Mathematical Programs with Equilibrium Constraints: A sequential optimality condition, new constraint qualifications and algorithmic consequences *

A. Ramos†

April 30, 2016 (Revised May 27, 2017)

Abstract

Mathematical programs with equilibrium (or complementarity) constraints, MPECs for short, are a difficult class of constrained optimization problems. The feasible set has a very special structure and violates most of the standard constraint qualifications (CQs). Thus, the Karush-Kuhn-Tucker (KKT) conditions are not necessarily satisfied by minimizers and the convergence assumptions of many methods for solving constrained optimization problems are not fulfilled. Therefore it is necessary, both from a theoretical and numerical point of view, to consider suitable optimality conditions, tailored CQs and specially designed algorithms for solving MPECs. In this paper, we present a new sequential optimality condition useful for the convergence analysis for several methods of solving MPECs, such as relaxations schemes, complementarity-penalty methods and interior-relaxation methods. We also introduce a variant of the augmented Lagrangian method for solving MPEC whose stopping criterion is based on this sequential condition and it has strong convergence properties. Furthermore, a new CQ for M-stationary which is weaker than the recently introduced MPEC relaxed constant positive linear dependence (MPEC-RCPLD) associated to such sequential condition is presented. Relations between the old and new CQs as well as the algorithmic consequences will be discussed.

1 Introduction

In this paper we study sequential optimality conditions and constraint qualifications for mathematical programs with equilibrium constraints. We consider mathematical programs with complementarity (or equilibrium) constraints given by

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0, \quad h(x) = 0 \\ & 0 \leq H(x) \perp G(x) \geq 0 \end{array} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^q$, $g : \mathbb{R}^n \to \mathbb{R}^p$, $H, G : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable functions. The notation $0 \leq u \perp v \geq 0$ for u and v in $\mathbb{R}^n$ is a shortcut for $u \geq 0, v \geq 0$ and $\langle u, v\rangle = 0$.

MPECs form an important class of optimization problems. MPEC has its origin in bilevel programming [22] and appears naturally in various applications of the Stackelberg game in economic sciences. It also plays an important role in many others fields, such as engineering design, robotics, multilevel game and transportation science, [48]. For further details, see [22, 40, 45, 20].

MPECs are known to be difficult constrained optimization problems. The main problem, both from a theoretical and a numerical point of view, comes from the complementarity constraints. In fact, many of the standard CQs as the linear independence CQ (LICQ) and the Mangasarian Fromovitz CQ (MFCQ) are violated at any feasible point. The only known standard CQ applicable in the context of MPECs is the Guignard CQ, see [24], which is the weakest CQ for nonlinear mathematical programming (NLP) problems; see [25]. In the failure

*This research was funded by CNPq grants 454798/2015-6.
†Department of Mathematics, Federal University of Paraná, Curitiba, PR, Brazil. Email: albertoramos@ufpr.br.

of CQs, the Karush-Kuhn-Tucker (KKT) conditions may not hold at minimizer (even in the case, where all constraint functions are linear) and the convergence assumptions for most all standard methods for the solution of constrained optimization problems are not satisfied.

For this reason, several notions of stationarity designed for MPECs have emerged over the years. The strong stationarity and the Clarke-stationarity (C-stationarity) are introduced in [51]. It is known that the notion of strong stationarity is equivalent to the KKT conditions of (1) seen as NLP, e.g. [24]. As a consequence, the strong stationarity is a necessary optimality condition only under strong assumptions. C-stationarity, by the other hand, is weaker than strong stationarity and a necessary optimality condition under very mild conditions. Using Mordukhovich's limiting calculus, [42], another stationarity concept emerged, called M-stationarity, see [43, 44, 57]. M-stationarity is stronger than C-stationarity and it can be shown that it is a necessary optimality condition under the same assumptions as C-stationarity.

In order to ensure that a local minimizer of the MPEC (1) is stationary in one of the above senses, we need the CQs. In view of the fact that many standard CQs do not hold for MPECs, a variety of tailored MPEC-CQs have been developed over the years, most of them analogues to standard CQs for NLPs, c.f. [56, 37, 35, 28, 16]. In the other hand, since standard CQs fail, we may search for strong optimality conditions that are valid independently of any CQ. For NLPs, a useful concept is the notion of sequential optimality conditions, [2, 4, 6, 12]. Sequential optimality conditions are genuine optimality conditions that do not depend of the fulfillment of any CQ. There are strong optimality conditions in the sense that they imply the KKT conditions, under weaker CQs, and more important, they provide theoretical tools to justify a stopping criteria for several NLP solvers. This property makes the sequential optimality conditions a useful tool for the improvement of global convergence analysis of several NLP methods under weak assumptions, [6, 7].

In this paper we introduce a new sequential optimality condition suitable for MPECs, called MPEC-AKKT. We will show that MPEC-AKKT is a proper optimality condition, strong in the sense that it implies the M-stationarity under weaker assumption and under mild assumptions several relaxation methods generate sequences whose limit points satisfy that condition. We also introduce a companion CQ for M-stationarity, based on the *cone-continuity property* (CCP) introduced in [6], called MPEC-CCP. Such CQ is strictly weaker than the recently introduced MPEC-RCPLD [27] and can be used in the global convergence analysis of certain methods for solving MPECs.

The paper is organized in the following way: In section 2 we set our standard notation, the basic definitions and stationarity concepts for MPECs. In section 3 we will introduce a new sequential optimality condition and a companion CQ. We also discuss the relation of this sequential optimality condition with the standard sequential optimality condition for NLPs, the AKKT condition, [2, 47]. Relationship between old and new CQs for M-stationarity will be discussed in the section 4. We will pay special attention for the MPEC-RCPLD, MPEC-Abadie CQ and MPEC-quasinormality (see the section 4 for definitions). Finally in section 5 we will present a simple scheme inspired on the augmented Lagrangian method whose stopping criterion is based on the sequential condition MPEC-AKKT. Additionally, we will use the MPEC-AKKT to improve the convergence analysis of several algorithms under standard assumptions. Conclusions will be given in the section 6.

2 Preliminaries and basic assumptions

Our notation is standard in optimization and variational analysis; cf. [50, 15]. $\mathbb{R}^n$ stands for the n-dimensional real Euclidean space, $n \in \mathbb{N}$. $\mathbb{R}_+$ is the set of positive scalars and $\mathbb{R}_-$ the set of negative numbers. Set $a^+ = \max\{0, a\}$, the positive part of $a \in \mathbb{R}$ and put $a_- := -\min\{0, a\} = (-a)_+$. We use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product, $\|\cdot\|$ the associated norm. Given a differentiable mapping $\Gamma : \mathbb{R}^s \to \mathbb{R}^d$, we use $\nabla\Gamma(x)$ to denote the Jacobian matrix of Γ at x when $d > 1$ and the gradient vector when $d = 1$. For every $a \in \mathbb{R}^s$, the support of a is $\text{supp}(a) := \{i : a_i \neq 0\}$. Given a set-valued mapping $\mathit{\Gamma} : \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the *sequential Painlevé-Kuratowski outer limit* of $\mathit{\Gamma}(z)$ as $z \to z^*$ is denoted by

$$\limsup_{z \to z^*} \mathit{\Gamma}(z) := \{w^* \in \mathbb{R}^d : \exists \ (z^k, w^k) \to (z^*, w^*) \text{ with } w^k \in \mathit{\Gamma}(z^k)\}. \tag{2}$$

We say that Γ is *outer semicontinuous* (osc) at z^* if $\limsup_{z\to z^*}\Gamma(z)\subset\Gamma(z^*)$. The *sequential Painlevé-Kuratowski inner limit* of $\Gamma(z)$ as $z\to z^*$ is given by

$$\liminf_{z\to z^*}\Gamma(z):=\{w^*\in\mathbb{R}^d:\forall z^k\to z^*,\ \ \exists w^k\to w^* \text{ with } w^k\in\Gamma(z^k)\} \tag{3}$$

and Γ is *inner semicontinuous* (isc) at z^* iff $\Gamma(z^*)\subset\liminf_{z\to z^*}\Gamma(z)$. We denote by cl X the closure of X, and for conv X the convex hull of X. For a cone $\mathcal{K}\subset\mathbb{R}^s$, its polar is $\mathcal{K}^\circ:=\{v\in\mathbb{R}^s|\langle v,k\rangle\le 0 \text{ for all } k\in\mathcal{K}\}$. In this case, we always have $\mathcal{K}^{\circ\circ}=\text{cl conv}\,\mathcal{K}$. The notation $o(t)$ means any real function $\phi(t)$ such that $\limsup_{t\to 0_+}t^{-1}\phi(t)=0$. Given $X\subset\mathbb{R}^n$ and $z^*\in X$, we define the *tangent/contingent cone* to X at z^* by

$$T_X(z^*):=\limsup_{t\downarrow 0}\frac{X-z^*}{t}=\{d\in\mathbb{R}^n:\exists\ \ t_k\downarrow 0, d^k\to d \text{ with } z^*+t_kd^k\in X\}. \tag{4}$$

The *regular normal cone* to X at $z^*\in X$ is

$$\widehat{N}_X(z^*):=\{w\in\mathbb{R}^n:\langle w,z-z^*\rangle\le o(\|z-z^*\|) \text{ for } z\in X\}. \tag{5}$$

The (Mordukhovich) *limiting normal cone* to X at $z^*\in X$ is defined by

$$N_X(z^*):=\limsup_{z\to z^*,z\in X}\widehat{N}_X(z). \tag{6}$$

The *Clarke's normal cone* to X at $z^*\in X$ is $N_X^C(z^*):=\text{cl conv}N_X(z^*)$. From the definitions, we always have $\widehat{N}_X(z)\subset N_X(z)\subset N_X^C(z)$, $\forall z\in X$. When X is a closed convex set, all these normal cones coincide to the classical normal cone of convex analysis, [50].

In order to describe geometrically the complementary constraints, we define

$$\mathcal{C}:=\{(c_1,c_2)\in\mathbb{R}^2:0\le -c_1\perp -c_2\ge 0\}. \tag{7}$$

Observe that $((-H_1(x),-G_1(x)),\dots,(-H_m(x),-G_m(x)))\in\mathcal{C}^m$ is equivalent to the complementary constraints $0\le H(x)\perp G(x)\ge 0$. The minus signs are used only for convenience of our analysis.

Proposition 2.1. *[26, Proposition 2.1] For every $(c_1,c_2)\in\mathcal{C}$, we have*

a). The tangent cone

$$T_{\mathcal{C}}((c_1,c_2))=\left\{d=(d_1,d_2):\begin{array}{ll} d_1=0,d_2\in\mathbb{R} & \text{if } c_1=0,c_2<0\\ d_1\in\mathbb{R},d_2=0 & \text{if } c_1<0,c_2=0\\ (d_1,d_2)\in\mathcal{C} & \text{if } c_1=0,c_2=0\end{array}\right\}; \tag{8}$$

b). The regular normal cone

$$\widehat{N}_{\mathcal{C}}((c_1,c_2))=\left\{d=(d_1,d_2):\begin{array}{ll} d_1\in\mathbb{R},d_2=0 & \text{if } c_1=0,c_2<0\\ d_1=0,d_2\in\mathbb{R} & \text{if } c_1<0,c_2=0\\ d_1\ge 0,d_2\ge 0 & \text{if } c_1=0,c_2=0\end{array}\right\}; \tag{9}$$

c). The limiting normal cone

$$N_{\mathcal{C}}((c_1,c_2))=\left\{d=(d_1,d_2):\begin{array}{ll} d_1\in\mathbb{R},d_2=0 & \text{if } c_1=0,c_2<0\\ d_1=0,d_2\in\mathbb{R} & \text{if } c_1<0,c_2=0\\ \text{either } d_1d_2=0 & \text{if } c_1=0,c_2=0\\ \ \text{or } d_1>0,d_2>0 & \end{array}\right\}; \tag{10}$$

d). The Clarke's normal cone

$$N^C_{\mathcal{C}}((c_1,c_2))=\left\{d=(d_1,d_2):\begin{array}{ll} d_1\in\mathbb{R},d_2=0 & \text{if } c_1=0,c_2<0\\ d_1=0,d_2\in\mathbb{R} & \text{if } c_1<0,c_2=0\\ d_1\in\mathbb{R},d_2\in\mathbb{R} & \text{if } c_1=0,c_2=0\end{array}\right\}. \tag{11}$$

By [50, Proposition 6.41], we obtain

Lemma 2.2. *Let* $\Lambda := \mathbb{R}^p_- \times \{0\}^q \times \mathcal{C}^m$ *and* $z := (a, b, (c_1^1, c_2^1), \dots, (c_1^m, c_2^m)) \in \Lambda$. *We may write the limiting normal cone as*

$$N_\Lambda(z) = \prod_{j=1}^{p} N_{\mathbb{R}_-}(a_j) \times \prod_{j=1}^{p} N_{\{0\}}(b_j) \times \prod_{i=1}^{m} N_{\mathcal{C}}((c_1^i, c_2^i)). \tag{12}$$

Similar formula holds for the Clarke's normal cone [19, Exercise 10.33]. We end with the following lemma, which is a variation of Caratheódory's lemma.

Lemma 2.3. *[3, Lemma 1] Let* $v = \sum_{i\in\mathcal{B}} \alpha_i p_i + \sum_{j\in\mathcal{D}} \beta_j q_j$ *with* $p_i, q_j \in \mathbb{R}^n$, $i \in \mathcal{B}, j \in \mathcal{D}$, *such that* $\{p_i : i \in \mathcal{B}\}$ *is a linearly independent set and* $\beta_j \neq 0$, $\forall j \in \mathcal{D}$. *Then, there is a subset* $\mathcal{D}' \subset \mathcal{D}$ *and scalars* $\hat{\alpha}_i, \hat{\beta}_j$, $i \in \mathcal{B}$, $j \in \mathcal{D}'$ *with* $\beta_j \hat{\beta}_j > 0$, $j \in \mathcal{D}'$ *such that* $v = \sum_{i\in\mathcal{B}} \hat{\alpha}_i p_i + \sum_{j\in\mathcal{D}'} \hat{\beta}_j q_j$ *and the set* $\{p_i, q_j : i \in \mathcal{B}, j \in \mathcal{D}'\}$ *is linearly independent.*

2.1 Mathematical programs with equilibrium constraints.

For NLPs, the KKT conditions are the most common notion of stationarity. In contrast to MPECs, several different stationarity concepts have emerged over the years. Before continuing, in order to exploit the very special structure of the complementary constraints, we rewrite the MPEC problem (1) as a optimization problem with geometric constraints:

$$\text{Minimize } f(x) \text{ subject to } F(x) \in \Lambda, \tag{13}$$

where

$$\begin{gathered} F(x) := (g(x), h(x), \Psi(x)) \\ \Psi(x) := ((-H_1(x), -G_1(x)), \dots, (-H_m(x), -G_m(x))) \\ \Lambda := \mathbb{R}^p_- \times \{0\}^q \times \mathcal{C}^m \end{gathered} \tag{14}$$

The feasible region of the optimization problem with geometric constraints (13) is $\Omega := \{x \in \mathbb{R}^n : F(x) \in \Lambda\}$. Now, we will define some crucial index sets that will occur frequently in the subsequent analysis. Since, we will deal with several MPECs, these index sets must be explicitly dependent of the feasible constraint sets. Consider a set of the form, $\Lambda = \mathbb{R}^p_- \times \{0\}^q \times \mathcal{C}^m$ for some $p, q, m \in \mathbb{N}$. Now, for every point, $z := (a, b, -(c_1^1, c_2^1), \dots, -(c_1^m, c_2^m))$ in $\Lambda = \mathbb{R}^p_- \times \{0\}^q \times \mathcal{C}^m$, we use the notation

$$\begin{aligned} \mathcal{I}(z, \Lambda) &:= \{i \in \{1, \dots, m\} : c_1^i = 0, c_2^i > 0\}, \\ \mathcal{J}(z, \Lambda) &:= \{i \in \{1, \dots, m\} : c_1^i = 0, c_2^i = 0\}, \\ \mathcal{K}(z, \Lambda) &:= \{i \in \{1, \dots, m\} : c_1^i > 0, c_2^i = 0\}. \end{aligned} \tag{15}$$

When, Λ is clear for the context, we write $\mathcal{I}(z)$, $\mathcal{J}(z)$ and $\mathcal{K}(z)$ instead of $\mathcal{I}(z, \Lambda)$, $\mathcal{J}(z, \Lambda)$ and $\mathcal{K}(z, \Lambda)$ respectively. For $x^* \in \Omega = \{x \in \mathbb{R}^n : F(x) \in \Lambda\}$, we let

$$\begin{aligned} A(x^*, \Omega) :=& \quad \{j \in \{1, \dots, p\} : g_j(x^*) = 0\}, \\ \mathcal{I}(x^*, \Omega) := \mathcal{I}(F(x^*), \Lambda) =& \quad \{i \in \{1, \dots, m\} : H_i(x^*) = 0, G_i(x^*) > 0\}, \\ \mathcal{J}(x^*, \Omega) := \mathcal{J}(F(x^*), \Lambda) =& \quad \{i \in \{1, \dots, m\} : H_i(x^*) = 0, G_i(x^*) = 0\}, \\ \mathcal{K}(x^*, \Omega) := \mathcal{K}(F(x^*), \Lambda) =& \quad \{i \in \{1, \dots, m\} : H_i(x^*) > 0, G_i(x^*) = 0\}. \end{aligned} \tag{16}$$

Similarly, when Ω is clear in the context, we write $A(x^*)$, $\mathcal{I}(x^*)$, $\mathcal{J}(x^*)$ and $\mathcal{K}(x^*)$ instead of $A(x^*, \Omega)$, $\mathcal{I}(x^*, \Omega)$, $\mathcal{J}(x^*, \Omega)$ and $\mathcal{K}(x^*, \Omega)$ respectively. There is no risk of confusion, between $\mathcal{I}(x)$ and $\mathcal{I}(z)$ since we reserve the letter z for elements of Λ. The same considerations for the other index sets.

The set $A(x^*)$ is index set of active inequalities and the index sets $\mathcal{I}(x^*)$, $\mathcal{J}(x^*)$ and $\mathcal{K}(x^*)$ form a partition of $\{1, \dots, m\}$ for every $x^* \in \Omega$. The set $\mathcal{J}(x^*)$ is called the bi-active set. Now, we are now able to define the next stationarity concepts.

Definition 2.1. Let x^* be a feasible point for the MPEC (1). Suppose that there are multipliers $\mu \in \mathbb{R}^p_+$, $\lambda \in \mathbb{R}^q$, $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^m$ with $\text{supp}(\mu) \subset A(x^*)$ such that

$$\nabla f(x^*) + \sum_{j=1}^{p} \mu_j \nabla g_j(x^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x^*) - \sum_{\imath=1}^{m} u_\imath \nabla H_\imath(x^*) - \sum_{\jmath=1}^{m} v_\jmath \nabla G_\jmath(x^*) = 0. \tag{17}$$

Then, x^* is said to be

1. Strongly stationary (S-STATIONARY), if $u_i = 0$, $i \in \mathcal{K}(x^*)$, $v_j = 0$, $i \in \mathcal{I}(x^*)$ and $u_i \geq 0$, $v_i \geq 0$ for all $i \in \mathcal{J}(x^*)$;
2. M-STATIONARY, if $u_i = 0$, $i \in \mathcal{K}(x^*)$, $v_j = 0$, $i \in \mathcal{I}(x^*)$ and either $u_i > 0, v_i > 0$ or $u_i v_i = 0$ for all $i \in \mathcal{J}(x^*)$;
3. C-STATIONARY, if $u_i = 0$, $i \in \mathcal{K}(x^*)$, $v_j = 0$, $i \in \mathcal{I}(x^*)$ and $u_i v_i \geq 0$ for all $i \in \mathcal{J}(x^*)$;
4. WEAKLY STATIONARY, if $u_i = 0$, $i \in \mathcal{K}(x^*)$ and $v_j = 0$, $i \in \mathcal{I}(x^*)$.

The different notions differ, basically, in how the multipliers u_i and v_i act over the bi-active set $\mathcal{J}(x^*)$. All these stationary concepts coincide when the bi-active set is an empty set. From the definitions the following chain of implications holds: S-stationary ⇒ M-stationary ⇒ C-stationary ⇒ weak stationary. These stationary concepts can be state in a geometric way, using Proposition 2.1 and Lemma 2.2. See [26, Proposition 2.2]

Proposition 2.4. *Let x^* be a feasible point for the MPEC 1. We have the following statements:*

1. S-STATIONARY is equivalent to $0 \in \nabla f(x^*) + \nabla F(x^*)^\top \widehat{N}_\Lambda(F(x^*))$;
2. M-STATIONARY is equivalent to $0 \in \nabla f(x^*) + \nabla F(x^*)^\top N_\Lambda(F(x^*))$;
3. C-STATIONARY is equivalent to $0 \in \nabla f(x^*) + \nabla F(x^*)^\top \widetilde{N}_\Lambda(F(x^*))$, where $\widetilde{N}_\Lambda(z) := N_{\mathbb{R}^p_-}(a) \times N_{\{0\}^q}(b) \times \prod_{i=1}^m (\widehat{N}_\mathcal{C}((c_1^i, c_2^i)) \cup -\widehat{N}_\mathcal{C}((c_1^i, c_2^i)))$, for $z = (a, b, c) \in \mathbb{R}^p_- \times \{0\}^q \times \mathcal{C}^m$;
4. WEAKLY STATIONARY is equivalent to $0 \in \nabla f(x^*) + \nabla F(x^*)^\top N^C_\Lambda(F(x^*))$.

3 Sequential optimality condition and a new CQ

This section is devoted to the study of sequential optimality conditions suitable for MPECs. We will introduce a new sequential optimality condition called MPEC-AKKT. Relations with the AKKT condition (a standard sequential optimality condition for NLPs) will be discussed. We also present a new CQ for M-stationarity.

First, we consider the role of sequential optimality condition for NLPs. Consider the NLP: minimize $f(x)$ subject to $x \in X$, where

$$X := \{x \in \mathbb{R}^n : g_j(x) \leq 0, j \in \{1, \ldots, p\}, h_i(x) = 0, i \in \{1, \ldots, q\}\}. \tag{18}$$

Since, it is usually not possible to "solve" the NLPs exactly; most of the standard NLPs method stops when the KKT conditions are satisfied approximately, although KKT conditions are not necessary satisfied by minimizers. The *Approximate KKT* (AKKT) condition justifies this practice [2, 47].

Definition 3.1. The Approximate-Karush-Kuhn-Tucker (AKKT) condition holds at $x^* \in \mathbb{R}^n$ if there are sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^q$, $\{\mu^k\} \subset \mathbb{R}^p_+$ and $\{\varepsilon_k\} \subset \mathbb{R}_+$, such that $x^k \to x^*$, $\varepsilon_k \to 0^+$,

$$\|h(x^k)\| \leq \varepsilon_k, \quad \|\max\{0, g(x^k)\}\| \leq \varepsilon_k, \tag{19}$$

$$\|\nabla f(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^{q} \lambda_i^k \nabla h_i(x^k)\| \leq \varepsilon_k, \qquad \text{and} \tag{20}$$

$$\mu_j^k = 0 \text{ if } g_j(x^k) < -\varepsilon_k. \tag{21}$$

Certainly, we can used different ε_k for the each different parts of the definition of AKKT. To keep, the notation simple, we decided to take the same ε_k. The notion of AKKT condition is independent of this choice, in fact, the AKKT condition is equivalent to say that there are sequence $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^q$ and $\{\mu^k\} \subset \mathbb{R}^p_+$ with $x^k \to x^*$, $\text{supp}(\mu^k) \subset A(x^*)$ such that $\nabla f(x^k) + \sum_{j \in A(x^*)} \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^q \lambda_i^k \nabla h_i(x^k) \to 0$.

Remark 1. Each point x^k is called of ε_k-stationarity point. Several variants of the AKKT condition have been proposed in the literature. They basically differ in the way the complementarity conditions are treated. For instance, if additionally to (19), (20) and (21) we require $\sum_{i=1}^p |\lambda_i^k h_i(x^k)| + \sum_{j=1}^q |\mu_j^k g_j(x^k)| \leq \varepsilon_k$, we have the *complementary* AKKT condition (CAKKT) introduced in [8]. If instead of $\sum_{i=1}^p |\lambda_i^k h_i(x^k)| + \sum_{j=1}^q |\mu_j^k g_j(x^k)| \leq \varepsilon_k$ we only require $\sum_{j=1}^q |\mu_j^k g_j(x^k)| \leq \varepsilon_k$ we get the notion of ε-stationarity considered in [37].

The AKKT condition justifies the stopping criteria for several NLP methods. In fact, it attempts to catch a property of several NLPs solvers: NLPs solvers are devised to find a primal sequence and approximate multipliers They basically differ in the way the complementarity conditions are treated for which the KKT residual goes to zero. The AKKT condition, as other sequential optimality conditions, shares three important properties. First, it is a true necessary optimality condition independently of the fulfillment of any CQ [2]. Second, it is strong, in the sense that it implies necessary optimality conditions as "KKT or not CQ" for weak CQs as MFCQ, RCPLD or CPG, see [3, 4, 6]. Third, there are many algorithms that generate sequences whose limit points satisfy it. In the case of AKKT, we have that some augmented Lagrangian methods [1, 12], some Sequential Quadratic Programming (SQP) algorithms [47], interior point methods [17] and inexact restoration methods generate primal sequences $\{x^k\}$ with approximate multipliers $\{\mu^k, \lambda^k\}$ for a given error tolerance $\{\varepsilon_k\}$ for which (19), (20) and (21) are fulfilled, see [4]. This makes possible improve the global converge analysis of those methods under weaker assumptions, [4, 6, 7]. The sequence $\{x^k\}$ is called an AKKT sequence and we say that these methods generate AKKT sequences.

Motivated by AKKT, we propose a sequential optimality condition suitable for optimization problems with geometric constraints (13).

Definition 3.2. We say the Approximate stationarity condition holds for the feasible point x^* for the problem (13), if there are sequences $\{x^k\}$, $\{z^k\}$ and $\{\gamma^k\}$ such that $x^k \to x^*$, $z^k \to F(x^*)$, $z^k \in \Lambda$

$$\nabla f(x^k) + \nabla F(x^k)^\top \gamma^k \to 0 \text{ and } \gamma^k \in N_\Lambda(z^k). \tag{22}$$

If $\gamma^k = (\mu^k, \lambda^k, u^k, v^k) \in \mathbb{R}^p_+ \times \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^m$. Then, using the explicit form of $N_\Lambda(z^k)$, Lemma 2.2, (22) can be written, for k sufficiently large, as

$$\nabla f(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^q \lambda_i^k \nabla h_i(x^k) - \sum_{\imath=1}^m u_\imath^k \nabla H_\imath(x^k) - \sum_{\jmath=1}^m v_\jmath^k \nabla G_\jmath(x^k) \to 0 \tag{23}$$

where $\text{supp}(\mu^k) \subset A(x^*)$, $\text{supp}(u^k) \subset \mathcal{I}(z^k) \cup \mathcal{J}(z^k)$, $\text{supp}(v^k) \subset \mathcal{K}(z^k) \cup \mathcal{J}(z^k)$ and either $u_\ell^k v_\ell^k = 0$ or $u_\ell^k > 0, v_\ell^k > 0$ for each $\ell \in \mathcal{J}(z^k)$.

In the absence of complementary constraints, the definition above reduces to the AKKT condition. For that motive, we call every feasible x^* that conforms the Definition 3.2 of a MPEC-AKKT point and each sequence $\{x^k\}$ of a MPEC-AKKT sequence.

A small variations of this definition have appeared implicitly stated in the literature, for example in [56, 10, 33] or even in [35] in the proof that each local minimizer of (1) satisfies a MPEC-tailored Fritz-John point. However, we will use the version described above as it is better suited for our purposes. Note that MPEC-AKKT is a property of the optimization problem, rather than a property of the constraint set, since it depends on the objective function f. As we will see in section 5, MPEC-AKKT will be an useful theoretical tool in the analysis of convergence of some algorithms for solving MPECs.

Remark 2. In the definition of MPEC-AKKT, there is no loss of generality if we choose γ^k in $\widehat{N}_\Lambda(z^k)$ (maybe for another z^k with $z^k \to F(x^*)$). This is a consequence of the definition of N_Λ as outer limit of regular normal cones. Thus, each MPEC-AKKT point can be approximated by a sequence of approximate M-stationary points and also by a sequence of approximate S-stationary points. Note that if $0 \in \nabla f(x^k) + \nabla F(x^k)^\top \gamma^k$, $\gamma^k \in N_\Lambda(z^k)$ for $k \in \mathbb{N}$, $\{x^k\}$ is a sequence of M-stationary points and if $0 \in \nabla f(x^k) + \nabla F(x^k)^\top \gamma^k$, $\gamma^k \in \widehat{N}_\Lambda(z^k)$ for $k \in \mathbb{N}$, $\{x^k\}$ is a sequence of S-stationary points.

MPEC-AKKT is a proper optimality condition and independent of any assumption over the constraints as Theorem 3.1 will shows. The proof use classical penalty approach and it can be derived from the proofs of [56, 35] or [50, Theorem 6.14]. We state it here for sake of completeness.

Theorem 3.1. *Every local minimizer is a MPEC-AKKT point.*

Proof Let $\delta > 0$ such that $f(x^*) \leq f(x)$, $\forall x \in \Omega \cap \mathbb{B}(x^*, \delta)$ and $\rho_k \uparrow \infty$. For each k, consider the problem

$$\text{Min } f(x) + \frac{1}{2}\|(x - x^*, z - F(x^*))\|^2 + \frac{1}{2}\rho_k\|F(x) - z\|^2 \text{ s.t. } (x, z) \in U, \tag{24}$$

where $U := \{(x, z) \in \mathbb{R}^n \times \Lambda : \|(x, z) - (x^*, F(x^*))\| \leq \delta\}$. Let (x^k, z^k) be a global solution of (24), which is well-defined by the compactness of U and continuity of the functions. We will show that the sequence $\{(x^k, z^k)\}$ converges to $(x^*, F(x^*))$. In fact, due to the optimality, we have

$$f(x^k) + \frac{1}{2}\|(x^k - x^*, z^k - F(x^*))\|^2 + \frac{1}{2}\rho_k\|F(x^k) - z^k\|^2 \leq f(x^*) \tag{25}$$

Let $(\hat{x}, \hat{z})$ be a limit point of $\{(x^k, z^k)\}$. From (25) follows that $\|F(x^k) - z^k\| \to 0$ and so $F(\hat{x}) = \hat{z}$. As consequence $\hat{x}$ is a feasible point. Evenmore, by (25), we have $\|(\hat{x} - x^*, \hat{z} - F(x^*))\|^2 \leq 2(f(x^*) - f(\hat{x}))$. But, since $f(x^*) \leq f(\hat{x})$ we get that $\hat{x} = x^*$ and that $\{(x^k, z^k)\}$ converge since it has a unique limit point, namely, $(x^*, F(x^*))$. Now, for k sufficiently large, $\|(x, z) - (x^*, F(x^*))\| < \delta$ and then, by the optimality of (x^k, z^k), [50, Theorem 6.12], we obtain

$$0 = r^k + \nabla f(x^k) + \nabla F(x^k)^\top \gamma^k \text{ with } \gamma^k \in N_\Lambda(z^k), \tag{26}$$

where $r^k := -\nabla F(x^k)^\top(F(x^*) - z^k) + (x^k - x^*)$ and $\gamma^k := \rho_k(F(x^k) - z^k) + (F(x^*) - z^k)$. From the continuity, we get $r^k \to 0$. Thus, from (26), x^* is a MPEC-AKKT point. □

Remark 3. Note that the AKKT condition holds at x^* for (1) seen as NLP iff there is a sequence $x^k \in \mathbb{R}^n$ with $x^k \to x^*$ such that

$$\begin{aligned} &\nabla f(x^k) + \textstyle\sum_{j \in A(x^*)} \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^q \lambda_i^k \nabla h_i(x^k) \\ &- \textstyle\sum_{i=1}^m u_i^k \nabla H_i(x^k) - \sum_{j=1}^m v_j^k \nabla G_j(x^k) + \sum_{i=1}^m \rho_i^k \nabla(G_i(x^k) H_i(x^k)) \to 0 \end{aligned} \tag{27}$$

for some approximate multipliers $(\mu^k, \lambda^k, u^k, v^k, \rho^k) \in \mathbb{R}_+^p \times \mathbb{R}^q \times \mathbb{R}_+^m \times \mathbb{R}_+^m \times \mathbb{R}^m$ with $\text{supp}(\mu^k) \subset A(x^*)$, $\text{supp}(u^k) \subset \{i \in \{1, \ldots, m\} : H_i(x^*) = 0\}$ and $\text{supp}(v^k) \subset \{i \in \{1, \ldots, m\} : G_i(x^*) = 0\}$. We observe that (27) and (23) differ in the way how we deal with the complementarity constraints.

The problem (1) can be view as a optimization problem with geometrical constraints (13). Thus, each local minimizer of (1) is a MPEC-AKKT point. In the other hand, (1) can be also see as a NLP. So, a local minimizer is also an AKKT point. Now, if we try to solve (1) by using NLP algorithms (as augmented lagrangian methods or some SQP methods), we get a sequence of iterates such that every feasible limit point satisfies the AKKT condition (a nontrivial optimality condition). We can consider this fact as a possible reason why, in general, NLP algorithms are successful when they are applied to MPECs, [5]. Thus, since AKKT and MPEC-AKKT are both optimality conditions, one with clear algorithmically implications, the exact relation between MPEC-AKKT and AKKT becomes relevant.

First, MPEC-AKKT does not imply AKKT as the following example shows.

Example 3.1 (MPEC-AKKT does not imply AKKT)**.** In $\mathbb{R}^2$, consider the point $x^* := (0, 0)$, the function $f(x_1, x_2) := x_1$ and the complementary constraints given by $H_1(x_1, x_2) := x_2 \exp(-x_1 x_2)$, $G_1(x_1, x_2) := \exp(x_1 x_2)$, $H_2(x_1, x_2) := -x_2 \exp(x_1 x_2)$ and $G_2(x_1, x_2) := \exp(-x_1 x_2)$. Clearly, x^* is a feasible point. Now, we will show that x^* is not an AKKT point. Indeed, from the AKKT condition, $\nabla f(x_1^k, x_2^k) - u_1^k \nabla H_1(x_1^k, x_2^k) - u_2^k \nabla H_2(x_1^k, x_2^k) + \rho_1^k \nabla(H_1 G_1)(x_1^k, x_2^k) + \rho_2^k \nabla(H_2 G_2)(x_1^k, x_2^k) \to (0, 0)$ for some sequence $x^k = (x_1^k, x_2^k) \to (0, 0)$ and multipliers $u_1^k, u_2^k \geq 0$, $\rho_1^k, \rho_2^k \in \mathbb{R}$, see Remark 3. Since $(H_1 G_1)(x_1, x_2) = x_2$ and $(H_2 G_2)(x_1, x_2) = -x_2$, from the AKKT condition, we must get $1 + u_1^k (x_2^k)^2 \exp(-x_1^k x_2^k) + u_2^k (x_2^k)^2 \exp(x_1^k x_2^k) \to 0$ which is impossible because $u_1^k, u_2^k \geq 0$. Now, we will show that MPEC-AKKT holds at x^*. Take $x_1^k := 0$, $x_2^k := 1/k$ and $z^k := (-(0, 1), -(0, 1))$. Then, $(x_1^k, x_2^k) \to (0, 0)$, $z^k \to F(x^*) = (-(0, 1), -(0, 1))$, $\mathcal{I}(z^k) = \{1, 2\}$, $\mathcal{K}(z^k) = \emptyset$

and $\mathcal{J}(z^k) = \emptyset$. Put $u_1^k := -((x_2^k)^2 \exp(-x_1^k x_2^k) + (x_2^k)^2 \exp(x_1^k x_2^k))^{-1}$, $u_2^k := u_1^k$, $v_2^k := v_1^k = 0$. By calculations, $\nabla f(x_1^k, x_2^k) - u_1^k \nabla H_1(x_1^k, x_2^k) - u_2^k \nabla H_2(x_1^k, x_2^k) = (0, 0)$. Thus, (23) holds at x^* and hence x^* is a MPEC-AKKT point.

Under additional assumptions, MPEC-AKKT implies AKKT

Theorem 3.2. *Let x^* be a MPEC-AKKT point and $\{(x^k, z^k, \gamma^k)\}$ be a MPEC-AKKT sequence associated to x^* with $\gamma^k = (\mu^k, \lambda^k, (u_1^k, v_1^k), \ldots, (u_m^k, v_m^k))$ and $z^k := (a^k, b^k, -c^k) \in \Lambda$. If, we assume that*

1. *$u_j^k \geq 0$, $v_j^k \geq 0$ for all $j \in \mathcal{J}(z^k)$ and;*
2. *The sequence $\{\max\{0, \max_{j \in \mathcal{I}(z^k)}(-u_j^k / c_{2j}^k), \max_{j \in \mathcal{K}(z^k)}(-v_j^k / c_{1j}^k)\} : k \in \mathbb{N}\}$ has a subsequence bounded.*

Then, x^* is an AKKT point for the problem 1 considered as a NLP.

Proof. Define $\rho_k := \max\{0, \max_{j \in \mathcal{I}(z^k)}(-u_j^k / c_{2j}^k), \max_{j \in \mathcal{K}(z^k)}(-v_j^k / c_{1j}^k)\}$. We assume without loss of generality (after take an adequate subsequence) that $\{\rho_k\}$ itself is bounded. To show, that under the boundedness of $\{\rho_k\}$, MPEC-AKKT implies AKKT, we only focus on the complementary part of (23). Define

$$\omega^k := -\sum_{\imath \in \mathcal{I}(z^k) \cup \mathcal{J}(z^k)} u_\imath^k \nabla H_\imath(x^k) - \sum_{\jmath \in \mathcal{K}(z^k) \cup \mathcal{J}(z^k)} v_\jmath^k \nabla G_\jmath(x^k). \tag{28}$$

Let us recall that the sets $\mathcal{I}(z^k)$, $\mathcal{K}(z^k)$ and $\mathcal{J}(z^k)$ are a partition of $\{1, \ldots, m\}$. Take $\hat{u}_j^k := u_j^k + \rho_k c_{2j}^k \geq 0$, $j \in \mathcal{I}(z^k)$, $\hat{u}_j^k := u_j^k$, $j \in \mathcal{J}(z^k) \cup \mathcal{K}(z^k)$ and $\hat{v}_j^k := v_j^k + \rho_k c_{1j}^k \geq 0$, $j \in \mathcal{K}(z^k)$, $\hat{v}_j^k := v_j^k$, $j \in \mathcal{J}(z^k) \cup \mathcal{I}(z^k)$. Note that $\mathcal{I}(z^k) \cup \mathcal{J}(z^k) = \{i : c_{1i}^k = 0\}$ and $\mathcal{K}(z^k) \cup \mathcal{J}(z^k) = \{j : c_{2j}^k = 0\}$. Substituting $\bar{u}^k$ and $\bar{v}^k$ into (28), we obtain that ω^k is equal to

$$-\sum_{i : c_{1i}^k = 0} \hat{u}_i^k \nabla H_i(x^k) - \sum_{j : c_{2j}^k = 0} \hat{v}_j^k \nabla G_j(x^k) + \sum_{i=1}^m \rho_k \nabla(G_i(x^k) H_i(x^k)) + \Delta^k \tag{29}$$

where $\Delta^k := \sum_{i=1}^m \rho_k (c_{2i}^k - G_i(x^k)) \nabla H_i(x^k) + \rho_k (c_{1i}^k - H_i(x^k)) \nabla G_i(x^k)$.

Put $\rho_i^k := \rho_k$, $\forall i \in \{1, \ldots, m\}$. Since $\{\rho_k\}$ is bounded and $(c_{1i}^k, c_{2i}^k) \to (H_i(x^*), G_i(x^*))$, $\forall i$, we get $\Delta^k \to 0$. To show that x^* is an AKKT point, we only rest to show that $\hat{u}_i^k \geq 0$, $\forall i : H_i(x^*) = 0$ and $\hat{v}_i^k \geq 0$, $\forall i : G_i(x^*) = 0$. But, it holds, since $\{i \in \{1, \ldots, m\} : c_{1i}^k = 0\} \subset \{i \in \{1, \ldots, m\} : H_i(x^*) = 0\}$ and $\{i \in \{1, \ldots, m\} : c_{2i}^k = 0\} \subset \{i \in \{1, \ldots, m\} : G_i(x^*) = 0\}$, for k large enough. □

By the other hand, it is not true, that AKKT always implies MPEC-AKKT.

Example 3.2 (AKKT does not imply MPEC-AKKT)**.** In $\mathbb{R}^2$, take $f(x_1, x_2) := -x_2$, $0 \leq H_1(x_1, x_2) := x_1 \perp G_1(x_1, x_2) := x_2 \geq 0$ and $x^* := (0, 1)$. Clearly, x^* is a feasible point. Now, we will see that AKKT holds at x^*. Take $x^k = (x_1^k, x_2^k) := (1/k, 1)$, $\rho_k := k$, $\lambda_1 := k$ and $\lambda_2 := 0$. By straightforward calculations, $\nabla f(x^k) - \lambda_1 \nabla H_1(x^k) - \lambda_2 \nabla G_1(x^k) + \rho_k \nabla(H_1(x^k) G_1(x^k)) = 0$. Thus, AKKT holds at x^*. However, MPEC-AKKT fails at x^*. In fact, since there is only one complementary constraint and $\mathcal{I}(x^*) = \{1\}$, expression (23) holds iff $\nabla f(x^k) - u^k \nabla H_1(x^k) = (0, -1) - u^k(1, 0) = -(u^k, 1) \to (0, 0)$ for some $u^k \in \mathbb{R}$, which is impossible. Thus, MPEC-AKKT fails.

Under some assumptions, AKKT implies MPEC-AKKT as Theorem 3.3 shows

Theorem 3.3. *Let x^* be an AKKT point. If there exists an AKKT sequence $\{x^k\}$ with $x^k \in \Omega$, $k \in \mathbb{N}$ such that $u_j^k H_j(x^k) = 0$ and $v_i^k G_i(x^k) = 0$ hold for every $k \in \mathbb{N}$ and for every $i \in \{1, \ldots, m\}$. Then, x^* is a MPEC-AKKT point.*

Proof. Let $\{x^k\}$ be the feasible sequence satisfying the hypothesis. Denote by $(\mu^k, \lambda^k, u^k, v^k, \rho^k)$ the approximate multipliers associated with $\{x^k\}$. Now, we focus on the complementarity part. Denote by ω^k the next expression

$$-\sum_{\{i : H_i(x^*) = 0\}} u_i^k \nabla H_i(x^k) - \sum_{\{i : G_i(x^*) = 0\}} v_i^k \nabla G_i(x^k) + \sum_{i=1}^m \rho_i^k \nabla(H_i(x^k) G_i(x^k)) \tag{30}$$

with $(u^k, v^k, \rho^k) \in \mathbb{R}^m_+ \times \mathbb{R}^m_+ \times \mathbb{R}^m$, $\text{supp}(u^k) \subset \{i : H_i(x^*) = 0\}$, $\text{supp}(v^k) \subset \{i : G_i(x^*) = 0\}$, $\langle u^k, H(x^k)\rangle = 0$ and $\langle v^k, G(x^k)\rangle = 0$. Put $z^k := F(x^k)$, $k \in \mathbb{N}$. Clearly, $z^k \in \Lambda$, $\forall k$ and $\sum_{i\in\mathcal{I}(z^k)} \rho_i^k G_i(x^k)\nabla H_i(x^k) + \sum_{i\in\mathcal{K}(z^k)} \rho_i^k H_i(x^k)\nabla G_i(x^k)$ is equal to $\sum_{i=1}^m \rho_i^k \nabla(H_i(x^k)G_i(x^k))$. Since $u_i^k H_i(x^k) = 0$, $v_i^k G_i(x^k) = 0$, $\forall i$, we have that $u_i^k = 0$, $i \in \mathcal{K}(z^k)$ and $v_i^k = 0$, $i \in \mathcal{I}(z^k)$. Moreover, for k large enough, $\mathcal{I}(z^k) = \mathcal{I}(z^k) \cap \{i : H_i(x^*) = 0\}$ and $\mathcal{K}(z^k) = \mathcal{K}(z^k) \cap \{i : G_i(x^*) = 0\}$. Then, (30) can be written as

$$\omega^k = -\sum_{\{i:H_i(x^*)=0\}} \bar{u}_i^k \nabla H_i(x^k) - \sum_{\{i:G_i(x^*)=0\}} \bar{v}_i^k \nabla G_i(x^k) \tag{31}$$

where $(\bar{u}^k, \bar{v}^k) \in \mathbb{R}^m \times \mathbb{R}^m$ is defined as follows

$$\bar{u}_i^k := \begin{cases} u_i^k - \rho_i^k G_i(x^k) & \text{if } i \in \{i : H_i(x^*) = 0\} \cap \mathcal{I}(z^k) \\ u_i^k & \text{if } i \in \{i : H_i(x^*) = 0\} \cap \mathcal{J}(z^k) \end{cases} \tag{32}$$

$$\bar{v}_i^k := \begin{cases} v_i^k - \rho_i^k H_i(x^k) & \text{if } i \in \{i : G_i(x^*) = 0\} \cap \mathcal{K}(z^k) \\ v_i^k & \text{if } i \in \{i : G_i(x^*) = 0\} \cap \mathcal{J}(z^k) \end{cases} \tag{33}$$

with $\text{supp}(\bar{u}^k) \subset \{i : H_i(x^*) = 0\}$ and $\text{supp}(\bar{v}^k) \subset \{i : G_i(x^*) = 0\}$. Note that $\bar{u}_\ell^k \geq 0$, $\bar{v}_\ell^k \geq 0$ for $\ell \in \mathcal{J}(z^k)$. Thus, $\{x^k\}$ is a MPEC-AKKT sequence. □

Recently, in [6] the authors introduced a new CQ intimately related with the AKKT condition, called CCP. This CQ is the most accurate measure of strength of the sequential optimality condition AKKT, in fact, under CCP, every AKKT point is actually a KKT point and when CCP fails, it is possible to find an AKKT point which is not a KKT point as the proof of the [6, Theorem 3.2] shows. Unfortunately, as others standard CQs, CCP may not hold for MPECs. So, when we try to solve MPECs problems using NLPs methods, such methods (as the augmented lagrangian methods) can generate an AKKT point, accepted as possible solution, but which is not a stationary KKT-point. Motivated by CCP, we define the next MPEC-type CCP condition.

Definition 3.3. Let x^* be a feasible point. We say that the MPEC-Cone Continuity Property (MPEC-CCP) holds at x^* if set-valued mapping $\mathbb{R}^n \times \Lambda \ni (x, z) \rightrightarrows \nabla F(x)^\top N_\Lambda(z)$ is outer semicontinuous at the point $(x^*, F(x^*))$, i.e.

$$\limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z) \subset \nabla F(x^*)^\top N_\Lambda(F(x^*)). \tag{34}$$

From $\limsup_{z\to z^*} \widehat{N}_\Lambda(z) = N_\Lambda(z^*)$, the next lemma easily follows.

Lemma 3.4. *We always have*

$$\limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z) = \limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top \widehat{N}_\Lambda(z). \tag{35}$$

From [6], CCP is equivalent to state that every AKKT point is a KKT point. A similar result holds for MPEC-CCP. The precise statement is given in the next theorem.

Theorem 3.5. *Let x^* be a feasible point. Then, MPEC-CCP holds at x^* iff x^* is a M-stationary point whenever x^* is a MPEC-AKKT point.*

Proof. Let us show first that, if MPEC-CCP holds, the sequential MPEC-AKKT condition implies the M-stationarity condition independently of the objective function. Let f be an objective function such that the MPEC-AKKT condition holds at x^*. Then, there are sequences $\{x^k\} \to x^*$, $z^k \to F(x^*)$ and $\{\gamma^k\} \in N_\Lambda(z^k)$ such that $r^k := \nabla f(x^k) + \nabla F(x^k)^\top \gamma^k \to 0$. Define $\omega^k := \nabla F(x^k)^\top \gamma^k$, we see that $\omega^k \in \nabla F(x^k)^\top N_\Lambda(z^k)$ and $\omega^k = r^k - \nabla f(x^k)$. By the continuity of $\nabla f(x)$ and $r^k \to 0$, we get

$$-\nabla f(x^*) = \lim \omega^k \in \limsup_{(x,z)\to(x^*,F(x^*)} \nabla F(x)^\top N_\Lambda(z) \subset \nabla F(x^*)^\top N_\Lambda(F(x^*)), \tag{36}$$

where the last inclusion follows from the MPEC-CCP. Thus, $-\nabla f(x^*)$ belongs to $\nabla F(x^*)^\top N_\Lambda(F(x^*))$, which is equivalent, by Proposition 2.4(2), to say that x^* is a M-stationary point.

Now, let us prove that, if the MPEC-AKKT condition implies the M-stationarity for every objective function, then MPEC-CCP holds. Let ω^* be an element of $\limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z)$. Thus, there are sequences $\{x^k\}$, $\{z^k\}$, $\{\gamma^k\}$ and $\{\omega^k\}$ such that $x^k \to x^*$, $z^k \to F(x^*)$, $\omega^k \to \omega^*$ and $\omega^k = \nabla F(x^k)^\top \gamma^k$ where $\gamma^k \in N_\Lambda(z^k)$. Define $f(x) = -\langle w^*, x\rangle$, $x \in \mathbb{R}^n$. Note that MPEC-AKKT holds at x^* for f, since $\nabla f(x^k) + \omega^k = -\omega^* + \omega^k \to 0$. So, by hypothesis x^* is a M-stationary point, that is, $-\nabla f(x^*) = \omega^* \in \nabla F(x^*)^\top N_\Lambda(F(x^*))$. □

We will show that MPEC-CCP is a CQ for M-stationary. For this purpose we need the next lemma.

Lemma 3.6. *[50, Theorem 6.11] For every $v \in T^\circ_\Omega(\bar{x})$, there is a smooth function ϕ such that $-\nabla\phi(\bar{x}) = v$ and attains its minimum relative to Ω uniquely at $\bar{x}$.*

Theorem 3.7. *We always have*

$$N_\Omega(x^*) \subset \limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z). \tag{37}$$

If in addition, MPEC-CCP holds at x^*, then $N_\Omega(x^*) \subset \nabla F(x^*)^\top N_\Lambda(F(x^*))$.

Proof Let $\omega \in N_\Omega(x^*)$, so by definition of normal cone, there are sequences $\{x^k\} \in \Omega$, $\{v^k\}$ such that $x^k \to_k x^*$, $v^k \to_k \omega$ and $v^k \in T^\circ_\Omega(x^k)$.

Using the Lemma 3.6, for each $v^k \in T^\circ_\Omega(x^k)$, we have a smooth function ϕ_k such that $-\nabla\phi_k(x^k) = v^k$ and attains its global minimum relative to Ω uniquely at x^k. Since MPEC-AKKT is an optimality condition, by Theorem 3.1, there are sequences $\{x^{k,s}\}$, $\{z^{k,s}\}$, $\{v^{k,s}\}$ and $\{\gamma^{k,s}\}$ satisfying $x^{k,s} \to_s x^k$, $z^{k,s} \to_s F(x^k)$, $v^{k,s} := -\nabla\phi_k(x^{k,s}) \to_s v^k$ and $-v^{k,s} + \nabla F(x^{k,s})^\top\gamma^{k,s} \to_s 0$ with $\gamma^{k,s} \in N_\Lambda(z^{k,s})$. Thus, for each $k \in \mathbb{N}$, there exists $s(k)$ such that:

- $\|x^k - x^{k,s(k)}\| + \|F(x^k) - z^{k,s(k)}\| < 1/2^k$;
- $\|v^k - \nabla F(x^{k,s})^\top\gamma^{k,s(k)}\| < 1/2^k$ with $\gamma^{k,s(k)} \in N_\Lambda(z^{k,s(k)})$.

Thus, we have found sequences such that $x^{k,s(k)} \to_k x^*$ $z^{k,s(k)} \to_k F(x^*)$ and $\nabla F(x^{k,s})^\top\gamma^{k,s(k)} \to_k \omega$ with $\nabla F(x^{k,s})^\top\gamma^{k,s(k)} \in \nabla F(x^{k,s})^\top N_\Lambda(z^{k,s(k)})$. Thus, $\omega \in \limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z)$. If in addition, we suppose that MPEC-CCP holds at x^*, we obtain $N_\Omega(x^*) \subset \nabla F(x^*)^\top N_\Lambda(F(x^*))$. □As the consequence of the Theorem 3.7, we get

Corollary 3.8. *If x^* is a local minimizer with MPEC-CCP holding at x^*. Then, x^* is a M-stationarity point.*

Proof Let x^* be a local minimizer of (13) for a smooth objective function f. Due to the optimality, [50, Theorem 6.12], $0 \in \nabla f(x^*) + N_\Omega(x^*)$ but since MPEC-CCP holds at x^*, we get $0 \in \nabla f(x^*) + \nabla F(x^*)^\top N_\Lambda(F(x^*))$ which is equivalent to state that x^* is a M-stationary point, by Proposition 2.4 (2). □

4 Relationship between old and new MPEC-CQs

In this section we show the relationship between the recently proposed MPEC-CQ and the other MPEC-CQs found in the literature. We will focus on MPEC-RCPLD, MPEC-Abadie CQ and MPEC-quasinormality.

4.1 MPEC-RCPLD and MPEC-CCP

In this subsection, we will show that MPEC-CCP is strictly weaker than MPEC-RCPLD. First some notations. Given $x \in \mathbb{R}^n$ and index subsets $I_1 \subset \{1,\ldots,q\}$, $I_2 \subset \{1,\ldots,m\}$ and $I_3 \subset \{1,\ldots,m\}$, following [26], we define

$$\mathcal{G}(x; I_1, I_2, I_3) := \{\nabla h_i(x), \nabla H_\imath(x), \nabla G_\jmath(x) : i \in I_1, \imath \in I_2, \jmath \in I_3\}. \tag{38}$$

Denote by span $\mathcal{G}(x; I_1, I_2, I_3)$ the linear subspace generated by $\mathcal{G}(x; I_1, I_2, I_3)$. Put $\mathcal{E} := \{1,\ldots,q\}$. Now, we proceed with the definition of MPEC-RCPLD.

Definition 4.1. Let x^* be a feasible point and $\mathcal{E}' \subset \mathcal{E}$, $\mathcal{I}' \subset \mathcal{I}(x^*)$, $\mathcal{K}' \subset \mathcal{K}(x^*)$ be index sets such that $\mathcal{G}(x^*; \mathcal{E}', \mathcal{I}', \mathcal{K}')$ is a basis for span $\mathcal{G}(x^*; \mathcal{E}, \mathcal{I}(x^*), \mathcal{K}(x^*))$. We say that MPEC relaxed constant positive linear dependence CQ (MPEC-RCPLD) holds at x^* iff there is a $\delta > 0$ such that

1. $\mathcal{G}(x; \mathcal{E}, \mathcal{I}(x^*), \mathcal{K}(x^*))$ has the same rank for each $x \in \mathbb{B}(x^*, \delta)$;
2. For each $A' \subset A(x^*)$ and $\mathcal{J}_H', \mathcal{J}_G' \subset \mathcal{J}(x^*)$, if there are multipliers $\{\lambda, \mu, u, v\}$ which are not all zero with $\mu_j \geq 0$ for each $j \in A'$, and either $u_\ell v_\ell = 0$ or $u_\ell > 0, v_\ell > 0$ for each $\ell \in \mathcal{J}(x^*)$, such that

$$\sum_{i \in A'} \mu_i \nabla g_i(x^*) + \sum_{j \in \mathcal{E}'} \lambda_j \nabla h_j(x^*) + \sum_{\imath \in \mathcal{I}' \cup \mathcal{J}_H'} u_\imath \nabla H_\imath(x^*) + \sum_{\jmath \in \mathcal{K}' \cup \mathcal{J}_G'} v_\jmath \nabla G_\jmath(x^*) = 0. \tag{39}$$

Then, the set $\{\mathcal{G}(x; \mathcal{E}', \mathcal{I}' \cup \mathcal{J}_H', \mathcal{K}' \cup \mathcal{J}_G'), \nabla g_j(x) : j \in A'\}$ is linearly independent for every $x \in \mathbb{B}(x^*, \delta)$.

Now, we will continue with the next theorem.

Theorem 4.1. *MPEC-RCPLD implies MPEC-CCP.*

Proof Take $\omega^* \in \limsup_{(x,z) \to (x^*, F(x^*))} \nabla F(x)^\top N_\Lambda(z)$. Then, there are sequences $\{x^k\}$, $\{z^k\}$, $\{\omega^k\}$, $\{\gamma^k\}$ such that $x^k \to x^*$, $z^k \to F(x^*)$, $\omega^k \to \omega^*$ with $\omega^k := DF(x^k)^\top \gamma^k$ and $\gamma^k \in N_\Lambda(z^k)$. By Lemma 3.4, there is no loss of generality, if we take $\gamma^k \in \widehat{N}_\Lambda(z^k)$. Set $\gamma^k := (\mu^k, \lambda^k, (u_1^k, v_1^k), \ldots, (u_m^k, v_m^k))$. Then, for k sufficiently large, we have that ω^k is equal to

$$\sum_{j \in A(x^*)} \mu_j^k \nabla g_j(x^k) + \sum_{i \in \mathcal{E}} \lambda_i^k \nabla h_i(x^k) - \sum_{\imath \in \mathcal{I}(z^k) \cup \mathcal{J}(z^k)} u_\imath^k \nabla H_\imath(x^k) - \sum_{\jmath \in \mathcal{K}(z^k) \cup \mathcal{J}(z^k)} v_\jmath^k \nabla G_\jmath(x^k) \tag{40}$$

where $\mu^k \in \mathbb{R}_+^p$, $\text{supp}(\mu^k) \subset A(x^*)$ $\text{supp}(u^k) \subset \mathcal{I}(z^k) \cup \mathcal{J}(z^k)$ $\text{supp}(v^k) \subset \mathcal{K}(z^k) \cup \mathcal{J}(z^k)$ and $u_\ell^k \geq 0, v_\ell^k \geq 0$, $\forall \ell \in \mathcal{J}(z^k)$ since $\gamma \in \widehat{N}_\Lambda(z^k)$. Taking an adequate subsequence, we assume that $\mathcal{I}(x^*) \subset \mathcal{I}(z^k)$ and $\mathcal{K}(x^*) \subset \mathcal{K}(z^k)$, $\forall k \in \mathbb{N}$. Now, we decompose each ω^k into two parts ω_1^k and ω_2^k, where

$$\omega_1^k := \sum_{i \in \mathcal{E}} \lambda_i^k \nabla h_i(x^k) - \sum_{\imath \in \mathcal{I}(z^*)} u_\imath^k \nabla H_\imath(x^k) - \sum_{\jmath \in \mathcal{K}(z^*)} v_\jmath^k \nabla G_\jmath(x^k) \tag{41}$$

and ω_2^k is given by

$$\sum_{j \in A(x^*)} \mu_j^k \nabla g_j(x^k) - \sum_{\imath \in (\mathcal{I}(z^k) \setminus \mathcal{I}(x^*)) \cup \mathcal{J}(z^k)} u_\imath^k \nabla H_\imath(x^k) - \sum_{\jmath \in (\mathcal{K}(z^k) \setminus \mathcal{K}(x^*)) \cup \mathcal{J}(z^k)} v_\jmath^k \nabla G_\jmath(x^k). \tag{42}$$

Take index sets $\mathcal{E}' \subset \mathcal{E}$, $\mathcal{I}' \subset \mathcal{I}(x^*)$ and $\mathcal{K}' \subset \mathcal{K}(x^*)$ such that $\mathcal{G}(x^*; \mathcal{E}', \mathcal{I}', \mathcal{K}')$ is a basis of span $\mathcal{G}(x^*; \mathcal{E}, \mathcal{I}(x^*), \mathcal{K}(x^*))$. By MPEC-RCPLD, $\mathcal{G}(x^k; \mathcal{E}', \mathcal{I}', \mathcal{K}')$ is a basis of span $\mathcal{G}(x^k; \mathcal{E}, \mathcal{I}(x^*), \mathcal{K}(x^*))$, for $k \in \mathbb{N}$. Thus, we can write ω_1^k as

$$\omega_1^k = \sum_{i \in \mathcal{E}'} \hat{\lambda}_i^k \nabla h_i(x^k) - \sum_{\imath \in \mathcal{I}'} \hat{u}_\imath^k \nabla H_\imath(x^k) - \sum_{\jmath \in \mathcal{K}'} \hat{v}_\jmath^k \nabla G_\jmath(x^k) \tag{43}$$

for some $(\hat{\lambda}^k, \hat{u}^k, \hat{v}^k) \in \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^m$ with $\text{supp}(\hat{\lambda}^k) \subset \mathcal{E}'$, $\text{supp}(\hat{u}^k) \subset \mathcal{I}'$ and $\text{supp}(\hat{v}^k) \subset \mathcal{K}'$. Using Lemma 2.3 for each $k \in \mathbb{N}$, we find index subsets $A'(k) \subset A(x^*)$, $\mathcal{I}_+'(k) \subset \mathcal{I}(x^k) \setminus \mathcal{I}(x^*)$, $\mathcal{K}_+'(k) \subset \mathcal{K}(x^k) \setminus \mathcal{K}(x^*)$, $\mathcal{J}_H'(k), \mathcal{J}_G'(k) \subset \mathcal{J}(x^k)$ such that

$$\omega_2^k = \sum_{j \in A'(k)} \tilde{\mu}_j^k \nabla g_j(x^k) - \sum_{\imath \in \mathcal{I}_+'(k) \cup \mathcal{J}_H'(k)} \tilde{u}_\imath^k \nabla H_\imath(x^k) - \sum_{\jmath \in \mathcal{K}_+'(k) \cup \mathcal{J}_G'(k)} \tilde{v}_\jmath^k \nabla G_\jmath(x^k) \tag{44}$$

for multipliers $(\tilde{\mu}^k, \tilde{u}^k, \tilde{v}^k) \in \mathbb{R}_+^p \times \mathbb{R}^m \times \mathbb{R}^m$ with $\text{supp}(\tilde{\mu}^k) \subset A'(k)$, $\text{supp}(\tilde{u}^k) \subset \mathcal{I}_+'(k) \cup \mathcal{J}_H'(k)$, $\text{supp}(\tilde{v}^k) \subset \mathcal{K}_+'(k) \cup \mathcal{J}_G'(k)$, and $\tilde{u}_\ell^k, \tilde{v}_\ell^k \geq 0$, $\ell \in \mathcal{J}(z^k)$ such that for each $k \in \mathbb{N}$, the vectors

$$\mathcal{G}(x^k; \mathcal{E}', \mathcal{I}' \cup \mathcal{I}_+'(k) \cup \mathcal{J}_H'(k), \mathcal{K}' \cup \mathcal{K}_+'(k) \cup \mathcal{J}_G'(k)) \text{ and } \{\nabla g_j(x^k) : j \in A'(k)\} \tag{45}$$

are linearly independent. Since there are only a finite number of possible index subset, we can assume, after taken an adequate subsequence, that $A^{'}(k)$, $\mathcal{I}^{'}_{+}(k)$, $\mathcal{K}^{'}_{+}(k)$, $\mathcal{J}^{'}_{H}(k)$ and $\mathcal{J}^{'}_{G}(k)$ are independent of k. Denote them by $A^{'}$, $\mathcal{I}^{'}_{+}$, $\mathcal{K}^{'}_{+}$, $\mathcal{J}^{'}_{H}$ and $\mathcal{J}^{'}_{G}$ respectively. Substituting (44) and (43) into (40), we get that ω^k is equal to

$$\sum_{j\in A'} \bar{\mu}^k_j \nabla g_j(x^k) + \sum_{i\in\mathcal{E}'} \bar{\lambda}^k_i \nabla h_i(x^k) - \sum_{\imath\in\mathcal{I}'\cup\mathcal{I}'_+\cup\mathcal{J}'_H} \bar{u}^k_\imath \nabla H_\imath(x^k) - \sum_{\jmath\in\mathcal{K}'\cup\mathcal{K}'_+\cup\mathcal{J}'_G} \bar{v}^k_\jmath \nabla G_\jmath(x^k) \tag{46}$$

where the multipliers $(\bar{\mu}^k, \bar{\lambda}^k, \bar{u}^k, \bar{v}^k) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^m$ are given by

$$\begin{array}{lll} \bar{\mu}^k_j := \tilde{\mu}^k_j(j \in A^{'}), & \bar{u}^k_\imath := \hat{u}^k_\imath(\imath \in \mathcal{I}^{'}), & \bar{v}^k_\jmath := \hat{v}^k_\jmath(\jmath \in \mathcal{K}^{'}), \\ \bar{\lambda}^k_i := \hat{\lambda}^k_i(i \in \mathcal{E}^{'}), & \bar{u}^k_\imath := \tilde{u}^k_\imath(\imath \in \mathcal{I}^{'}_+ \cup \mathcal{J}^{'}_H) & \bar{v}^k_\jmath := \tilde{v}^k_\jmath(\jmath \in \mathcal{K}^{'}_+ \cup \mathcal{J}^{'}_G)) \end{array} \tag{47}$$

with $\mathrm{supp}(\bar{\mu}^k) \subset A^{'}$, $\mathrm{supp}(\bar{\lambda}^k) \subset \mathcal{E}^{'}$, $\mathrm{supp}(\bar{u}^k) \subset \mathcal{I}^{'} \cup \mathcal{I}^{'}_+ \cup \mathcal{J}^{'}_H$, $\mathrm{supp}(\bar{v}^k) \subset \mathcal{K}^{'} \cup \mathcal{K}^{'}_+ \cup \mathcal{J}^{'}_G$ and $\bar{u}^k_\ell, \bar{v}^k_\ell \geq 0$, for $\ell \in \mathcal{J}(z^k)$. Furthermore, we can see that

$$\bar{\gamma}^k := (\bar{\mu}^k, \bar{\lambda}^k, (\bar{u}^k_1, \bar{v}^k_1), \dots, (\bar{u}^k_m, \bar{v}^k_m)) \in N_\Lambda(z^k). \tag{48}$$

Now, the sequence $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{u}^k, \bar{v}^k)\}$ has a bounded subsequence, otherwise, dividing (46) by $M_k := \|(\bar{\mu}^k, \bar{\lambda}^k, \bar{u}^k, \bar{v}^k)\|$ and considering an adequate convergent subsequence of $M_k^{-1}(\bar{\mu}^k, \bar{\lambda}^k, \bar{u}^k, \bar{v}^k)$, says $\{(\bar{\mu}, \bar{\lambda}, \bar{u}, \bar{v})\}$, we obtain that

$$\sum_{j\in A'} \bar{\mu}_j \nabla g_j(x^*) + \sum_{i\in\mathcal{E}'} \bar{\lambda}_i \nabla h_i(x^*) - \sum_{\imath\in\mathcal{I}'\cup\mathcal{I}'_+\cup\mathcal{J}'_H} \bar{u}_\imath \nabla H_\imath(x^*) - \sum_{\jmath\in\mathcal{K}'\cup\mathcal{K}'_+\cup\mathcal{J}'_G} \bar{v}_\jmath \nabla G_\jmath(x^*) = 0 \tag{49}$$

where $\{(\bar{\mu}, \bar{\lambda}, \bar{u}, \bar{v})\}$ is a nonzero vector. Furthermore, $\bar{\mu} \geq 0$, $\mathrm{supp}(\bar{\mu}) \subset A^{'}$ and either $\bar{u}_\ell \bar{v}_\ell = 0$ or $\bar{u}_\ell > 0, \bar{v}_\ell > 0$ for each $\ell \in \mathcal{J}(x^*)$ since N_Λ is outer semicontinuous, [50, Proposition 6.6]. Now, (45) and (49) are not compatible with the MPEC-RCPLD assumption. Thus, the sequence $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{u}^k, \bar{v}^k)\}$ has a convergent subsequence. Assume, without loss of generality, that the sequence $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{u}^k, \bar{v}^k)\}$ itself converges to some vector $(\bar{\mu}, \bar{\lambda}, \bar{u}, \bar{v})$. By (48) and the outer semicontinuity of N_Λ, we get that $(\bar{\mu}, \bar{\lambda}, (\bar{u}_1, \bar{v}_1), \dots, (\bar{u}_m, \bar{v}_m))$ is in $N_\Lambda(F(x^*))$. Taking limit in (46), ω^* can be written as

$$\sum_{j\in A'} \bar{\mu}_j \nabla g_j(x^*) + \sum_{i\in\mathcal{E}'} \bar{\lambda}_i \nabla h_i(x^*) - \sum_{\imath\in\mathcal{I}'\cup\mathcal{I}'_+\cup\mathcal{J}'_H} \bar{u}_\imath \nabla H_\imath(x^*) - \sum_{\jmath\in\mathcal{K}'\cup\mathcal{K}'_+\cup\mathcal{J}'_G} \bar{v}_\jmath \nabla G_\jmath(x^*), \tag{50}$$

where $\mathrm{supp}(\bar{\mu}) \subset A^{'}$, $\mathrm{supp}(\bar{\lambda}) \subset \mathcal{E}^{'}$, $\mathrm{supp}(\bar{u}) \subset \mathcal{I}^{'} \cup \mathcal{I}^{'}_+ \cup \mathcal{J}^{'}_H$, $\mathrm{supp}(\bar{v}) \subset \mathcal{K}^{'} \cup \mathcal{K}^{'}_+ \cup \mathcal{J}^{'}_G$ $\bar{\mu} \in \mathbb{R}^n_+$ and either $\bar{u}_\ell \bar{v}_\ell = 0$ or $\bar{u}_\ell > 0, \bar{v}_\ell > 0$ for each $\ell \in \mathcal{J}(x^*)$. From (50) and since $\mathcal{I}^{'} \cup \mathcal{I}^{'}_+$ and $\mathcal{K}^{'} \cup \mathcal{K}^{'}_+$ are disjoint index sets, we conclude that ω^* is an element of $\nabla F(x^*)^\top N_\Lambda(F(x^*))$. Thus, MPEC-CCP holds at x^*. □MPEC-CCP is strictly weaker than MPEC-RCPLD as the next example shows.

Example 4.1 (MPEC-CCP does not imply MPEC-RCPLD)**.** Consider in $\mathbb{R}^2$, the point $x^* = (0, 0)$ and the constraint system defined by $g_1(x_1, x_2) := x_1 + x_2$; $g_2(x_1, x_2) := -x_1 - x_2$; $g_3(x_1, x_2) := x_1^2 + x_2^2$ and $0 \leq H_1(x_1, x_2) := x_1 \perp G_1(x_1, x_2) := x_2 \geq 0$. MPEC-RCPLD does not hold, since $\nabla g_3(x_1, x_2)$ is positive linearly dependent at $x^* = (0, 0)$, but not in any neighborhood of x^*. By the other hand, MPEC-CCP holds at x^*, since $\nabla F(x^*)^\top N_\Lambda(F(x^*)) = \mathbb{R}^2$.

4.2 MPEC-CCP and MPEC-Abadie CQ

MPEC-CQs come from several different approaches and in many cases, clarifying the relations among them is a difficult task. Here, we show the MPEC-CCP implies the MPEC-Abadie CQ under certain assumption. Now, we continue with the definition of MPEC-Abadie CQ.

Definition 4.2. We say that the MPEC-Abadie CQ (MPEC-ACQ) holds at x^* iff $T_\Omega(x^*) = L_\Omega(x^*)$, where $L_\Omega(x^*) := \{d \in \mathbb{R}^n : \nabla F(x^*)d \in T_\Lambda(F(x^*))\}$.

Using the Proposition 2.1, $L_\Omega(x^*)$ can be written as:

$$L_\Omega(x^*) = \left\{ d \in \mathbb{R}^n : \quad \begin{array}{l} \nabla g_j(x^*)^\top d \le 0, j \in A(x^*); \quad \nabla h_j(x^*)^\top d = 0, j \in \mathcal{E} \\ \nabla H_j(x^*)^\top d = 0, j \in \mathcal{I}(x^*); \quad \nabla G_j(x^*)^\top d = 0, j \in \mathcal{K}(x^*) \\ 0 \le \nabla H_j(x^*)^\top d \perp \nabla G_j(x^*)^\top d \ge 0, j \in \mathcal{J}(x^*) \end{array} \right\}. \tag{51}$$

Note that always $T_\Omega(x^*) \subset L_\Omega(x^*)$. Now, we proceed with the introduction of the set-valued mapping $\mathbb{R}^n \times \Lambda \ni (x, z) \rightrightarrows L_\Omega(x, z)$ where $L_\Omega(x, z)$ is given by

$$L_\Omega(x, z) := \left\{ d \in \mathbb{R}^n : \quad \begin{array}{l} \nabla g_j(x)^\top d \le 0, j \in A(x^*); \quad \nabla h_j(x)^\top d = 0, j \in \mathcal{E} \\ \nabla H_j(x)^\top d = 0, j \in \mathcal{I}(z); \quad \nabla G_j(x)^\top d = 0, j \in \mathcal{K}(z) \\ 0 \le \nabla H_j(x)^\top d \perp \nabla G_j(x)^\top d \ge 0, j \in \mathcal{J}(z) \end{array} \right\}. \tag{52}$$

Note that $L_\Omega(x, z)$ coincides $L_\Omega(x^*)$ when $(x, z) = (x^*, F(x^*))$. Thus, $L_\Omega(x, z)$ can be considered as a perturbation of $L_\Omega(x^*)$. Now, we continue with the relations between $\nabla F(x)^\top \widehat{N}_\Lambda(z)$, $L^\circ_\Omega(x, z)$ and $\nabla F(x)^\top N_\Lambda(z)$. Since $\widehat{N}_\Lambda(z) = T^\circ_\Lambda(z)$, a simple inspection shows that $\nabla F(x)^\top \hat{N}_\Lambda(z) \subset L^\circ_\Omega(x, z)$. By the other hand, adapting the proof of [26, Theorem 3.3], we see that $L^\circ_\Omega(x, z) \subset \nabla F(x)^\top N_\Lambda(z)$. We summarize these results in the next proposition.

Proposition 4.2. *We always have* $\nabla F(x)^\top \widehat{N}_\Lambda(z) \subset L^\circ_\Omega(x, z) \subset \nabla F(x)^\top N_\Lambda(z)$.

The next theorem can be seen as a primal version of Theorem 3.7

Theorem 4.3. *We always have,* $\liminf_{(x,z)\to(x^*,F(x^*))} L^{\circ\circ}_\Omega(x, z) \subset T_\Omega(x^*)$.

Proof. By Theorem 3.7, we have $N_\Omega(x^*) \subset \limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z)$. Then, we get

$$[\limsup_{(x,z)\to(x^*,F(x^*))} L^\circ_\Omega(x, z)]^\circ = [\limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z)]^\circ \subset N^\circ_\Omega(x^*) \subset T_\Omega(x^*) \tag{53}$$

where the first inclusion follows from Proposition 4.2 and the last inclusion from [50, Theorems 6.28(b) and 6.26]. By the duality theorem [11, Theorem 1.1.8], $[\limsup_{(x,z)\to(x^*,F(x^*))} L^\circ_\Omega(x, z)]^\circ = \liminf_{(x,z)\to(x^*,F(x^*))} L^{\circ\circ}_\Omega(x, z)$. From the last equality, we obtain the desired result. ☐

Motivated by Theorem 4.3, we define the next property

Definition 4.3. Let x^* be a feasible point. We say that MPEC-Continuity of the Linearized Cone (MPEC-CLC) holds at x^* if the set-valued mapping $L_\Omega(x, z)$ is isc at $(x^*, F(x^*))$.

From the inclusion $L_\Omega(x, z) \subset \text{cl conv } L_\Omega(x, z) = L^{\circ\circ}_\Omega(x, z)$ and the Theorem 4.3, we get that $L_\Omega(x^*) \subset T_\Omega(x^*)$ if MPEC-CLC holds. Thus, MPEC-CLC implies MPEC-ACQ and as consequence MPEC-CLC is a CQ for M-stationarity. The next theorem shows that MPEC-CLC implies MPEC-CCP.

Theorem 4.4. *MPEC-CLC always implies MPEC-CCP.*

Proof. From MPEC-CLC, we get that $L_\Omega(x^*) \subset \liminf_{(x,z)\to(x^*,F(x^*))} L^{\circ\circ}_\Omega(x, z)$. By polarity theorem, we have $[\liminf_{(x,z)\to(x^*,F(x^*))} L^{\circ\circ}_\Omega(x, z)]^\circ \subset L^\circ_\Omega(x^*) \subset \nabla F(x^*)^\top N_\Lambda(F(x^*))$, where the last inclusion follows from Proposition 4.2. By [11, Theorem 1.1.8], $\limsup_{(x,z)\to(x^*,F(x^*))} L^\circ_\Omega(x, z)$ ($= \limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z)$), is a subset of $[\liminf_{(x,z)\to(x^*,F(x^*))} L^{\circ\circ}_\Omega(x, z)]^\circ$. Thus, MPEC-CCP holds. ☐

Remark 4. In the absence of complementary constraints, $N_\Lambda(z) = \widehat{N}_\Lambda(z)$, $\forall z \in \Lambda$. Thus by Proposition 4.2, we get that $L^\circ_\Omega(x, z) = \nabla F(x^*)^\top N_\Lambda(z)$ and $L^{\circ\circ}_\Omega(x, z) = L_\Omega(x, z)$. Thus, the osc of $\nabla F(x)^\top N_\Lambda(z)$ is equivalent to the isc of $L^\circ_\Omega(x, z)$. So, MPEC-CLC can be seeing as a *primal* version of MPEC-CCP.

An interesting question is the exact relation between MPEC-RCPLD and MPEC-ACQ. Several partial answers have been provided in the literature. In [26, Example 4.1], the authors showed an example where MPEC-ACQ holds but not MPEC-RCPLD. In the other hand, in [18], the authors proved that in the absence of inequality constraints and the presence of only one complementary constraint, MPEC-RCPLD always implies MPEC-ACQ, c.f. [18, Theorem 3.3]. A simple observation shows that under $L^\circ_\Omega(x^*) = \nabla F(x^*)^\top N_\Lambda(F(x^*))$, MPEC-CCP (and as consequence MPEC-RCPLD) is sufficient to guarantee the validity of MPEC-ACQ.

Proposition 4.5. *If we have that $L_\Omega^\circ(x^*) = \nabla F(x^*)^\top N_\Lambda(F(x^*))$, Then, MPEC-CCP implies MPEC-ACQ.*

Proof. Since MPEC-CCP holds at x^*, then $N_\Omega(x^*) \subset \nabla F(x^*)^\top N_\Lambda(F(x^*))$. Now, by assumption, this implies that $N_\Omega(x^*) \subset L_\Omega^\circ(x^*)$. By polarity theorem and [50, Theorems 6.28(b) and 6.26], we have

$$L_\Omega(x^*) \subset \text{cl conv } L_\Omega(x^*) = L_\Omega^{\circ\circ}(x^*) \subset N_\Omega^\circ \subset T_\Omega(x^*). \tag{54}$$

Since, $T_\Omega(x^*)$ is always included in $L_\Omega^\circ(x^*)$, MPEC-ACQ holds at x^*. □

Without complementary constraints, $L_\Omega^\circ(x^*) = \nabla F(x^*)^\top N_\Lambda(F(x^*))$. So, MPEC-CCP (in this case, CCP) always implies Abadie CQ. The example 4.2 will show that MPEC-ACQ is not sufficient to guarantee MPEC-CCP,

Remark 5. It is possible that $L_\Omega^\circ(x^*) = \nabla F(x^*)^\top N_\Lambda(F(x^*))$ even if $\mathcal{J}(x^*)$ is not empty (i.e. x^* has degenerate index). For instance, consider in $\mathbb{R}$, the constraints $G(x) = H(x) = x^2$, $g(x) = h(x) = 0$ and $x^* = 0$.

4.3 MPEC-CCP, MPEC-quasinormality and MPEC-pseudonormality

Here we will show that MPEC-CCP is independent of MPEC-pseudonormality and MPEC-quasinormality. Let us recall the definition of MPEC-quasinormality,

Definition 4.4. We say that MPEC-quasinormality holds at $x^* \in \Omega$ if whenever

$$\sum_{j\in A(x^*)} \mu_j \nabla g_j(x^*) + \sum_{i\in\mathcal{E}} \lambda_i \nabla h_i(x^*) - \sum_{\imath=1}^{m} u_\imath \nabla H_\imath(x^*) - \sum_{\jmath=1}^{m} v_\jmath \nabla G_\jmath(x^*) = 0 \tag{55}$$

for some nonzero multipliers $\{(\mu, \lambda, u, v)\}$ with $\mu \in \mathbb{R}^p_+$, $\text{supp}(\mu) \subset A(x^*)$, $\text{supp}(u) \subset \mathcal{I}(x^*) \cup \mathcal{J}(x^*)$, $\text{supp}(v) \subset \mathcal{K}(x^*) \cup \mathcal{J}(x^*)$ and either $u_\ell v_\ell = 0$ or $u_\ell > 0, v_\ell > 0$, $\ell \in \mathcal{J}(x^*)$, there is no sequence $x^k \to x^*$ such that, for each k, $\mu_i g_i(x^k) > 0$ when $\mu_j > 0$, $\lambda_i h_i(x^k) > 0$ if $\lambda_i > 0$, $-u_\imath H_i(x^k) > 0$ when $u_\imath > 0$ and $-v_\jmath G_\jmath(x^k) > 0$ if $v_\jmath > 0$.

The definition of MPEC-pseudonormality is the following one

Definition 4.5. We say that MPEC-pseudonormality holds at $x^* \in \Omega$ if whenever (55) holds for some nonzero multipliers $\{(\mu, \lambda, u, v)\}$ with $\mu \in \mathbb{R}^p_+$, $\text{supp}(\mu) \subset A(x^*)$, $\text{supp}(u) \subset \mathcal{I}(x^*) \cup \mathcal{J}(x^*)$, $\text{supp}(v) \subset \mathcal{K}(x^*) \cup \mathcal{J}(x^*)$ and either $u_\ell v_\ell = 0$ or $u_\ell > 0, v_\ell > 0$, $\ell \in \mathcal{J}(x^*)$, there is no sequence $x^k \to x^*$, with $\sum_{j=1}^p \mu_j g_j(x^k) + \sum_{i=1}^q \lambda_i \nabla h_i(x^k) - \sum_{i=1}^m u_i H_i(x^k) - \sum_{i=1}^m v_i G_i(x^k) > 0$, $\forall k \in \mathbb{N}$.

From [35], we get that MPEC-pseudonormality implies MPEC-quasinormality and MPEC-ACQ. Both conditions are sufficient to guarantee that every minimizer is a M-stationary point and when x^* satisfies the strict complementarity condition, (i.e. $\mathcal{J}(x^*) = \emptyset$) then MPEC-quasinormality implies MPEC-ACQ. See [58, Theorem 3.1]. The next examples will show that MPEC-CCP is independent of the MPEC-pseudonormality and MPEC-quasinormality.

Example 4.2 (MPEC-pseudonormality does not imply MPEC-CCP)**.** In $\mathbb{R}^2$, consider $x^* := (0,0)$, $g_1(x_1,x_2) = -x_1$, $g_2(x_1,x_2) = x_1 - x_2^2 x_1^2$, $H_1(x_1,x_2) := x_1$ and $G_1(x_1,x_2) := 1$. MPEC-pseudonormality holds at x^*. Indeed, assume by contradiction that there are non zero multipliers (μ_1, μ_2, u) such that $\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) - u \nabla H_1(x^*) = 0$. Then, $u = \mu_2 - \mu_1$ with $\mu_1 \geq 0$ and $\mu_2 \geq 0$. Take any sequence $\{x^k = (x_1^k, x_2^k)\}$ with $x^k \to x^*$. If, for $k \in \mathbb{N}$, we have $\mu_1 g_1(x^k) + \mu_2 g_2(x^k) - u H_1(x^k) > 0$, the last inequality implies that $-\mu_2 (x_2^k)^2 (x_1^k)^2 > 0$ which is impossible. Now, we will show that MPEC-CCP fails. By some calculations, we have $\nabla F(x^*)^\top N_\Lambda(F(x^*)) = \mathbb{R} \times \{0\}$. Define $x_1^k := 1/k$, $x_2^k := x_1^k$, $u^k := 0$, $\mu_2^k := (2x_2^k (x_1^k)^2)^{-1}$, $\mu_1^k := \mu_2 (1 - 2x_1^k (x_2^k)^2)$ and $z^k := (0, 0, (0, -1))$. Clearly, $(x_1^k, x_2^k) \to (0,0)$, $\mu_1^k, \mu_2^k \geq 0$. Set $w^k := \mu_1^k \nabla g_1(x^k) + \mu_2^k \nabla g_2(x^k) - u^k \nabla H_1(x^k) = (0, -1)$, $\forall k \in \mathbb{N}$. Thus, $(0,-1) \in \limsup_{(x,z)\to(x^*,F(x^*))} \nabla F(x)^\top N_\Lambda(z)$ but $(0,-1)$ does not belong to $\nabla F(x^*)^\top N_\Lambda(F(x^*)) = \mathbb{R} \times \{0\}$. So, MPEC-CCP fails.

Since MPEC-pseudonormality implies MPEC-ACQ, the above example also shows that MPEC-ACQ also does not imply MPEC-CCP.

Example 4.3 (MPEC-quasinormality fails but MPEC-CCP holds)**.** Consider in $\mathbb{R}^2$, the point $x^* := (0,0)$ and the constrained system given by $g_1(x_1,x_2) = -x_1 + x_2^2 \exp(x_1)$; $g_2(x_1,x_2) = x_1 \exp(x_2)$; $0 \le H_1(x_1,x_2) := 1 \perp G_1(x_1,x_2) := x_2 \ge 0$. MPEC-CCP holds at x^*. In fact, it follows from $\nabla F(x^*)^\top N_\Lambda(F(x^*)) = \mathbb{R}^2$. Now, we will see that MPEC-quasinormality fails. Take $x_1^k := 1/k$, $x_2^k := \sqrt{2x_1^k \exp(-x_1^k)}$, $k \in \mathbb{N}$, and multipliers $\mu_1 := 1$, $\mu_2 := 1$, $v := 0$. For that choice, $\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) - v \nabla G_1(x^*) = 0$. Evenmore, $\mu_2 g_2(x^k) = \mu_2 x_1^k \exp(x_2^k) > 0$ and $\mu_1 g_1(x^k) = \mu_1(-x_1^k + (x_2^k)^2 \exp(x_1^k)) = \mu_1 x_1^k > 0$. Thus, MPEC-quasinormality cannot hold at x^*

The above example also shows that MPEC-CCP does not implies MPEC-pseudonormality. From these examples, we get that MPEC-CCP is independent of MPEC-pseudonormality and MPEC-quasinormality. Another important MPEC-CQ is the following one introduced in [56] under a different name, namely, MPEC-*no nonzero abnormal multiplier CQ* (MPEC-NNAMCQ).

Definition 4.6. We say that MPEC *generalized MFCQ* (MPEC-GMFCQ) holds at $x^* \in \Omega$ if there are no nonzero multipliers $\{(\mu, \lambda, u, v)\}$ with $\mu \in \mathbb{R}^p_+$, $\text{supp}(\mu) \subset A(x^*)$, $\text{supp}(u) \subset \mathcal{I}(x^*) \cup \mathcal{J}(x^*)$, $\text{supp}(v) \subset \mathcal{K}(x^*) \cup \mathcal{J}(x^*)$ and either $u_\ell v_\ell = 0$ or $u_\ell > 0, v_\ell > 0$ for each $\ell \in \mathcal{J}(x^*)$ such that (55) holds.

Observe that MPEC-GMFCQ is equivalent to state that $\gamma = 0$ is the unique solution of $\nabla F(x^*)^\top \gamma = 0$, $\gamma \in N_\Lambda(F(x^*))$, where F is given by (14). Furthermore, it is not difficult to see that MPEC-GMFCQ is weaker than MPEC-MFCQ condition; which state that there is no nonzero multipliers $\{\mu, \lambda, u, v\}$ such that (55) holds with $\mu \in \mathbb{R}^p_+$, $\text{supp}(\mu) \subset A(x^*)$, $\text{supp}(u) \subset \mathcal{I}(x^*) \cup \mathcal{J}(x^*)$ and $\text{supp}(v) \subset \mathcal{K}(x^*) \cup \mathcal{J}(x^*)$. Under MPEC-GMFCQ (and as consequence MPEC-MFCQ), we have that $\{\gamma \in N_\Lambda(F(x^*)) : \nabla f(x^*) + \nabla F(x^*)^\top \gamma = 0\}$ is bounded, for every smooth function f. The figure 1 shows the relations among several CQs for M-stationarity involved in this paper. For further informations about other MPEC-CQs, see [35, 27, 58, 26].

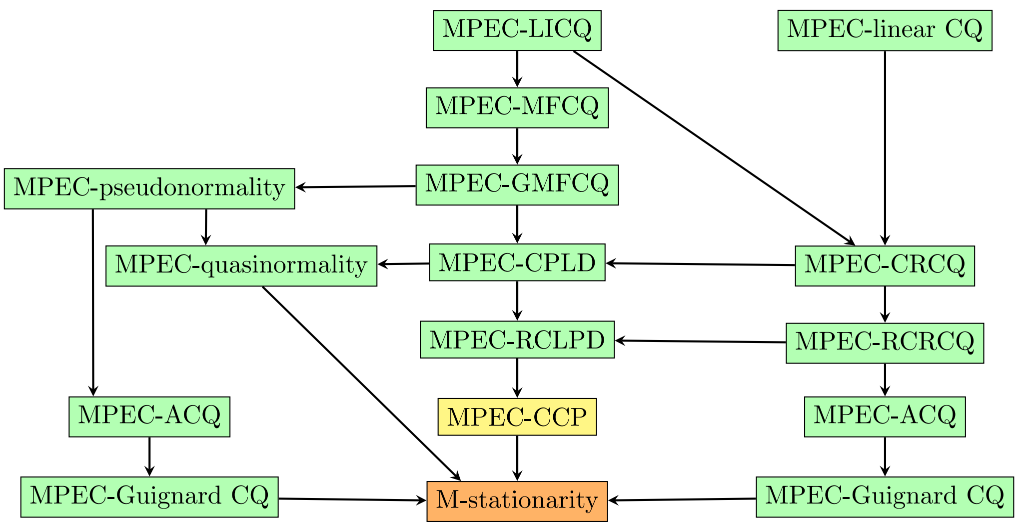

Figure 1: Relations among several CQs for M-stationarity.

5 Algorithmic applications of MPEC-CCP

In this section, we will show that MPEC-CCP can be used in the convergence analysis of algorithm for solving for MPECs. We will show that MPEC-CCP can replace other more stringent MPEC-CQs for M-stationarity in the assumptions to ensure the convergence of several algorithms under mild assumptions.

Since the complementarity constraints of the MPECs is the cause of difficulties from a numerical and theoretical point of view. A number of specially designed methods have been suggested to deal with it. For instance, we have the relaxations schemes [52, 32, 30, 54, 21, 39], the complementary-penalty methods [31, 49, 9], interior-penalty method [38]. We pay special attention to the complementary-penalty method, the interior-relation method of Leyffer et al [38], the L-relaxation method of Kanzow and Schwartz [36], the nonsmooth relaxation of Kadrani et al. [32] and the relaxation of Scholtes [52] and we will show that MPEC-CCP can be used to extend convergence results of such algorithms above under the same assumptions. Furthermore, we present a variant of the augmented Lagrangian method with strong convergence properties which generates sequences of iterates whose limit points conform the MPEC-AKKT condition.

5.1 A new simple variant of the Augmented Lagrangian method for MPEC

Here we propose a new method based on the method called sequential equality-constraints optimization (SECO) recently introduced in [13], where an each iteration, an equality-constrained optimization problem is solved approximately. In order to describe the method, we need first some notations. For $x \in \mathbb{R}^n$, a penalty parameter $\rho \in \mathbb{R}_{++}$ and multipliers $(\mu, \lambda, u, v) \in \mathbb{R}^p_+ \times \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^m$, we define the Augmented Lagrangian function $L_\rho(x, \lambda, \mu, u, v)$ [29, 46, 12] as

$$L_\rho(x, \lambda, \mu, u, v) := f(x) + \frac{\rho}{2}\left(\left\|H(x) + \frac{\lambda}{\rho}\right\|^2 + \left\|\left(G(x) + \frac{\mu}{\rho}\right)_+\right\|^2 + \left\|\left(\frac{u}{\rho} - H\right)_+\right\|^2 + \left\|\left(\frac{v}{\rho} - G\right)_+\right\|^2\right).$$

The main model algorithm is now describe, which can general framework. See also [12, 13].

Algorithm 5.1. *Let $\{\varepsilon_k\} \subset \mathbb{R}_+$ be a sequence of tolerance parameters with $\varepsilon_k \to 0$.*

Step 1 (Initialization). *Set parameters $\lambda_{min} < \lambda_{max}$, $\mu_{max} > 0$, $u_{max} > 0$, $v_{max} > 0$, $\gamma > 1$, $\tau \in [0, 1)$ and an iteration counter $k := 1$. Choose $x^0 \in \mathbb{R}^n$, $\bar{\lambda}^1 \in \mathbb{R}^m$, $\bar{\mu}^1 \in \mathbb{R}^p_+$ $\bar{\lambda}^1 \in \mathbb{R}^m_+$ $\bar{\lambda}^1 \in \mathbb{R}^m_+$ and a penalty parameter $\rho_1 > 0$. Set $V^0 = \max\{0, G(x^0)\}$, $V^0_g := \max\{g(x^0), 0\}$, $V^0_H := \max\{-H(x^0), 0\}$ and $V^0_G := \max\{-G(x^0), 0\}$.*

Step 2 (Solving the sub-problems). *Compute an ε_k-stationary point x^k, that is, compute (if possible) $x^k \in \mathbb{R}^n$ such that there exist $\eta^k \in \mathbb{R}^m$ satisfying*

$$\|\nabla_x L_{\rho_k}(x^k, \bar{\lambda}^k, \bar{\mu}^k, \bar{u}^k, \bar{v}^k) + \sum_{i=1}^m \eta^k_i \left(H^2_i(x^k)G_i(x^k)\nabla_x G(x^k) + G^2(x^k)H_i(x^k)\nabla_x H_i(x^k)\right)\| \leq \varepsilon_k \tag{56}$$

$$\max\{|\eta^k_i H_i(x^k)G_i(x^k)| : i = 1, \ldots, m\} \leq \varepsilon_k \tag{57}$$

$$\max\{|H_i(x^k)G_i(x^k)| : i = 1, \ldots, m\} \leq \varepsilon_k \tag{58}$$

If it is not possible to find x^k satisfying (56), (57) and (58) we stop the execution of the method.

Step 3 (Estimate of multipliers). *Compute*

$$\lambda^{k+1} := \bar{\lambda}^k + \rho_k h(x^k) \qquad u^{k+1} := \max\{0, \bar{u}^k - \rho_k H(x^k)\} \tag{59}$$

$$\mu^{k+1} := \max\{0, \bar{\mu}^k + \rho_k g(x^k)\} \qquad v^{k+1} := \max\{0, \bar{v}^k - \rho_k G(x^k)\} \tag{60}$$

Then, set $\bar{\mu}^{k+1}$, $\bar{u}^{k+1}$ and $\bar{v}^{k+1}$ as the projections of μ^{k+1}, u^{k+1} and v^{k+1} onto the safeguarding intervals $[0, \mu_{\max}]^p$, $[0, u_{\max}]^m$ and $[0, v_{\max}]^m$ respectively.

Step 4 (Update the penalty parameter).

Define $V^k_g := \max\{g(x^k), -\bar{\mu}^k/\rho_k\}$, $V^k_H := \max\{-H(x^k), -\bar{u}^k/\rho_k\}$ and $V^k_G := \max\{-G(x^k), -\bar{v}^k/\rho_k\}$.

If $\max\{\|h(x^k)\|, \|V^k_g\|, \|V^k_H\|, \|V^k_G\|\} \leq \tau \max\{\|h(x^{k-1})\|, \|V^{k-1}_g\|, \|V^{k-1}_H\|, \|V^{k-1}_G\|\}$. Then, set $\rho_{k+1} := \rho_k$. Otherwise, define $\rho_{k+1} := \gamma\rho_k$.

Step 4 (New iteration). *Put $k \leftarrow k + 1$. Go to Step 2.*

Remark 6. (1) Different error tolerances ε_k can used in each (56), (57) and (58) as long they approach to zero. Similarly, different updates for the penalty parameter can be considered without change the global convergence theory. See Theorem 5.1. (2) As Augmented Lagrangian algorithms, the efficient of such method, depends on the achievement of the Step 2. For instance, we can assure the fulfilment of the Step 2 if we solve exactly the sub-problems

$$\text{Minimize } L_{\rho_k}(x, \lambda^k, \mu^k, u^k, v^k) \text{ s.t. } H_i(x)G_i(x) = 0, \forall i$$

and if some weak CQ (as Guignard's CQ) holds for $\mathcal{X} := \{x \in \mathbb{R}^n : H_i(x)G_i(x) = 0, \forall i\}$ at the solution (for instance, it will hold if $\{x \in \mathbb{R}^n : H(x) = 0\}$ and $\{x \in \mathbb{R}^n : G(x) = 0\}$ are transverse manifolds, [53]). Since we cannot expect to solve the subproblems exactly, several approaches can be taken for fulfill (56), (57) and (58). For instance, we can solve the sub-problems by using a Newton's method for the nonlinear system : $\mathcal{F}(x, \theta) := \nabla_x L_{\rho_k}(x, \lambda^k, \mu^k, u^k, v^k) + \sum_{i=1}^m \theta_i \mathcal{H}_i(x) \nabla_x \mathcal{H}_i(x) = 0$ and $\mathcal{H}_i(x) := H_i(x)G_i(x) = 0$, $\forall i$. Another approach, it is to solve the sub-problems using a quadratic penalty method for the constraints $H_i^2(x)G_i^2(x) = 0$, $i = 1, \dots, m$, in such a way that the sequence $\{\max\{\rho_k H_i(x^k)G_i(x^k) : \forall i\}\}$ or the sequence $\{\max\{\rho_k H_i^2(x^k)G_i^2(x^k) : \forall i\}\}$ is bounded.

Now, we will investigate the limit point of the sequence of iterates generated by Algorithm 5.1.

Theorem 5.1. *Let $\{x^k\}$ be a sequence of iterative generated by the algorithm. Then*

1. *If a limit point of $\{x^k\}$ is feasible such point conforms the MPEC-AKKT condition.*
2. *If, in a feasible limit point of $\{x^k\}$, MPEC-CCP holds. Then, such point is M-stationary.*
3. *If MPEC-MFCQ (or MPEC-NNAMCQ) holds in a feasible limit point of $\{x^k\}$. Then, such limit point is S-stationary.*

Proof. We assume without loss of generality that x^k converge to x^*. Since $\varepsilon_k \to 0$, using (58), we get that $H_i(x^*)G_i(x^*) = 0$, for all $i = 1, \dots, m$. Now, let us analyse the behaviour of the sequence $\{x^k\}$. From (56) and the definition of the multipliers (60), we obtain that

$$\|\nabla f(x^k) + \sum_{i=1}^p \lambda_i^{k+1} \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^{k+1} \nabla g_j(x^k) - \sum_{j=1}^m \hat{u}_j^{k+1} \nabla H_j(x^k) - \sum_{j=1}^m \hat{v}_j^{k+1} \nabla G_j(x^k)\| \le \varepsilon_k,$$

where

$$\hat{u}_j^{k+1} := u_j^{k+1} - \eta_j^k G_j^2(x^k) H_j(x^k) \quad \text{and} \quad \hat{v}_j^{k+1} := v_j^{k+1} - \eta_j^k G_j(x^k) H_j^2(x^k).$$

Now, from (57) and the continuity of the constraints, we have that $|\hat{u}_j^{k+1} - u_j^{k+1}| \to 0$ and $|\hat{v}_j^{k+1} - v_j^{k+1}| \to 0$. Then, using the continuity of the gradients,

$$\|\nabla f(x^k) + \sum_{i=1}^p \lambda_i^{k+1} \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^{k+1} \nabla g_j(x^k) - \sum_{j=1}^m u_j^{k+1} \nabla H_j(x^k) - \sum_{j=1}^m v_j^{k+1} \nabla G_j(x^k)\| \to 0.$$

Now, let us see assume that x^* is feasible. Set $z^k := F(x^*)$, for $k \in \mathbb{N}$. Let us see that, for k large enough $\mu_j^{k+1} = 0$, if $g_j(x^*) < 0$, $u_j^{k+1} = 0$ if $H_j(x^*) > 0$ and $u_j^{k+1} = 0$, if $G_j(x^*) > 0$. Indeed, take j such that $g_j(x^*) < 0$. If ρ_k is unbounded, $-\rho_k^{-1} \bar{\mu}_j^k \to 0$. Then, for k large enough the inequality: $g_j(x^k) < 2^{-1} g_j(x^*) < -\rho_k^{-1} \bar{\mu}_j^k$ holds and hence $\mu_j^{k+1} := \max\{0, \bar{\mu}_j^k + \rho_k g_j(x^k)\} = 0$. If ρ_k is bounded. Then, by Step 4, we see that $(V_g^k)_j = \max\{g_j(x^k), -\rho_k^{-1} \bar{\mu}_j^k\} \to 0$ and hence $-\rho_k^{-1} \bar{\mu}_j^k \to 0$. Then, for k sufficiently large, $g_j(x^k) < 2^{-1} g_j(x^*) < -\rho_k^{-1} \bar{\mu}_j^k$ which implies $\mu_j^{k+1} = 0$. Similarly, we get $u_j^{k+1} = 0$, if $H_j(x^*) > 0$ and $u_j^{k+1} = 0$, if $G_j(x^*) > 0$.

Thus, $u_j^{k+1} = 0$ for $j \in \mathcal{K}(x^*)$, $v_j^{k+1} = 0$ for $j \in \mathcal{I}(x^*)$ and $\mu_j^{k+1} = 0$ for $j \notin A(x^*)$. Furthermore, using the Step 3, we have that $u_j^{k+1} \ge 0$ and $v_j^{k+1} \ge 0$ for all $j = 1, \dots, m$. Then, x^k is an ε_k approximate S-stationary point and by Remark 2, we conclude that x^* is a MPEC-AKKT point.

If, we assume that MPEC-CCP holds at the limit point x^*, by Theorem 3.5, we get that x^* is a M-stationary.

Now, assume that MPEC-MFCQ holds at x^*. By MPEC-MFCQ, the sequence $(\lambda^{k+1}, \mu^{k+1}, u^{k+1}, v^{k+1})$ must be bounded. Taking an adequate subsequence we can assume that $(\lambda^{k+1}, \mu^{k+1}, u^{k+1}, v^{k+1})$ converges to (λ, μ, u, v). Since $u^{k+1} \in \mathbb{R}_+^m$, $v^{k+1} \in \mathbb{R}_+^m$, $u_j^{k+1} = 0$ for $j \in \mathcal{K}(x^*)$, $v_j^{k+1} = 0$ for $j \in \mathcal{I}(x^*)$ and $\mu_j^{k+1} = 0$ for $j \notin A(x^*)$. We have that $(u, v) \in \mathbb{R}_+^m \times \mathbb{R}_+^m$, $v^{k+1} \in \mathbb{R}_+^m$, $u_j = 0$ for $j \in \mathcal{K}(x^*)$, $v_j = 0$ for $j \in \mathcal{I}(x^*)$ and $\mu_j = 0$ for $j \notin A(x^*)$, i.e., x^* is S-stationary. Similar result holds if we assume that MPEC-NNAMCQ holds. □

A desired property for every method used for solving non-linear mathematical programs is the feasibility of the limit points, since this may be impossible in general, a result above the infeasibility of iterates is important. See [55, 23, 14]. The next theorem says that the limit points of the iterates given by Algorithm 5.1 are stationary for some infeasibility measure. For the classical Augmented Lagrangian methods see [12, Theorem 6.3]

Theorem 5.2. *Let $\{x^k\}$ be a sequence of iterative generated by the algorithm. Then, every limit point of $\{x^k\}$ is a KKT point of the optimization problem*

$$\begin{array}{ll} \textit{Minimize} & \|h(x)\|^2 + \|(g(x))_+\|^2 + \|(H(x))_-\|^2 + \|(G(x))_-\|^2 \\ \textit{subject to} & H_i(x)G_i(x) = 0, \ \textit{for } i = 1, \dots, m. \end{array} \tag{61}$$

Proof. We can assume that x^k converge to x^*. Since $\varepsilon_k \to 0$, by (58) we get that $H_i(x^*)G_i(x^*) = 0$, for all $i = 1, \dots, m$. If x^* is feasible for the NLP problem (1), the KKT conditions holds for (61). Now, assume that x^* is not feasible. Hence ρ_k must be unbounded. Following the proof of the Theorem 5.1, we have that

$$\delta_k := \nabla f(x^k) + \sum_{i=1}^{p} \lambda_i^{k+1} \nabla h_i(x^k) + \sum_{j=1}^{p} \mu_j^{k+1} \nabla g_j(x^k) - \sum_{j=1}^{m} u_j^{k+1} \nabla H_j(x^k) - \sum_{j=1}^{m} v_j^{k+1} \nabla G_j(x^k) \to 0. \tag{62}$$

Following [12, Theorem 6.3], we obtain that $|u^{k+1}/\rho_k - (-H(x^k))_+| \to 0$, $|v^{k+1}/\rho_k - (-G(x^k))_+| \to 0$ and $|\mu^{k+1}/\rho_k - (g(x^k))_+| \to 0$. Substituting into (62), using $\delta_k/\rho_k \to 0$, the continuity of the constraints and $x^k \to x^*$ we get

$$\sum_{i=1}^{p} h_i(x^*) \nabla h_i(x^*) + \sum_{j=1}^{p} (g_j(x^*))_+ \nabla g_j(x^*) - \sum_{j=1}^{m} (-H_j(x^*))_+ \nabla H_j(x^*) - \sum_{j=1}^{m} (-G_j(x^*))_+ \nabla G_j(x^*) = 0.$$

In other words, the limit point x^* is a KKT point with no null Lagrange multipliers corresponding to the constraints $H_i(x)G_i(x) = 0$, for all $i = 1, \dots, m$. □

As we just see this method have strong convergence properties: it converges to M-stationary points under weak assumptions and to S-stationary under MPEC-MFCQ. The practical implementation, efficiency, robustness and numerical tests of this method is out of the scope of this article and it will subject of future research.

5.2 The complementary-penalty method of Leyffer et al.

The complementary-penalty method consist in solving

$$\begin{array}{ll} \text{minimize} & f(x) + \sum_{i=1}^{m} \pi_i G_i(x) H_i(x) \\ \text{subject to} & g(x) \leq 0, \quad h(x) = 0 \\ & 0 \leq H(x), \quad 0 \leq G(x) \end{array} \tag{63}$$

where $\pi = (\pi_1, \dots, \pi_m) \in \mathbb{R}^m_+$ is a penalty parameter. This kind of regularization technique is based on the penalization approach. The idea is remove the complementary constraints by adding them into the objective function through a penalty function. In [38], the authors used an interior-point method for solving (63). That algorithm is called Algorithm I. Given an error tolerance $\varepsilon_k > 0$ and a barrier parameter $\zeta_k \in \mathbb{R}_+$, Algorithm I tries to find a penalty parameter π^k and vectors $(x^k, s^k, \lambda_E^k, \lambda_I^k) \in \mathbb{R}^n \times \mathbb{R}^p_+ \times \mathbb{R}^q \times \mathbb{R}^p_+$ satisfying $G_i(x^k) > 0$, $H_i(x^k) > 0$, $\forall i$, $s^k > 0$, $\lambda_I^k > 0$ such that

$$\|\nabla f(x^k) + \sum_{j=1}^{p} \lambda_{Ij}^k \nabla g_j(x^k) + \sum_{i=1}^{q} \lambda_{Ei}^k \nabla h_i(x^k) - \sum_{i=1}^{m} u_i^k \nabla H_i(x^k) - \sum_{i=1}^{m} v_i^k \nabla G_i(x^k)\| \leq \varepsilon_k, \tag{64}$$

$$\|h(x^k)\| \leq \varepsilon_k, \quad \|g(x^k) + s^k\| \leq \varepsilon_k, \quad \|\zeta_k - \lambda_{Ii}^k s_i^k\| \leq \varepsilon_k \tag{65}$$

and

$$\|\min\{H_i(x^k), G_i(x^k)\}\| \leq \varepsilon_k, \quad \forall i. \tag{66}$$

where $u_i^k := \zeta_k/H_i(x^k) - \pi_i^k G_i(x^k)$, $v_i^k := \zeta_k/G_i(x^k) - \pi_i^k H_i(x^k)$, $\forall i = 1, \dots, m$. The iterate (x^k, s^k) (if exits) is called of a ε_k-stationary point with approximate multipliers $(\lambda_E^k, \lambda_I^k)$ and parameters $\{\pi^k, \zeta_k\}$.

The Theorem 3.4 of [38] guarantees convergence to C-stationarity point if MPEC-LICQ holds at the limit point and to S-stationarity point x^* if additionally we require that $(\pi_i^k G_i(x^k), \pi_i^k H_i(x^k)) \to (0, 0)$, $\forall i \in \mathcal{J}(x^*)$. Here, we show convergence to M-stationarity point using MPEC-CCP (strictly weaker than MPEC-LICQ) under some additional assumption.

Theorem 5.3. *Suppose that Algorithm I generates an infinite sequence of ε_k-stationary point (x^k, s^k) with approximate multipliers (λ^k, μ^k) and parameters $\{\pi^k, \zeta_k\}$ with $\{\varepsilon_k\}$ and $\{\zeta_k\}$ converging to zero. If x^* is a limit point of x^k, then x^* is a feasible point. If, in addition, $\varepsilon_k = o(\max\{\pi_i^k\}^{-1})$ with MPEC-CCP holding at x^*. Then, x^* is M-stationary.*

Proof. The feasibility of the limit points follows from [38, Theorem 3.4]. Now, we will show that x^* is a MPEC-AKKT point. Assume that $\{x^k\}$ converges to x^*. To simplify the notation, we only focus on the complementarity part. Consider

$$u_i^k := \frac{\zeta^k}{H_i(x^k)} - \pi_i^k G_i(x^k); \quad v_i^k := \frac{\zeta^k}{G_i(x^k)} - \pi_i^k H_i(x^k). \tag{67}$$

Set $z^k := (g(x^*), h(x^*), -(H_1(x^*), G_1(x^*)), \dots, -(H_m(x^*), G_m(x^*)))$, $k \in \mathbb{N}$. Clearly, $\mathcal{I}(x^*) = \mathcal{I}(z^k)$, $\mathcal{J}(x^*) = \mathcal{J}(z^k)$ and $\mathcal{K}(x^*) = \mathcal{K}(z^k)$. Our aim is to find a subsequence $\mathcal{N} \subset \mathbb{N}$ and vectors $\hat{u}^k$ and $\hat{v}^k$ such that for every $k \in \mathcal{N}$, $\text{supp}(\hat{u}^k) \subset \mathcal{I}(z^k) \cup \mathcal{J}(z^k)$, $\text{supp}(\hat{v}^k) \subset \mathcal{K}(z^k) \cup \mathcal{J}(z^k)$, either $\hat{u}_\ell^k \hat{v}_\ell^k = 0$ or $\hat{u}_\ell^k > 0, \hat{v}_\ell^k > 0$ for $\ell \in \mathcal{J}(z^k)$ and

$$\| \sum_{i=1}^m u_i^k \nabla H_i(x^k) + \sum_{i=1}^m v_i^k \nabla G_i(x^k) - \sum_{i=1}^m \hat{u}_i^k \nabla H_i(x^k) - \sum_{i=1}^m \hat{v}_i^k \nabla G_i(x^k) \| \to_{\mathcal{N}} 0. \tag{68}$$

First, note that $u_i^k \to 0$, for all $i \in \mathcal{K}(z^k) = \mathcal{K}(x^*)$. In fact, for $i \in \mathcal{K}(z^k)$, we see that $H_i(x^*) > 0$. Thus, $G_i(x^k) \to 0$, $\min\{H_i(x^k), G_i(x^k)\} = G_i(x^k)$ and $\pi_i^k G_i(x^k) \to 0$ where in the last line we use (66). Then, from (67), $u_i^k \to 0$. Thus, we can define $\hat{u}_i^k := 0$, $\forall i \in \mathcal{I}(x^*)$. Clearly, $\|\hat{u}_i^k - u_i^k\| \to 0$, $\forall i \in \mathcal{K}(x^*)$.

Similarly, we have that $v_i^k \to 0$, for $i \in \mathcal{I}(z^k) = \mathcal{I}(x^*)$. Thus, define $\hat{v}_i^k := 0$, $i \in \mathcal{I}(x^*)$. Obviously, $|\hat{v}_i^k - v_i^k| \to 0$, $\forall i \in \mathcal{I}(x^*)$.

Now take $i \in \mathcal{J}(x^*)$. Since $\min\{G_i(x^k), H_i(x^k)\}$ can take only two alternatives, we have the following cases: (a) there is a subsequence $\mathcal{N} \subset \mathbb{N}$ such that $\min\{G_i(x^k), H_i(x^k)\} = H_i(x^k)$, $\forall k \in \mathcal{N}$. Thus, from $\varepsilon_k = o(\max\{\pi_i^k\}^{-1})$, $\pi_i^k H_i(x^k) \to 0$. Now, we may assume here that $\zeta^k/G_i(x^k) \to_{\mathcal{N}} \beta$ for some $\beta \in [0, \infty]$. If $\beta = 0$, from (67), $v_i^k \to 0$. Thus, put $\hat{v}_i^k = 0$, $\hat{u}_i^k = u_i^k$. Clearly, $\hat{v}_i^k \hat{u}_i^k = 0$ and $\|(\hat{v}_i^k, \hat{u}_i^k - (v_i^k, u_i^k)\| = |\hat{v}_i^k - v_i^k| \to_{\mathcal{N}} 0$. If $\beta \neq 0$. Then, there is $\delta > 0$ such that $v_i^k = \zeta^k/G_i(x^k) - \pi_i^k H_i(x^k) > \delta$, $k \in \mathcal{N}$ large enough. It implies that $u_i^k > 0$, $k \in \mathcal{N}$ large enough. Set $\hat{v}_i^k = v_i^k$, $\hat{u}_i^k = u_i^k$. Clearly, $\hat{v}_i^k > 0$, $\hat{u}_i^k > 0$ and $\|(\hat{v}_i^k, \hat{u}_i^k - (v_i^k, u_i^k)\| \to_{\mathcal{N}} 0$; (b) there is a subsequence $\mathcal{N} \subset \mathbb{N}$ such that $\min\{G_i(x^k), H_i(x^k)\} = G_i(x^k)$, $\forall k \in \mathcal{N}$. As in the previous case, we can find $\hat{v}_i^k$, $\hat{u}_i^k$ with $\hat{v}_i^k > 0$, $\hat{u}_i^k > 0$ or $\hat{v}_i^k \hat{u}_i^k = 0$ such that $\|(\hat{v}_i^k, \hat{u}_i^k - (v_i^k, u_i^k)\| \to_{\mathcal{N}} 0$.

Summing up, we always can find $\mathcal{N} \subset \mathbb{N}$ and vectors $\hat{u}^k$ and $\hat{v}^k$ such that for every $k \in \mathcal{N}$, $\text{supp}(\hat{u}^k) \subset \mathcal{I}(z^k) \cup \mathcal{J}(z^k)$, $\text{supp}(\hat{v}^k) \subset \mathcal{K}(z^k) \cup \mathcal{J}(z^k)$, either $\hat{u}_\ell^k \hat{v}_\ell^k = 0$ or $\hat{u}_\ell^k > 0, \hat{v}_\ell^k > 0$ for $\ell \in \mathcal{J}(z^k)$ and (68) holds. Thus, $\{x^k\}_{k \in \mathbb{N}}$ is a MPEC-AKKT sequence converging to x^*. Since, MPEC-CCP holds, x^* must be a M-stationary point. □

The example below shows that we cannot ensure M-stationarity of the limit point x^* without the assumption $\varepsilon_k = o(\max\{\pi_i^k\}^{-1})$ used in Theorem 5.3, even if stronger MPEC-CQs are used.

Example 5.1. In $\mathbb{R}^2$, consider the function $f(x_1, x_2) = -x_1 - x_2$, the complementary constraints $H_1(x_1, x_2) := x_1$, $G_1(x_1, x_2) := x_2$ and the point $x^* = (0, 0)$. Take $x_1^k := 1/k$, $x_2^k := 1/k$, $\pi^k := k$ and $\varepsilon_k := 1/k$. Clearly, $\pi^k \varepsilon_k = 1$, $x^k = (x_1^k, x_2^k)$ goes to x^* and it is easy to verify that x^k is an ε_k-stationary point generated by Algorithm I. Furthermore, MPEC-LICQ (and hence MPEC-CCP) holds at x^* but x^* is a C-stationary point which is not a M-stationary.

In [38], the authors also considered an interior-relaxation method. We refer for it as the Algorithm II: Given ζ^k and θ^k, find variables $(x^k, s^k, s_c^k, \lambda_E^k, \lambda_I^k, \xi^k) \in \mathbb{R}^n \times \mathbb{R}_+^p \times \mathbb{R}_+^m \times \mathbb{R}^q \times \mathbb{R}_+^p \times \mathbb{R}_+^m$ with $G_i(x^k) > 0$, $H_i(x^k) > 0$,

$\forall i,\ s^k > 0,\ \lambda_I^k > 0,\ s_c^k > 0,\ \xi^k > 0$ such that

$$\|\nabla f(x^k) + \sum_{j=1}^{p} \lambda_{Ij}^k \nabla g_j(x^k) + \sum_{i=1}^{q} \lambda_{Ei}^k \nabla h_i(x^k) - \sum_{i=1}^{m} u_i^k \nabla H_i(x^k) - \sum_{i=1}^{m} v_i^k \nabla G_i(x^k)\| \le \varepsilon_k, \tag{69}$$

where $u_i^k := \zeta_k / H_i(x^k) - \pi_i^k G_i(x^k)$, $v_i^k := \zeta_k / G_i(x^k) - \pi_i^k H_i(x^k)$, $\forall i = 1, \ldots, m$.

$$\max\{\|h(x^k)\|, \|g(x^k) + s^k\|\} \quad \le \varepsilon_k, \tag{70}$$

$$\max\{\|\zeta_k - \lambda_{Ii}^k s_i^k\|, \|\zeta_k - \xi_i^k s_{ci}^k\|, \|\theta_k - G_i(x^k) H_i(x^k) - s_{ci}^k\|\} \quad \le \varepsilon_k\ \forall i \tag{71}$$

Similarly to Theorem 5.3 we have the next result.

Theorem 5.4. *Assume that Algorithm II generates an infinite sequence of ε_k-stationary point (x^k, s^k, s_c^k) with approximate multipliers $(\lambda^k, \mu^k, \xi^k)$ and parameters $\{\zeta_k, \theta_k\}$ satisfying conditions (seila), with $\{\zeta_k\}$, $\{\varepsilon_k\}$, $\{\theta_k\}$ converging to zero. If x^* is a limit point of x^k, then x^* is a feasible point. If, in addition, $\varepsilon_k = o(\max\{\xi_i^k\}^{-1})$, $\theta_k = o(\max\{\xi_i^k\}^{-1})$ with MPEC-CCP holding at x^*. Then, x^* is M-stationary.*

Proof. Following the proof of [38, Theorem 3.6], we have that $\{x^k, s^k\}$ is an $\widehat{\varepsilon}^k$-stationary point with approximate multipliers (λ^k, μ^k) and parameters $\{\zeta_k, \pi^k := \xi^k\}$ where $\widehat{\varepsilon}_k := \max\{\varepsilon_k, (\varepsilon_k + \theta_k)^{1/2}\}$, satisfying SEILA. The conclusions came from a direct application of Theorem 5.3, since $\varepsilon_k = o(\max\{\xi_i^k\}^{-1})$, $\theta_k = o(\max\{\xi_i^k\}^{-1})$ implies $\widehat{\varepsilon}_k = o(\max\{\pi_i^k\}^{-1})$. □

5.3 L-relaxation of Kanzow and Schwartz

The relaxation scheme established by Kanzow and Schwartz in [36] is:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \le 0, \quad h(x) = 0, \quad 0 \le H(x), \quad 0 \le G(x), \\ & \Phi_i^{KS}(x;t) \le 0,\ \forall i \in \{1, \ldots, m\} \end{array} \tag{72}$$

where $\Phi_i^{KS}(x;t)$ is defined as

$$\begin{cases} (H_i(x) - t)(G_i(x) - t) & \text{if } H_i(x) + G_i(x) \ge 2t, \\ -\frac{1}{2}((H_i(x) - t)^2 + (G_i(x) - t)^2 & \text{if } H_i(x) + G_i(x) < 2t. \end{cases} \tag{73}$$

For a given $t > 0$, we use $NLP^{KS}(t)$ to denote the NLP (72). The feasible set of $NLP^{KS}(t)$ is given by the figure 2. By direct calculations, we get that $\nabla \Phi_i^{KS}(x;t)$ is equal to

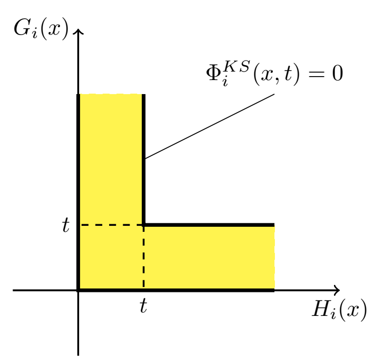

Figure 2: Feasible set of the relaxation of Kanzow and Schwartz.

$$\begin{cases} (H_i(x) - t)\nabla G_i(x) + (G_i(x) - t)\nabla H_i(x) & \text{if } H_i(x) + G_i(x) \ge 2t, \\ (t - G_i(x))\nabla G_i(x) + (t - H_i(x))\nabla H_i(x) & \text{if } H_i(x) + G_i(x) < 2t. \end{cases} \tag{74}$$

We improve the main convergence result for this relaxation [30] by using the weaker MPEC-CCP instead of MPEC-CPLD.

Theorem 5.5. *Let $\{t_k\} \downarrow 0$ and x^k be a KKT point of $NLP^{KS}(t_k)$. If $x^k \to x^*$ and MPEC-CCP holds at x^*. Then, x^* is an M-stationary point for* (1).

Proof We will show that x^* is a MPEC-AKKT point. For this, it will be sufficient to show that there is subsequence of $\{x^k\}$ which is a MPEC-AKKT sequence. Since x^k is a stationary point for $NLP^{KS}(t_k)$, we see that using the gradient of $\nabla\Phi^{KS}(x^k, t_k)$, (74), the KKT conditions are equivalent to

$$\nabla f(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^{q} \lambda_i^k \nabla h_i(x^k) - \sum_{i=1}^{m} \bar{u}_i^k \nabla H_i(x^k) - \sum_{i=1}^{m} \bar{v}_i^k \nabla G_i(x^k) = 0, \tag{75}$$

where $\bar{u}^k$ and $\bar{v}^k$ are defined as follows

$$\bar{u}_i^k := \begin{cases} u_i^k - \rho_i^k(G_i(x^k) - t_k) & \text{if } H_i(x^k) + G_i(x^k) \geq 2t_k, \\ u_i^k + \rho_i^k(H_i(x^k) - t_k) & \text{if } H_i(x^k) + G_i(x^k) < 2t_k. \end{cases} \tag{76}$$

$$\bar{v}_i^k := \begin{cases} v_i^k - \rho_i^k(H_i(x^k) - t_k) & \text{if } H_i(x^k) + G_i(x^k) \geq 2t_k, \\ v_i^k + \rho_i^k(G_i(x^k) - t_k) & \text{if } H_i(x^k) + G_i(x^k) < 2t_k. \end{cases} \tag{77}$$

with the constraints

$$\begin{array}{rll} g_j(x^k) \leq 0 & \mu_j^k \geq 0 & \mu_j^k g_j(x^k) = 0 \;\; \forall j \in \{1, \dots, p\}, \\ H_i(x^k) \geq 0 & u_i^k \geq 0 & u_i^k H_i(x^k) = 0 \;\; \forall i \in \{1, \dots, m\}, \\ G_i(x^k) \geq 0 & v_i^k \geq 0 & v_i^k G_i(x^k) = 0 \;\; \forall i \in \{1, \dots, m\}, \\ (H_i(x^k) - t_k)(G_i(x^k) - t_k) \leq 0 & \rho_i^k \geq 0 & \rho_i^k (H_i(x^k) - t_k)(G_i(x^k) - t_k) = 0 \;\; \forall i. \end{array} \tag{78}$$

Set $z^k := (g(x^*), h(x^*), -(H_1(x^*), G_1(x^*)), \dots, -(H_m(x^*), G_m(x^*)))$, $k \in \mathbb{N}$. Clearly, $\mathcal{I}(x^*) = \mathcal{I}(z^k)$, $\mathcal{J}(x^*) = \mathcal{J}(z^k)$ and $\mathcal{K}(x^*) = \mathcal{K}(z^k)$. Let us show that for k large enough, $\bar{u}^k$ and $\bar{v}^k$ conform the definition of MPEC-AKKT, i.e., $\bar{u}_i^k = 0$, $i \in \mathcal{K}(z^k)$; $\bar{v}_i^k = 0$, $i \in \mathcal{I}(z^k)$ and either $\bar{u}_i^k \bar{v}_i^k = 0$ or $\bar{u}_i^k > 0$, $\bar{v}_i^k > 0$, $i \in \mathcal{J}(z^k)$.

First, we will show that $\bar{v}_i^k = 0$, $i \in \mathcal{I}(z^k) = \mathcal{I}(x^*)$ for k large enough. For this purpose, we decompose the index set $\mathcal{I}(x^*)$ into a partition of four subsets, namely: $\mathcal{I}_1(x^*, k) := \{i \in \mathcal{I}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) \geq t_k\}$, $\mathcal{I}_2(x^*.k) := \{i \in \mathcal{I}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) < t_k\}$, $\mathcal{I}_3(x^*, k) := \{i \in \mathcal{I}(x^*) : G_i(x^k) < t_k, H_i(x^k) \geq t_k\}$, $\mathcal{I}_4(x^*, k) := \{i \in \mathcal{I}(x^*) : G_i(x^k) < t_k, H_i(x^k) < t_k\}$. There is no loss of generality, possibly after taking an adequate subsequence, if we assume that each element of the partition is independent of k. Thus, we have the next partition

$$\begin{array}{l} \mathcal{I}_1(x^*) := \{i \in \mathcal{I}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) \geq t_k\}, \\ \mathcal{I}_2(x^*) := \{i \in \mathcal{I}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) < t_k\}, \\ \mathcal{I}_3(x^*) := \{i \in \mathcal{I}(x^*) : G_i(x^k) < t_k, H_i(x^k) \geq t_k\}, \\ \mathcal{I}_4(x^*) := \{i \in \mathcal{I}(x^*) : G_i(x^k) < t_k, H_i(x^k) < t_k\}. \end{array} \tag{79}$$

Now, we proceed showing that $\bar{v}_i^k = 0$, $i \in \mathcal{I}(z^k) = \mathcal{I}(x^*)$. Since $G_i(x^*) > 0$, $i \in \mathcal{I}(x^*)$, the sets $\mathcal{I}_3(x^*)$ and $\mathcal{I}_4(x^*)$ must be empty sets for k large enough, otherwise taking limit in $G_i(x^k) < t_k$, we will get $G_i(x^*) = 0$, a contradiction. Now, take $i \in \mathcal{I}_1(x^*)$, then $G_i(x^k) + H_i(x^k) \geq 2t_k$ and from (77) we have $\bar{v}_i^k = v_i^k - \rho_i^k(H_i(x^k) - t_k)$. Since $G_i(x^*) > 0$, we have for k large enough, that, $G_i(x^k) > t_k > 0$. From the KKT conditions, $v_i^k = 0$ and $\rho_i^k(G_i(x^k) - t_k) = 0$. Thus, $\bar{v}_i^k = 0$, for every $i \in \mathcal{I}_1(x^*)$. Carry out the same analysis for $i \in \mathcal{I}_2(x^*)$, we get $\bar{v}_i^k = 0$. Therefore, we conclude that $\bar{v}_i^k = 0$ for every $i \in \mathcal{I}(z^k) = \mathcal{I}(x^*)$.

Now, following the same arguments, we get that $\bar{u}_i^k = 0$ for $i \in \mathcal{K}(x^*)$. We continue analyzing $\bar{u}_i^k$ and $\bar{v}_i^k$ for $i \in \mathcal{J}(x^*)$. Following the same arguments we decompose the index set $\mathcal{J}(x^*)$ into a partition of four subsets, namely: $\mathcal{J}_1(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) \geq t_k\}$, $\mathcal{J}_2(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) < t_k\}$, $\mathcal{J}_3(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) < t_k, H_i(x^k) \geq t_k\}$ and $\mathcal{J}_4(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) < t_k, H_i(x^k) < t_k\}$, which we assume independent of k, maybe after taking an adequate subsequence. For each part, we will obtain that either $\bar{u}_i^k \bar{v}_i^k = 0$ or $\bar{u}_i^k > 0$, $\bar{v}_i^k > 0$, $i \in \mathcal{J}(z^k) = \mathcal{J}(x^*)$, for k sufficiently large enough. We have the next cases:

- If $i \in \mathcal{J}_1(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) \geq t_k\}$. Here, $\bar{u}_i^k = u_i^k - \rho_i^k(G_i(x^k) - t_k)$ and $\bar{v}_i^k = v_i^k - \rho_i^k(H_i(x^k) - t_k)$. From the KKT conditions, since $t_k > 0$, we get that $u_i^k = 0$ and $v_i^k = 0$. Now, if $G_i(x^k) = t_k$ or $H_i(x^k) = t_k$, we have that $\bar{u}_i^k = 0$ or $\bar{v}_i^k = 0$ respectively. If $G_i(x^k) > t_k$ and $H_i(x^k) > t_k$, we get $\rho_i^k = 0$, $\bar{u}_i^k = 0$, $\bar{v}_i^k = 0$. In both cases, $\bar{v}_i^k \bar{u}_i^k = 0$.

- If $i \in \mathcal{J}_2(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) < t_k\}$. From the KKT conditions, we see that $v_i^k = 0$, $\rho_i^k(G_i(x^k) - t_k) = 0$. If $H_i(x^k) + G_i(x^k) \geq 2t_k$, we have that $\bar{v}^k = -\rho_i^k(H_i(x^k) - t_k) \geq 0$ and $\bar{u}_i^k = u_i^k \geq 0$. Now, if $H_i(x^k) + G_i(x^k) < 2t_k$, we get that $\bar{v}_i^k = v_i^k + \rho_i^k(G_i(x^k) - t_k) = 0$. Thus, for $i \in \mathcal{J}_2(x^*)$, we have that either $\bar{v}_i^k \bar{u}_i^k = 0$ or $\bar{v}_i^k > 0$ $\bar{u}_i^k > 0$.

- If $i \in \mathcal{J}_3(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) < t_k, H_i(x^k) \geq t_k\}$. By symmetry, we see that either $\bar{v}_i^k \bar{u}_i^k = 0$ or $\bar{v}_i^k > 0$ $\bar{u}_i^k > 0$.

- If $i \in \mathcal{J}_4(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) < t_k, H_i(x^k) < t_k\}$. From the KKT conditions, we see that $\rho_i^k = 0$. Then, $\bar{u}_i^k = u_i^k \geq 0$ and $\bar{v}_i^k = v_i^k \geq 0$.

Thus, x^* is an MPEC-AKKT point. So, if MPEC-CCP holds at x^*, then x^* is M-stationary point. □

When we try to solve NLPs, we usually end up in an approximate KKT point, instead of a true KKT point. In fact, the stopping criteria for solving NLPs basically check whether an approximate KKT point has been found (in addition, maybe, to other criteria). So, we rarely end up in a KKT point. This fact has practical relevance. For example, the method of Kansow and Schwartz has stronger convergence properties: All limit points are M-stationary points, under weak CQ for M-stationary (as MPEC-CPLD, see [30, Theorem 3.3]), if each iterative is a KKT point. However, if we consider approximate KKT points instead of KKT points, we lost most of this advantage. Indeed, without additional assumptions, only convergence to weakly stationary points can be obtained. We can improve the result of [37, Theorem 13] under the MPEC-CCP assumption. The precise statement is the following.

Theorem 5.6. *Let $t_k \downarrow 0$, $\varepsilon_k = o(t_k)$, $\{x^k\}$ be a sequence of ε_k-stationary points of $NLP^{KS}(t_k)$ with approximate multipliers $(\mu^k, \lambda^k, u^k, v^k, \rho^k) \in \mathbb{R}^p_+ \times \mathbb{R}^q \times \mathbb{R}^m_+ \times \mathbb{R}^m_+ \times \mathbb{R}^m_+$, such that $\max\{|u_i^k H_i(x^k)|, |v_i^k G_i(x^k)|, |\rho_i^k \Phi_i^{KS}(x^k, t_k)| : i \in \{1, \ldots, m\}\} \leq \varepsilon_k$. Assume that $x^k \to x^*$ with MPEC-CCP holding in $x^* \in \Omega$. Suppose that there is a constant $c > 0$ such that, for all $i \in \mathcal{J}(x^*)$ and all k sufficiently large, the iterates $(G_i(x^k), H_i(x^k))$ satisfy*

$$(G_i(x^k), H_i(x^k)) \notin \begin{array}{l} [(t_k, (1+c)t_k) \times ((1-c)t_k, t_k)] \\ \cup[((1-c)t_k, t_k) \times (t_k, (1+c)t_k)] \\ \cup(t_k, (1+c)t_k)^2 \cup ((1-c)t_k, t_k)^2. \end{array} \tag{80}$$

Then, x^* is a M-stationary point.

Proof We will proof that under the expression (80), x^* is a MPEC-AKKT point. For this purpose, it will be sufficient to show that there is a subindex $\mathcal{N} \subset \mathbb{N}$ such that $\{x^k\}_{k \in \mathcal{N}}$ is a MPEC-AKKT sequence. Since x^k is an ε_k- stationary point for $NLP^{KS}(t_k)$, we have, following [37], that for k large enough

$$\|\nabla f(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^{q} \lambda_i^k \nabla h_i(x^k) - \sum_{i=1}^{m} \bar{u}_i^k \nabla H_i(x^k) - \sum_{i=1}^{m} \bar{v}_i^k \nabla G_i(x^k)\| \leq \varepsilon_k \tag{81}$$

with $\text{supp}(\mu^k) \subset A(x^*)$, where $\bar{u}^k$ and $\bar{v}^k$ are defined as follows

$$\bar{u}_i^k := \begin{cases} u_i^k - \rho_i^k(G_i(x^k) - t_k) & \text{if } H_i(x^k) + G_i(x^k) \geq 2t_k, \\ u_i^k + \rho_i^k(H_i(x^k) - t_k) & \text{otherwise} \end{cases} \tag{82}$$

$$\bar{v}_i^k := \begin{cases} v_i^k - \rho_i^k(H_i(x^k) - t_k) & \text{if } H_i(x^k) + G_i(x^k) \geq 2t_k, \\ v_i^k + \rho_i^k(G_i(x^k) - t_k) & \text{otherwise} \end{cases} \tag{83}$$

and

$$\begin{array}{rlll} g_j(x^k) \leq \varepsilon_k, & \mu_j^k \geq 0 & |\mu_j^k g_j(x^k)| \leq \varepsilon_k & \forall j \in \{1, \ldots, p\}, \\ H_i(x^k) \geq -\varepsilon_k, & u_i^k \geq 0 & |u_i^k H_i(x^k)| \leq \varepsilon_k & \forall i \in \{1, \ldots, m\}, \\ G_i(x^k) \geq -\varepsilon_k, & v_i^k \geq 0 & |v_i^k G_i(x^k)| \leq \varepsilon_k & \forall i \in \{1, \ldots, m\}, \\ \Phi_i^{KS}(x^k, t_k) \leq \varepsilon_k, & \rho_i^k \geq 0 & |\rho_i^k \Phi_i^{KS}(x^k, t_k)| \leq \varepsilon_k & \forall i \in \{1, \ldots, m\}, \end{array} \tag{84}$$

Put $z^k := (g(x^*), h(x^*), -(H_1(x^*), G_1(x^*)), \ldots, -(H_m(x^*), G_m(x^*)))$, $k \in \mathbb{N}$. Clearly, $\mathcal{I}(z^k) = \mathcal{I}(x^*)$, $\mathcal{K}(z^k) = \mathcal{K}(x^*)$ and $\mathcal{J}(z^k) = \mathcal{J}(x^*)$. Our aim is to find a subsequence $\mathcal{N} \subset \mathbb{N}$ and vectors $\hat{u}^k$ and $\hat{v}^k$ such that for every

$k \in \mathcal{N}$, $\text{supp}(\hat{u}^k) \subset \mathcal{I}(z^k) \cup \mathcal{J}(z^k)$, $\text{supp}(\hat{v}^k) \subset \mathcal{K}(z^k) \cup \mathcal{J}(z^k)$, either $\hat{u}^k_\ell \hat{v}^k_\ell = 0$ or $\hat{u}^k_\ell > 0, \hat{v}^k_\ell > 0$ for $\ell \in \mathcal{J}(z^k)$ and

$$\|\sum_{i=1}^m \bar{u}^k_i \nabla H_i(x^k) + \sum_{i=1}^m \bar{v}^k_i \nabla G_i(x^k) - \sum_{i=1}^m \hat{u}^k_i \nabla H_i(x^k) - \sum_{i=1}^m \hat{v}^k_i \nabla G_i(x^k)\| \to_{\mathcal{N}} 0. \tag{85}$$

Now, similarly as the proof of Theorem 5.5, we decompose $\mathcal{I}(x^*)$ into a four disjoint subsets (which we assume independent of k, after possibly taking an adequate subsequence), $\mathcal{I}_1(x^*) := \{i \in \mathcal{I}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) \geq t_k\}$, $\mathcal{I}_2(x^*) := \{i \in \mathcal{I}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) < t_k\}$, $\mathcal{I}_3(x^*) := \{i \in \mathcal{I}(x^*) : G_i(x^k) < t_k, H_i(x^k) \geq t_k\}$, $\mathcal{I}_4(x^*) := \{i \in \mathcal{I}(x^*) : G_i(x^k) < t_k, H_i(x^k) < t_k\}$.

We will show that there is a subsequence $\mathcal{N} \subset \mathbb{N}$, such that $\bar{v}^k_i \to_{\mathcal{N}} 0$, $\forall i \in \mathcal{I}(x^*) = \mathcal{I}(z^k)$. Thus, we can define $\hat{v}^k_i := 0$ for all $i \in \mathcal{I}(x^*)$, $k \in \mathcal{N}$. Note that, $\|\hat{v}^k_i - \bar{v}^k_i\| \to_{\mathcal{N}} 0$, $\forall i \in \mathcal{I}(x^*)$. Now, take $i \in \mathcal{I}_1(x^*)$. Thus, $G_i(x^k) + H_i(x^k) \geq 2t_k$. Due to the ε_k-stationarity of x^k, we have that $v^k_i G_i(x^k) \to 0$ and $\rho^k_i (G_i(x^k) - t_k)(H_i(x^k) - t_k) \to 0$. But, since $G_i(x^k) \to G_i(x^*) > 0$, we have that $v^k_i \to 0$, $\rho^k_i(H_i(x^k) - t_k) \to 0$ and as consequence $\bar{v}^k_i = v^k_i - \rho^k_i(H_i(x^k) - t_k) \to 0$. Now, if $i \in \mathcal{I}_2(x^*)$. We have two cases: if there is a subsequence $K_1 \subset \mathbb{N}$, such that $G_i(x^k) + H_i(x^k) \geq 2t_k$ for $k \in K_1$. Here, as we just have seen, $\bar{v}^k_i \to_{K_1} 0$. In the other case, if there is a subsequence $K_2 \subset \mathbb{N}$, such that $G_i(x^k) + H_i(x^k) < 2t_k$ for $k \in K_2$, we have $\bar{v}^k_i = v^k_i + \rho^k_i(G_i(x^k) - t_k)$. But, from the ε_k-stationarity of x^k, we get $t_k \leq G_i(x^k) < 2t_k - H_i(x^k) \leq 2t_k + \varepsilon_k$ for $k \in K_2$, which is a contradiction, since $G_i(x^k) \to G_i(x^*) > 0$. Since $G_i(x^*) > 0$, $i \in \mathcal{I}(x^*)$, the sets $\mathcal{I}_3(x^*)$ and $\mathcal{I}_4(x^*)$ must be empty sets for k large enough, otherwise taking limit in $G_i(x^k) < t_k$, we will get $G_i(x^*) = 0$, a contradiction. So, in every case, there is subsequence $\mathcal{N} \subset \mathbb{N}$ such that $\bar{v}^k_i \to_{\mathcal{N}} 0$, $\forall i \in \mathcal{I}(x^*)$.

Following the same above arguments for $\mathcal{K}(x^*) = \mathcal{K}(z^k)$, given any subsequence of $\{x^k\}$, we can find another subsequence $\mathcal{N} \subset \mathbb{N}$ such that $\bar{u}^k_i \to_{\mathcal{N}} 0$. Thus, we can define $\hat{u}^k_i := 0$ for all $i \in \mathcal{K}(z^k)$, $k \in \mathcal{N}$ with the property $\|\hat{u}^k_i - \bar{u}^k_i\| \to_{\mathcal{N}} 0$, $\forall i \in \mathcal{K}(z^k) = \mathcal{K}(x^*)$.

Now, we will focus on the index subset $\mathcal{J}(z^k)$. We will find a subsequence $\mathcal{N} \subset \mathbb{N}$ and scalars $\hat{u}^k_i$, $\hat{v}^k_i$ for $i \in \mathcal{J}(z^k)$ such that either $\hat{u}^k_i \hat{v}^k_i = 0$ or $\hat{u}^k_i > 0$ and $\hat{v}^k_i > 0$ with $\|(\hat{u}^k_i, \hat{v}^k_i) - (\bar{u}^k_i, \bar{v}^k_i)\| \to_{\mathcal{N}} 0$, $\forall i \in \mathcal{J}(z^k) = \mathcal{J}(x^*)$. For this purpose, decompose $\mathcal{J}(x^*)$ into a partition of four subsets, namely: $\mathcal{J}_1(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) \geq t_k\}$, $\mathcal{J}_2(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) < t_k\}$, $\mathcal{J}_3(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) < t_k, H_i(x^k) \geq t_k\}$ and $\mathcal{J}_4(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) < t_k, H_i(x^k) < t_k\}$. We have the next sub-cases.

- If $i \in \mathcal{J}_1(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) \geq t_k\}$. Certainly, $G_i(x^k) + H_i(x^k) \geq 2t_k$. In that case, $\bar{u}^k_i = u^k_i - \rho^k_i(H_i(x^k) - t_k)$ and $\bar{v}^k_i = v^k_i - \rho^k_i(G_i(x^k) - t_k)$. We will show that $\bar{u}^k_i \to 0$ or $\bar{v}^k_i \to 0$ for an adequate subsequence. First, since $H_i(x^k) \geq t_k$, $G_i(x^k) \geq t_k$, $|u^k_i H_i(x^k)| \leq o(t_k)$ and $|v^k_i G_i(x^k)| \leq o(t_k)$, we get that $u^k_i \to 0$ and $v^k_i \to 0$. Furthermore, from $H_i(x^k) \geq t_k$ and $G_i(x^k) \geq t_k$, there is a subsequence $K \subset \mathbb{N}$ such that $H_i(x^k)/t_k \to_K \alpha$ and $G_i(x^k)/t_k \to_K \beta$, for some scalars $\alpha, \beta \in [1, \infty]$. If $\alpha \neq 1$. Then, $|\rho^k_i(G_i(x^k) - t_k)(H_i(x^k)/t_k - 1)| \leq |\rho^k_i \Phi^{KS}_i(x^k, t_k)|/t_k \leq o(t_k)/t_k$, implies that $\rho^k_i(G_i(x^k) - t_k) \to_K 0$ and hence $\bar{u}^k_i = u^k_i - \rho^k_i(H_i(x^k) - t_k) \to_K 0$. Similarly, if $\beta \neq 1$, we conclude that $\bar{v}^k_i \to_K 0$. Now, only rest to analyze when $\alpha = \beta = 1$. For this case, we get $(1 + c)t_k > H_i(x^k) \geq t_k$ and $(1 + c)t_k > G_i(x^k) \geq t_k$ for $k \in K$ large enough. Now, if $H_i(x^k) > t_k$ for infinite k, (80) implies $G_i(x^k) = t_k$ and hence $\bar{u}^k_i = u^k_i \to 0$. If $G_i(x^k) > t_k$ for infinite k, from (80) we get $H_i(x^k) = t_k$ and hence $\bar{v}^k_i = v^k_i \to 0$. From, all the case, we conclude that there is subsequence $\mathcal{N} \subset \mathbb{N}$, such that $\bar{u}^k_i \to_{\mathcal{N}} 0$ or $\bar{v}^k_i \to_{\mathcal{N}} 0$. When $\bar{u}^k_i \to_{\mathcal{N}} 0$, we define $\hat{u}^k_i := 0$, $\hat{v}^k_i := \bar{v}^k_i$, $k \in \mathcal{N}$. Thus, $\|(\hat{u}^k_i, \hat{v}^k_i) - (\bar{u}^k_i, \bar{v}^k_i)\| = |\bar{u}^k_i| \to_{\mathcal{N}} 0$. In the case, $\bar{v}^k_i \to_{\mathcal{N}} 0$, define $\hat{v}^k_i := 0$, $\hat{u}^k_i := \bar{u}^k_i$, $k \in \mathcal{N}$. Clearly, $\|(\hat{u}^k_i, \hat{v}^k_i) - (\bar{u}^k_i, \bar{v}^k_i)\| = |\bar{v}^k_i| \to_{\mathcal{N}} 0$. In both cases, we always have $\hat{u}^k_i \bar{v}^k_i = 0$ for all $k \in \mathcal{N}$, $i \in \mathcal{J}_1(z^k)$.

- Take $i \in \mathcal{J}_2(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) \geq t_k, H_i(x^k) < t_k\}$. We will show that there is a subsequence $\mathcal{N} \subset \mathbb{N}$ and scalars $\{\hat{u}^k_i, \hat{v}^k_i\}$, such that $\|(\bar{u}^k_i, \bar{v}^k_i) - (\hat{u}^k_i, \hat{v}^k_i)\| \to_{\mathcal{N}} 0$ and either $\hat{u}^k_i \hat{v}^k_i = 0$ or $\hat{u}^k_i > 0$ and $\hat{v}^k_i > 0$. Now, since $G_i(x^k) \geq t_k$, we conclude $v^k_i \to 0$. Now, we have two alternatives, that $G_i(x^k) + H_i(x^k) \geq 2t_k$ holds for infinite many k or $G_i(x^k) + H_i(x^k) < 2t_k$ holds for infinite many k.

  - If $G_i(x^k) + H_i(x^k) \geq 2t_k$ holds for infinite many k. In this case, we get $\bar{v}^k_i = v^k_i - \rho^k_i(H_i(x^k) - t_k)$ and $\bar{u}^k_i = u^k_i - \rho^k_i(G_i(x^k) - t_k)$. From, $-\varepsilon_k \leq H_i(x^k) < t_k$ and $G_i(x^k) \geq t_k$, there is a subsequence $K \subset \mathbb{N}$ such that $H_i(x^k)/t_k \to_K \alpha$ and $G_i(x^k)/t_k \to_K \beta$, for some scalars $\alpha \in [0, 1]$, $\beta \in [1, \infty]$. If $\alpha \neq 1$. Then, the inequality $|\rho^k_i(G_i(x^k) - t_k)(1 - H_i(x^k)/t_k)| = |\rho^k_i(G_i(x^k) - t_k)(H_i(x^k) - t_k)|/2t_k \leq |\rho^k_i \Phi^{KS}_i(x^k, t_k)|/t_k$ implies that $\rho^k_i(G_i(x^k) - t_k) \to_K 0$. Now, we define $\hat{v}^k_i := \bar{v}^k_i$ and $\hat{u}^k_i := u^k_i$.

Clearly, $\hat{v}_i^k$, $\hat{u}_i^k \geq 0$ and $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| = |\rho_i^k(G_i(x^k) - t_k)| \to_K 0$. Now, if $\beta \neq 1$, we get $\rho_i^k(H_i(x^k) - t_k) \to 0$ and hence $\bar{v}_i^k \to_K 0$. So, define $\hat{v}_i^k := 0$ and $\hat{u}_i^k := \bar{u}_i^k$. Here, $\hat{v}_i^k \hat{u}_i^k = 0$ and $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| = |\bar{v}_i^k| \to_K 0$. Now, when $\alpha = \beta = 1$, we get $(1+c)t_k < H_i(x^k) < t_k$ and $(1+c)t_k > G_i(x^k) \geq t_k$ for $k \in K$ large enough. From (80), the only possibility is $G_i(x^k) = t_k$. Thus, $\bar{u}_i^k = u_i^k \geq 0$. Now, we define, $\hat{v}_i^k := \bar{v}_i^k$ and $\hat{u}_i^k := u_i^k$. Clearly, $\hat{v}_i^k, \hat{u}_i^k \geq 0$ and $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| = |\rho_i^k(G_i(x^k) - t_k)| = 0$.

  - If $G_i(x^k) + H_i(x^k) < 2t_k$ holds for infinite many k. In this case $\bar{v}_i^k \to 0$. From $-\varepsilon_k \leq H_i(x^k) < t_k$ and $G_i(x^k) \geq t_k$, there is a $K \subset \mathbb{N}$ such that $H_i(x^k)/t_k \to_K \alpha$ and $G_i(x^k)/t_k \to_K \beta$, for some scalars $\alpha \in [0,1]$, $\beta \in [1,\infty]$. If $\alpha \neq 1$, from $|\rho_i^k \Phi_i^{KS}(x^k, t_k)| \leq o(t_k)$, the expression $|\rho_i^k(G_i(x^k) - t_k)(H_i(x^k)/t_k - 1)| \leq \rho_i^k|(G_i(x^k) - t_k)^2 + (H_i(x^k) - t_k)^2|/2t_k = |\rho_i^k \Phi_i^{KS}(x^k, t_k)|/t_k$, implies that $\rho_i^k(G_i(x^k) - t_k) \to_K 0$ and hence $\bar{v}_i^k = v_i^k + \rho_i^k(G_i(x^k) - t_k) \to_K 0$. Now, we define $\hat{v}_i^k := 0$ and $\hat{u}_i^k := \bar{u}_i^k$. Clearly, $\hat{v}_i^k \hat{u}_i^k = 0$ and $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| = \bar{v}_i^k \to_K 0$. Now, if $\beta \neq 1$, we get $|\rho_i^k(G_i(x^k) - t_k)(G_i(x^k)/t_k - 1)| \leq \rho_i^k|(G_i(x^k) - t_k)^2 + (H_i(x^k) - t_k)^2|/t_k = 2|\rho_i^k \Phi_i^{KS}(x^k, t_k)|/t_k$ which implies $\rho_i^k(G_i(x^k) - t_k) \to 0$ and hence $\bar{v}_i^k \to_K 0$. So, define $\hat{v}_i^k := 0$ and $\hat{u}_i^k := \bar{u}_i^k$. Hence, $\hat{v}_i^k \hat{u}_i^k = 0$ and $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| = |\bar{v}_i^k| \to_K 0$. Now, we will analyze when $\alpha = \beta = 1$. In this case, $(1-c)t_k < H_i(x^k) < t_k$ and $(1+c)t_k > G_i(x^k) \geq t_k$ for $k \in K$ large enough. From (80), we get $G_i(x^k) = t_k$. Thus, $\bar{v}_i^k = v_i^k + \rho_k(G_i(x^k) - t_k) \to_K 0$. Here, define $\hat{v}_i^k := 0$ and $\hat{u}_i^k := \bar{u}_i^k$. Hence, $\hat{v}_i^k \hat{u}_i^k = 0$ and $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| = |v_i^k| \to_K 0$.

- For $i \in \mathcal{J}_3(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) < t_k, H_i(x^k) \geq t_k\}$. Similarly, as item anterior, we can find a subsequence $K \subset \mathbb{N}$ and points $\{\hat{u}_i^k, \hat{v}_i^k\}$, such that either $\hat{u}_i^k \hat{v}_i^k = 0$ or $\hat{u}_i^k > 0$ and $\hat{v}_i^k > 0$ with $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| \to 0$.

- For $i \in \mathcal{J}_4(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) < t_k, H_i(x^k) < t_k\}$. Clearly, $G_i(x^k) + H_i(x^k) < 2t_k$. In that case, $\bar{u}_i^k = u_i^k + \rho_i^k(H_i(x^k) - t_k)$ and $\bar{v}_i^k = v_i^k + \rho_i^k(G_i(x^k) - t_k)$. From $-\varepsilon_k \leq H_i(x^k) < t_k$ and $-\varepsilon_k \leq G_i(x^k) < t_k$, there is a subsequence $K \subset \mathbb{N}$ such that $H_i(x^k)/t_k \to_K \alpha$ and $G_i(x^k)/t_k \to_K \beta$, for some scalars $\alpha, \beta \in [0,1]$. If $\alpha \neq 1$, then $|\rho_i^k(G_i(x^k) - t_k)(H_i(x^k)/t_k - 1)| \leq \rho_i^k|(G_i(x^k) - t_k)^2 + (H_i(x^k) - t_k)^2|/2t_k = |\rho_i^k \Phi_i^{KS}(x^k, t_k)|/t_k \to 0$ implies that $\rho_i^k(G_i(x^k) - t_k) \to_K 0$. Furthermore, from $|\rho_i^k(H_i(x^k) - t_k)(H_i(x^k)/t_k - 1)| \leq \rho_i^k|(G_i(x^k) - t_k)^2 + (H_i(x^k) - t_k)^2|/t_k = 2|\rho_i^k \Phi_i^{KS}(x^k, t_k)|/t_k \to 0$, we get $\rho_i^k(H_i(x^k) - t_k) \to_K 0$. In this case, define $\hat{v}_i^k := v_i^k$ and $\hat{u}_i^k := u_i^k$. Clearly, $\hat{v}_i^k \geq 0$, $\hat{u}_i^k \geq 0$ and $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\|^2 = |\rho_i^k(G_i(x^k) - t_k)|^2 + |\rho_i^k(H_i(x^k) - t_k)|^2 \to_K 0$. By symmetry, we obtain the same result if $\beta \neq 1$. If $\alpha = \beta = 1$, we get $(1-c)t_k < H_i(x^k) < t_k$ and $(1-c)t_k < G_i(x^k) < t_k$ for $k \in K$ large enough, which is impossible by (80). Thus, we can find a subsequence $\mathcal{N} \subset \mathbb{N}$ and vector $\{\hat{u}_i^k, \hat{v}_i^k\}$, such that $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| \to_{\mathcal{K}} 0$ and either $\hat{u}_i^k \hat{v}_i^k = 0$ or $\hat{u}_i^k > 0$ and $\hat{v}_i^k > 0$ for $i \in \mathcal{J}_4(z^k)$.

Summarizing, we conclude, from all the cases, that there is a subsequence $\mathcal{N} \subset \mathbb{N}$ and points $\{\hat{u}_i^k, \hat{v}_i^k\}$, $k \in \mathcal{N}$, $i \in \{1, \ldots, m\}$, with $\hat{u}_i^k = 0$ for $i \in \mathcal{K}(z^k)$, $\hat{v}_i^k = 0$ for $i \in \mathcal{I}(z^k)$ and either $\hat{u}_i^k \hat{v}_i^k = 0$ or $\hat{u}_i^k > 0$ and $\hat{v}_i^k > 0$ for $i \in \mathcal{J}(z^k)$ such that for all $i \in \{1, \ldots, m\}$, $\|(\bar{u}_i^k, \bar{v}_i^k) - (\hat{u}_i^k, \hat{v}_i^k)\| \to_{\mathcal{N}} 0$ and $\|\sum_{i=1}^m \bar{u}_i^k \nabla H_i(x^k) + \sum_{i=1}^m \bar{v}_i^k \nabla G_i(x^k) - \sum_{i=1}^m \hat{u}_i^k \nabla H_i(x^k) - \sum_{i=1}^m \hat{v}_i^k \nabla G_i(x^k)\| \to_{\mathcal{N}} 0$. Thus, the subsequence $\{x^k : k \in \mathcal{N}\}$ with the approximate multipliers $(\mu^k, \lambda^k, \hat{u}^k, \hat{v}^k)$ is a MPEC-AKKT sequence with $x^k \to_{\mathcal{N}} x^*$. Since by hyphotesis, MPEC-CCP holds at x^*, the point x^* is a M-stationary point. □

Remark 7. MPEC-MFCQ implies that the sequence of approximate multipliers $\{(\mu^k, \lambda^k, u^k, v^k)\}$ is bounded. In fact, most of the CQs for M-stationary used in the convergence analysis of several MPECs algorithms, as MPEC-MFCQ, tries to bound or indirectly control the sequence of approximate multipliers. But, as we just see, we can obtain convergence to M-stationary points even if the sequence of approximate multipliers is unbounded. In the Theorem 5.5 or in the Theorem 5.6 (under (80)), we can guarantee convergence to M-stationary points with the less stringent MPEC-CCP without require boundedness of the multipliers.

5.4 The nonsmooth relaxation by Kadrani et al.

The relaxation scheme of Kadrani et al [32] is given by

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0, \quad h(x) = 0 \quad 0 \leq H_i(x) + t, \quad 0 \leq G_i(x) + t \\ & \Phi_i^{KDB}(x; t) \leq 0 \quad \forall i \in \{1, \ldots, m\} \end{array} \tag{86}$$

where $\Phi_i^{KDB}(x;t) := (H_i(x)-t)(G_i(x)-t)$. The NLP (86) is denoted by $NLP^{KDB}(t)$. The figure 3 shows the feasible set of $NLP^{KDB}(t)$ for a given $t > 0$. By straightforward calculations, we have

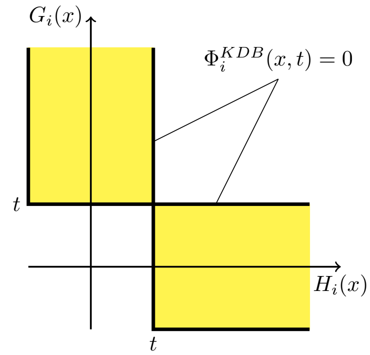

Figure 3: Feasible set of the relaxation of Kadrani et al.

$$\nabla\Phi_i^{KDB}(x;t) := (H_i(x)-t)\nabla G_i(x) + (G_i(x)-t)\nabla H_i(x) \quad \forall i \in \{1,\dots,m\}. \tag{87}$$

Using MPEC-CCP instead of MPEC-CPLD, we improve the result of Hoheisel et al [30]. The proof follows similar arguments as the theorem 5.5.

Theorem 5.7. *Let* $\{t_k\} \downarrow 0$ *and* x^k *be a KKT point of* $NLP^{KDB}(t_k)$. *If* $x^k \to x^*$ *and MPEC-CCP holds in* x^*. *Then,* x^* *is an M-stationary point.*

If we replace the sequence of KKT points by a sequence of ε_k-stationary points, following a similar line of arguments as the theorem 5.6, we have the next result.

Theorem 5.8. *Let* $t_k \downarrow 0$, $\varepsilon_k = o(t_k)$, x^k *be a sequence of* ε_k*-stationary points of* $NLP^{KDB}(t_k)$. *Assume that* $(\mu^k,\lambda^k,u^k,v^k,\rho^k) \in \mathbb{R}^p_+ \times \mathbb{R}^q \times \mathbb{R}^m_+ \times \mathbb{R}^m_+ \times \mathbb{R}^m_+$, *satisfy* $\max\{|u_i^k(H_i(x^k)+t_k)|, |v_i^k(G_i(x^k)+t_k)|, |\rho_i^k\Phi_i^{KDB}(x^k,t_k)| : i \in \{1,\dots,m\}\} \le \varepsilon_k$ *and* $x^k \to x^*$. *Suppose that* x^* *conforms MPEC-CCP and suppose further that there is a constant* $c > 0$ *such that, for all* $i \in \mathcal{J}(x^*)$ *and all* k *sufficiently large, the iterates* $(G_i(x^k), H_i(x^k))$ *satisfy*

$$\begin{aligned}(G_i(x^k),H_i(x^k)) \notin \quad & [(t_k,(1+c)t_k) \times ((1-c)t_k,t_k)] \\ & \cup[((1-c)t_k,t_k) \times (t_k,(1+c)t_k)] \cup (t_k,(1+c)t_k)^2.\end{aligned} \tag{88}$$

Then, x^* *is a M-stationary point.*

Proof We will show that, under the hypothesis (88), x^* is a MPEC-AKKT point. Since x^k is an ε_k- stationary point for $NLP^{KDB}(t_k)$, we have, using the gradient of $\nabla\Phi^{KDB}(x^k,t_k)$ (see (87)) that for k large enough

$$\|\nabla f(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^q \lambda_i^k \nabla h_i(x^k) - \sum_{i=1}^m \bar{u}_i^k \nabla H_i(x^k) - \sum_{i=1}^m \bar{v}_i^k \nabla G_i(x^k)\| \le \varepsilon_k, \tag{89}$$

where $\text{supp}(\mu^k) \subset A(x^*)$, $\bar{u}_i^k := u_i^k - \rho_i^k(G_i(x^k)-t_k)$, $\bar{v}_i^k := v_i^k - \rho_i^k(H_i(x^k)-t_k)$, $\forall i \in \{1,\dots,m\}$ and the approximate multipliers satisfy the relations

$$\begin{array}{rll} g_j(x^k) \le \varepsilon_k, & \mu_j^k \ge 0 & |\mu_j^k g_j(x^k)| \le \varepsilon_k \quad \forall j \in \{1,\dots,p\}, \\ H_i(x^k)+t_k \ge -\varepsilon_k, & u_i^k \ge 0 & |u_i^k(H_i(x^k)+t_k)| \le \varepsilon_k \quad \forall i \in \{1,\dots,m\}, \\ G_i(x^k)+t_k \ge -\varepsilon_k, & v_i^k \ge 0 & |v_i^k(G_i(x^k)+t_k)| \le \varepsilon_k \quad \forall i \in \{1,\dots,m\}, \\ \Phi_i^{KDB}(x^k,t_k) \le \varepsilon_k, & \rho_i^k \ge 0 & |\rho_i^k \Phi_i^{KDB}(x^k,t_k)| \le \varepsilon_k \quad \forall i \in \{1,\dots,m\}, \end{array} \tag{90}$$

Put $z^k := (g(x^*), h(x^*), -(H_1(x^*),G_1(x^*)),\dots,-(H_m(x^*),G_m(x^*)))$, $k \in \mathbb{N}$. Clearly, $\mathcal{I}(z^k) = \mathcal{I}(x^*)$, $\mathcal{K}(z^k) = \mathcal{K}(x^*)$ and $\mathcal{J}(z^k) = \mathcal{J}(x^*)$. Our aim is to find a subsequence $\mathcal{N} \subset \mathbb{N}$ and vectors $\hat{u}^k$ and $\hat{v}^k$ such that for every

$k \in \mathcal{N}$, $\text{supp}(\hat{u}^k) \subset \mathcal{I}(z^k) \cup \mathcal{J}(z^k)$, $\text{supp}(\hat{v}^k) \subset \mathcal{K}(z^k) \cup \mathcal{J}(z^k)$, either $\hat{u}^k_\ell \hat{v}^k_\ell = 0$ or $\hat{u}^k_\ell > 0, \hat{v}^k_\ell > 0$ for $\ell \in \mathcal{J}(z^k)$ and

$$\lim_{k \in \mathcal{N}} \| \sum_{i=1}^m \bar{u}^k_i \nabla H_i(x^k) + \sum_{i=1}^m \bar{v}^k_i \nabla G_i(x^k) - \sum_{i=1}^m \hat{u}^k_i \nabla H_i(x^k) - \sum_{i=1}^m \hat{v}^k_i \nabla G_i(x^k) \| = 0. \tag{91}$$

Now, we will show that $\bar{u}^k_i \to 0$, $\forall i \in \mathcal{K}(z^k)$ and $\bar{v}^k_i \to 0$, $\forall i \in \mathcal{I}(z^k)$. In this case, we can define $\hat{u}^k_i = 0$, $\forall i \in \mathcal{K}(z^k)$ and $\hat{v}^k_i = 0, \forall i \in \mathcal{I}(z^k)$. Clearly, $|\hat{u}^k_i - \bar{u}^k_i| = |\bar{u}^k| \to 0, \forall i \in \mathcal{K}(z^k)$ and $|\hat{v}^k_i - \bar{v}^k_i| = |\bar{v}^k| \to 0, \forall i \in \mathcal{I}(z^k)$. To prove $\bar{u}^k_i \to 0$, $\forall i \in \mathcal{K}(z^k)$, observe that for $i \in \mathcal{K}(z^k) = \mathcal{K}(x^*)$, we Since $|u^k_i(H_i(x^k) + t_k)| \le \varepsilon_k$ and $|\rho^k_i(G_i(x^k) - t_k)(H_i(x^k) - t_k)| \le \varepsilon_k$, we get $\rho^k_i(G_i(x^k) - t_k) \to 0$ and $u^k_i \to 0$. Thus, $\bar{u}^k_i = u^k_i - \rho^k_i(G_i(x^k) - t_k) \to 0$. Similarly, we get $\bar{v}^k_i = v^k_i - \rho^k_i(H_i(x^k) - t_k) \to 0$ for all $i \in \mathcal{I}(z^k)$.

A continuation, we will analyze $\mathcal{J}(z^k)$. Decompose $\mathcal{J}(x^*)$ into a partition, namely: $\mathcal{J}_1(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) > t_k, H_i(x^k) > t_k\}$, $\mathcal{J}_2(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) > t_k, H_i(x^k) \le t_k\}$, $\mathcal{J}_3(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) \le t_k, H_i(x^k) > t_k\}$ and $\mathcal{J}_4(x^*) := \{i \in \mathcal{J}(x^*) : G_i(x^k) \le t_k, H_i(x^k) \le t_k\}$. As always, we can assume that each element of the partition independent of k. We will only analyze the set $\mathcal{J}_1(x^*)$, the other cases can be proven similarly. Take $i \in \mathcal{J}_1(x^*) = \{i \in \mathcal{J}(x^*) : G_i(x^k) > t_k, H_i(x^k) > t_k\}$. Then, there is a subsequence $K \subset \mathbb{N}$ such that $H_i(x^k)/t_k \to_K \alpha$ and $G_i(x^k)/t_k \to_K \beta$, for some scalars $\alpha, \beta \in [1, \infty]$. If $\alpha \ne 1$. Then, from $\varepsilon_k = o(t_k)$ we get $|u^k_i(H_i(x^k)/t_k + 1)| \le o(t_k)/t_k \to 0$ and $|\rho^k_i(G_i(x^k) - t_k)(H_i(x^k)/t_k - 1)| \le o(t_k)/t_k \to 0$. Thus, $u^k_i \to 0$, $\rho^k_i(G_i(x^k) - t_k) \to 0$ and hence $\bar{u}^k_i = u^k_i - \rho^k_i(G_i(x^k) - t_k) \to 0$. Similarly, if $\beta \ne 1$, from $|v^k_i(G_i(x^k) + t_k)| \le \varepsilon_k = o(t_k)$ and $|\rho^k_i(G_i(x^k) - t_k)(H_i(x^k) - t_k)| \le \varepsilon_k = o(t_k)$, we get $\bar{u}^k_i = v^k_i - \rho^k_i(H_i(x^k) - t_k) \to 0$. Now, if $\alpha = \beta = 1$, from $H_i(x^k)/t_k \to_K 1$ and $G_i(x^k)/t_k \to_K 1$, we get that $(1+c)t_k > H_i(x^k) > t_k$ and $(1+c)t_k > G_i(x^k) > t_k$ for k large enough, which is impossible by (88). Thus, we see that there is subsequence $\mathcal{N} \subset \mathbb{N}$, such that $\bar{u}^k_i \to_{\mathcal{N}} 0$ or $\bar{v}^k_i \to_{\mathcal{N}} 0$. When $\bar{u}^k_i \to_{\mathcal{N}} 0$, define $\hat{u}^k_i := 0$, $\hat{v}^k_i := \bar{v}^k_i$, $k \in \mathcal{N}$, and when $\bar{v}^k_i \to_{\mathcal{N}} 0$, define $\hat{v}^k_i := 0$, $\hat{u}^k_i := \bar{u}^k_i$, $k \in \mathcal{N}$. In any case, $\|(\hat{u}^k_i, \hat{v}^k_i) - (\bar{u}^k_i, \bar{v}^k_i)\| \to_{\mathcal{N}} 0$ and $\hat{u}^k_i \bar{v}^k_i = 0$, $k \in \mathcal{N}$, $i \in \mathcal{J}_1(z^k) = \mathcal{J}_1(x^*)$.

In resume, there is a subsequence $\mathcal{N} \subset \mathbb{N}$ and vectors $\{(\hat{u}^k, \hat{v}^k) : k \in \mathcal{N}\}$ with $\hat{u}^k_i = 0$; $i \in \mathcal{K}(z^k)$, $\hat{v}^k_i = 0$; $i \in \mathcal{I}(z^k)$ and either $\hat{u}^k_i \hat{v}^k_i = 0$ or $\hat{u}^k_i > 0$ and $\hat{v}^k_i > 0$ $\forall i \in \mathcal{J}(z^k)$ such that $\|(\bar{u}^k_i, \bar{v}^k_i) - (\hat{u}^k_i, \hat{v}^k_i)\| \to_{\mathcal{N}} 0$, $\forall i$ and $\| \sum_{i=1}^m (\bar{u}^k_i - \hat{u}^k_i) \nabla H_i(x^k) + \sum_{i=1}^m (\bar{v}^k_i - \hat{v}^k_i) \nabla G_i(x^k)\| \to_{\mathcal{N}} 0$. Thus, $\{x^k : k \in \mathcal{N}\}$ with approximate multipliers $(\mu^k, \lambda^k, \hat{u}^k, \hat{v}^k)$ is a MPEC-AKKT sequence with $x^k \to_{\mathcal{N}} x^*$. Since MPEC-CCP holds, x^* is a M-stationary point. □

5.5 The global relaxation of Scholtes

The global relaxation scheme of Scholtes [52] is:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \le 0, \quad h(x) = 0, \quad 0 \le H(x), \quad 0 \le G(x), \\ & \Phi^S_i(x; t) := H_i(x) G_i(x) - t \le 0, \ \forall i \in \{1, \dots, m\} \end{array} \tag{92}$$

The above NLP is denoted by $NLP^S(t)$. Similarly to the relaxations or penalization schemes mentioned above, we can use the MPEC-CCP to replace more stringent MPEC-CQs in order to guarantee convergence to M-stationary points, as the next theorem shows.

Theorem 5.9. *Let $t_k \downarrow 0$, $\varepsilon_k = o(t_k)$, x^k be a sequence of ε_k-stationary points of $NLP^S(t_k)$. Assume that $(\mu^k, \lambda^k, u^k, v^k, \rho^k) \in \mathbb{R}^p_+ \times \mathbb{R}^q \times \mathbb{R}^m_+ \times \mathbb{R}^m_+ \times \mathbb{R}^m_+$, satisfy $\max\{|u^k_i H_i(x^k)|, |v^k_i G_i(x^k)|, |\rho^k_i \Phi^S_i(x^k, t_k)| : i \in \{1, \dots, m\}\} \le \varepsilon_k$ and $x^k \to x^*$. Suppose that x^* conforms MPEC-CCP and that there is a constant $c > 0$ such that, for all $i \in \{1, \dots, m\}$ and all k large enough, the iterates $(G_i(x^k), H_i(x^k))$ do not belong to $\{(a, b) \in \mathbb{R}^2 : (1-c)t_k < ab < (1+c)t_k\}$. Then, x^* is a M-stationary point.*

Proof. Hence MPEC-CCP holds at x^*, it will be sufficient to show that x^* is a MPEC-AKKT point. Since $\{x^k\}$ is a sequence of ε_k-stationary point of $\text{NLP}^S(t_k)$, we get that

$$\|\nabla f(x^k) + \sum_{j=1}^p \mu^k_j \nabla g_j(x^k) + \sum_{i=1}^q \lambda^k_i \nabla h_i(x^k) - \sum_{i=1}^m \bar{u}^k_i \nabla H_i(x^k) - \sum_{i=1}^m \bar{v}^k_i \nabla G_i(x^k)\| \le \varepsilon_k \tag{93}$$

where $\bar{v}^k_i := v^k_i - \rho^k_i H_i(x^k)$, $\bar{u}^k_i := u^k_i - \rho^k_i G_i(x^k)$, $\text{supp}(\mu^k) \subset A(x^*)$ and $g_j(x^k) \le \varepsilon_k$, $\mu^k_j \ge 0$, $|\mu^k_j g_j(x^k)| \le \varepsilon_k$, $\forall j \in \{1, \dots, p\}$; $H_i(x^k) \ge -\varepsilon_k$, $u^k_i \ge 0$, $|u^k_i H_i(x^k)| \le \varepsilon_k$, $\forall i \in \{1, \dots, m\}$; $G_i(x^k) \ge -\varepsilon_k$, $v^k_i \ge 0$, $|v^k_i G_i(x^k)| \le \varepsilon_k$, $\forall i \in \{1, \dots, m\}$; $\max\{\Phi^S_i(x^k, t_k), \rho^k_i \Phi^{KS}_i(x^k, t_k)\} \le \varepsilon_k$, $\rho^k_i \ge 0$, $\forall i \in \{1, \dots, m\}$.

Now, set $z^k := (g(x^*), h(x^*), -(H_1(x^*), G_1(x^*)), \dots, -(H_m(x^*), G_m(x^*)))$, $k \in \mathbb{N}$. Now, we will show that there is a subsequence $\mathcal{N} \subset \mathbb{N}$ such that $\bar{u}_i^k \to_{\mathcal{N}} 0$, $\forall i \in \mathcal{K}(z^k)$ and $\bar{v}_i^k \to_{\mathcal{N}} 0$, $\forall i \in \mathcal{I}(z^k)$. Indeed, take $i \in \mathcal{K}(z^k) = \mathcal{K}(x^*)$. Since x^k is a ε_k-stationarity point, we get $u_i^k \to 0$, $G_i(x^k)H_i(x^k) \geq -\varepsilon_k H_i(x^k)$ and $\rho_i^k(G_i(x^k)H_i(x^k)/t^k - 1) \to 0$. Then, there is a subsequence $\mathcal{N} \subset \mathbb{N}$ such that $G_i(x^k)H_i(x^k)/t^k \to_{\mathcal{N}} \alpha$ for some $\alpha \in [0,1]$. If $\alpha \neq 1$, then $\rho_i^k \to 0$ and $\bar{u}_i^k = u_i^k - \rho_i^k G_i(x^k) \to 0$. If $\alpha = 1$, we get $(1-c)t_k < H_i(x^k)G_i(x^k) < (1+c)t_k$, $k \in \mathcal{N}$, which is a contradiction. Then as well for $i \in \mathcal{I}(z^k) = \mathcal{I}(x^k)$. Thus, we can define $\hat{u}_i^k = 0$, $\forall i \in \mathcal{K}(z^k)$ and $\hat{v}_i^k = 0$, $\forall i \in \mathcal{I}(z^k)$.

Take $i \in \mathcal{J}(z^k) = \mathcal{J}(x^*)$. From, $H_i(x^k)G_i(x^k) - t_k \leq o(t_k)$, we conclude that there is a subsequence $\mathcal{N} \subset \mathbb{N}$ such that $G_i(x^k)H_i(x^k)/t^k \to_{\mathcal{N}} \beta$ for some $\beta \in [-\infty, 1]$. If $\beta \neq 1$, $\rho_i^k \to_{\mathcal{N}} 0$. Define $\hat{u}_i^k := u_i^k$, $\hat{v}_i^k := v_i^k$. Note that $\hat{u}_i^k \geq 0$, $\hat{v}_i^k \geq 0$ and $\|(\hat{u}_i^k, \hat{v}_i^k) - (\bar{u}_i^k, \bar{v}_i^k)\| \to_{\mathcal{N}} 0$. If $\beta = 1$, we get $(1-c)t_k < H_i(x^k)G_i(x^k) < (1+c)t_k$, $k \in \mathcal{N}$, which is impossible for assumption.

Thus, we can find a subsequence $\mathcal{N} \subset \mathbb{N}$ and vectors $(\hat{u}^k, \hat{v}^k)$ such that $\text{supp}(\hat{u}^k) \subset \mathcal{I}(z^k) \cup \mathcal{J}(z^k)$, $\text{supp}(\hat{v}^k) \subset \mathcal{K}(z^k) \cup \mathcal{J}(z^k)$, $\hat{u}_\ell^k \hat{v}_\ell^k = 0$ or $\hat{u}_\ell^k > 0, \hat{v}_\ell^k > 0$ for $\ell \in \mathcal{J}(z^k)$ and $\|(\hat{u}_i^k, \hat{v}_i^k) - (\bar{u}_i^k, \bar{v}_i^k)\| \to_{\mathcal{N}} 0$. Then, x^* is a MPEC-AKKT point which is a M-stationary point, since MPEC-CCP holds. □

Note that we have obtain convergence to M-stationarity point for the relaxation of Scholtes [52] using only first-order information and a weak MPEC-CQ, This compares with the Theorem 3.3 of [52] where MPEC-LICQ is assume to be valid, together with a second-order optimality condition holding for all iterates and an additional condition about the bi-active index set. The assumption $\varepsilon_k = o(t_k)$ is, in general, necessary in order to get convergence to a M-stationary point, despite that a stronger MPEC-CQs may be considered and the iterates do not belong to the $\{(a,b) \in \mathbb{R}^2 : (1-c)t_k < ab < (1+c)t_k\}$ for some $c > 0$, as the next example shows.

Example 5.2. Take the system considered in the example 5.1. Define $t_k = 1/k$, $\varepsilon_k := (t_k^{1/4} - t_k^{3/4})/(1 - t_k^{1/4} + t_k^{3/4})$, $x_1^k = x_2^k := t_k^{1/4}$, $u^k := 0$, $v^k := 0$ and $\rho^k := (1+\varepsilon_k)/x_1^k$. By some calculations, we see that $x^k = (x_1^k, x_2^k)$ is an ε_k-stationaty point. Clearly, x^k goes to $x^* := (0,0)$, the iterates do not belong to $\{(a,b) \in \mathbb{R}^2 : (1-c)t_k < ab < (1+c)t_k\}$ for any $c > 0$, since $x_1^k x_2^k / t_k = 1/t_k^{1/2} \to \infty$. Note that MPEC-LICQ (and hence MPEC-CCP) holds at x^* and x^* is not a M-stationary. The assumption $\varepsilon_k = o(t_k)$ fails, since $\varepsilon_k / t_k \to \infty$.

6 Conclusions

Many constrained optimization algorithms, in the search for optimal solutions, usually end it up at a point where the KKT condition holds approximately. This motivates the study of the possible limit points generated by those methods and their relationship with optimality conditions. Sequential optimality conditions analyzed in [2, 41, 8, 34] serve for that purpose and they can be seen as sequential counter-parts of KKT. In the presence of complementarity constraints, it is well-known that standard constrained optimization algorithms may converge to non-KKT points from which a descent direction arises and that different many alternatives to point-wise optimality conditions such as weakly stationary, C-stationary or M-stationary were stated. Thus, we introduced the sequential counter-part of the optimality condition M-stationarity (which is the same of S-stationarity) called MPEC-AKKT and we settled its main properties. The non-triviality of that sequential optimality condition is guaranteed by its companion CQ (MPEC-CCP). Out of all the constrained optimization algorithms studied here and found in the literature, which are designed to solve MPECs, we see that the convergence to M-stationary points is achieved if additional assumptions about the iterates are required, which we believe is a reflection of the deep relation between AKKT and the recently introduced MPEC-AKKT. Furthermore, using the notion of MPEC-AKKT in combination with the companion MPEC-CQ, we prove convergence to M-stationary points for several schemes, under weaker assumptions than the previous stated, even in the case, where the set of multipliers is unbounded. Additionally, guided by MPEC-AKKT, we propose a new method whose stopping criterion are based on the MPEC-AKKT condition and we show that such method has nice convergence properties. Finally, we hope that, this kind of analysis can be useful in the development and derivation of new methods for solving MPECs with strong convergence properties, specially in the search for algorithms whose stopping criteria are based on the MPEC-AKKT condition instead of the AKKT condition and variations.

References

[1] Andreani, R., Birgin, E., Martínez, J.M., Schuverdt, M.L.: Augmented lagragrian methods under the constant positive linear dependence constraint qualification. Mathematical Programming **111(1-2)**, 5–32 (2008)

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

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

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

[5] Andreani, R., Martínez, J.M.: On the solution of mathematical programming problems with equilibrium constraints. Mathematical Methods of Operations Research **54**, 345–358 (2001)

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

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

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

[9] Anitescu, M.: On using the elastic mode in nonlinear programming approaches to mathematical programming with complementarity constraints. SIAM Journal on Optimization **15**, 1203–1236 (2005)

[10] Anitescu, M., Tseng, P., Wright, S.J.: Elastic-mode algorithms for mathematical programs with equilibrium constraints: global convergence and stationarity properties. Mathematical Programming **110**(2), 337–371 (2007)

[11] Aubin, J.P., Frankowska, H.: Set-Valued Analysis. Birkhauser (1990)

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

[13] Birgin, E.G., Bueno, L.F., Martinez, J.M.: Sequential equality-constrained optimization for nonlinear programming. Computational Optimization and Applications **65**, 699–721 (2016)

[14] Birgin, E.G., Martinez, J.M., Prudente, L.F.: Optimality properties of an augmented lagrangian method on infeasible problems. Computational Optimization and Applications **60**, 609–631 (2015)

[15] Borwein, J.M., Lewis, A.S.: Convex Analysis and Nonlinear Optimization: Theory and Examples. CMS Books in Mathematics (2006)

[16] Cervinka, M., Outrata, J.V., Pistek, M.: On stability of M-stationary points in MPCCs. Set-valued and variational analysis **22(3)**, 575–595 (2014)

[17] Chen, L., Goldfarb, D.: Interior-point 2-penalty methods for nonlinear programming with strong global convergence properties. Mathematical Programming **108**, 1026 (2006)

[18] Chieu, N.H., Lee, G.M.: Constraint qualifications for mathematical programs with equilibrium constraints and their local preservation property. Journal of Optimization Theory and Applications **163**, 755–776 (2014)

[19] Clarke, F.: Functional Analysis, Calculus of Variations and Optimal Control. Graduate text in Mathematics, 264 (2012)

[20] Cottle, R.W., Pang, J.S., Stone, R.E.: The Linear Complementary Problems. Academic press (1992)

[21] Demiguel, A.V., Friedlander, M.P., Scholtes, F.J.N.S.: A two-sided relaxation scheme for mathematical programs with equilibrium constraints. SIAM Journal of Optimization **16**, 587–609 (2005)

[22] Dempe, S.: Foundations of Bilevel Programming. Kluwer Academic (2002)

[23] Facchinei, F., Kansow, C.: Penalty methods for the solution of generalized nash equilibrium problems. SIAM Journal on Optimization **20**, 2228–2253 (2010)

[24] Flegel, M.L., Kanzow, C.: On the Guignard constraint qualification for mathematical programs with equilibrium constraints. Optimization **54(6)**, 517–534 (2005)

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

[26] Guo, L., Lin, G.H.: Notes on some constraint qualifications for mathematical programs with equilibrium constraints. Journal of Optimization Theory and Applications **156**, 600–616 (2013)

[27] Guo, L., Lin, G.H., Ye, J.J.: Second order optimality conditions for mathematical programs with equilibrium constraints. Journal of Optimization Theory and Applications **158**, Issue 1, pp 33–64 (2013)

[28] Guo, L., Lin, G.H., Ye, J.J., .Zhang, J.: Sensitivity analysis of the value function for parametric mathematical programs with equilibrium constraints. SIAM Journal on Optimization **24(3)**, 1206–1237 (2014)

[29] Hestenes, M.R.: Multiplier and gradient methods. Journal of Optimization Theory and Applications **4**, 303–320 (1969)

[30] Hoheisel, T., Kanzow, C., Schwartz, A.: Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. Mathematical Programming **137**, 257–288 (2013)

[31] Hu, X.M., Ralph, D.: Convergence of a penalty method for mathematical programming with complementarity constraints. Journal of Optimization Theory and Applications **123**(2), 365–390 (2004)

[32] Kadrani, A., Dussault, J.P., Benchakroun, A.: A new regularization method scheme for mathematical programs with complementary constraints. SIAM Journal of Optimization **20**, 23, pp 78–103 (2009)

[33] Kadrani, A., Dussault, J.P., Benchakroun, A.: A globally convergent algorithm for mpcc. EURO Journal on Computational Optimization **3**(3), 263–296 (2015)

[34] Kansow, C., Steck, D., Waschsmuth, D.: An augmented lagrangian method for optimization problems in banach spaces. Preprint, Institute of Mathematics, University of Wrzburg (December 2016)

[35] Kanzow, C., Schwartz, A.: Mathematical programs with equilibrium constraints: enhanced Fritz-John conditions, new constraint qualifications and improving exact penalty results. SIAM Journal of Optimization **20**, 2730–2753 (2010)

[36] Kanzow, C., Schwartz, A.: A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. SIAM Journal on Optimization **23**, 770–798 (2013)

[37] Kanzow, C., Schwartz, A.: The price of inexactness: Convergence properties of relaxation methods for mathematical programs with complementary constraints revisited. Mathematics of Operations Research **40 (21)**, 253–275 (2015)

[38] Leyffer, S., López-Calva, G., Nocedal, J.: Interior methods for mathematical programs with complementarity constraints. SIAM Journal on Optimization **17**(1), 52–77 (2006). DOI 10.1137/040621065. URL http://dx.doi.org/10.1137/040621065

[39] Lin, G.H., Fukushima, M.: A modified relaxation scheme for mathematical programs with complementarity constraints. Annals of Operation Reseach **133**, 63–84 (2005)

[40] Luo, Z.Q., Pang, J.S., Ralph, D.: Mathematical programs with equilibrium constraints. Cambridge University Press (1996)

[41] Martinez, J.M., Svaiter, B.F.: A practical optimality condition without constraint qualifications for nonlinear programming. Journal of Optimization Theory and Applications **118 No 1**, 117–133 (2003)

[42] Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I: Basic Theory. A series of comprehensive studies in mathematics, 330 (2006)

[43] Outrata, J.V.: Optimality conditions for a class of mathematical programs with equilibrium constraints. Mathematics of Operations Research **24(3)**, 627–644 (1999)

[44] Outrata, J.V.: A generalized mathematical program with equilibrium constraints. SIAM Journal on Control and Optimization **38**, 1623–1638 (2000)

[45] Outrata, J.V., Kocvara, M., Zowe, J.: Nonsmooth approach to optimization problems with equilibrium constraints: Theory, Applications and Numerical Results. Kluwer Academic (1998)

[46] Powell, M.J.D.: A method for nonlinear constraints in minimization problems

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

[48] Ralph, D.: Mathematical programs with complementarity constraints in traffic and telecommunications networkss. Royal Society of London. Philosophical Transactions. Mathematical, Physical and Engineering Sciences **52**, 366(1872): 1973–1987 (20083). DOI 10.1098/rsta.2008.0026

[49] Ralph, D., Wright, S.J.: Some properties of regularization and penalization schemes for mpecs. Optimization Methods and Software **19**(5), 527–556 (2004)

[50] Rockafellar, R.T., Wets, R.: Variational Analysis. Series: Grundlehren der mathematischen Wissenschaften, Vol. 317 (2009)

[51] Scheel, H., Scholtes, S.: Mathematical programs with complementary constraints: Stationarity, optimality and sensitivity. Mathematics of Operations Research **25(1)**, 1–22 (2000)

[52] Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM Journal of Optimization **11**, 918–936 (2001)

[53] Spivak, M.: Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus. Addison-Wesley Publishing Company (1965)

[54] Steffensen, S., Ulbrich, M.: A new regularization scheme for mathematical programs with equilibrium constraints. SIAM Journal of Optimization **20**, 2504–2539 (2010)

[55] Tseng, P.: A convergent infeasible interior-point trust-region method for constrained minimization. SIAM Journal on Optimization **13**, 432–469 (2002)

[56] Ye, J.: Necessary and sufficient optimality conditions for mathematical problems with equilibrium constraints. Journal of Mathematical Analysis and Applications pp. 307, pp 350–369 (2005)

[57] Ye, J.J., Ye, X.Y.: Necessary optimality conditions for optimization problems with variational inequality constraints. Mathematics of Operations Research **22(4)**, 977–997 (1997)

[58] Ye, J.J., Zhang, J.: Enhanced Karush-Kuhn-Tucker conditions for mathematical programs with equilibrium constraints. Journal of Optimization Theory and Applications **163**, 777–794 (2014)