ν is given by:

$$\nabla_{x^\nu} L_H^\nu(x, \lambda^\nu; \rho_\nu) = \nabla_{x^\nu} f^\nu(x) + \sum_{i=1}^{m_\nu} \lambda_i^\nu \left(1 + \frac{\rho_\nu \lambda_i^\nu g_i^\nu(x)}{\sqrt{(\rho_\nu \lambda_i^\nu g_i^\nu(x))^2 + 1}} \right) \nabla_{x^\nu} g_i^\nu(x).$$

To simplify the notation, we consider the increasing bijection $h : \mathbb{R} \to (0, 2)$ defined by

$$h(t) := 1 + \frac{t}{\sqrt{1 + t^2}}, \quad t \in \mathbb{R}, \tag{24}$$

with this notation, the partial gradient of L_H^ν can be written as

$$\nabla_{x^\nu} L_H^\nu(x, \lambda^\nu; \rho^\nu) = \nabla_{x^\nu} f^\nu(x) + \sum_{i=1}^{m_\nu} \lambda_i^\nu h(\rho_\nu \lambda_i^\nu g_i^\nu(x)) \nabla g_i^\nu(x).$$

An important feature of the augmented Lagrangian method is to decide when the parameter penalization will increase. Several approaches are possible; see [3] and references therein. Our HALA method considers a novel measure of infeasibility and complementarity, inspired by the function of Fischer-Burmeister (an important tool for developing efficient numerical methods for solving throught semi-smooth Newton methods), see [17].

We proceed by explaining our method. For each player $\nu = 1, \ldots, N$, we define the infeasibility-complementarity measure with respect to ν as the function $\mathbb{V}^\nu : \mathbb{R}^n \times \mathbb{R}^{m_\nu} \times \mathbb{R}_+ \to \mathbb{R}_+$ defined by

$$\mathbb{V}^\nu(x, u, \rho) := \max_{i=1,\ldots,m_\nu} \left| \sqrt{(u_i g_i^\nu(x))^2 + \frac{1}{\rho^2}} + u_i g_i^\nu(x) - \frac{1}{\rho} \right|. \tag{25}$$

Note that, as a consequence of the Fischer-Burmeister function [17], we get

$$\sqrt{\left[(u_i g_i^\nu(x))^2 + \frac{1}{\rho^2}\right]} + u_i g_i^\nu(x) - \frac{1}{\rho} = 0 \iff u_i g_i^\nu(x) \leq 0, \ \rho u_i g_i^\nu(x) = 0 \ \text{ and } \rho \geq 0. \tag{26}$$

The expression $\mathbb{V}^\nu(x,u;\rho)=0$, given in (25), is a way to measure the complementarity of the pair (x,u). In fact, from the above we get $\mathbb{V}^\nu(x,u;\rho)=0$ if and only if $u_i g_i^\nu(x)=0$, $\forall i=1,\ldots,m_\nu$. In particular, if $u_i>0$, then $g_i^\nu(x)\le 0$. We would like to mention that under reasonable assumptions, our method HALA reaches feasible points even if the corresponding multipliers go to zero; see Proposition 6.6 for the proper statement.

We proceed by stating our algorithm framework. Our hyperbolic augmented Lagrangian method is presented in Algorithm 1.

Algorithm 1 Hyperbolic Augmented Lagrangian (HALA) method

Inputs: Parameters $\tau_\nu>1$, $\sigma_\nu\in(0,1)$, and $\rho_0>0$; and initial points $(x^0,\lambda^0)\in\mathbb{E}\times\mathbb{R}_{++}$. Set $k:=0$.

1. If (x^k,λ^k) satisfies some stopping criteria of the GNEP: STOP
2. Compute an approximate solution of $x^{k+1}\in C=C_1\times\cdots\times C_N$ of NEP consisting of minimization subproblems

$$\min_{x^\nu} L_H^\nu\left(x^\nu,x^{-\nu},\lambda^{\nu,k};\rho_{\nu,k}\right) \text{ subject to } x^\nu\in C_\nu, \tag{27}$$

   for each player $v=1,...,N$.

3. For $\nu=1,\ldots,N$, update the vector of multipliers to $\lambda^{\nu,k}=(\lambda_1^{\nu,k},\ldots,\lambda_{m_\nu}^{\nu,k})$ by the rule:

$$\lambda_i^{\nu,k+1}:=\lambda_i^{\nu,k}\left(1+\frac{\rho_{\nu,k}\lambda_i^{\nu,k}g_i^v\left(x^{k+1}\right)}{\sqrt{\left(\rho_{\nu,k}\lambda_i^{\nu,k}g_i^v\left(x^{k+1}\right)\right)^2+1}}\right), \text{ for all } i=1,\ldots,m_\nu. \tag{28}$$

4. For $\nu=1,\ldots,N$. If $k=0$ or

$$\mathbb{V}^\nu(x^{k+1},\lambda^{\nu,k+1};\rho_{\nu,k})\le\sigma_\nu\mathbb{V}^\nu(x^k,\lambda^{\nu,k};\rho_{\nu,k-1}). \tag{29}$$

   Set $\rho_{\nu,k+1}:=\rho_{\nu,k}$. Otherwise, take $\rho_{\nu,k+1}\ge\tau_\nu\rho_{\nu,k}$.

5. Set $k\leftarrow k+1$, and go to Step 1.

Now, some comments concerning Algorithm 1:

1. In Step 2, we require the computation of an approximate solution x^{k+1} for the GNEP problem given by (27). In our case, it requires finding $x^{k+1}\in\mathbb{R}^n$ and $e^{\nu,k+1}\in\mathbb{R}^n$ such that

$$0\in\nabla_{x^\nu}L_H^\nu(x^{k+1},\lambda^{\nu,k};\rho_{\nu,k})+e^{\nu,k+1}+N_{C_\nu}(x^{\nu,k+1}) \text{ with } \|e^{\nu,k+1}\|\le\epsilon_k, \tag{30}$$

   for each player $\nu=1,\ldots,N$. Here, $\{\epsilon_k\}\subset\mathbb{R}_{++}$ represents the degree of inexactness of the approximate solution. In our optimality theorems, we require $\epsilon_k\to 0$.

2. As usual the update rule of the multipliers ensure that

$$\nabla_{x^\nu}L_H^\nu(x^{k+1},\lambda^{\nu,k};\rho_{\nu,k})=\nabla_{x^\nu}L^\nu(x^{k+1},\lambda^{\nu,k+1}).$$

where L^ν is the Lagrangian function associated to the player ν defined by

$$L^\nu(x,\lambda) = f^\nu(x) + \sum_{i=1}^{m_\nu} \lambda_i^\nu g_i^\nu(x).$$

We note that the Lagrange multiplier update is multiplicative, and thus, as a consequence of the update rule of the multipliers, $\lambda^{\nu,k}$ is always positive, see Proposition 6.1. Following [13], a positive effect of considering non-safeguarded updates is that the penalty parameter does not need to increase to infinity.

3. Clearly, the overall behavior of the augmented Lagrangian method depends on the ability to compute an approximate solution in Step 2. Here, we follow the approach given by [19, 15], but unlike them, we can use general-purpose nonlinear equation due to the smoothness of the gradient of L_H^ν.

6 Convergence analysis of the HALA method

We proceed to analyze the convergence of Algorithm 1. Throughout the analysis, we implicitly assume that the algorithm does not terminate after a finite number of iterations.

We start by showing that the vector of approximate multipliers λ^k is always positive.

Proposition 6.1. *For every $\nu = 1,\dots,N$, we always have $\{\lambda^{\nu,k}\} \subset \mathbb{R}^{m_\nu}_{++}$.*

Proof. Let $\nu \in \{1,\dots,N\}$ and $i = 1,\dots,m_\nu$. Then, by multiplying (28) by $\lambda_i^{\nu,k}$, one has

$$\lambda_i^{\nu,k+1}\lambda_i^{\nu,k} = (\lambda_i^{\nu,k})^2 \left(1 + \frac{\rho_{\nu,k}\lambda_i^{\nu,k} g_i^v\left(x^{k+1}\right)}{\sqrt{\left(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^v\left(x^{k+1}\right)\right)^2 + 1}}\right) = (\lambda_i^{\nu,k})^2 h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^v\left(x^{k+1}\right)), \tag{31}$$

where h is given by (24). As h is strictly positive, we obtain $\lambda_i^{\nu,k+1}\lambda_i^{\nu,k} > 0$ whenever $\lambda_i^{\nu,k} \neq 0$. Thus, from $\lambda_i^{\nu,0} > 0$, we get $\lambda_i^{\nu,k} > 0$, $\forall k \in \mathbb{N}$. ☐

As a consequence of the above proposition, we conclude that all multipliers in the algorithm are strictly positive. This feature differs of other augmented Lagrangian methods, as the safeguarded augmented Lagrangian method presented in [1]. Another important statement about the multipliers will be stated in the following proposition.

Proposition 6.2. *Suppose that there exists a subsequence of multipliers $\{\lambda_i^{\nu,k+1}\}_{k\in\mathcal{N}_0}$ with $\mathcal{N}_0 \subset \mathbb{N}$, such that $\lambda_i^{\nu,k+1} \to_{\mathcal{N}_0} +\infty$ for some $\nu = 1,\dots,N$ and some $i = 1,\dots,m_\nu$. Then, there exists a subsequence $\mathcal{N}_+ \subset \mathbb{N}$ such that $g_i^\nu\left(x^{\ell+1}\right) > 0$, $\forall \ell \in \mathcal{N}_+$.*

Proof. Take $\nu = 1\dots,N$ and $i = 1,\dots,m_\nu$ such that $\|\lambda_i^{k,v}\| \to_{\mathcal{N}_0} +\infty$. As $\lambda_i^{k,v} \to_{\mathcal{N}_0} +\infty$, for every $k \in \mathcal{N}_0$ there must exist a $\ell_k \in \mathbb{N}$ such that $\lambda_i^{\nu,\ell_k} < \lambda_i^{\nu,\ell_k+1}$. By using the update rule of the multipliers, one has

$$\lambda_i^{\nu,k} < \lambda_i^{\nu,k} \left(1 + \frac{\rho_{\nu,k}\lambda_i^{\nu,k} g_i^v\left(x^{k+1}\right)}{\sqrt{\left(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^v\left(x^{k+1}\right)\right)^2 + 1}}\right) = \lambda_i^{\nu,k+1}.$$

The above expression is valid only if $\rho_{\nu,\ell_k}\lambda_i^{\nu,\ell_k} g_i^\nu(x^{\ell_k+1}) > 0$ and, therefore, $g_i^\nu(x^{\ell_k+1}) > 0$. Thus, $\mathcal{N}_+ := \{\ell_k : k \in \mathcal{N}_0\} \subset \mathbb{N}$ satisfies the required condition. □

6.1 On the optimality of the HALA method

We proceed by discussing the optimality of the limit point of Algorithm 1 applied to the GNEP. We will assume that $\epsilon_k \to 0$. Our main result in this section is to show that Algorithm 1 generates AGP-GNEP points. We divide the main proof into several propositions. The following proposition shows that the approximate complementarity condition CAKKT-GNEP (7) holds when the penalty parameter is bounded. Here, we do not require the existence of a limit point of $\{x^{k+1}\}$.

Proposition 6.3. *Suppose that the sequence $\{\rho_{\nu,k}\} \subset \mathbb{R}_{++}$ is bounded for some $\nu \in \{1,\dots,N\}$. Then, there exists a subsequence $\mathcal{N}$ such that $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to_{\mathcal{N}} 0$, $\forall i = 1,\dots,m_\nu$.*

Proof. Let $\nu \in \{1,\dots,N\}$ such that the sequence $\{\rho_{\nu,k}\}_{k\in\mathbb{N}} \subset \mathbb{R}_{++}$ is bounded. So, by the update rule of the penalty parameter (1), we get $\mathbb{V}^\nu(x^{k+1},\lambda^{\nu,k+1},\rho_{\nu,k}) \to 0$. As $\{\rho_{\nu,k}\} \subset \mathbb{R}_{++}$ is bounded, we can assume that $\rho_{\nu,k} \to \rho_\nu$ after an adequate subsequence. Furthermore, without relabeling, we assume that $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to a_i^\nu \in \mathbb{R} \cup \{\pm\infty\}$, $\forall i$. We consider the following cases:

- If $a_i^\nu \in \mathbb{R}$. So, since $\mathbb{V}^\nu(x^{k+1},\lambda^{\nu,k+1},\rho_{\nu,k}) \to 0$, from (25) and (26), we get $\rho_\nu a_i^\nu = 0$, $a_i^\nu \le 0$. As $\rho_v > 0$ (since $\{\rho_{\nu,k}\}$ is a nondecreasing sequence), we get $a_i^\nu = 0$ and therefore, $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to a_i^\nu = 0$;

- If $a_i^\nu = \infty$ then $\lambda_i^{\nu,k+1} g_i(x^{k+1}) \to \infty$ and hence $g_i(x^{k+1}) > 0$. From (25) and after some computation, we obtain

$$|\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1})| + \lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) - \frac{1}{\rho_{\nu,k}^2} \le \mathbb{V}^\nu(x^{k+1},\lambda^{\nu,k+1},\rho_{\nu,k}).$$

  From above, $0 \le 2\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) = 2\max\{\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}), 0\} \le \mathbb{V}^\nu(x^{k+1},\lambda^{\nu,k+1},\rho_{\nu,k}) + (1/\rho_{\nu,k}^2)$. The latter implies that the sequence $\{\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1})\}$ is bounded which it is a contradiction since $\lambda_i^{\nu,k+1} g_i(x^{k+1}) \to \infty$;

- If $a_i^\nu = -\infty$ then $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to -\infty$ and hence $\rho_{\nu,k}\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to -\infty$. Now, from $|\sqrt{(\lambda^{\nu,k+1} g_i^\nu(x^{k+1}))^2 + \rho_{\nu,k}^{-2}} + \lambda^{\nu,k+1} g_i^\nu(x^{k+1}) - \rho_{\nu,k}^{-1}| \le \mathbb{V}^\nu(x^{k+1},\lambda^{\nu,k+1},\rho_{\nu,k})$ and after straightforward calculations, we get

$$\left|\sqrt{(\rho_{\nu,k}\lambda^{\nu,k+1} g_i^\nu(x^{k+1}))^2 + 1} + \rho_{\nu,k}\lambda^{\nu,k+1} g_i^\nu(x^{k+1}) - 1\right| \le \rho_{\nu,k}\mathbb{V}^\nu(x^{k+1},\lambda^{\nu,k+1},\rho_{\nu,k}) \to 0.$$

  The above is a contradiction, since $\sqrt{[(\rho_{\nu,k}\lambda^{\nu,k+1} g_i^\nu(x^{k+1}))^2 + 1]} + \rho_{\nu,k}\lambda^{\nu,k+1} g_i^\nu(x^{k+1}) \to 0$.

From all the above cases, there must exists $\mathcal{N} \subset \mathbb{N}$ such that $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to_{\mathcal{N}} 0$. □

We now proceed by analyzing the fulfillment of the approximate complementarity condition when the penalty parameter is unbounded and the corresponding vector of multipliers is bounded. We start with the following statement.

Proposition 6.4. *Let $\bar{x}$ be the limit of the sequence $\{x^k\}_{k\in\mathbb{N}}$. We assume that $\bar{x}$ is feasible, that is, $\bar{x} \in \Omega$. Suppose that the sequence $\{\rho_{\nu,k}\}_{k\in\mathbb{N}} \subset \mathbb{R}_{++}$ is an unbounded sequence and that the subsequence $\{\lambda^{\nu,k+1}\}_{k\in\mathcal{N}}$ is bounded for some $\nu \in \{1,\dots,N\}$. Then, there exists a subsequence $\mathcal{N}_b \subset \mathcal{N}$, such that*

$$\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to_{\mathcal{N}_b} 0, \ \forall i = 1,\dots,m_\nu.$$

Proof. First, since $\{\rho_{\nu,k}\}_{k\in\mathbb{N}}$ is a nondecreasing sequence, any subsequence of $\{\rho_{\nu,k}\}_{k\in\mathbb{N}}$ must be unbounded. We now after taking an adequate subsequence (and without relabeling), we assume that $\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1}) \to_{\mathcal{N}} a_i^\nu \in \mathbb{R} \cup \{\pm\infty\}$, $\forall i$. We consider the following three cases

(a) If $a_i^\nu \in \mathbb{R}$. Then $\lambda_i^{\nu,k} g_i(x^{k+1}) = \rho_{\nu,k}\lambda_i^{\nu,k} g_i(x^{k+1})/\rho_{\nu,k} \to_{\mathcal{N}} 0$ since $\rho_{\nu,k} \to \infty$. From the update rule of the multipliers (28), we obtain

$$\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) = \lambda_i^{\nu,k} g_i^\nu(x^{k+1}) \underbrace{\left(1 + \frac{\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu\left(x^{k+1}\right)}{\sqrt{\left(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu\left(x^{k+1}\right)\right)^2 + 1}}\right)}_{bounded} \to 0. \tag{32}$$

(b) If $a_i^\nu = \infty$. Then, $g_i^\nu(x^{k+1}) > 0$ along an adequate subsequence. As $x^k \to_{\mathbb{N}} \bar{x}$, we get $g_i^\nu(x^{k+1}) \to g_i^\nu(\bar{x}) = 0$ since the limit $\bar{x}$ is feasible. Now, from the boundedness of $\{\lambda^{\nu,k+1}\}_{k\in\mathcal{N}}$, we get $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to 0$;

(c) If $a_i^\nu = -\infty$ then we subdivide this case into two subcases:

(c.1) If $\{\lambda_i^{\nu,k}\}_{k\in\mathcal{N}}$ is bounded. Then, as $\rho_{\nu,k}\lambda_i^{\nu,k} g_i(x^{k+1}) \to_{\mathcal{N}} -\infty$, from the definition of h, we get $h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu\left(x^{k+1}\right)) \to 0$. Thus, by the update rule of the multipliers (28), $\lambda_i^{\nu,k+1} \to 0$ and hence $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to 0$.

(c.2) If $\{\lambda_i^{\nu,k}\}_{k\in\mathcal{N}}$ is unbounded. In this case, by Proposition (6.2), we find a subsequence $\mathcal{N}_1$ such that $g_i^\nu(x^\ell) > 0$, $\forall \ell \in \mathcal{N}_1$. The latter implies that $g_i^\nu(\bar{x}) = 0$ (since $\bar{x}$ is feasible). As $\{\lambda^{\nu,k+1}\}_{k\in\mathcal{N}}$ is bounded, after taking an adequate subsequence, we get $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to 0$ along an adequate subsequence.

From all the above cases, there must exists $\mathcal{N}_b \subset \mathcal{N}$ such that $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to_{\mathcal{N}} 0$. □

The above statements show that the sequence of iteratives $\{x^k\}$ fulfills the CAKKT conditions when penalty parameter or the multipliers is bounded. In the case, that both sequences are unbounded, we get AGP-GNEP given by (6).

Before continue, we mention that the function $x \to x + \sqrt{x^2+1}$, $x \in \mathbb{R}$ is convex. Due to the convexity, we have

$$b + \sqrt{b^2+1} \geq a + \sqrt{a^2+1} + \left[1 + \frac{a}{\sqrt{a^2+1}}\right](b-a), \quad \text{for all } a, b \in \mathbb{R}. \tag{33}$$

We have the next proposition.

Proposition 6.5. *Let $\bar{x} \in \Omega$ be the limit of the sequence $\{x^k\}_{k\in\mathbb{N}}$. Suppose that the sequence $\{\rho_{\nu,k}\}_{k\in\mathbb{N}} \subset \mathbb{R}_{++}$ and $\{\lambda^{\nu,k+1}\}_{k\in\mathcal{N}}$ are unbounded for some $\nu \in \{1,\dots,N\}$. Then, there exists a subsequence $\mathcal{N}_{ub} \subset \mathcal{N}$, such that*

$$\lambda_i^{\nu,k+1} \min\{g_i^\nu(x^{k+1}), 0\} \to_{\mathcal{N}_{ub}} 0, \ \forall i = 1,\dots,m_\nu.$$

Proof. Let $\nu \in \{1,\dots,N\}$ such that $\{\lambda^{\nu,k+1}\}_{k\in\mathcal{N}}$ is unbounded. First, since $\lambda_i^{\nu,k+1} \le 2\lambda_i^{\nu,k}$, we get that $\{\lambda^{\nu,k}\}_{k\in\mathcal{N}}$ is also unbounded. So, without loss of generality, we assume that $\lambda^{\nu,k+1} \to_{\mathcal{N}} \infty$ and $\lambda^{\nu,k} \to_{\mathcal{N}} \infty$. We now suppose, after taking an adequate subsequence, that $\lambda_i^{\nu,k} g_i^\nu(x^{k+1}) \to a_i^\nu \in \mathbb{R} \cup \{\pm\infty\}$. We consider the following three cases:

(a) If $a_i^\nu = \infty$ then $\lambda_i^{\nu,k} g_i^\nu(x^{k+1}) \to \infty$. As a consequence, we get $\min\{g_i^\nu(x^{k+1}), 0\} = 0$ for k large enough. The latter implies that $\lambda_i^{\nu,k} \min\{g_i^\nu(x^{k+1}), 0\} \to 0$. We now by using (32), we get $\lambda_i^{\nu,k+1} \min\{g_i^\nu(x^{k+1}), 0\} \to 0$.

(b) If $a_i^\nu = -\infty$ then $\lambda_i^{\nu,k} g_i(x^{k+1}) \to -\infty$. Thus, $\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1}) \to -\infty$. Let us recall that

$$h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1})) = \left(1 + \frac{\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu\left(x^{k+1}\right)}{\sqrt{\left(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu\left(x^{k+1}\right)\right)^2 + 1}}\right).$$

Now, by choosing $a = \rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1})$ and $b = 1$ into expression (33), we get

$$1+\sqrt{2} \ge \rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1}) + \sqrt{[(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1}))^2+1]} + [1-\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1})]h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1})).$$

Now, after some calculations and since $a + \sqrt{a^2+1} \ge \max\{a, 0\}$, $\forall a \in \mathbb{R}$, we get

$$\begin{aligned} \frac{1+\sqrt{2}}{\rho_{\nu,k}} + h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i(x^{k+1}))\lambda_i^{\nu,k} g_i^\nu(x^{k+1}) &\ge \frac{h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1}))}{\rho_{\nu,k}} + \max\{\lambda_i^{\nu,k} g_i^\nu(x^{k+1}), 0\} \\ \frac{1+\sqrt{2}}{\rho_{\nu,k}} + \lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) &\ge \frac{h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1}))}{\rho_{\nu,k}} + \max\{\lambda_i^{\nu,k} g_i^\nu(x^{k+1}), 0\}, \end{aligned} \tag{34}$$

where in the last inequality we use $\lambda_i^{\nu,k} h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1})) = \lambda_i^{\nu,k+1}$. Now, as $\rho_{\nu,k}$ goes to infinity and $\lambda_i^{\nu,k} g_i^\nu(x^{k+1}) \to -\infty$, and after taking an adequate subsequence (without relabeling), (34) implies that $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to b_i^\nu \ge 0$, for some $b_i^\nu \in \mathbb{R}_+ \cup \{+\infty\}$. On the other hand, since $g_i^\nu(x^{k+1}) < 0$ for k infinite large in a proper subsequence and $\lambda_i^{\nu,k+1} \ge 0$, we get $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to b_i^\nu = 0$.

(c) If $a_i^\nu \in \mathbb{R}$. Then, we subdivide this case into three subcases:

(c.1) If $a_i^\nu > 0$. Then, in this case, as $\lambda_i^{\nu,k} g_i(x^{k+1}) \to a_i^\nu > 0$, we get $g_i^\nu(x^{k+1}) > 0$ and hence $\min\{g_i^\nu(x^{k+1}), 0\} = 0$ for k large enough. So, $\lambda_i^{\nu,k+1} \min\{g_i^\nu(x^{k+1}), 0\} \to 0$.

(c.2) If $a_i^\nu < 0$. Then, as $\lambda_i^{\nu,k} g_i(x^{k+1}) \to a_i^\nu < 0$, we get $\max\{\lambda_i^{\nu,k} g_i(x^{k+1}), 0\} = 0$. Furthermore, as $\rho_{\nu,k} \to \infty$, one has $\rho_{\nu,k}\lambda_i^{\nu,k} g_i(x^{k+1}) \to -\infty$. Now, similar to item (b) above, we can show that $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to 0$ for an adequeate subsequence.

(c.3) If $a_i^\nu = 0$, there is nothing to show.

For all cases, we can find a subset $\mathcal{N}_{ub} \subset \mathcal{N}$ such that $\lambda_i^{\nu,k+1} \min\{g_i^\nu(x^{k+1}), 0\} \to_{\mathcal{N}_{ub}} 0$. □

Finally, we obtain the main result of this section.

Theorem 6.1. *Let $\bar{x}$ be a feasible point which is the limit of the sequence $\{x^k\}$. Then, we have the following statements:*

1. *If $\{\min\{\rho_{\nu,k}; \|\lambda^{\nu,k}\| : \nu = 1, \dots, N\}\}$ has a bounded subsequence, then $\bar{x}$ is a CAKKT-GNEP point. In this case, $\bar{x}$ is a KKT point if CAKKT-GNEP regularity condition holds*
2. *If $\{\min\{\rho_{\nu,k}; \|\lambda^{\nu,k}\| : \nu = 1, \dots, N\}\}$ does not contain a bounded subsequence, then $\bar{x}$ is a AGP-GNEP point. In this case, $\bar{x}$ is a KKT point if AGP-GNEP regularity condition holds*

Proof. Clearly, from Step 2 and from $\nabla_{x^\nu} L_H^\nu(x^{k+1}, \lambda^{\nu,k}; \rho_{\nu,k}) = \nabla_{x^\nu} L^\nu(x^{k+1}, \lambda^{\nu,k+1})$, one has

$$0 \in \nabla_{x^\nu} L^\nu(x^{k+1}, \lambda^{\nu,k+1}) + e^{\nu,k+1} + N_{C_\nu}(x^{\nu,k+1}) \text{ with } \|e^{\nu,k+1}\| \le \epsilon_k, \tag{35}$$

From above, we only need to analyze the approximate fulfillment of the complementarity conditions. Note that the first item follows from Propositions 6.3 and 6.4. The second item is a consequence of Proposition 6.5. □

6.2 On the feasibility of the HALA method

We proceed by analyzing the feasibility problem; in short, we will describe the best that we can expect regarding the feasibility of the limit point.

Proposition 6.6. *Let $\bar{x}$ be the limit of the sequence $\{x^k\}_{k\in\mathbb{N}}$. If one of the following conditions holds:*

(a) $\{\rho_{\nu,k}\}$ is bounded;

(b) $\{\rho_{\nu,k}\}$ is unbounded and there exists some $B > 0$ such that $L_H^\nu(x^{k+1}, \lambda^{\nu,k}; \rho_{\nu,k}) \le B$, $\forall k \in \mathbb{N}$.

Then, the point $\bar{x}$ satisfies $g_i^\nu(\bar{x}) \le 0$, $\forall i = 1, \dots, m_\nu$.

Proof. We start with item (a). Suppose that $\{\rho_{\nu,k}\}$ is bounded. From Proposition 6.3, there is a subsequence $\mathcal{N}$ such that $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to_{\mathcal{N}} 0$, $\forall i = 1, \dots, m_\nu$. Observe that Proposition 6.3 holds even if $\{x^{k+1}\}$ does not converge.

Take $i \in \{1, \dots, m_\nu\}$. We consider two sub-cases:

Case (a): If $\liminf_{\mathcal{N}} \lambda_i^{\nu,k+1} > 0$ then, from $\lambda_i^{\nu,k+1} g_i^\nu(x^{k+1}) \to_{\mathcal{N}} 0$, we get $\liminf_{\mathcal{N}} g_i^\nu(x^{k+1}) = 0$, that is, $g_i^\nu(\bar{x}) = 0$.

Case (b): If $\lambda_i^{\nu,k+1} \to_{\mathcal{N}} 0$ for some adequate subsequence. We now suppose by contradiction that $g_i^\nu(\bar{x}) = 2a > 0$ for some $a > 0$. Thus, there is $k_0 \in \mathbb{N}$ such that $g_i^\nu(x^{k+1}) \ge a > 0$, $\forall k \ge k_0$. The latter implies that $\lambda_i^{\nu,k+1} \ge \lambda_i^{\nu,k}$ for every $k \ge k_0$ contradicting the fact that $\lambda_i^{\nu,k+1} \to 0$ along an adequate subsequence. Thus, we conclude that $g_i^\nu(\bar{x}) \le 0$.

We now proceed by proving item (b). Considering the above case, we only need to analyze the case when the sequence $\{\rho_{\nu,k}\}$ is unbounded. We split the proof into two cases:

1. Suppose that $\{\lambda_i^{\nu,k}\}$ goes to infinity. Here, from $L_H^\nu(x^{k+1},\lambda^k,\rho_k)\le B$, one has

$$2\lambda_i^{\nu,k}\max\{g_i(x^{k+1}),0\}\le L_H^\nu(x^{k+1},\lambda^{\nu,k};\rho_{\nu,k})-f^\nu(x^{k+1})\le B-f^\nu(x^{k+1})$$

   Rearranging terms yields

$$\max\{g_i(x^{k+1}),0\}\le \frac{1}{2\lambda_i^{\nu,k}}(B-f^\nu(x^{k+1})),\ \forall k\in\mathbb{N}.$$

   Now, by taking the limit in the above expression, we get $\max\{g_i(\bar{x}),0\}\le 0$.

2. Suppose that $\{\lambda_i^{\nu,k}\}$ is bounded with $\liminf\lambda_i^{\nu,k}>0$. We now assume by contradiction that $g_i^\nu(\bar{x})>0$. Now, as $\liminf\lambda_i^{\nu,k}>0$ and since $\{\rho_{\nu,k}\}$ is a non-decreasing sequence, we can find a scalar $a>0$ such that $\rho_{\nu,k}\lambda_i^{\nu,k}g_i^\nu(x^{k+1})\ge a>0$, for every k large enough. From the update rule for the multipliers, we get $\lambda_i^{\nu,k+1}\ge\lambda_i^{\nu,k}(1+\frac{a}{\sqrt{a^2+1}})$ for every k sufficiently large, and thus $\{\lambda_i^{\nu,k}\}$ must be unbounded, contradicting the assumption. Thus, we get $g_i^\nu(\bar{x})\le 0$.

3. Suppose that $\{\lambda_i^{\nu,k}\}$ is bounded with $\liminf\lambda_i^{\nu,k}=0$. We now assume by contradiction that $g_i^\nu(\bar{x})>0$ for some $a>0$. Similarly, to the case when $\{\rho_{\nu,k}\}$ is bounded, we can find $k_0\in\mathbb{N}$ such that $\lambda_i^{\nu,k+1}>\lambda_i^{\nu,k}$ for every $k\ge k_0$, since $g_i^\nu(x^{k+1})>0$, $\forall k\ge k_0$. Therefore, $\liminf\lambda_i^{\nu,k}>0$ which is a contradiction.

From all the above cases, we conclude that $g_i^\nu(\bar{x})\le 0$. □

Verifying conditions (a) and (b) in Proposition 6.6 is, in general, not straightforward to perform a priori. Nevertheless, one situation in which condition (b) is satisfied is when the iterate x^{k+1} is almost a global minimizer of GNEP, in the sense that there exists a feasible point $\hat{x}$ such that

$$L_H^\nu(x^{k+1},\lambda^{\nu,k};\rho_{\nu,k})\le L_H^\nu(\hat{x},\lambda^{\nu,k};\rho_{\nu,k})+M, \tag{36}$$

for some $M>0$. Let us show this statement. In fact, by definition, we obtain

$$\begin{aligned}L_H^\nu\left(\hat{x},\lambda^{\nu,k};\rho_{\nu,k}\right)&=f^\nu(\hat{x})+\sum_{i=1}^{m_\nu}\left(\lambda_i^{\nu,k}g_i^\nu(\hat{x})+\sqrt{\left(\lambda_i^{\nu,k}g_i^\nu(\hat{x})\right)^2+\frac{1}{\rho_{\nu,k}^2}}\right)\\&\le f^\nu(\hat{x})+\sum_{i:g_i^\nu(\hat{x})<0}\left(\lambda_i^{\nu,k}g_i^\nu(\hat{x})+\sqrt{\left(\lambda_i^{\nu,k}g_i^\nu(\hat{x})\right)^2+\frac{1}{\rho_{\nu,k}^2}}\right)+\sum_{i:g_i^\nu(\hat{x})=0}\frac{1}{\rho_{\nu,0}}.\end{aligned}$$

Let us see that the above expression is bounded. First, straightforward computations show that

$$\lambda_i^{\nu,k}g_i^\nu(\hat{x})+\sqrt{\left(\lambda_i^{\nu,k}g_i^\nu(\hat{x})\right)^2+\frac{1}{\rho_{\nu,k}^2}}=\frac{\rho_{\nu,k}\lambda_i^{\nu,k}g_i^\nu(\hat{x})+\sqrt{(\rho_{\nu,k}\lambda_i^{\nu,k}g_i^\nu(\hat{x}))^2+1}}{\rho_{\nu,k}},\quad\forall i. \tag{37}$$

Take $i\in\{1,\ldots,m_\nu\}$ satisfying $g_i^\nu(\hat{x})<0$ and suppose that (37) is not bounded. Thus, there exists a subsequence such that (37) goes to infinity, and thus $\rho_{\nu,k}\lambda_i^{\nu,k}g_i^\nu(\hat{x})+\sqrt{(\rho_{\nu,k}\lambda_i^{\nu,k}g_i^\nu(\hat{x}))^2+1}\to\infty$. We will show that is not possible. Clearly, it is not possible if $\rho_{\nu,k}\lambda_i^{\nu,k}\to a_i^\nu\in\mathbb{R}_+$ along a subsequence. So, we need to analyze the case when $\rho_{\nu,k}\lambda_i^{\nu,k}\to\infty$ along a subsequence. As $g_i^\nu(\hat{x})<0$, it yields that $\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(\hat{x}) \to -\infty$, but it is well-known that $a + \sqrt{a^2+1} \to 0$ as $a \to -\infty$. Thus, from the all above cases, we get that (37) is bounded. Finally, we point out that (36) holds if x^{k+1} is a global solution of (27).

Proposition 6.7. *Let $\bar{x}$ be the limit of the sequence $\{x^k\}$ not necessary feasible. Then, for every $\nu = 1, \ldots, N$, we have that $\bar{x}$ satisfies $g_i^\nu(\bar{x}) \le 0$, $\forall i = 1, \ldots, m_\nu$, or there exist a non-null vector of multipliers $\bar{\lambda}^\nu \in \mathbb{R}_+^{m_\nu}$ such that*

$$\sum_{i=1}^{m_\nu} \bar{\lambda}_i^\nu \nabla_{x^\nu} g_i^\nu(\bar{x}) = 0 \quad \textit{with} \quad \bar{\lambda}_i^\nu \min\{g_i^\nu(\bar{x}), 0\} = 0, \ \forall i = 1, \ldots, m_\nu.$$

Proof. Let $\nu \in \{1, \ldots, N\}$. We split the proof into several cases:

(a) Suppose that $\{\rho_{\nu,k}\}$ is bounded. So, by Proposition 6.6, the point $\bar{x}$ satisfies $g_i^\nu(\bar{x}) \le 0$, $\forall i = 1, \ldots, m_\nu$;

(b) Assume that $\{\rho_{\nu,k}\}$ is unbounded. Here, we denote $M_{k+1}^\nu := \max\{\lambda_i^{\nu,k+1} : i = 1, \ldots, m_\nu\}$. Now, we consider two cases:

(b.1) Assume that $\{M_{k+1}^\nu\}$ is bounded. Take $i \in \{1, \ldots, m_\nu\}$ assume by contradiction that $g_i^\nu(\bar{x}) > 0$. Now, if $\liminf \lambda_i^{\nu,k+1} > 0$ and since $\{\rho_{\nu,k}\}$ is a non-decreasing, there exists $a > 0$ such that $\rho_{\nu,k+1}\lambda_i^{\nu,k+1} g_i^\nu(x^{k+2}) \ge a > 0$, for every k large enough. Therefore, by the update rule for the multipliers, we get $\lambda_i^{\nu,k+2} \ge \lambda_i^{\nu,k+1}(1 + \frac{a}{\sqrt{a^2+1}})$ for every k sufficiently large, and thus $\{\lambda_i^{\nu,k+1}\}$ must be unbounded, contradicting the assumption.

Now, if $\liminf \lambda_i^{\nu,k+1} = 0$ then, we can find $k_0 \in \mathbb{N}$ such that $\lambda_i^{\nu,k+1} > \lambda_i^{\nu,k}$ for every $k \ge k_0$, since $g_i^\nu(x^{k+1}) > 0$, $\forall k \ge k_0$. Thus, we have $\liminf \lambda_i^{\nu,k} > 0$ which is a contradiction. For all the case, we have that $g_i^\nu(\bar{x}) \le 0$, $\forall i = 1, \ldots, m_\nu$;

(b.2) Suppose that $\{M_{k+1}^\nu\}$ is a unbounded sequence. So, we can assume, without loss of generality, that $M_{k+1}^\nu \to \infty$. Since x^{k+1} is an approximate solution, expression (35) holds and by dividing this expression with M_{k+1}^ν, one has:

$$0 \in \frac{\nabla_{x^\nu} f^\nu(x^{k+1})}{M_{k+1}^\nu} + \sum_{i=1}^{m_\nu} \frac{\lambda_i^{\nu,k+1}}{M_{k+1}^\nu} \nabla_{x^\nu} g_i^\nu(x^{k+1}) + \frac{e^{\nu,k+1}}{M_{k+1}^\nu} + N_{C_\nu}(x^{\nu,k+1}) \text{ with } \|e^{\nu,k+1}\| \le \epsilon_k,$$

Furthermore, without loss of generality, we assume that $\lambda_i^{\nu,k+1}/M_{k+1}^\nu \to \bar{\lambda}_i^\nu$, $\forall i$, and by taking the limit in the above expression, one has

$$0 \in \sum_{i=1}^{m_\nu} \bar{\lambda}_i^\nu \nabla_{x^\nu} g_i^\nu(\bar{x}) + N_{C_\nu}(\bar{x}^\nu).$$

We now prove that $\bar{\lambda}_i^\nu \min\{g_i^\nu(\bar{x}), 0\} = 0$, $\forall i$. Following the proof of Proposition 6.5 and (34), it is possible to get

$$\frac{1+\sqrt{2}}{\rho_{\nu,k} M_{k+1}^\nu} + \frac{\lambda_i^{\nu,k+1}}{M_{k+1}^\nu} g_i^\nu(x^{k+1}) \ge \frac{h(\rho_{\nu,k}\lambda_i^{\nu,k} g_i^\nu(x^{k+1}))}{\rho_{\nu,k} M_{k+1}^\nu} + \frac{\max\{\lambda_i^{\nu,k} g_i^\nu(x^{k+1}), 0\}}{M_{k+1}^\nu}.$$

The above yields that $\bar{\lambda}_i^\nu \min\{g_i^\nu(\bar{x}), 0\} = 0$, $\forall i$. ∎

6.3 On the boundedness of multipliers of the HALA method

In this subsection, we study the assumptions under which we can ensure the boundedness of the dual sequence generated by augmented Lagrangian methods for smooth functions.

Definition 6.1. *We say that $\bar{x}$ satisfies the GNEP strong-quasinormality condition, if for any $\nu = 1, \dots, N$, and any scalars $\lambda^\nu \in \mathbb{R}^{m_\nu}_+$ satisfying*

$$\sum_{i=1}^{m_\nu} \lambda_i^\nu \nabla_{x^\nu} g_i^\nu(\bar{x}) = 0 \quad \textit{and} \quad \lambda_i^\nu g_i^\nu(\bar{x}) = 0, \forall i = 1, \dots, m_\nu,$$

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

Our main proposition is given by the following proposition.

Proposition 6.8. *Let $\bar{x}$ be the limit of the sequence $\{x^k\}$ generated by Algorithm 1 with $\bar{x} \in \Omega$. Suppose that $\bar{x}$ satisfies the GNEP-strong quasinormality, then the sequences $\{\lambda^{\nu,k}\}$, $\nu = 1, \dots, N$ are bounded. In particular, $\bar{x}$ is a KKT point of* (GNEP).

Proof. Let $\nu = 1, \dots, N$, and set $M_{k+1}^\nu := \|(1, \lambda^{\nu,k+1})\|$. Suppose that M_{k+1}^ν is an unbounded sequence. So, we can assume, without loss of generality, that $M_{k+1}^\nu \to \infty$. Since x^{k+1} is an approximate solution, expression (35) holds and by dividing this expression with M_{k+1}^ν, one has:

$$0 \in \frac{\nabla_{x^\nu} f^\nu(x^{k+1})}{M_{k+1}^\nu} + \sum_{i=1}^{m_\nu} \frac{\lambda_i^{\nu,k+1}}{M_{k+1}^\nu} \nabla_{x^\nu} g_i^\nu(x^{k+1}) + \frac{e^{\nu,k+1}}{M_{k+1}^\nu} + N_{C_\nu}(x^{\nu,k+1}) \text{ with } \|e^{\nu,k+1}\| \le \epsilon_k,$$

Furthermore, without loss of generality, we assume that $\lambda_i^{\nu,k+1}/M_{k+1}^\nu \to \bar{\lambda}_i^\nu \ge 0$, $\forall i$. Following the proof of item (b.2) of Proposition 6.7 and by taking the limit in the above expression, one has

$$0 \in \sum_{i=1}^{m} \bar{\lambda}_i^\nu \nabla_{x^\nu} g_i^\nu(\bar{x}) + N_{C_\nu}(\bar{x}^\nu) \text{ and } \bar{\lambda}_i^\nu \min\{g_i^\nu(\bar{x}), 0\} = 0, \forall i.$$

Clearly, as $\bar{x}$ is feasible, we have $\bar{\lambda}_i^\nu g_i^\nu(\bar{x}) = 0, \forall i$. Now, for every $i = 1, \dots, m_\nu$ satisfying $\bar{\lambda}_i^\nu > 0$, we must have that $\lambda_i^{\nu,k+1} \to \infty$. By Proposition 6.2, there exists a subsequence $\mathcal{N}_+(i) \subset \mathbb{N}$ such that $g_i^\nu\left(x^{\ell+1}\right) > 0$, $\forall \ell \in \mathcal{N}_+(i)$. Clearly, $x^{\ell+1} \to_{\mathcal{N}_+(i)} \bar{x}$. The existence of such sequences together with $\bar{\lambda}_i^\nu > 0$ contradicts the definition of GNEP strong-quasinormality. Thus, $\left\{\lambda^{\nu,k+1}\right\}$ is a bounded sequence. □

7 A Safeguarded Augmented Lagrangian method

In the previous section, we showed that the sequential approximate conditions are useful for analyzing the global convergence of the HALA method. A natural question is whether these conditions can also be employed in the analysis of other methods for (GNEP). In this section, we address this question by refining the general convergence theory of the scaled augmented Lagrangian method applied to (GNEP). In particular, we extend and strengthen the results established in [19] and [14]

For each player $\nu = 1, ..., N$, and given $u \in \mathbb{R}^{m_\nu}_+$ and $\rho_\nu > 0$, the *augmented Lagrangian* function for the player ν is defined as

$$L_a^\nu(x,u;\rho) := f^\nu(x) + \frac{\rho}{2}\sum_{i=1}^{m_\nu} \max\left\{g_i^\nu(x) + \frac{u}{\rho}, 0\right\}^2.$$

We consider the safeguarded augmented Lagrangian method presented in Algorithm 2.

Algorithm 2 A safeguarded Augmented Lagrangian (SALA) method

Inputs: Parameters $\tau_\nu > 1$, $\sigma \in (0,1)$, and $\rho_0 > 0$; a bounded set $\mathbb{B}_\nu$, and initial points $(x^0, \lambda^0, u^0) \in \mathbb{E} \times \mathbb{R}_{++}$ with $u^{\nu,0} \in \mathbb{B}_\nu$. Set $k := 0$.

1. If (x^k, λ^k) satisfies some stopping criteria of the GNEP: STOP

2. Compute an approximate solution of $x^{k+1} \in C = C_1 \times \cdots \times C_N$ of GNEP consisting of minimization subproblems

$$\min_{x^\nu} \; L_a^\nu\left(x^\nu, x^{-\nu}, u^{\nu,k}; \rho_{\nu,k}\right) \text{ subject to } x^\nu \in C_\nu, \tag{38}$$

   for each player $\nu = 1, ..., N$. In our case, we must find $x^{k+1} \in \mathbb{R}^n$ and $e^{\nu,k+1} \in \mathbb{R}^n$ such that

$$0 \in \nabla_{x^\nu} L_a^\nu(x^{k+1}, u^{\nu,k}; \rho_{\nu,k}) + e^{\nu,k+1} + N_{C_\nu}(x^{\nu,k+1}) \text{ with } \|e^{\nu,k+1}\| \le \epsilon_k, \; \forall\nu. \tag{39}$$

3. For $\nu = 1, \ldots, N$, update the vector of multipliers to $\lambda^{\nu,k} = (\lambda_1^{\nu,k}, \ldots, \lambda_{m_\nu}^{\nu,k})$ in an additive way by the rule:

$$\lambda_i^{\nu,k+1} := \max\{u_i^{\nu,k} + \rho_{\nu,k} g_i^\nu(x^{k+1}), 0\}, \text{ for all } i = 1, \ldots, m_\nu. \tag{40}$$

4. For $\nu = 1, \ldots, N$. If $k = 0$ or $\|\min\{-g^\nu(x^{k+1}), \lambda^{\nu,k+1}\}\|_\infty \le \sigma_\nu \|\min\{-g^\nu(x^k), \lambda^{\nu,k}\}\|_\infty$. Set $\rho_{\nu,k+1} := \rho_{\nu,k}$. Otherwise, take $\rho_{\nu,k+1} \ge \tau_\nu \rho_{\nu,k}$.

5. Set $u^{\nu,k+1} \in \mathbb{B}_\nu$, $k \leftarrow k+1$, and go to Step 1.

We start our analysis of the convergence with the following statement.

Theorem 7.1. *Let $\bar{x}$ be a feasible limit point of the sequence $\{x^k\}$. Then, $\bar{x}$ is a AGP-GNEP and PAKKT-GNEP point.*

Proof. First, due to the update rule in Step 3, it is not difficult to see that $\nabla_{x^\nu} L_a^\nu(x^{k+1}, u^{\nu,k}; \rho_{\nu,k}) = \nabla_{x^\nu} L^\nu(x^{k+1}, \lambda^{\nu,k+1})$. We now will show that $\bar{x}$ is a AGP-GNEP point. From (39), one has

$$0 \in \nabla_{x^\nu} L^\nu(x^{k+1}, \lambda^{\nu,k+1}) + e^{\nu,k+1} + N_{C_\nu}(x^{\nu,k+1}) \text{ with } \|e^{\nu,k+1}\| \le \epsilon_k, \; \forall\nu.$$

So, we need to show that $\lambda_i^{\nu,k+1} \min\{g_i^\nu(x^{k+1}), 0\} \to 0$, $\forall i = 1, \ldots, m_\nu$. We split the proof into two cases depending on if $\{\rho_{\nu,k}\}$ is bounded or not.

(a) If $\{\rho_{\nu,k}\}$ is bounded then $\{\lambda^{\nu,k+1}\}$ is bounded and $\min\{-g_i^\nu(x^{k+1}), \lambda_i^{\nu,k+1}\} \to 0$, $\forall i$. The above yields that $\lambda_i^{\nu,k+1} \to 0$ if $g_i^\nu(\bar{x}) < 0$ and so $\lambda_i^{\nu,k+1} \min\{g_i^\nu(x^{k+1}), 0\} \to 0$. Now, if $g_i^\nu(\bar{x}) = 0$ then $\min\{g_i^\nu(x^{k+1}), 0\} \to 0$, and as $\lambda_i^{\nu,k+1}$ is bounded, we get $\lambda_i^{\nu,k+1} \min\{g_i^\nu(x^{k+1}), 0\} \to 0$.

(b) If $\{\rho_{\nu,k}\}$ is unbounded then, without loss of generality, we assume that $\rho_{\nu,k} \to \infty$. Thus, as $u^{\nu,k}$ is bounded and hence $u^{\nu,k}/\rho_{\nu,k} \to 0$, one has $\lambda_i^{\nu,k+1} = \max\left\{u^{\nu,k}\rho_{\nu,k}^{-1} + g_i^\nu(x^{k+1}), 0\right\} \to 0$ whenever $g_i^\nu(\bar{x}) < 0$. We now analyze the case when $g_i^\nu(\bar{x}) = 0$. Thus, we observe that

$$
\begin{aligned}
\lambda_i^{\nu,k+1}|\min\{g_i^\nu(x^{k+1}),0\}| &= \max\left\{u^{\nu,k} + \rho_{\nu,k} g_i^\nu(x^{k+1}), 0\right\} |\min\{g_i^\nu(x^{k+1}),0\}| \\
&= \left|\rho_{\nu,k}\max\left\{\frac{u^{\nu,k}}{\rho_{\nu,k}} + g_i^\nu(x^{k+1}), 0\right\}\min\{g_i^\nu(x^{k+1}),0\} - \rho_{\nu,k}\max\left\{g_i^\nu(x^{k+1}),0\right\}\min\left\{g_i^\nu(x^{k+1}),0\right\}\right| \\
&\le \rho_{\nu,k}\left|\max\left\{\frac{u^{\nu,k}}{\rho_{\nu,k}} + g_i^\nu(x^{k+1}), 0\right\} - \max\left\{g_i^\nu(x^{k+1}),0\right\}\right| \left|\min\left\{g_i^\nu(x^{k+1}),0\right\}\right| \\
&\le \rho_{\nu,k}\left\|\frac{u^{\nu,k}}{\rho_{\nu,k}}\right\| \left|\min\left\{g_i^\nu(x^{k+1}),0\right\}\right| = \left\|u^{\nu,k}\right\| \left|\min\left\{g_i^\nu(x^{k+1}),0\right\}\right| \to 0,
\end{aligned}
$$

since $u^{\nu,k}$ is bounded, $|\max\{u,0\} - \max\{v,0\}| \le \|u - v\|$ $(\forall u, v)$, and $|\min\left\{g_i^\nu(x^{k+1}),0\right\}| \to 0$.

To prove that $\bar{x}$ is PAKKT-GNEP point, we assume that $\delta_k := \|(1, \lambda^{\nu,k+1})\| \to \infty$. Suppose that $\lambda_i^{\nu,k+1}/\delta_{k+1}^\nu \to \bar{\lambda}_i^\nu > 0$. So, for every k large enough, we have

$$
\begin{aligned}
0 < \frac{1}{2}\bar{\lambda}_i^\nu \le \frac{\lambda_i^{\nu,k+1}}{\delta_{k+1}^k} = \frac{\max\{u^{\nu,k} + \rho_{\nu,k} g_i^\nu(x^{k+1}), 0\}}{\delta_{k+1}^\nu} &= \max\left\{\frac{u^{\nu,k}}{\delta_{k+1}^\nu} + \frac{\rho_{\nu,k}}{\delta_{k+1}^\nu} g_i^\nu(x^{k+1}), 0\right\} \\
&\le \max\left\{\frac{\rho_{\nu,k}}{\delta_{k+1}^\nu} g_i^\nu(x^{k+1}), 0\right\} + \frac{\|u^{\nu,k}\|}{\delta_{k+1}^\nu}.
\end{aligned}
$$

The latter implies that $g_i^\nu(x^{k+1}) > 0$ for k sufficiently large. Thus, $\bar{x}$ is a PAKKT-GNEP point. □

For nonlinear optimization problems, it is known that the safeguarded augmented Lagrangian method converges to CAKKT points under an assumption related to the Łojasiewicz (GL) property of the associated infeasibility measure. In the absence of this property, convergence to CAKKT points cannot, in general, be guaranteed. For further details, we refer the reader to [4, 6, 7] and the references therein.

For (GNEP), we establish an analogous convergence result under a corresponding assumption tailored for (GNEP), see (GNEP-KL). Thus, for every $\nu = 1, \ldots, N$, we consider the following measure of infeasible given by the function

$$
\Phi^\nu(x) := \frac{1}{2}\sum_{i=1}^{m_\nu}(\max\{g_i^\nu(x), 0\})^2.
$$

We emphasize that, unlike in previous works, we explicitly consider abstract constraints C_ν in its formulation. Our main result is given in the following statement.

Theorem 7.2. *Let $\bar{x}$ be a feasible limit point of the sequence $\{x^k\}$ generated by the safeguarded augmented Lagrangian method. Suppose that for all $\nu = 1, \ldots, N$, there exists a continuous function ϕ_ν with $\phi_\nu(x) \to 0$ as $x \to \bar{x}$ and a neighborhood of $\bar{x}$ such that*

$$
|\Phi^\nu(x) - \Phi^\nu(\bar{x})| \le \phi_\nu(x)\, dist(\nabla_{x^\nu}\Phi^\nu(x) + N_{C_\nu}(x), 0). \qquad \text{(GNEP-KL)}
$$

Then $\bar{x}$ is a PCAKKT-GNEP point.

Proof. The proof is similar to [4, Theorem 5.1] but with the proper modifications. Let us show that $\lambda_i^{\nu,k} g_i^\nu(x^k) \to 0$, $\forall i$ and $\forall \nu$. We observe that

$$0 = \nabla_{x^\nu} f^\nu(x^k) + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^\nu(x^k) + e^{\nu,k} + \eta^{\nu,k} \text{ with } \eta^{\nu,k} \in N_{C_\nu}(x^{\nu,k}). \tag{41}$$

First, we will show that $\rho_{\nu,k}(\max\{g_i^\nu(x^k), 0\})^2 \to 0,\ \forall i = 1, \dots, m_\nu$. We start with the following observation:

$$|\lambda_i^{\nu,k} - \rho_{\nu,k} \max\{g_i^\nu(x^k), 0\}| = |\max\{u_i^{\nu,k} + \rho_{\nu,k} g_i^\nu(x^k), 0\} - \max\{\rho_{\nu,k} g_i^\nu(x^k), 0\}| \leq \|u^{\nu,k}\| \tag{42}$$

So, let M be a scalar such that

$$\left\| \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^\nu(x^k) - \sum_{i=1}^{m_\nu} \rho_{\nu,k} \max\{g_i^\nu(x), 0\} \nabla_{x^\nu} g_i^\nu(x^k) \right\| \leq \sum_{i=1}^{m_\nu} \|u^{\nu,k}\| \|\nabla_{x^\nu} g_i^\nu(x^k)|\| \leq M$$

for every k large enough. Now, by multiplying $\rho_{\nu,k}$ on both sides of (GNEP-KL), we get

$$\begin{aligned} \left| \sum_{i=1}^{m_\nu} \rho_{\nu,k} (\max\{g_i^\nu(x^k), 0\})^2 \right| &\leq 2\phi_\nu(x^k) \left\| \sum_{i=1}^{m_\nu} \rho_{\nu,k} \max\{g_i^\nu(x), 0\} \nabla_{x^\nu} g_i^\nu(x^k) + \eta_\nu \right\| \\ &\leq 2\phi_\nu(x^k) \left(\left\| \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^\nu(x^k) + \eta_\nu \right\| + M \right) \\ &\leq 2\phi_\nu(x^k) \left(\|\nabla_{x^\nu} f^\nu(x^k) + e^{\nu,k}\| + M \right). \end{aligned}$$

From $\phi_\nu(x^k) \to 0$, the above yields $\rho_{\nu,k}(\max\{g_i^\nu(x^k), 0\})^2 \to 0,\ \forall i = 1, \dots, m_\nu$.

We split the proof into two cases:

1. If $\{\rho_{\nu,k}\}$ is bounded for some ν. Then in this case, by item (a) of the proof of Theorem 7.1, we get $\lambda_i^{\nu,k} g_i^\nu(x^k) \to 0$, $\forall i$;

2. If $\{\rho_{\nu,k}\}$ is unbounded for some ν. Without loss of generality, we assume that $\rho_{\nu,k} \to 0$. As the proof of Theorem 7.1, we see $\lambda_i^{\nu,k} \to 0$ if $g_i^\nu(\bar{x}) < 0$. So, we will analyze the case that $g_i^\nu(\bar{x}) = 0$. From (42) and $\min\{g_i^\nu(x^k), 0\} \max\{g_i^\nu(x^k), 0\} = 0$, $\forall k$, we get

$$\begin{aligned} &|\lambda_i^{\nu,k} \max\{g_i^\nu(x^k), 0\}| \leq \rho_{\nu,k} \max\{g_i^\nu(x^k), 0\}^2 + \|u^{\nu,k}\| \max\{g_i^\nu(x^k), 0\} \to 0, \text{ and} \\ &|\lambda_i^{\nu,k} \min\{g_i^\nu(x^k), 0\}| \leq \|u^{\nu,k}\| |\min\{g_i^\nu(x^k), 0\}| \to 0. \end{aligned}$$

   The above implies $\lambda_i^{\nu,k} g_i^\nu(x^k) = \lambda_i^{\nu,k} \max\{g_i^\nu(x^k), 0\} + \lambda_i^{\nu,k} \min\{g_i^\nu(x^k), 0\} \to 0$.

In any case, we have that $\bar{x}$ is a CAKKT-GNEP point. □

In particular, if every feasible limit point $\bar{x}$ of the sequence generated by the safeguarded augmented Lagrangian method satisfies (GNEP-KL) then $\bar{x}$ is a CAKKT point.

8 Numerical experience

In this section, we report computational results for the proposed HALA algorithm. We also compare its performance with that of other methods available in the literature. To this end, we implemented HALA in Python, and all experiments were carried out on a Windows PC equipped with an Intel (R) Core (TM) i5-1035G1 CPU @ 1.00GHz (1.20GHz), 8.00 GB of RAM, and a 64-bit operating system.

Parameter initialization

In [13] (Section 4.2), it is suggested that suitable initial parameters for HALA are, a penalty parameter of 0.01 and initial Lagrange multipliers of 100. This choice is motivated by the fact that the multiplier update in HALA is multiplicative (see Algorithm 2.1 in [13] and equation (28) in this work), rather than additive as in the quadratic penalty method (see equation (13) in [19], and equation (40) in this work).

Nevertheless, in our experiments, we adopt initial parameter values similar to those used in [19], where they are analyzed and shown to be suitable for the quadratic penalty method. This choice also facilitates a direct comparison with results reported in the literature under the same initial conditions. Unless stated otherwise, the penalty parameters and Lagrange multipliers are initialized according to the guideline in [19], that is,

$$\rho_{\nu,0} = 10, \qquad \lambda^{\nu,0} = 1, \qquad \forall \nu = 1, \ldots, N,$$

where the penalty parameters are updated independently for each player We follow the recommendations in the benchmark description, and thus, the update parameters are chosen based on the total dimension n as

$$\sigma_\nu = \begin{cases} 0.1, & \text{if } n \le 100, \\ 0.5, & \text{if } n > 100, \end{cases} \qquad \tau_\nu = \begin{cases} 10, & \text{if } n \le 100, \\ 2, & \text{if } n > 100. \end{cases}$$

This choice yields a more aggressive penalization mechanism for small-scale problems, while inducing a milder update strategy for large-scale instances. In our implementation, this rule is enforced uniformly across all test problems. The initial primal points x^0 are chosen as specified in the benchmark definitions (see Table 1 in [19]). When necessary to avoid numerical singularities (e.g., zero or boundary starting points in the presence of logarithmic or ratio-type expressions), we adopt a warm-start strategy by slightly perturbing the initial point to obtain a strictly interior point.

Stopping criteria and iteration counter

At each outer iteration, three residual measures are monitored:

- a primal feasibility residual $R_f = \max_{\nu=1,\ldots,N} \|max\{g^\nu(x), 0\}\|_\infty$, measuring the maximum constraint violation;
- a stationarity residual $R_o = \max_{\nu=1,\ldots,N} \|\nabla_{x^\nu} f^\nu(x) + \sum_{i=1}^{m_\nu} \lambda_i^{\nu,k} \nabla_{x^\nu} g_i^\nu(x)\|_\infty$,, measuring the infinity norm of the concatenated KKT stationarity conditions;

- a complementarity residual $R_c = \max_{\nu=1,\dots,N} |\langle g^\nu(x), \lambda^k \rangle|$, measuring the maximum magnitude of the multiplier–constraint products.

The algorithm is declared converged only if

$$R_f \le \varepsilon, \quad R_o \le \varepsilon \;\text{ and }\; R_c \le \varepsilon,$$

with $\varepsilon = 10^{-6}$ in all experiments. To compare computational effort across problems of different sizes, we report a uniform iteration counter i_{tot}, defined as the cumulative number of effective inner iterations of the nonlinear solver. Concretely, if n_{fev} denotes the number of function evaluations performed by the root solver at a given outer iteration and n is the dimension of the concatenated system, then

$$i_{\text{tot}} \;\mathrel{+}= \left\lceil \frac{n_{\text{fev}}}{n+1} \right\rceil,$$

which provides an approximation of the total number of inner Newton-type steps, consistent with the reporting style adopted in the benchmark reference.

Solution of the subproblems

At each outer iteration of HALA, the nonlinear system associated with the concatenated stationarity conditions (Step 2) is solved using Powell's hybrid method (`hybr`) provided by `scipy.optimize.root` (see [22]). This solver combines Newton steps with a trust-region safeguard and finite-difference Jacobian approximations, offering a practical balance between robustness and fast local convergence.

In several challenging instances (notably A.4, A.6, A.7, A.8, and A.10), Powell's hybrid method was observed to stagnate or to produce poor search directions when the penalty parameters became large. In such cases, a fallback to the Levenberg–Marquardt variant (`lm`) was also tested; however, no consistent improvement in the convergence behavior was observed. For this reason, and to ensure a uniform basis for comparison, all results reported in Table 1 correspond to runs performed with the hybrid method.

8.1 Benchmark results and discussion

In this section, we report the results obtained by our algorithm and compare them with those produced by the following methods:

- Facchinei and Kanzow (2010), [16]: based on an exact penalty algorithm. The results for the following cases are not reported: A.2 with $x^0 = 1$ and the A.7 with $x^0 = 10$;
- Jordan et al. (2023), [18]: based on the Accelerated Mirror-Prox Augmented Lagrangian Method (AMPAL). The results for the following cases are not reported: A.10, A.2 with $x^0 = 1$, A.3 with $x^0 = 0$ and $x^0 = 1$, A.8 with $x^0 = 0$ and A.16d with $x^0 = 10$.
- Kanzow and Steck (2016), [19]: based on a quadratic augmented Lagrangian algorithm. The results for A.8 with $x^0 = 0$ are not reported.

The methods proposed in [18] and [19] are essentially based on a quadratic penalty framework. Our computational comparison of HALA with other algorithms therefore focuses primarily on the test problems for which HALA produced results. We now discuss the numerical results and the corresponding comparisons reported in the tables below From Table 1, we observe the following:

- We report the numerical performance of HALA on the complete benchmark set. For each problem, we provide the number of players N, the total dimension n, the initial point x^0, the number of outer iteration k, the total inner iteration i_{tot}, the final residuals (R_f, R_o, R_c), and the maximum penalty parameter $\rho_{\max}$ attained during the run.

- The experiments indicate that HALA converges reliably on a substantial portion of the test set, including Problems A.1, A.2, A.3, A.5, A.9, A.11, A.12, A.13, A.14, A.15, A.16, A.17, and A.18, attaining residuals close to machine precision. In these instances, the smooth subproblem formulation allows for an efficient solution of the stationarity system and stable updates of the penalty parameters.

- On the other hand, several challenging instances remain unsolved under the present configuration. In Problems A.4, A.6, A.7, A.8, and A.10, the stationarity residual R_o fails to decrease despite repeated penalty updates, leading to unbounded growth of the penalty parameters and eventual stagnation. These observations motivate a more in-depth investigation aimed at identifying suitable parameter choices for the proposed algorithm.

- Overall, the numerical study indicates that HALA is a viable and accurate method for a broad class of smooth GNEPs. However, additional stabilization mechanisms or alternative penalty update strategies may be required to handle highly nonlinear or ill-conditioned benchmark problems in a reliable manner.

From Table 2, we observe the following:

- This table reports the primal residual R_f, which measures the maximum constraint violation. For Problems A.1, A.3, A.5, A.9a, A.11, A.12, A.13, A.14, A.15, A.16 and A.18, HALA achieves feasibility values close to zero, in many cases reaching machine precision. In contrast, the methods proposed in [18] and [16] exhibit significantly larger residuals for several instances. Overall, our results are highly competitive with those reported in [19].

From Table 3, we observe the following:

- It reports the optimality residual R_o associated with the concatenated KKT stationarity conditions. HALA consistently attains values on the order of 10^{-13} to 10^{-16} for several problems, generally outperforming first-order methods and yielding results comparable to, and in some cases better than, those reported for the algorithm in [19].

9 Conclusions and remarks

In recent years, significant progress has been made in the convergence analysis of augmented Lagrangian methods for nonlinear programming (NLP) based on so-called sequential optimality conditions. However, comparatively little is known for (GNEP). In this work, we extend these concepts to (GNEP) and show how they can be employed to analyze the global convergence of several augmented Lagrangian methods. Although our stationarity concepts are not, in general, optimality conditions, we prove that they lead to convergence to a KKT point under very weak CQs, thereby improving upon previous results.

We proposed a novel augmented Lagrangian scheme based on the hyperbolic penalty function for solving (GNEP), referred to as the HALA method. By employing our approximate stationarity

Table 1: Results reported by the HALA algorithm

Ex.	N	n	x^0	k	i_{tot}	R_f	R_o	R_c	$\rho_{\max}$
A.1	10	10	0.01	7	32	0.00e+00	9.65e-13	4.20e-07	1.00e+07
			0.1	7	29	0.00e+00	9.67e-13	4.20e-07	1.00e+07
			1.0	7	36	0.00e+00	9.28e-13	4.20e-07	1.00e+07
A.2	10	10	0.01	18	41	4.7e-07	1.1e-09	8.8e-07	7e+05
			0.1	18	40	4.7e-07	1.1e-09	8.8e-07	7e+05
			1.0						
A.3	3	7	0	8	16	0.00e+00	4.76e-15	3.63e-07	1.00e+06
			1	8	16	0.00e+00	4.44e-15	3.63e-07	1.00e+06
			10	8	16	0.00e+00	4.22e-15	3.63e-07	1.00e+06
A.5	3	7	0	9	24	1.04e-13	4.44e-16	1.30e-07	1.00e+08
			1	9	24	1.04e-13	4.44e-16	1.30e-07	1.00e+08
			10	9	25	1.04e-13	4.44e-16	1.30e-07	1.00e+08
A.9a	7	56	0	18	41	7.39e-12	1.93e-08	7.42e-07	1.95e+06
A.9b	7	112	0						
A.11	2	2	0	9	24	0.00e+00	2.22e-16	9.26e-07	1.00e+01
A.12	2	2	0	13	29	0.00e+00	1.44e-16	7.42e-07	7.81e+05
A.13	3	3	0	8	39	0.0000e+00	3.5200e-09	2.4742e-07	1000000.0
A.14	10	10	0.01	9	18	0.00e+00	7.71e-17	3.55e-07	1.00e+06
A.15	3	6	0	9	22	0.00e+00	7.85e-14	1.09e-07	1.00e+07
A.16a	5	5	10	9	91	0.00e+00	1.36e-07	2.25e-07	1.00e+06
A.16b	5	5	10	11	91	0.00e+00	1.49e-07	2.41e-07	1.00e+06
A.16c	5	5	10	10	47	0.00e+00	1.51e-08	2.42e-07	1.00e+06
A.16d	5	5	10	10	30	0.00e+00	4.63e-09	2.40e-07	1.00e+06
A.17	2	3	0	18	82	2.0604e-11	7.4633e-09	8.8097e-07	327680.0
A.18	2	12	0	8	218	0.00e+00	5.84e-07	1.87e-07	1.00e+06
			1	8	218	0.00e+00	4.86e-07	1.87e-07	1.00e+06
			10	8	239	0.00e+00	5.74e-07	1.87e-07	1.00e+06

Table 2: Comparison of feasibility value R_f with other algorithms

Ex.	N	n	x^0	**[16]**	**HALA**	**[18]**	**[19]**
A.1	10	10	0.01	7.6e-4	0.00e+00	4.7e-05	1.5e-10
			0.1	7.6e-4	0.00e+00	2.9e-05	8e-09
			1.0	7.6e-4	0.00e+00	4.0e-08	1.5e-10
A.2	10	10	0.01	6.2e-4	4.7e-07	2.9e-05	4.7e-09
			0.1	9.1e-4	4.7e-07	9.6e-05	2.9e-09
			1.0				4.9e-10
A.3	3	7	0	0	0.00e+00		0
			1	0	0.00e+00		0
			10	0	0.00e+00	4.3e-07	0
A.5	3	7	0	4.3e-4	1.04e-13	2.3e-05	2e-10
			1	4.3e-4	1.04e-13	9.0e-06	3.5e-10
			10	4.3e-4	1.04e-13	3.7e-05	6.9e-09
A.9a	7	56	0	3.7e-3	7.39e-12	8.2e-06	2.3e-09
A.9b	7	112	0	9.6e-4		4.2e-06	2.8e-10
A.11	2	2	0	9.6e-5	0.00e+00	8.0e-05	6.4e-09
A.12	2	2	0	0	0.00e+00	0	0
A.13	3	3	0	1.5e-4	0.0000e+00	0	3.3e-09
A.14	10	10	0.01	0	0.00e+00	0	0
A.15	3	6	0	0	0.00e+00	0	0
A.16a	5	5	10	3.9e-4	0.00e+00	1.1e-06	1.3e-10
A.16b	5	5	10	3.1e-4	0.00e+00	1.5e-06	6.1e-11
A.16c	5	5	10	2.9e-4	0.00e+00	7.1e-06	9e-10
A.16d	5	5	10	2.3e-4	0.00e+00		4e-09
A.17	2	3	0	1.9e-4	2.0604e-11	0	4.5e-11
A.18	2	12	0	0.001	0.00e+00	4.5e06	1.3e-11
			1	5.2e-4	0.00e+00	4.5e-06	1.3e-11
			10	5.7e-4	0.00e+00	4.5e-06	1.3e-11

Table 3: Comparison of stationarity value R_0 with other algorithms

Ex.	N	n	x^0	**HALA**	**[18]**	**[19]**
A.1	10	10	0.01	9.65e-13	1.9e-06	8.9e-16
			0.1	9.67e-13	1.9e-06	5.9e-13
			1.0	9.28e-13	1.3e-06	2.9e-16
A.2	10	10	0.01	1.1e-09	1.0e-06	2.3e-09
			0.1	1.1e-09	2.8e-06	4.1e-14
			1.0			5.6e-14
A.3	3	7	0	4.76e-15		1e-09
			1	4.44e-15		3.6e-15
			10	4.22e-15	9.9e-07	1.7e-10
A.5	3	7	0	4.44e-16	9.9e-07	1.7e-13
			1	4.44e-16	9.9e-07	4.9e-13
			10	4.44e-16	1.1e-06	1e-13
A.9a	7	56	0	1.93e-08	1.0e-06	8e-15
A.9b	7	112	0		8.2e-07	1e-14
A.11	2	2	0	2.22e-16	2.3e-06	2.9e-15
A.12	2	2	0	1.44e-16	2.9e-06	8.9e-16
A.13	3	3	0	3.5200e-09	8.9e-05	7.6e-12
A.14	10	10	0.01	7.71e-17	1.2e-06	8.2e-14
A.15	3	6	0	7.85e-14	1.6e-05	2.8e-14
A.16a	5	5	10	1.36e-07	1.0e-06	6e-14
A.16b	5	5	10	1.49e-07	2.4e-06	3.6e-15
A.16c	5	5	10	1.51e-08	1.7e-06	1.5e-13
A.16d	5	5	10	4.63e-09		2.1e-14
A.17	2	3	0	7.4633e-09	1.3e-06	3.4e-13
A.18	2	12	0	5.84e-07	1e-06	1.1e-11
			1	4.86e-07	1e-06	1.2e-11
			10	5.74e-07	1e-06	1.8e-11

concepts, we established convergence properties of the method with respect to both feasibility and optimality of its limit points. Moreover, we proved boundedness of the associated multipliers under a constraint qualification inspired by quasinormality.

Computational experiments were conducted to assess the practical performance of the proposed method. The results indicate robust convergence across a broad class of GNEPs, including nonconvex instances. Tables 2 and 3 suggest that HALA is a promising approach. Nevertheless, further computational investigations regarding the selection and tuning of its initial parameters are needed. These results provide a theoretical foundation for hyperbolic augmented Lagrangian techniques in the GNEP framework and contribute to the broader convergence theory of augmented Lagrangian methods for (GNEP) in nonconvex settings.

Funding

This research was supported by ANID Chile under Fondecyt Grant 1231188 as well as by UTA-Mayor (University of Tarapacá, Chile). FAPERJ Brazil (grant E-26/205.684/2022).

Conflict of interest

The author has no relevant financial or non-financial interests to disclose.

References

[1] R. Andreani, E. G. Birgin, J. M. Martinez, and M. L. Schuverdt. Second-order negative-curvature methods for box-constrained and general constrained optimization. *Computational Optimization and Applications*, 45:209–236, 2010.

[2] R. Andreani, N. S. Fazzio, M. L. Schuverdt, and L. D. Secchin. A sequential optimality condition related to the quasi-normality constraint qualification and its algorithmic consequences. *SIAM Journal on Optimization*, 29 (1):743–766, 2019.

[3] R. Andreani, G. Haeser, L. M. Mito, A. Ramos, and L. D. Secchin. On the best achievable quality of limit points of augmented lagrangian schemes. *Numerical Algorithms*, 90 (2):851–877, 2022.

[4] R. Andreani, J. M. Martinez, and B. F. Svaiter. A new sequencial optimality condition for constrained optimization and algorithmic consequences. *SIAM Journal of Optimization*, 20 No 6:3533–3554, 2010.

[5] R. Andreani, J.M. Martínez, A. Ramos, and P.J.S. Silva. Strict constraint qualifications and sequential optimality conditions for constrained optimization. *Mathematics of Operations Research*, 43(3):693–717, 2018.

[6] R. Andreani, A. Ramos, A. A. Ribeiro, L. D. Secchin, and A. R. Velazco. On the convergence of augmented Lagrangian strategies for nonlinear programming. *IMA Journal of Numerical Analysis*, 42(2):1735–1765, 2022.

[7] R. Andreani, A. Ramos, and L. D. Secchin. Improving the global convergence of Inexact Restoration methods for constrained optimization problems. *SIAM Journal on Optimization*, 34 (4):3429–3455, 2024.

[8] A. M Bagirov, N. Karmitsa, and S. Taheri. Partitional clustering via nonsmooth optimization. *Cham, Switzerland: Springer Nature*, 2020.

[9] A. M Bagirov, B. Ordin, G. Ozturk, and A. E Xavier. An incremental clustering algorithm based on hyperbolic smoothing. *Computational Optimization and Applications*, 61(1):219–241, 2015.

[10] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982.

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

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

[13] E. G. Birgin, G. Haeser, N. Maculan, and L. M. Ramirez. On the global convergence of a general class of augmented Lagrangian methods. *Journal of Optimization Theory and Applications*, 206 (3):1–25, 2025.

[14] L.F. Bueno, G. Haeser, and F.N. Rojas. Optimality conditions and constraint qualifications for generalized Nash equilibrium problems and their practical implications. *SIAM Journal on Optimization*, 29 (1):31–54, 2019.

[15] A. Dreves, F. Facchinei, C. Kanzow, and S. Sagratella. On the solution of the KKT conditions of generalized nash equilibrium problems. *SIAM Journal on Optimization*, 21 (3):1082–1108, 2011.

[16] F. Facchinei and C. Kanzow. Penalty methods for the solution of generalized Nash equilibrium problems. *SIAM Journal on Optimization*, 20(5):2228–2253, 2010.

[17] A. Fischer. A special Newton-type optimization method. *Optimization*, 24 (34):269–284, 1992.

[18] M. I Jordan, T. Lin, and M. Zampetakis. First-order algorithms for nonlinear generalized Nash equilibrium problems. *Journal of Machine Learning Research*, 24(38):1–46, 2023.

[19] C. Kanzow and D. Steck. Augmented Lagrangian methods for the solution of generalized nash equilibrium problems. *SIAM Journal on Optimization*, 26 (4):2034–2058, 2016.

[20] B. S. Murdokhovich. *Variational Analysis and Generalized Differentiation I: Basic Theory*. Series: Grundlehren der mathematischen Wissenschaften, Vol. 330, 2005.

[21] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer, 2006.

[22] M. J. D Powell. A hybrid method for nonlinear equations. In *Numerical Methods for Nonlinear Algebraic Equations*. 1970.

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

[24] A. E. Xavier. *Penalização hiperbólica*. PhD thesis, COPPE-UFRJ, Rio de Janeiro, RJ, 1992.

[25] A. E. Xavier. The hyperbolic smoothing clustering method. *Pattern Recognition*, 43(3):731–737, 2010.

[26] W. I. Zangwill. Non-linear programming via penalty functions. *Management Science*, 13(5):344–358, 1967.