On the boundedness of multipliers in augmented Lagrangian methods for mathematical programs with complementarity constraints*

R. Andreani[†] M. da Rosa[†] L. D. Secchin[‡]

August 18, 2025

Abstract

In this paper, we present a theoretical analysis of augmented Lagrangian (AL) methods applied to mathematical programs with complementarity constraints (MPCCs). Our focus is on a variant that reformulates the complementarity constraints using slack variables, where these constraints are handled directly in the subproblems rather than being penalized. We introduce specialized constraint qualifications (CQs) of the quasi-normality type and establish a global convergence result under these assumptions. In addition, we analyze the behavior of the associated dual sequences and prove their boundedness under the CQs introduced. These conditions also yield a new, simple, tailored error bound property. Finally, we compare this approach with the standard AL framework in terms of theoretical properties and numerical stability on problems from the MacMPEC collection. Our results indicate that maintaining complementarities within the subproblems leads to improved numerical stability.

1 Introduction

In this paper, we consider the *mathematical programs with complementarity constraints* (MPCC)

$$\min_{w} f(w) \quad \text{s.t.} \quad h(w) = 0, \quad g(w) \leq 0, \quad 0 \leq H(w) \perp G(w) \geq 0, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^s$, $g : \mathbb{R}^n \to \mathbb{R}^p$, $G, H : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable functions. Here, $0 \leq H(w) \perp G(w) \geq 0$ means $H(w) \geq 0$, $G(x) \geq 0$ and $H_i(w)G_i(w) = 0$ for all i. These are the *complementarity constraints*.

MPCCs are known to violate most of established constraint qualifications (CQs) for standard nonlinear programming. In particular, the Mangasarian–Fromovitz CQ (MFCQ) is violated at all feasible points [43], and even weaker CQs like the relaxed constant positive linear dependence (RCPLD) or Abadie's CQs are generally not satisfied for MPCCs [26]. As a consequence, the KKT conditions are usually inapplicable to MPCCs. The main difficulty lies in the complementarity constraints, specifically when strict complementarity does not hold, that is, when $H_i(w) = G_i(w) = 0$ for some i.

In light of these issues, several stationary concepts weaker than KKT, such as M-, C-, and W-stationarity, have been developed in conjunction with CQs tailored to MPCCs. These developments have guided the analysis of several strategies designed to address the pathological nature of MPCCs. Among the proposed methods to solve (1) are: dealing with complementarities by penalizing a parameter that relax them [9, 10, 11]; solving a sequence of problems in which relaxed versions

*This work has been partially supported by CEPID-CeMEAI (FAPESP 2013/07375-0), FAPESP (grants 2018/24293-0, 2017/18308-2, 2022/06745-7 and 2024/12967-8), CNPq (grants 306988/2021-6, 407147/2023-3 and 302520/2025-2), PRONEX - CNPq/FAPERJ (grant E-26/010.001247/2016) and UNICAMP/PRP/FAEPEX (grant 3319/23).

†Department of Applied Mathematics, University of Campinas, Campinas, SP, Brazil. Email: andreani@ime.unicamp.br, m265027@dac.unicamp.br

‡Department of Applied Mathematics, Federal University of Espírito Santo, ES, Brazil. Email: leonardo.secchin@ufes.br

of complementarity are successively tightened [34]; adapting well-known strategies to (1) [32, 35, 39, 42, 36]; and apply standard optimization methods such as augmented Lagrangian [31] directly, considering (1) as a standard nonlinear problem. These approaches typically guarantee convergence only to C-stationary. A key research objective was to ensure convergence to M-stationary points, which are tighter than C-stationary points.

The SQP method developed in [12] guarantees M-stationary points by assuming the boundedness of the dual sequence associated with quadratic subproblems with complementarity constraints. Andreani et al. [8] showed that an AL method that uses second-order information achieves convergence to M-stationary points under the *MPCC-linear independence CQ* (MPCC-LICQ) without further hypotheses on the sequences generated by the algorithm. Guo and Deng [28] introduced a first-order AL method that converges to M-stationary points under the less stringent *MPCC-relaxed positive linear dependence* (MPCC-RCPLD) CQ. Their approach introduces slack variables to handle complementarity constraints directly in the subproblems, rather than incorporating them into the AL function. This enables efficient solution of the subproblems via a nonmonotone projected gradient method, while preserving the feasibility in relation to complementarity constraints. Such structural preservation has theoretical and numerical implications that set this approach apart from traditional AL methods.

In traditional AL methods, where all constraints are penalized, the multipliers may diverge even under strong theoretical assumptions, as illustrated by the following simple example.

Example 1. *Consider the MPCC problem*

$$\min_{y,z} -y - z \quad s.t. \quad 0 \leq y \perp z \geq 0.$$

As discussed in [31], the safeguarded quadratic penalty-like AL method presented in [1] may converges to the point $(0, 0)$ *via the sequence* $\{(y^k, z^k) = (1/k, 1/k)\}$ *and the penalty sequence* $\{\rho_k = k^3\}$*. Approximating the first order optimality of this problem viewed as a standard model with constraints* $y, z \geq 0$ *and* $yz = 0$*, we see that*

$$\begin{bmatrix} -1 \\ -1 \end{bmatrix} + \mu_y^k \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \mu_z^k \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \lambda^k \begin{bmatrix} z^k \\ y^k \end{bmatrix} \to \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where $\{(\mu_y^k, \mu_z^k, \lambda^k)\}$ *is a sequence of multipliers. As* $\mu^k \geq 0$*, the only possibility for* $(y^k, z^k) \to (0, 0)$ *is that the multiplier* λ^k *associated with the complementarity tends to infinity, exactly what happens with the aforementioned AL method. Note that in this problem, the gradients of* y *and* z *are linearly independent everywhere, that is, the counterpart of LICQ to MPCC, MPCC-LICQ, holds at all feasible points.*

This highlights a key issue: when convergence is to a point where strict complementarity fails, it can only be achieved by forcing the feasibility of complementarity constraints through increasingly larger penalty parameters, which necessarily lead to an unbounded corresponding dual sequence. In contrast, the AL method proposed by Guo and Deng preserves the complementarity structure in the subproblems, which lead us to think that this behavior is avoided. However, to the best of our knowledge, there are no consistent theory that guarantees the boundedness of multipliers in any AL method applied to MPCCs. The only result in this line is [2, Corollary 4.9], but it refers to the so-called MPCC-multipliers, not the multiplier estimates treated within the method (in time, due to Example 1, the conclusion in this corollary regarding the boundedness of the dual sequence "$\{\gamma^{0,k}\}$" associated with the complementarity constraint is incorrect).

In this work, we address this gap by analyzing the dual sequence generated by the Guo and Deng's AL method, which is presented in Algorithm 1. Our analysis is inspired by recent advances in nonlinear programming, particularly the very recent *relaxed quasi-normality* (RQN) CQ proposed in [4] and the classical notion of *quasi-normality* (QN), originally introduced by Hestenes and later extensively studied by Bertesekas, Ozdaglar and collaborators; see e.g. [13, 14].

Global convergence of the safeguarded AL method called ALGENCAN [1] under QN was first established in [2]. In this last work, all constraints are penalized, so AL subproblems are unconstrained. In [5], box-constrained subproblems are considered in the theory and the corresponding version of QN that involves the projection onto the box was used. Later, a general QN definition stated in terms of an abstract set [14, Definition 3.1] was evoked to handle an AL method where an abstract set is maintained in the subproblems [23, 25]. Aiming to state the boundedness of dual sequences generated by the Guo and Deng's AL method, and considering the boundedness result for ALGENCAN stated earlier [2], it becomes natural to consider QN with the abstract set consisting of the complementarity constraints. We refer to this condition as *specialized MPCC-QN* (sMPCC-QN, see Definition 6).

Based on this foundation, we propose a new and even weaker CQ: *specialized MPCC-relaxed quasi-normality* (sMPCC-RQN). In the context of general mathematical programs without an abstract set, RQN [4] is the weakest CQ that ensures boundedness of the multiplier sequence in AL methods [3].

It is known that QN and RQN imply an (local) error bound property; see [38] and [3], respectively. In the context of MPCCs, an error bound was proposed in [24], which we refer here as MPCC-EB (see section 3.1). We then prove that the tailored MPCC-QN and MPCC-RQN CQs imply MPCC-EB. This is done by proving that these CQs imply a new notion of error bound for MPCCs (see Definition 8). The new error bound is equivalent to MPCC-EB; however, it is clearer than MPCC-EB as complementarity constraints are required to be fulfilled, so they are not treated explicitly. This is in accordance with the Guo and Deng's method.

Summarizing, our contributions are:

- *Introduce new QN-type tailored CQs for MPCCs* We propose the specialized MPCC-RQN, extending the relaxed quasi-normality (RQN) condition from standard nonlinear programming [4] to the MPCC context. We also show that both the specialized MPCC-RQN and the specialized MPCC-QN are valid MPCC CQs for M-stationarity;

- *Establish the boundedness of the multipliers sequence in the Guo and Deng's AL framework.* Unlike the standard AL method, we prove that the dual sequence produced by the AL method defined in [28], which preserves the complementarity structure in the subproblems, remains bounded under either the specialized MPCC-RQN or the specialized MPCC-QN. Furthermore, we show a similar result with the generalized MPCC-QN condition introduced in [33]. In addition, our theory extends the convergence results previously established in [28];

- *Define an error bound with respect to the complementarity set.* We formulate an error bound that enforces feasibility regarding to the set of complementarity constraints, allowing the distance to the feasible set to be measured directly from constraint violations. We prove that our definition is equivalent to the more intricate MPCC-EB defined in [24] and that it is implied by both the specialized MPCC-RQN and MPCC-QN conditions;

- *Illustrate the computational benefits over the standard AL approach.* We present numerical experiments that support our theoretical findings and show that maintaining complementarities in the subproblems enhances numerical stability, which may become a new paradigm for the development of inner solvers within the AL framework. The code used is available for free download.

The paper is organized as follows. In section 2, we review stationarity concepts for MPCCs and describe the AL framework of [28], that preserves the complementarity constraints in the subproblems. Section 3 introduces the proposed specialized MPCC-RQN CQ and compare it with the generalized MPCC-QN condition. The new error bound for MPCCs is introduced and the relationship of it with the new CQs are addressed. Section 4 presents the main theoretical result, establishing the boundedness of multipliers sequence of the AL method under the introduced CQs. We also illustrate the practical implications through numerical experiments, using instances from the MacMPEC collection. Finally, concluding remarks are given in section 5.

Notation Given $x \in \mathbb{R}^n$, $\|x\|$ is its Euclidean norm and $x_+ \in \mathbb{R}^n$ is the vector $(\max\{0, x_1\}, \dots, \max\{0, x_n\})$. The Euclidean closed ball centred at x with radius $\delta > 0$ is denoted by $B(x, \delta)$. Given $x = (w, y, z)$, we define the index sets $I_g(x) = \{j \mid g_j(x) = 0\}$, $I_{+0}(x) = \{i \mid y_i > 0,\ z_i = 0\}$, $I_{0+}(x) = \{i \mid y_i = 0,\ z_i > 0\}$, and $I_{00}(x) = \{i \mid y_i = 0,\ z_i = 0\}$. Given two sequences $\{a_k\}, \{b_k\} \subseteq \mathbb{R}^n$, we write $a_k = o(b_k)$ to say that there is a real positive sequence $\{m_k\}$ converging to zero such that $|a_k| \leq m_k b_k$ for all k large enough. For $\Gamma : \mathbb{R}^n \to \mathbb{R}$, we denote the transpose of its Jacobian at x by $\nabla \Gamma(x)$.

2 Preliminaries

The complementarity constraints in (1) can be equivalently written using slack variables $y, z \in \mathbb{R}^m$, leading to the problem

$$\begin{aligned} \min_{w,y,z} \; & f(w) \\ \text{s.t. } & h(w) = 0, \quad g(w) \le 0, \\ & y - G(w) = 0, \quad z - H(w) = 0, \quad 0 \le y \perp z \ge 0. \end{aligned}$$

To simplify notation, we write $x = (w, y, z) \in \mathbb{R}^n \times \mathbb{R}^{2m}$ and assume, by an abuse of notation, that the data functions f, h, g, H, G are defined on the enlarged space $\mathbb{R}^n \times \mathbb{R}^{2m}$ by the correspondences

$$f(x) := f(w), \quad h(x) := (h(w), y - H(w), z - G(w)), \quad g(x) := g(w).$$

With this, "$h(x) = 0$" encompasses the new equality constraints that define the slacks y and z. Thus, in the remainder of the paper we consider the problem

$$\min f(x) \quad \text{s.t.} \quad h(x) = 0, \quad g(x) \le 0, \quad x \in X, \tag{MPCC}$$

where $X = \{(w, y, z) \in \mathbb{R}^n \times \mathbb{R}^{2m} \mid y_i \ge 0,\; z_i \ge 0,\; y_i z_i = 0,\; i = 1, \ldots, m\}$. The feasible set of (MPCC) is denoted by

$$\Omega = \{x \in X \mid h(x) = 0,\; g(x) \le 0\}.$$

As in [28], we also consider the functions $c, d : \mathbb{R}^n \times \mathbb{R}^{2m} \to \mathbb{R}^m$ defined by $c(x) := y$, $d(x) := z$. So, following X, we have the constraints $c(x) \ge 0$, $d(x) \ge 0$ and $c_i(x) d_i(x) = 0$ for all i. This allows us to write (MPCC) in the form of (1).

As we mentioned in the introduction, MPCCs have a highly level of degeneracy [31] since most standard CQs fail to hold. Therefore, weaker stationary concepts than KKT and tailored CQs have been studied in the literature for this class of problems. Given a target point $x^* \in \Omega$, a common approach is to consider the *tightened nonlinear problem* (TNLP)

$$\begin{aligned} \min_{x} \; & f(x) \\ \text{s.t. } & h(x) = 0, \quad g(x) \le 0, \\ & c_i(x) = 0 \quad \forall i \in I_{0+}(x^*) \cup I_{00}(x^*), \\ & d_i(x) = 0 \quad \forall i \in I_{+0}(x^*) \cup I_{00}(x^*), \\ & c_i(x) \ge 0 \quad \forall i \in I_{+0}(x^*), \\ & d_i(x) \ge 0 \quad \forall i \in I_{0+}(x^*). \end{aligned} \tag{TNLP}$$

In this model, active complementary constraints ($c_i(x) \ge 0$ or $d_i(x) \ge 0$) at x^* are transformed into equalities, thus fulfilling $c_i(x) d_i(x) = 0$ automatically. Geometrically, the feasible set of (TNLP) is the "smallest face" of Ω that contains x^*.

The main notions of stationarity for (MPCC) are derived from (TNLP). We define the *MPCC-Lagrangian function* for (MPCC) $\mathcal{L} : \mathbb{R}^{n+2m} \times \mathbb{R}^q \times \mathbb{R}^p \times \mathbb{R}^{2m} \to \mathbb{R}$ by

$$\mathcal{L}(x, \lambda, \mu, u, v) = f(x) + \lambda^t h(x) + \mu^t g(x) - c(x)^t u - d(x)^t v.$$

This is the Lagrangian associated with (TNLP), where u_i and v_i are the Lagrange multipliers associated with $c_i(x) \ge (=) 0$ and $d_i(x) \ge (=) 0$, respectively.

The weakest stationary (w.r.t. x^*) is simply KKT for (TNLP) [37]:

Definition 1. *We say that a feasible point x^* of* (MPCC) *is* weakly stationary *(W-stationary) if there exist multipliers $(\lambda, \mu, u, v) \in \mathbb{R}^q \times \mathbb{R}^p_+ \times \mathbb{R}^{2m}$ such that*

$$\nabla_x \mathcal{L}(x^*, \lambda, \mu, u, v) = 0, \;\; \mu^t g(x^*) = 0, \;\; u_i = 0 \textit{ for } i \in I_{+0}(x^*), \;\; v_i = 0 \textit{ for } i \in I_{0+}(x^*).$$

Other known stationary concepts are obtained by strengthening the multipliers of complementary constraints that vanish simultaneously [41, 43].

Definition 2. *Let x^* be a W-stationary point with associated vector of multipliers (λ, μ, u, v). We say that x^* is*

- *Clarke-stationary (C-stationary) if $u_i v_i \geq 0$ for all $i \in I_{00}(x^*)$;*
- *Mordukhovich-stationary (M-stationary) if $u_i, v_i > 0$ or $u_i v_i = 0$ for all $i \in I_{00}(x^*)$;*
- *Strongly-stationary (S-stationary) if $u_i \geq 0$ and $v_i \geq 0$ for all $i \in I_{00}(x^*)$.*

We have the following strict implications: S-stationarity $\Rightarrow$ M-stationarity $\Rightarrow$ C-stationarity $\Rightarrow$ W-stationarity. W-stationarity is not considered a good characterization of local optimality, while S-stationarity is unexpected when strict complementarity fails (indeed, it is equivalent to KKT [27]). In turn, C- and M-stationarity are widely accepted in algorithmic contexts; see [6, 28, 29, 34, 44] and references therein.

As occurs with KKT, the validity of the above stationarity conditions at minimizers of (MPCC) requires appropriate CQs. For our purposes, we recall the definition of MPEC generalized quasi-normality introduced in [33], which we call MPCC-quasi-normality here. For other MPCC-CQs, see [6, 30] for instance.

Definition 3. *We say that a feasible point x^* conforms to the* MPCC-quasi-normality *(MPCC-QN) CQ if there is no $(\lambda, \mu, u, v) \in \mathbb{R}^q \times \mathbb{R}^p_+ \times \mathbb{R}^{2m}$ such that*

1. *$\nabla h(x^*)\lambda + \nabla g(x^*)\mu - \nabla c(x^*)u - \nabla d(x^*)v = 0$;*
2. *$(\lambda, \mu, u, v) \neq 0$;*
3. *$\mu_i = 0$ for all $i \notin I_g(x^*)$, $\quad u_i = 0$ for all $i \in I_{+0}(x^*)$, $\quad v_i = 0$ for all $i \in I_{0+}(x^*)$, and for all $i \in I_{00}(x^*)$, either $u_i > 0$, $v_i > 0$, or $u_i v_i = 0$;*
4. *Let $I_{\neq} = \{i \mid \lambda_i \neq 0\}$, $J_+ = \{j \mid \mu_j > 0\}$, $I_u = \{i \mid u_i \neq 0\}$ and $I_v = \{i \mid v_i \neq 0\}$. There exists a sequence $\{x^k\}$ converging to x^* such that*

$$\lambda_i h_i(x^k) > 0 \text{ for all } i \in I_{\neq}, \quad g_j(x^k) > 0 \text{ for all } j \in J_+,$$
$$-u_i c_i(x^k) > 0 \text{ for all } i \in I_u \quad \text{and} \quad -v_i d_i(x^k) > 0 \text{ for all } i \in I_v.$$

Usually, MPCC-QN is defined directly to (1). However, it is straightforward to show that it is equivalent to the validity of MPCC-QN after adding slacks as done in (MPCC).

For the sake of completeness, we recall the definitions of QN and RQN for standard nonlinear programming, which inspired MPCC-QN and our sMPCC-RQN (Definition 7). They are used in the results of section 3.1.

Let us consider the set

$$S = \{\tilde{x} \mid \tilde{h}(\tilde{x}) = 0,\ \tilde{g}(\tilde{x}) \leq 0\}, \tag{2}$$

where $\tilde{h}$ and $\tilde{g}$ are continuously differentiable functions.

Definition 4. *We say that* quasi-normality *(QN) holds at $\tilde{x}^* \in S$ if there is no (λ, μ), $\mu \geq 0$, such that the following conditions hold:*

1. *$\nabla \tilde{h}(\tilde{x}^*)\lambda + \nabla \tilde{g}(\tilde{x}^*)\mu = 0$;*
2. *$(\lambda, \mu) \neq 0$;*
3. *Let $I_{\neq} = \{i \mid \lambda_i \neq 0\}$ and $J_+ = \{j \mid \mu_j > 0\}$. There exists a sequence $\{\tilde{x}^k\}$ converging to x^* such that $\lambda_i \tilde{h}_i(\tilde{x}^k) > 0$ for all $i \in I_{\neq}$ and $\tilde{g}_j(\tilde{x}^k) > 0$ for all $j \in J_+$.*

Definition 5 ([4])**.** *We say that* relaxed quasi-normality *(RQN) holds at $\tilde{x}^* \in S$ if there is no (λ, μ), $\mu \geq 0$, such that, in addition to items 1, 2 and 3 of Definition 4, it holds that $|\tilde{h}_i(\tilde{x}^k)| = o(\tilde{w}(\tilde{x}^k))$ for all $i \notin I_{\neq}$ and $\tilde{g}_j(\tilde{x}^k)_+ = o(\tilde{w}(\tilde{x}^k))$ for all $j \notin J_+$ for some $\{\tilde{x}^k\}$ converging to $\tilde{x}^*$, where*

$$\tilde{w}(\tilde{x}^k) = \min\left\{\min_{i \in I_{\neq}} |\tilde{h}_i(\tilde{x}^k)|,\ \min_{j \in J_+} \tilde{g}_j(\tilde{x}^k)_+\right\}.$$

2.1 Specialized augmented Lagrangian method for MPCCs

In contrast with standard AL methods, that penalize complementarity constraints, we consider in this section those AL methods that preserve such constraints in subproblems, notably the one proposed in [28]. We refer to this method as specialized AL. In the sequel, we describe it.

Given a penalty parameter $\rho > 0$ and *projected Lagrange multipliers* $\bar{\lambda} \in \mathbb{R}^q$ and $\bar{\mu} \in \mathbb{R}^p_+$, we define the augmented Lagrangian function

$$L_\rho(x, \bar{\lambda}, \bar{\mu}) = f(x) + \frac{\rho}{2}\left[\left\| h(x) + \frac{\bar{\lambda}}{\rho} \right\|^2 + \left\| \left(g(x) + \frac{\bar{\mu}}{\rho} \right)_+ \right\|^2 \right].$$

This is simply the quadratic penalty-like AL function without penalizing constraints within X in (MPCC). So, the subproblems are

$$\min_x \; L_\rho(x, \bar{\lambda}, \bar{\mu}) \quad \text{s.t.} \quad x \in X. \tag{3}$$

The gradient of L_ρ with respect to x is

$$\nabla_x L_\rho(x, \bar{\lambda}, \bar{\mu}) = \nabla f(x) + \nabla h(x)(\bar{\lambda} + \rho h(x)) + \nabla g(x)(\bar{\mu} + \rho g(x))_+.$$

The specialized AL method is presented in Algorithm 1. Basically, it is ALGENCAN [1] with subproblems (3): in fact, the penalty update rule and the use of bounded safeguarded Lagrange multipliers estimates are the same.

The global convergence to M-stationary points under the MPCC-relaxed constant positive linear dependence (MPCC-RCPLD) was stated in [28].

Algorithm 1 Specialized augmented Lagrangian method for MPCCs

---

Set $\tau \in [0, 1)$, $\theta > 1$, $\lambda_{\min} < \lambda_{\max}$, $\mu_{\max} > 0$ and $\rho_1 > 0$. Let the initial Lagrange multipliers estimates $\bar{\lambda}^1 \in [\lambda_{\min}, \lambda_{\max}]^q$ and $\bar{\mu}^1 \in [0, \mu_{\max}]^p$. Take $\{\varepsilon_k\} \downarrow 0$ and initialize $k \leftarrow 1$.

Step 1. (Solving the subproblem) Solve approximately (3), obtaining $x^k \in X$ satisfying

$$\pi_k \in \nabla_x L_{\rho_k}(x^k, \bar{\lambda}^k, \bar{\mu}^k) + N_{\lim}(x^k) \quad \text{and} \quad \|\pi_k\| \le \varepsilon_k$$

for some π_k.

Step 2. (Multiplier estimates) Compute $\lambda^k = \bar{\lambda}^k + \rho_k h(x^k)$ and $\mu^k = \left(\bar{\mu}^k + \rho_k g(x^k)\right)_+$.

Step 3. (Update the penalty parameter) Define $V^k = \min\{-g(x^k), \bar{\mu}^k/\rho_k\}$. If $k = 1$ or

$$\max\left\{\|h(x^k)\|_\infty, \|V^k\|_\infty\right\} \le \tau \max\left\{\|h(x^{k-1})\|_\infty, \|V^{k-1}\|_\infty\right\},$$

choose $\rho_{k+1} = \rho_k$. Otherwise, define $\rho_{k+1} = \theta\rho_k$.

Step 4. (New projected multipliers) Compute $\bar{\lambda}^{k+1} \in [\lambda_{\min}, \lambda_{\max}]^q$ and $\bar{\mu}^{k+1} \in [0, \mu_{\max}]^p$.

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

The condition required in Step 1 of Algorithm 1 is equivalent to computing an approximate M-stationary point. In fact, the notion of M-stationarity can be formally stated as follows: x^* is an M-stationarity point of (MPCC) if there are $\lambda \in \mathbb{R}^q, \mu \in \mathbb{R}^p_+$ such that

$$\mu^t g(x^*) = 0 \quad \text{and} \quad 0 \in \nabla f(x^*) + \nabla h(x^*)\lambda + \nabla g(x^*)\mu + N_{\lim}(x^*),$$

where $N_{\lim}(x^*)$ denotes the (Mordukhovich) *limiting normal cone* to $\mathbb{R}^n \times X$ at x^* and is given by

$$N_{\lim}(x^*) = \{0\}^n \times \prod_{i=1}^m \left\{ (u_i, v_i) \in \mathbb{R}^2 \;\middle|\; \begin{array}{ll} u_i \in \mathbb{R},\ v_i = 0 & \text{if } i \in I_{0+}(x^*) \\ u_i = 0,\ v_i \in \mathbb{R} & \text{if } i \in I_{+0}(x^*) \\ u_i < 0,\ v_i < 0 \text{ or } u_i v_i = 0 & \text{if } i \in I_{00}(x^*) \end{array} \right\}. \tag{4}$$

Analogously, the notion of S-stationarity can be formulated by replacing $N_{\lim}(x^*)$ with the "usual" normal cone to $\mathbb{R}^n \times X$, denoted by $N(x^*)$, and given by

$$N(x^*) = \{0\}^n \times \prod_{i=1}^{m} \left\{ (u_i, v_i) \in \mathbb{R}^2 \;\middle|\; \begin{array}{ll} u_i \in \mathbb{R},\ v_i = 0 & \text{if } i \in I_{0+}(x^*) \\ u_i = 0,\ v_i \in \mathbb{R} & \text{if } i \in I_{+0}(x^*) \\ u_i \leq 0,\ v_i \leq 0 & \text{if } i \in I_{00}(x^*) \end{array} \right\}.$$

Expression (4) is the explicit computation of

$$N_{\lim}(x^*) = \limsup_{X \ni x \to x^*} N(x) \tag{5}$$

[30, Proposition 2.1], which comes from the application of the Mordukhovich calculus [40, Definition 1.1]. Similarly, C-stationarity can be stated in the language of cones by considering "$u_i v_i \geq 0$ if $i \in I_{00}(x^*)$" in the last row of (4).

Finally, we remark that, to solve the AL subproblems in Step 1 of Algorithm 1, Guo and Deng [28] proposed an extension of the non-monotone spectral gradient projection method (originally developed in [19] for problems over convex sets) to handle problems with the simple complementarity constraints in X. This makes Step 1, and thus the Algorithm 1, feasible in practice.

3 Specialized CQs for MPCCs and error bound

We now proceed to analyze the theoretical properties of Algorithm 1 under suitable CQs.

Most CQs from the literature are designed treating all constraints together, e.g., without putting some of them into an "abstract" set. In the context of MPCCs, they are often formulated via the tightened nonlinear problem (TNLP). However, such CQs may introduce difficulties in the convergence analysis of algorithms when the complementarity constraints are treated separately, as in Algorithm 1. For instance, the CQs introduced in [33] (among them MPCC-QN, Definition 3) require a specific control over the sign of the complementarity components, but no such control is performed in Algorithm 1 as the feasibility of these constraints is achieved exactly at each iteration in subproblems. Consequently, this type of CQ does not easily fit to Algorithm 1.

We then introduce two specialized CQs derived from the enhanced Fritz John necessary optimality conditions that takes into account the abstract set X.

Theorem 1 ([14, Proposition 2.1]). *Let x^* be a local minimum of problem* (MPCC). *Then there exists $(\mu_0, \lambda, \mu) \in \mathbb{R}_+ \times \mathbb{R}^q \times \mathbb{R}^p_+$ satisfying the following conditions:*

1. $-[\mu_0 \nabla f(x^*) + \nabla h(x^*)\lambda + \nabla g(x^*)\mu] \in N_{\lim}(x^*)$;
2. $(\mu_0, \lambda, \mu) \neq 0$;
3. *Let $I_{\neq} = \{i \mid \lambda_i \neq 0\}$ and $J_+ = \{j \mid \mu_j > 0\}$. If $I_{\neq} \cup J_+ \neq \emptyset$, there exists a sequence $\{x^k\} \subset X$ that converges to x^* and is such that, for all k,*
   (a) $f(x^k) < f(x^*)$, $\lambda_i h_i(x^k) > 0$ for all $i \in I_{\neq}$, and $g_j(x^k) > 0$ for all $j \in J_+$;
   (b) $|h_i(x^k)| = o(w(x^k))$ for $i \notin I_{\neq}$ and $g_j(x^k)_+ = o(w(x^k))$ for $j \notin J_+$, where

$$w(x^k) = \min\left\{ \min_{i \in I_{\neq}} |h_i(x^k)|,\ \min_{j \in J_+} g_j(x^k)_+ \right\}. \tag{6}$$

In particular, the result above implies that a minimizer of (MPCC) is M-stationary if items 1-3 hold with $\mu_0 > 0$. This motivates the formulation of the QN condition with an abstract set like in [14], that here we call specialized MPCC-QN:

Definition 6. *We say that* specialized MPCC-quasi-normality *(sMPCC-QN) holds at a feasible point x^* if there is no $(\lambda, \mu) \in \mathbb{R}^q \times \mathbb{R}^p_+$ such that the following conditions hold:*

1. $0 \in \nabla h(x^*)\lambda + \nabla g(x^*)\mu + N_{\lim}(x^*)$;
2. $(\lambda, \mu) \neq 0$;

3. *Let* $I_{\neq} = \{i \mid \lambda_i \neq 0\}$ *and* $J_+ = \{j \mid \mu_j > 0\}$*. There exists a sequence* $\{x^k\} \subset X$ *converging to* x^* *such that* $\lambda_i h_i(x^k) > 0$ *for all* $i \in I_{\neq}$ *and* $g_j(x^k) > 0$ *for all* $j \in J_+$*.*

MPCC-QN (Definition 3) and sMPCC-QN have some relationship. In fact, in view of (4), $0 \in \nabla h(x^*)\lambda + \nabla g(x^*)\mu + N_{\lim}(x^*)$ means that $\nabla h(x^*)\lambda + \nabla g(x^*)\mu + [-\nabla c(x^*)u - \nabla d(x^*)v] = 0$ for some u, v as in item 3 of Definition 3. These conditions differ in the control of multipliers signs: while in MPCC-QN this control is also applied over complementary pairs $c_i(x^k)$, $d_i(x^k)$, in sMPCC-QN it not present since complementarity constraints are satisfied exactly ($x^k \in X$ for all k). Even so, one might think that MPCC-QN and sMPCC-QN are related to each other, but surprisingly this is not the case, as the next examples show.

Example 2 (sMPCC-QN does not imply MPCC-QN)**.** *Consider the set defined by* $g(y, z) = y^2 z$ *and the complementarity constraints* $X = \{(y, z) \mid 0 \le y \perp z \ge 0\}$*. Taking the multipliers* $\mu = 1$*,* $u = v = 0$ *and the sequence defined by* $(y^k, z^k) = (k^{-1}, k^{-1})$*, we see that MPCC-QN does not hold at the point* $(0, 0)$*.*

On the other hand, there does not exist a sequence $\{(y^k, z^k)\} \subset X$ *such that* $g(y^k, z^k) > 0$*. Therefore, the sMPCC-QN condition holds at* $(0, 0)$*.*

Example 3 (MPCC-QN does not imply sMPCC-QN)**.** *Consider the set defined by the inequality constraint* $g(y, z) = y \le 0$ *and the complementarity condition* $X = \{(y, z) \mid 0 \le y \perp z \ge 0\}$*. The point* $(0, 0)$ *satisfies MPCC-QN. In fact, to satisfy*

$$\nabla g(0)\mu - e_1 u - e_2 v = 0, \quad \mu \ge 0, \quad (\mu, u, v) \neq 0,$$

we must have $\mu = u > 0$ *and* $v = 0$*. Since no sequence* $\{(y^k, z^k)\}$ *can satisfy* $g(y^k, z^k) = y^k > 0$ *and* $-uc(y^k, z^k) = -uy^k > 0$ *simultaneously, MPCC-QN holds at* $x^* = (0, 0)$*.*

On the other hand, sMPCC-QN fails. To see this, take $\mu = 1$ *and the sequence defined by* $(y^k, z^k) = (k^{-1}, 0) \in X$*. This sequence converges to* $(0, 0)$*, items 1 and 2 of Definition 6 are satisfied, and* $\mu g(y^k, z^k) = k^{-1} > 0$ *for all* k*.*

The next result show that the sMPCC-QN condition is a genuine CQ to the MPCC with slack variables, that is, it guarantees M-stationarity at local minimizers of (MPCC). It is a consequence of Theorem 1.

Corollary 1. *If* x^* *is a minimizer of* (MPCC) *and satisfies sMPCC-QN, then* x^* *is an M-stationary point.*

Proof. Let x^* be a local minimizer of (MPCC) and suppose, for contradiction, that it is not M-stationary. By Theorem 1, the only possible case in which this occurs is when $\mu_0 = 0$. Otherwise, we could divide the expression in item 1 by μ_0, concluding that x^* is M-stationary. Therefore, there exists $(\lambda, \mu) \in \mathbb{R}^q \times \mathbb{R}^p_+$ satisfying all items of Definition 6, contradicting the MPCC-QN assumption. Thus, x^* must be M-stationary. □

Inspired by recent developments in nonlinear programming [4], where a relaxed QN condition was introduced, currently the weakest known condition that ensures the boundedness of multipliers generated by AL methods for general nonlinear problems, we propose a relaxed version of the sMPCC-QN condition introduced above. As far as we are aware, this is the first relaxed QN condition specifically designed for MPCCs.

Definition 7. *A feasible point* x^* *satisfies the* specialized MPCC relaxed quasi-normality *(sMPCC-RQN) if there is no* $(\lambda, \mu) \in \mathbb{R}^q \times \mathbb{R}^p_+$ *such that*

1. $0 \in \nabla h(x^*)\lambda + \nabla g(x^*)\mu + N_{\lim}(x^*)$*;*

2. $(\lambda, \mu) \neq 0$*;*

3. *Let* $I_{\neq} = \{i \mid \lambda_i \neq 0\}$ *and* $J_+ = \{j \mid \mu_j > 0\}$*. There exists a sequence* $\{x^k\} \subset X$ *converging to* x^* *such that*

   (a) $\lambda_i h_i(x^k) > 0$ *for all* $i \in I_{\neq}$ *and* $g_j(x^k) > 0$ *for all* $j \in J_+$*;*

(b) $|h_i(x^k)| = o(w(x^k))$ *for all* $i \notin I_{\neq}$ *and* $g_j(x^k)_+ = o(w(x^k))$ *for all* $j \notin J_+$*, where* $w(x^k)$ *is as in* (6)*.*

It is immediate that the sMPCC-QN condition implies its relaxed counterpart. Moreover, similarly to Corollary 1, by applying Theorem 1 it is straightforward to show that sMPCC-RQN is also a CQ for M-stationarity. Although relatively weak, they retain enough structure to ensure key algorithmic properties, most notably, the boundedness of the dual sequence, as we will see in the next section.

Corollary 2. *If* x^* *is a minimizer of* (MPCC) *and satisfies sMPCC-RQN, then* x^* *is an M-stationary point.*

3.1 Error bound

We say that the set S (see (2)) satisfies a (local) error bound property at $\tilde{x}^*$ if there are constants $\delta > 0$ and $M > 0$ such that

$$d(\tilde{x}, S) \leq M \max\{\|\tilde{h}(\tilde{x})\|_\infty, \|\tilde{g}(\tilde{x})_+\|_\infty\} \quad \text{for all} \quad \tilde{x} \in B(\tilde{x}^*, \delta), \tag{7}$$

where

$$d(\tilde{x}, S) = \min_{s \in S} \|s - \tilde{x}\|$$

denotes the Euclidean distance between x and S. Error bound is a well known property linked to the stability of algorithms; it measures how much we are far from the feasible set w.r.t. the violation of the constraints.

In view of Algorithm 1, it becomes natural to consider an error bound for (MPCC) restricted to X since all iterates generated by the method satisfy the complementarity constraints.

Definition 8. *We say that* Ω *satisfies the* specialized MPCC-error bound *at* x^* *(sMPCC-EB) if there are constants* $\delta > 0$ *and* $M > 0$ *such that*

$$d(x, \Omega) \leq M \max\{\|h(x)\|_\infty, \|g(x)_+\|_\infty\} \quad \textit{for all} \quad x \in B(x^*, \delta) \cap X.$$

In standard nonlinear programming, it is known that both QN and RQN CQs (Definitions 4 and 5) imply error bound; see [38] and [3], respectively. We take advantage of these results to establish similar implications for sMPCC-(R)QN. Our strategy is to view the complementarity constraints in X as a finite union of faces, denoted by $X(J)$, which allow us to apply existing error bound results to each face separately. The argument therefore starts by writing X as

$$X = \bigcup_{J \subset \{1,\dots,m\}} X(J), \quad X(J) = \left\{ (w, y, z) \in \mathbb{R}^{n+2m} \;\middle|\; \begin{array}{ll} y_i = 0 \;\forall i \in J, & y_i \geq 0 \;\forall i \notin J \\ z_i \geq 0 \;\forall i \in J, & z_i = 0 \;\forall i \notin J \end{array} \right\}$$

(remember that $x = (w, y, z)$). Here, we frequently need to work with x and w, y, z separately. Remembering that the functions in constraints are translated from w, y, z to x by writing

$$h(x) = (h(w), y - H(w), z - G(w)), \quad g(x) = g(w),$$

we will refer to the functions $h(x)$, $g(x)$ associated with (MPCC) and the functions $h(w)$, $g(w)$, $H(w)$, $G(w)$ associated with (1) without further explanations to the rest of this section.

Given $x \in X$, we consider the set of all index subsets whose corresponding face of X contains the point x, namely,

$$\mathcal{X}(x) = \{J \mid x \in X(J)\}.$$

To address the usual error bound on each face, we partition the feasible set accordingly, but eliminating fixed variables at zero and inequalities:

$$\Omega(J) = \left\{ (w, y, z) \in \mathbb{R}^{n+2m} \;\middle|\; \begin{array}{ll} h(w) = 0, & g(w) \leq 0 \\ y_i^2 - G_i(w) = 0, & i \notin J \\ z_i^2 - H_i(w) = 0, & i \in J \end{array} \right\}.$$

Here, the error bound on $\Omega(J)$ refers only to the constraints $h(x) = 0$ and $g(x) \leq 0$, but with quadratic slacks variables.

Lemma 1. *Ω satisfies sMPCC-EB at x^* whenever $\Omega(J)$ satisfies the usual error bound property at x^* for all $J \in \mathcal{X}(x^*)$, that is, for each J with $x^* \in X(J)$, there are $\delta_J > 0$ and $M_J > 0$ such that*

$$\begin{aligned} &d(x, \Omega(J)) \\ &\leq M_J \max\{\|h(w)\|_\infty, \|g(w)_+\|_\infty, \|y_{\complement J} * y_{\complement J} - G_{\complement J}(w)\|_\infty, \|z_J * z_J - H_J(w)\|_\infty\} \end{aligned}$$

for all x with $\|x - x^*\| \leq \delta_J$, where $q * q$ denotes the vector whose components are q_i^2 and $\complement J = \{1, \ldots, m\} \backslash J$.

Proof. The proof proceeds by contradiction: if sMPCC-EB fails at x^*, there is a sequence $\{x^k = (w^k, y^k, z^k)\} \subset X$ converging to x^* such that

$$d(x^k, \Omega) > k \max\{\|h(x^k)\|_\infty, \|g(x^k)_+\|_\infty\}$$

for all k. Passing to a subsequence if necessary, we can suppose that $\{x^k\} \subset X(J)$ for some J. Thus, $y^k \geq 0$ and $z^k \geq 0$ for all k and we can define $\tilde{y}_i^k = \sqrt{y_i^k}$ and $\tilde{z}_i^k = \sqrt{z_i^k}$ for all i, k.

Now, consider $(\bar{w}^k, \bar{y}^k, \bar{z}^k) \in \Omega(J)$ such that

$$d((w^k, \tilde{y}^k, \tilde{z}^k), \Omega(J)) = \|(w^k, \tilde{y}^k, \tilde{z}^k) - (\bar{w}^k, \bar{y}^k, \bar{z}^k)\|.$$

Since $\{\bar{y}^k + \tilde{y}^k, \bar{z}^k + \tilde{z}^k)\}$ is convergent, there exists $C \geq 1$ such that

$$C \geq \max\{\|\bar{y}^k + \tilde{y}^k\|_\infty^2, \|\bar{z}^k + \tilde{z}^k\|_\infty^2\}$$

for all k. Then,

$$\begin{aligned} d(x^k, \Omega)^2 &\leq \|(\bar{w}^k, \bar{y}^k * \bar{y}^k, \bar{z}^k * \bar{z}^k) - (w^k, y^k, z^k)\|^2 \\ &= \|\bar{w}^k - w^k\|^2 + \sum_{i=1}^m ((\bar{y}_i^k)^2 - y_i^k)^2 + \sum_{i=1}^m ((\bar{z}_i^k)^2 - z_i^k)^2 \\ &= \|\bar{w}^k - w^k\|^2 + \sum_{i=1}^m ((\bar{y}_i^k)^2 - (\tilde{y}_i^k)^2)^2 + \sum_{i=1}^m ((\bar{z}_i^k)^2 - (\tilde{z}_i^k)^2)^2 \\ &= \|\bar{w}^k - w^k\|^2 + \sum_{i=1}^m \left(\bar{y}_i^k - \tilde{y}_i^k\right)^2 \left(\bar{y}_i^k + \tilde{y}_i^k\right)^2 + \sum_{i=1}^m \left(\bar{z}_i^k - \tilde{z}_i^k\right)^2 \left(\bar{z}_i^k + \tilde{z}_i^k\right)^2 \\ &\leq mC\|\bar{w}^k - w^k\|^2 + mC \sum_{i=1}^m \left(\bar{y}_i^k - \tilde{y}_i^k\right)^2 + mC \sum_{i=1}^m \left(\bar{z}_i^k - \tilde{z}_i^k\right)^2 \\ &\leq mC\|(w^k, \tilde{y}^k, \tilde{z}^k) - (\bar{w}^k, \bar{y}^k, \bar{z}^k)\|^2 = mCd((w^k, \tilde{y}^k, \tilde{z}^k), \Omega(J))^2. \end{aligned}$$

This implies

$$\begin{aligned} &d((w^k, \tilde{y}^k, \tilde{z}^k), \Omega(J)) \geq (mC)^{-1/2} d(x^k, \Omega) > k(mC)^{-1/2} \max\{\|h(x^k)\|_\infty, \|g(x^k)_+\|_\infty\} \\ &= k(mC)^{-1/2} \max\{\|h(w^k)\|_\infty, \|g(w^k)_+\|_\infty, \|\tilde{y}_{\complement J}^k * \tilde{y}_{\complement J}^k - G_{\complement J}(w^k)\|_\infty, \\ &\qquad\qquad \|\tilde{z}_J^k * \tilde{z}_J^k - H_J(w^k)\|_\infty\}, \end{aligned}$$

which completes the proof. □

The lemma above shows that verifying the error bound on all active faces suffices to guarantee sMPCC-EB. In what follows, we prove that, under sMPCC-(R)QN conditions, each $\Omega(J)$ satisfies the standard (R)QN; as a consequence, the error bound property holds.

Theorem 2. *sMPCC-QN at x^* implies QN at x^* regarding each $\Omega(J)$, $J \in \mathcal{X}(x^*)$. The same is valid for sMPCC-RQN and RQN.*

Proof. We prove the statement only for sMPCC-RQN and RQN since the constructed sequences also serves for sMPCC-QN and QN.

Suppose that there is $J \in \mathcal{X}(x^*)$ such that RQN fails at $x^* = (w^*, y^*, z^*)$ w.r.t. $\Omega(J)$. Then, there are $(\lambda, \mu) \neq 0$, $\lambda = (\lambda^h, \lambda^G, \lambda^H)$, $\mu \geq 0$, and $\{x^k = (w^k, y^k, z^k)\}$ converging to x^* satisfying, for all k,

$$\begin{bmatrix} \nabla h(w^*) \\ 0 \\ 0 \end{bmatrix} \lambda^h + \begin{bmatrix} \nabla g(w^*) \\ 0 \\ 0 \end{bmatrix} \mu + \sum_{i \in \complement J} \begin{bmatrix} -\nabla G_i(w^*) \\ 2y_i^* e_i \\ 0 \end{bmatrix} \lambda_i^G + \sum_{i \in J} \begin{bmatrix} -\nabla H_i(w^*) \\ 0 \\ 2z_i^* e_i \end{bmatrix} \lambda_i^H = 0$$

and

$$\lambda_i^h h_i(w^k) > 0 \text{ if } \lambda_i^h \neq 0, \quad \lambda_i^G((y_i^k)^2 - G_i(w^k)) > 0 \text{ if } \lambda_i^G \neq 0$$
$$\text{and} \quad \lambda_i^H((z_i^k)^2 - H_i(w^k)) > 0 \text{ if } \lambda_i^H \neq 0.$$

Defining $\tilde{x}^k \in \mathbb{R}^{n+2m}$ as $\tilde{x}^k = (w^k, \tilde{y}^k, \tilde{z}^k)$ where

$$\tilde{y}_i^k = \begin{cases} (y_i^k)^2, & i \notin J \\ 0, & \text{otherwise} \end{cases}, \qquad \tilde{z}_i^k = \begin{cases} (z_i^k)^2, & i \in J \\ 0, & \text{otherwise} \end{cases}$$

we can write simply $\lambda_i h_i(\tilde{x}^k) > 0$ for all $i \in I_{\neq}$.

With this notation, we have additionally that $g_j(\tilde{x}^k) > 0$ for all $j \in J_+$ and $|h_i(\tilde{x}^k)| = o(w(\tilde{x}^k))$ for all $i \notin I_{\neq}$ and $g_j(\tilde{x}^k)_+ = o(w(\tilde{x}^k))$ for all $j \notin J_+$, where $w(\tilde{x}^k)$ is as in (6). Also, note that $\tilde{x}^k \in X$ for all k and that, considering $\lambda_i^G = 0$ for $i \in J$ and $\lambda_i^H = 0$ for $i \notin J$, the first item can be written as

$$\begin{bmatrix} \nabla h(w^*) & -\nabla H(w^*) & -\nabla G(w^*) \\ 0 & I_m & 0 \\ 0 & 0 & I_m \end{bmatrix} \lambda + \begin{bmatrix} \nabla g(w^*) \\ 0 \\ 0 \end{bmatrix} \mu$$
$$+ \sum_{i \in \complement J} \begin{bmatrix} 0 \\ (2y_i^* - 1)e_i \\ 0 \end{bmatrix} \lambda_i^G + \sum_{i \in J} \begin{bmatrix} 0 \\ 0 \\ (2z_i^* - 1)e_i \end{bmatrix} \lambda_i^H = 0.$$

It is straightforward to verify that the two sums in the above equation is in $N_{\lim}(x^*)$. In other words, we have $0 \in \nabla h(x^*)\lambda + \nabla g(x^*) + N_{\lim}(x^*)$. Thus, all items of Definition 7 are fulfilled for x^* with (λ, μ) and $\{\tilde{x}^k\}$. This concludes the proof. □

Corollary 3. *Suppose that $\nabla h_i(x)$ and $\nabla g_j(x)$ are locally Lipschitz continuous around x^* for all i, j. If sMPCC-(R)QN holds at x^*, then Ω satisfies sMPCC-EB at x^*.*

Proof. It suffices to prove the statement for sMPCC-RQN, since sMPCC-QN implies it. sMPCC-RQN at x^* implies RQN at x^* w.r.t. $\Omega(J)$ for all $J \in \mathcal{X}(x^*)$ by the previous theorem, which in turn ensures the standard error bound property regarding $\Omega(J)$ [3, Corollary 4.1] for all $J \in \mathcal{X}(x^*)$. This, in turn, implies the validity of sMPCC-EB at x^* by Lemma 1. □

So far, our discussion has focused on the slack-variable reformulation (1). It is worth noting that an alternative error bound notion for MPCCs, directly stated in terms of the original variables, has proposed in the literature [24], which we refer here as MPCC-EB. The authors deal directly with the original constraints in (1) by shifting them:

$$\Omega(a, b, c, d) = \left\{ w \in \mathbb{R}^n \;\middle|\; \begin{array}{l} h(w) + a = 0, \quad g(w) + b \leq 0, \\ H(w) + c \geq 0, \quad G(w) + d \geq 0, \\ (H(w) + c)^t (G(w) + d) = 0 \end{array} \right\}.$$

We say that MPCC-EB is satisfied at w^* if there are $\delta > 0$ and $M > 0$ such that

$$d(w, \Omega(0, 0, 0, 0)) \leq M \|(a, b, c, d)\| \tag{8}$$

for all (a, b, c, d) and $w \in \Omega(a, b, c, d)$ with $(a, b, c, d) \in B(0, \delta)$ and $w \in B(w^*, \delta)$. We prove next that the shifts a, b, c, d work in some sense as the slacks in (MPCC), and the following theorem makes this connection precise by showing that both notions are equivalent.

Theorem 3. *Let w^* be a feasible point for (1) and $x^* = (w^*, y, z)$ feasible for (MPCC), where (y, z) is arbitrary. Then MPCC-EB holds at w^* if, and only if, Ω satisfies sMPCC-EB at x^*.*

Proof. As H and G are continuously differentiable, they are locally Lipschitz continuous around w^*. Therefore, there exist constants $L_H, L_G > 0$, $\delta_L > 0$ such that

$$\|H(w) - H(\tilde{w})\| \leq L_H \|w - \tilde{w}\| \quad \text{and} \quad \|G(w) - G(\tilde{w})\| \leq L_G \|w - \tilde{w}\| \tag{9}$$

for all $w, \tilde{w} \in B(w^*, \delta_L)$.

Suppose that MPCC-EB holds at w^*. Then there exist $\delta > 0$ and $M \geq 1$ such that (8) is satisfied for each $(a', b', c', d') \in B(0, \delta)$ and $w' \in \Omega(a', b', c', d') \cap B(w^*, \delta)$. By continuity of the constraint functions, we can obtain $\delta_\Omega > 0$ such that for all $x = (w, y, z) \in B(x^*, \delta_\Omega) \cap X$, it holds

$$\|(h(w),\ g(w)_+,\ y - H(w),\ z - G(w))\| \leq \delta.$$

Define $\bar{\delta} := \dfrac{1}{2} \min\{\delta, \delta_\Omega, \delta_L\}$.

Fix any $x = (w, y, z) \in B(x^*, \bar{\delta}) \cap X$. We have $\|w - w^*\| \leq \bar{\delta} \leq \delta$ and

$$(a, b, c, d) := (-h(w),\ -g(w)_+,\ y - H(w),\ z - G(w)) \in B(0, \delta).$$

It is easy to see that $w \in \Omega(a, b, c, d)$, and so

$$d(w, \Omega(0, 0, 0, 0)) \leq M \|(a, b, c, d)\|.$$

Let $\tilde{w} \in \Omega(0, 0, 0, 0)$ be such that $d(w, \Omega(0, 0, 0, 0)) = \|w - \tilde{w}\|$. Since $w^* \in \Omega(0, 0, 0, 0)$, we have

$$\|w^* - \tilde{w}\| \leq \|w^* - w\| + \|w - \tilde{w}\| \leq 2\|w^* - w\| \leq \delta_L,$$

thus, $\tilde{w} \in B(w^*, \delta)$. Then, using (9), we obtain

$$\begin{aligned}
d(x, \Omega) &\leq \|w - \tilde{w}\| + \|y - H(\tilde{w})\| + \|z - G(\tilde{w})\| \\
&\leq \|w - \tilde{w}\| + (\|y - H(w)\| + \|H(w) - H(\tilde{w})\|) \\
&\qquad\qquad + (\|z - G(w)\| + \|G(w) - G(\tilde{w})\|) \\
&\leq (\|w - \tilde{w}\| + \|y - H(w)\| + \|z - G(w)\|) + L_H \|w - \tilde{w}\| + L_G \|w - \tilde{w}\| \\
&= (1 + L_H + L_G)\, d(w, \Omega(0, 0, 0, 0)) + \|y - H(w)\| + \|z - G(w)\| \\
&\leq M(1 + L_H + L_G)\, \|(h(w), g(w)_+, y - H(w), z - G(w))\| \\
&\qquad\qquad + \|y - H(w)\| + \|z - G(w)\| \\
&\leq M_\Omega \max\{\|h(x)\|_\infty, \|g(x)_+\|_\infty\}
\end{aligned}$$

where $M_\Omega := nM(1 + L_H + L_G) > 1$. As x is arbitrary, this shows that sMPCC-EB holds at x^*.

For the converse, assume that sMPCC-EB holds at $x^* = (w^*, y^*, z^*)$ with constants δ_Ω and M as in Definition 8, that is,

$$d(x, \Omega) \leq M \max\{\|h(x)\|_\infty, \|g(x)_+\|_\infty\} \quad \text{for all} \quad x \in B(x^*, \delta_\Omega) \cap X. \tag{10}$$

We aim to show that MPCC-EB also holds at w^*.

Take

$$\delta = \min\left\{\frac{\delta_\Omega}{2},\ \frac{\delta_\Omega}{2(1 + L_H + L_G)},\ \delta_L\right\},$$

where the constants δ_L, L_H and L_G are given by (9), and consider any $(a, b, c, d) \in B(0, \delta)$, $w \in \Omega(a, b, c, d) \cap B(w^*, \delta)$. Define $y := H(w) + c$ and $z := G(w) + d$, so that $x = (w, y, z) \in X$. Let $\tilde{x} = (\tilde{w}, \tilde{y}, \tilde{z}) \in \Omega$ be such that $d(x, \Omega) = \|x - \tilde{x}\|$.

Let us check that x is sufficiently close to x^*. Since $\|(a, b, c, d)\| \leq \delta$, $\|w - w^*\| \leq \delta$, $y^* = H(w^*)$ and $z^* = G(w^*)$, we have

$$\begin{aligned}
\|x - x^*\| &\leq \|w - w^*\| + \|H(w) + c - y^*\| + \|G(w) + d - z^*\| \\
&= \|w - w^*\| + \|(H(w) - H(w^*)) + c\| + \|(G(w) - G(w^*)) + d\| \\
&\leq \|w - w^*\| + \|H(w) - H(w^*)\| + \|G(w) - G(w^*)\| + \|(a, b, c, d)\| \\
&\leq (1 + L_H + L_G)\|w - w^*\| + \|(a, b, c, d)\| \\
&\leq (1 + L_H + L_G)\delta + \delta \leq \delta_\Omega.
\end{aligned}$$

Thus $x \in B(x^*, \delta_\Omega)$, allowing us to use (10). Moreover, $\tilde{w} \in \Omega(0,0,0,0)$ because $\tilde{x} \in \Omega$. Therefore

$$\begin{aligned} d(w, \Omega(0,0,0,0))^2 &\leq \|w - \tilde{w}\|^2 \leq \|w - \tilde{w}\|^2 + \|y - \tilde{y}\|^2 + \|z - \tilde{z}\|^2 \\ &= \|x - \tilde{x}\|^2 = d(x, \Omega)^2 \\ &\leq M^2 \max\{\|h(x)\|_\infty, \|g(x)_+\|_\infty\}^2 \\ &= M^2 \max\{\|h(w)\|_\infty, \|g(w)_+\|_\infty, \|y - H(w)\|_\infty, \|z - G(w)\|_\infty\}^2 \\ &\leq M^2 \|(a,b,c,d)\|_\infty^2. \end{aligned} \tag{11}$$

As (a,b,c,d) and w are arbitrary, we conclude that MPCC-EB is satisfied at w^*. □

Having established the error bound property for the MPCC framework, we now turn our attention to the boundedness of multipliers in the specialized AL method.

4 Boundedness of multipliers in the specialized AL method

The following theorem presents the main result of this paper: sMPCC-RQN is sufficient to ensure the boundedness of the dual sequence generated by Algorithm 1, without requiring any additional assumptions.

Theorem 4. *Let $\{x^k\}$ be a sequence generated by Algorithm 1 and x^* a feasible accumulation point of it, let us say, $\lim_{k \in K} x^k = x^*$. If x^* satisfies sMPCC-RQN (or sMPCC-QN) then the associated dual subsequences $\{\lambda^k\}_{k\in K}$ and $\{\mu^k\}_{k\in K}$ given in Step 2 are bounded. In particular, x^* is an M-stationary point.*

Proof. The proof is similar to that of [4, Theorem 3]. Suppose that $\{\delta_k := \|(\lambda^k, \mu^k)\|_\infty\}_{k\in K}$ is unbounded. It follows from Step 1 of Algorithm 1 that

$$\pi_k \in \nabla f(x) + \nabla h(x)\lambda^k + \nabla g(x)\mu^k + N_{\lim}(x^k) \quad \text{and} \quad \|\pi_k\| \leq \varepsilon_k$$

for all $k \in K$. Dividing this expression by δ_k, using the fact that $N_{\lim}(\cdot)$ is outer-semicontinuous everywhere and taking the limit over K, we obtain $(\lambda, \mu) \neq 0$ such that $\mu \geq 0$ and

$$0 \in \nabla h(x^*)\lambda + \nabla g(x^*)\mu + N_{\lim}(x^*),$$

where

$$0 \neq (\lambda, \mu) = \lim_{k\in K} \frac{(\lambda^k, \mu^k)}{\delta_k} = \lim_{k \in K} \frac{(\bar{\lambda}^k + \rho_k h(x^k), [\bar{\mu}^k + \rho_k g(x^k)]_+)}{\delta_k}.$$

Let us now verify condition 3(a) of the RQN definition for (λ, μ). First, observe that since $\lim_{k\in K} \delta_k = \infty$ and the sequences $\{\bar{\lambda}^k\}_{k\in K}$ and $\{\bar{\mu}^k\}_{k\in K}$ are bounded (by Step 3 of Algorithm 1), if follows that

$$\lim_{k\in K_1} |\lambda_i^k| = \lim_{k \in K_1} |\bar{\lambda}_i^k + \rho_k h_i(x^k)| = \infty$$

for some $K_1 \subseteq K$ whenever $\lambda_i \neq 0$. In particular, $h_i(x^k) \neq 0$ for all $k \in K_1$ large enough and $\rho_k |h_i(x^k)| \to \infty$, because $\{\bar{\lambda}_i^k\}_{k\in K_1}$ is bounded. So,

$$\lambda_i^k h_i(x^k) = (\bar{\lambda}_i^k + \rho_k h_i(x^k)) h_i(x^k) > 0.$$

for all $k \in K_1$ large enough. In addition, we may assume that $\lambda_i^k \lambda_i \geq 0$ for all i and $k \in K_1$ large enough. Consequently, for all sufficiently large $k \in K_1$, we also have $\lambda_i h_i(x^k) > 0$. The same reasoning is valid for $\mu_j > 0$. Hence, applying this argument consecutively, we can assume that for all $k \in K$,

$$\lambda_i h_i(x^k) > 0, \; i \in I_{\neq}, \quad g_j(x^k) > 0, \; j \in J_+, \tag{12}$$

where $I_{\neq}$ and J_+ are as in item 3(a) of Definition 3.

Finally, observe that if the set $\{i \mid i \notin I_{\neq} \cup J_+\}$ is empty, then we conclude that x^* cannot satisfy sMPCC-RQN, yielding a contradiction. Therefore, assume that some index l is such that $\lambda_l = 0$. Take $i \in I_{\neq}$ and consider the sequence $\{A_k\}$ defined by

$$A_k = \frac{|\bar{\lambda}_l^k + \rho_k h_l(x^k)|}{|\rho_k h_i(x^k)|} = \left| \frac{\bar{\lambda}_l^k}{\rho_k h_i(x^k)} + \frac{h_l(x^k)}{h_i(x^k)} \right|,$$

which is well-defined in view of (12). If $A_k \geq \varepsilon > 0$ for all $k \in K$, we have $|\bar{\lambda}_l^k + \rho_k h_l(x^k)| \geq \varepsilon|\rho_k h_i(x^k)|$ and therefore

$$0 < \varepsilon|\lambda_i| = \lim_{k\in K} \varepsilon\frac{|\bar{\lambda}_i^k + \rho_k h_i(x^k)|}{\delta_k} = \lim_{k\in K} \varepsilon\frac{|\rho_k h_i(x^k)|}{\delta_k} \leq \lim_{k\in K} \frac{|\bar{\lambda}_l^k + \rho_k h_l(x^k)|}{\delta_k} = |\lambda_l| = 0,$$

which is a contradiction. Thus, $\liminf_{k\in K} A_k = 0$. Consequently, passing to a subsequence if necessary, we obtain

$$\lim_{k\in K} \frac{|h_l(x^k)|}{|h_i(x^k)|} = \lim_{k\in K} A_k = 0.$$

An analogous argument is valid for the other combinations:

- l such that $\lambda_l = 0$, i such that $\mu_i > 0$;
- l such that $\mu_l = 0$, i such that $\lambda_i \neq 0$;
- l such that $\mu_l = 0$, i such that $\mu_i > 0$.

By applying the same reasoning successively and passing to subsequences if necessary, we obtain

$$\lim_{k\in K} \frac{|h_l(x^k)|}{w(x^k)} = 0, \; l \notin I_{\neq}, \quad \text{and} \quad \lim_{k\in K} \frac{g_l(x^k)_+}{w(x^k)} = 0, \; l \notin J_+,$$

where $w(x^k)$ is as in (6). We have thus constructed (λ, μ), $\mu \geq 0$, and sequences satisfying the conditions in Definition 7, contradicting the validity of sMPCC-RQN at x^*. We then conclude that $\{\|(\lambda^k, \mu^k)\|_\infty\}_{k\in K}$ is bounded as we wanted. □

To the best of our knowledge, the above theorem is the first result establishing multiplier boundedness for an AL method applied to MPCCs. This contrasts with the direct application of standard AL methods, like ALGENCAN, to MPCCs, where boundedness of multipliers is not guaranteed even under MPCC-LICQ (Example 1).

We proceed by proving that, while MPCC-QN and the sMPCC-QN conditions are not directly comparable (Examples 2 and 3), MPCC-QN is enough to guarantee the boundedness of multiplies sequence generated by Algorithm 1. This results does not conflict with Theorem 4; on the contrary, it expands the cases where the multipliers in Algorithm 1 are bounded, connecting the theory with a well-known QN-type CQ from the literature. Before, we need the following technical result:

Lemma 2. *For all $\bar{x} \in X$ and $\bar{\omega} \in N(\bar{x})$, there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{u^k\} \subset \mathbb{R}^m$ and $\{v^k\} \subset \mathbb{R}^m$, such that*

- *$\{x^k\}$ converges to $\bar{x}$ and $\{\omega^k := -\nabla c(x^k)u^k - \nabla d(x^k)v^k\}$ converges to $\bar{\omega}$;*
- *$u_i^k = -kc_i(x^k)$ for all $i \in I_{0+}(\bar{x}) \cup I_{00}(\bar{x})$ and k;*
- *$v_i^k = -kd_i(x^k)$ for all $i \in I_{+0}(\bar{x}) \cup I_{00}(\bar{x})$ and k;*
- *$u_i^k = k[-c_i(x^k)]_+$ for all $i \in I_{+0}(\bar{x})$ and k;*
- *$v_i^k = k[-d_i(x^k)]_+$ for all $i \in I_{0+}(\bar{x})$ and k.*

Proof. Let $\bar{x} \in X$ and $\omega \in N(\bar{x})$. Let us consider the set obtained from TNLP at $\bar{x}$ by removing the constraints $h(x) = 0$ and $g(x) \leq 0$:

$$\begin{aligned} \Omega' = \{x \in \mathbb{R}^{n+2m} \mid\ & c_i(x) = 0 \quad \forall i \in I_{0+}(\bar{x}) \cup I_{00}(\bar{x}), \\ & d_i(x) = 0 \quad \forall i \in I_{+0}(\bar{x}) \cup I_{00}(\bar{x}), \\ & c_i(x) \geq 0 \quad \forall i \in I_{+0}(\bar{x}), \\ & d_i(x) \geq 0 \quad \forall i \in I_{0+}(\bar{x})\}. \end{aligned}$$

Remembering that $c(x) = y$ and $d(x) = z$, we immediately have $\Omega' \subset X$. So, the normal cone $N_{\Omega'}(\bar{x})$ to Ω' at $\bar{x}$ contains $N(\bar{x})$, and thus $\omega \in N_{\Omega'}(\bar{x})$. As $c(x)$ and $d(x)$ are linear, it is clear that the Guignard's CQ is valid at $\bar{x}$ w.r.t. Ω', that is, the polar of the tangent cone of Ω' at $\bar{x}$ is equal to $N_{\Omega'}(\bar{x})$. Thus, the statement follows from [7, Lemma 4.3] applied to Ω'. □

Theorem 5. *Let $\{x^k\}$ be a sequence generated by Algorithm 1 and x^* a feasible accumulation point of it, let us say, $\lim_{k\in K} x^k = x^*$. If x^* satisfies MPCC-QN (Definition 3) then the associated dual subsequences $\{\lambda^k\}_{k\in K}$ and $\{\mu^k\}_{k\in K}$ given in Step 2 are bounded. In particular, x^* is an M-stationary point.*

Proof. Proceeding as in the proof of Theorem 4, we assume by contradiction that $\{\delta^k := \|(\lambda^k, \mu^k)\|_\infty\}_{k\in K}$ is unbounded, obtaining a sequence $\{x^k\}_{k\in K} \subset X$ and non-zero multipliers (λ, μ) such that (12) holds.

From Step 1 of Algorithm 1, we have

$$\omega^k := -\big[\nabla f(x^k) + \nabla h(x^k)\lambda^k + \nabla g(x^k)\mu^k\big] + \pi_k \in N_{\lim}(x^k) \tag{13}$$

for each $k \in K$, where $\lim_{k\in K} \pi_k = 0$. From (5), we can write

$$N_{\lim}(x) = \big\{\omega \in \mathbb{R}^{n+2m} \mid \exists\{x^j\} \subset X,\ \{\omega^j\} \subset N(x^j) \text{ such that } x^j \to x,\ \omega^j \to \omega\big\}.$$

So, for each $k \in K$ and regarding to $\omega^k \in N_{\lim}(x^k)$, we can find sequences $\{x^{k,j}\} \subset X$ and $\{\omega^{k,j}\} \subset N(x^{k,j})$ with

$$x^{k,j} \xrightarrow{j} x^k \quad \text{and} \quad \omega^{k,j} \xrightarrow{j} \omega^k. \tag{14}$$

Now, we proceed to construct multipliers and sequences suitable for Definition 3. Using Lemma 2 on each pair $(x^{k,j}, \omega^{k,j})$, we obtain sequences $\{x^{k,j,l}\} \subset \mathbb{R}^{n+2m}$, $\{u^{k,j,l}\} \subset \mathbb{R}^m$ and $\{v^{k,j,l}\} \subset \mathbb{R}^m$ such that

$$x^{k,j,l} \xrightarrow{l} x^{k,j}, \quad \omega^{k,j,l} \xrightarrow{l} \omega^{k,j}, \quad \omega^{k,j,l} := -\nabla c(x^{k,j,l})u^{k,j,l} - \nabla d(x^{k,j,l})v^{k,j,l} \tag{15}$$

where $u^{k,j,l}$ and $v^{k,j,l}$ are given as Lemma 2. By (14), (15) and $\lim_{k\in K} x^k = x^*$, we can choose, for each $k \in K$, indices j_k, l_k large enough satisfying $\|x^{k,j_k,l_k} - x^*\| \le 1/k$ and $\|\omega^{k,j_k,l_k} - \omega^k\| \le 1/k$ for all $k \in K$. Therefore, defining $\tilde{x}^k := x^{k,j_k,l_k}$, $\tilde{\omega}^k := \omega^{k,j_k,l_k}$, $\tilde{u}^k := u^{k,j_k,l_k}$ and $\tilde{v}^k := v^{k,j_k,l_k}$ for each $k \in K$, we have

- $\lim_{k\in K} \tilde{x}^k = x^*$, $\quad \lim_{k\in K}(\tilde{\omega}^k - \omega^k) = 0$,
- $\lambda_i^k h_i(\tilde{x}^k) > 0$ whenever $\lambda_i^k h_i(x^k) > 0$, $k \in K$,
- $g_i(\tilde{x}^k) > 0$ whenever $g_i(x^k) > 0$, $k \in K$.

In view of the last expression in (15) and $\lim_{k\in K} \pi_k = 0$, (13) gives

$$\begin{aligned}
&\lim_{k\in K} \nabla f(x^k) + \nabla h(x^k)\lambda^k + \nabla g(x^k)\mu^k - \nabla c(\tilde{x}^k)\tilde{u}^k - \nabla d(\tilde{x}^k)\tilde{v}^k = \\
&\lim_{k\in K} \big[\nabla f(x^k) + \nabla h(x^k)\lambda^k + \nabla g(x^k)\mu^k - \pi_k\big] + \big[-\nabla c(\tilde{x}^k)\tilde{u}^k - \nabla d(\tilde{x}^k)\tilde{v}^k\big] = \\
&\lim_{k\in K} -\omega^k + \tilde{\omega}^k = 0.
\end{aligned}$$

Let $\{\tilde{\delta}_k := \|(\lambda^k, \mu^k, \tilde{u}^k, \tilde{v}^k)\|_\infty\}_{k\in K}$. We are supposing that $\{(\lambda^k, \mu^k)\}_{k\in K}$ is unbounded, so $\lim_{k\in K} \tilde{\delta}_k = \infty$. Dividing the previous expression by $\tilde{\delta}_k$ and taking the limit over K we obtain a non-zero vector (λ, μ, u, v) satisfying

$$\nabla h(x^*)\lambda + \nabla g(x^*)\mu - \nabla c(x^*)u - \nabla d(x^*)v = 0.$$

Moreover, the boundedness of $\{\omega^k/\tilde{\delta}^k\}$ allow us to write, passing to a subsequence if necessary, that $\omega = \lim_{k\in K} \omega^k/\tilde{\delta}_k$. Since $N_{\lim}(\cdot)$ is a cone and outer semi-continuos everywhere and $\omega^k \in N_{\lim}(x^k)$ for all $k \in K$, we have $\omega \in N_{\lim}(x^*)$. In addition, because of $\tilde{\omega}^k - \omega^k \to 0$, we have

$$\omega = \lim_{k\in K} \frac{\omega^k}{\tilde{\delta}^k} = \lim_{k\in K} \frac{\tilde{\omega}^k}{\tilde{\delta}^k} = -\nabla c(x^*)u - \nabla d(x^*)v,$$

and thus $-(u, v) \in N_{\lim}(x^*)$. In other words, we prove that items 1, 2 and 3 of Definition 3 are satisfied with (λ, μ, u, v).

Adapting the argument of Theorem 4, for large $k \in K$ we have $\lambda_i h_i(\tilde{x}^k) > 0$ for all $i \in I_{\neq}$ and $g_j(\tilde{x}^k) > 0$ for all $j \in J_+$. Moreover, $i \in I_u$ implies that $i \in I_{0+}(x^*) \cup I_{00}(x^*)$ since, as we already proved, $u_r = 0$ for all $r \in I_{+0}(x^*)$. So, the construction of $\tilde{u}$ from Lemma 2 guarantees that $\tilde{u}_i = -kc_i(\tilde{x}^k)$ for all $i \in I_u$. By the continuity of c_i, we can suppose passing to a subsequence if necessary that the sign of $c_i(\tilde{x}^k)$ does not change for all $k \in K$ large enough. Then,

$$-u_i c_i(\tilde{x}^k) = -\left(\lim_{k\in K} \frac{\tilde{u}_i^k}{\delta_k}\right) c_i(\tilde{x}^k) = -\left(\lim_{k\in K} \frac{-kc_i(\tilde{x}^k)}{\delta_k}\right) c_i(\tilde{x}^k) > 0.$$

Similarly, $-v_i d_i(\tilde{x}^k) > 0$ for all $i \in I_v$ and $k \in K$. That is, item 4 of Definition 3 holds with $\{\tilde{x}^k\}_{k\in K}$.

Since the conditions of Definition 3 are satisfied, MPCC-QN would fail at x^*, contradicting the hypothesis. Therefore $\{(\lambda^k, \mu^k)\}_{k\in K}$ is bounded. □

4.1 Numerical illustration

We implemented Algorithm 1 in Julia. Test were conducted in a computer equipped with an Intel(R) Xeon(R) Silver 4114 CPU 2.20GHz and Julia v1.11.5. The code identifies complementarities in the "AMPL style" through the package `AmplNLReader.jl` (`github.com/JuliaSmoothOptimizers/AmplNLReader.jl`). In this format, complementarities are written as $0 \le c(w) \perp v \ge 0$, where v is a variable of the model. We then rewrite it to $u = c(w)$, $0 \le u \perp v \ge 0$ to fit the model to the theory. In this case, subproblems are solved by the inner method as described in [28, Algorithm 2]. As we already mentioned, this is an extension of the non-monotone SPG method with projections onto the non-convex set X. To maintain the compatibility, if a model without complementarities is passed then the box-constrained subproblems, that does not have complementarities, are solved by the regular SPG for convex sets [19, 20]. It is worth mentioned, however, that quadratic interpolation is employed within the SPG, a technique that improves the line search, but it is not present in the SPG variant for complementarities. Nevertheless, we maintain the quadratic interpolation in the regular SPG as it is recommended [19, 22]. In summary, the code accomplishes two variants of the AL method:

- Algorithm 1, that maintain the complementarity in the subproblems;
- Standard AL method that treats the model as a standard NLP, penalizing complementarities (Standard AL).

The code is freely available at `github.com/leonardosecchin/SimpleAL`.

Our implementation is inspired in the consolidate package Algencan [1, 18], a mature Fortran implementation of the counterpart of Algorithm 1 for standard NLP. We take $\tau = 0.8$, $\theta = 5$, $\lambda_{\min} = -10^{20}$, $\lambda_{\max} = \mu_{\max} = 10^{20}$, $\bar{\lambda}^0 = 0$ and $\bar{\mu}^0 = 0$. Similar to suggested in [18, Section 12.4], we take

$$\rho_1 = \max\left\{10^{-8}, \min\left\{10^8, \frac{2\max\{1, |f(x^0)|\}}{\max\{1, \|h(x^0)\|_2^2 + \|g(x^0)_+\|_2^2\}}\right\}\right\}.$$

This differs from the strategy adopted in [28]. The stopping criteria are inherited from Algencan; see [16]. As only first-order methods are used for solving subproblems, we set the moderate tolerances for feasibility and optimality as $\varepsilon_{\text{feas}} = \varepsilon_{\text{opt}} = 10^{-5}$. The tolerances for the subproblems are taken as $\varepsilon_1 = \varepsilon_{\text{opt}}$ and, for $k > 1$, $\varepsilon_k = \max\{0.1\varepsilon_{\text{opt}}, \min\{\theta\,\text{opt}_k, 0.1\varepsilon_k\}\}$ if $\max\left\{\|h(x^k)\|_\infty, \|V^k\|_\infty\right\} \le \sqrt{\varepsilon_{\text{feas}}}$ and $\text{opt}_k \le \sqrt{\varepsilon_{\text{opt}}}$, and $\varepsilon_k = \varepsilon_{k-1}$ otherwise. Here, the optimality measure opt_k is

$$\text{opt}_k = \text{dist}(\nabla L_{\rho_k}(x^k, \lambda^k, \mu^k), N_{\lim}(x^k)).$$

Due to the expression (4) for $N_{\lim}$, we can compute opt_k in practice as the sup-norm of the vector $L^k = (-\nabla_w L_{\rho_k}(x^k, \lambda^k, \mu^k), \tilde{u}^k, \tilde{v}^k) \in \mathbb{R}^{n+2m}$ where

$$\tilde{u}_i^k = \begin{cases} 0 & \text{if } u_i^k = 0 \text{ and } v_i^k > \delta \\ L_i^u & \text{if } u_i^k > \delta \text{ and } v_i^k = 0 \\ \max\left\{(L_i^u)_+,\ (L_i^v)_+,\ \min\left\{|L_i^u|, |L_i^v|\right\}\right\} & \text{otherwise} \end{cases},$$

$$\tilde{v}_i^k = \begin{cases} 0 & \text{if } u_i^k > \delta \text{ and } v_i^k = 0 \\ L_i^u & \text{if } u_i^k = 0 \text{ and } v_i^k > \delta \\ \max\left\{(L_i^u)_+,\ (L_i^v)_+,\ \min\left\{|L_i^u|, |L_i^v|\right\}\right\} & \text{otherwise} \end{cases},$$

$L_i^u = -\frac{\partial L_{\rho_k}(x^k,\lambda^k,\mu^k)}{\partial u_i}$, $L_i^v = -\frac{\partial L_{\rho_k}(x^k,\lambda^k,\mu^k)}{\partial v_i}$ and $\delta > 0$ is a tolerance (we take $\delta = 10^{-6}$ in our tests). Note that the iterates of Algorithm 1 satisfy the complementarity exactly, so we always have $u_i^k \geq 0$, $v_i^k \geq 0$ and $u_i^k v_i^k = 0$ for all k. Thus, each row in $\tilde{u}_i^k$ and $\tilde{v}_i^k$ represents the cases $i \in I_{0+}(x^k)$, $i \in I_{+0}(x^k)$ and $i \in I_{00}(x^k)$, respectively. Expression $\text{opt}_k \approx 0$ can be viewed as the approximate counterpart of "$-\nabla L_{\rho_k}(x^k, \lambda^k, \mu^k) \in N_{\text{lim}}(x^k)$". In Standard AL, this measure becomes $\text{opt}_k = \|\nabla L_{\rho_k}(x^k, \lambda^k, \mu^k)\|_\infty$, as expected. The combination of "max" and "min" in $\tilde{u}_i^k$ and $\tilde{v}_i^k$ captures the control of signs in the last row in (4). A similar strategy was used in [6] to describe M-stationarity approximately. Using "$\min\left\{|L_i^u|, |L_i^v|\right\}$" is better than "$|L_i^u L_i^v|$" because it keeps the measurements on the same scale.

Following [18, Section 12.5], we scale the problem data before solving it; this is a common practice that increases the conditioning of the problem. Regarding (MPCC), the following stopping criteria are implemented:

st 0. optimality, complementarity and (unscaled) feasibility are achieved, that is,

$$\text{opt}_k \leq \varepsilon_{\text{opt}}, \quad \|V^k\|_\infty \leq \varepsilon_{\text{opt}}, \quad \text{infeas}_k = \max\{\|h(x^k)\|_\infty, \|g(x^k)_+\|_\infty\} \leq \varepsilon_{\text{feas}}.$$

In other words, an approximate M-stationary point was found. Here, objective function and constraints in $\nabla L_{\rho_k}(x^k, \lambda^k, \mu^k)$ and V^k are scaled;

st 1. $\text{infeas}_k > \sqrt{\varepsilon_{\text{feas}}}$ and

$$\left\|\nabla\left(\|h(x^k)\|_2^2 + \|g(x^k)_+\|_2^2\right)\right\|_\infty \leq \max\{10^{-12}, 10^{-4}\varepsilon_{\text{opt}}\}.$$

This indicates that the method converges to a stationary point of the infeasibility;

st 2. $\rho_k > 10^{20}$;

st 3. the number of outer iterations reached 100;

st 4. $\text{infeas}_k \leq \sqrt{\varepsilon_{\text{feas}}}$ and $f(x^k) < -10^{20}$, which indicates unlimited problem.

We consider test problems from the MacMPEC collection maintained by Sven Leyffer (https://wiki.mcs.anl.gov/leyffer/index.php/MacMPEC). Note that S2 does not necessarily indicate that the method diverges because the convergence to a non KKT point (C/M-stationary) can require that $\rho_k \to \infty$, at least for the Standard AL; see Example 1. This in fact was observed in numerical tests with Algencan [8]. Similar to [8], we thus consider a problem solved where S0 is satisfied or S2 occurs with feasibility attained and the objective not worse than the best known value reported in the MacMPEC. The objective f_k attained by the algorithm is considered not worse than the best known value f^* in MacMPEC if $f_k \leq f^* + 10^{-8}|f^*|$.

The implementation of the inner solvers with and without complementarities follow [28, Algorithm 2, (30)] and [19, Algorithm 2.2], respectively. The corresponding parameters for the latter are $M = 10$, $\sigma_1 = 0.1$, $\sigma_2 = 0.9$, $\alpha_{\min} = 10^{-20}$, $\alpha_{\max} = 10^{20}$ and $\gamma = 10^{-4}$. The initial spectral steplength was taken following [21] (see also [22]):

$$\alpha_0 = \begin{cases} \max\{\alpha_{\min}, \min\{s^t s / s^t y, \alpha_{\max}\}\} & \text{if } s^t s / s^t y > 0 \\ \alpha_{\max} & \text{otherwise} \end{cases},$$

where $s = x^* - x^k$, $y = \nabla_x L_{\rho_k}(x^*, \lambda^k, \mu^k) - \nabla_x L_{\rho_k}(x^k, \lambda^k, \mu^k)$ (k is the current outer iteration), $x^* = x^k - t_{\text{small}} \nabla_x L_{\rho_k}(x^k, \lambda^k, \mu^k)$ and $t_{\text{small}} = \max\{10^{-7}\|x^k\|_\infty, 10^{-10}\}$. This strategy tries to estimate the curvature of the Lagrangian from the current point x^k. In turn, the parameters for [28, Algorithm 2] are $\tau = 5$, $L_{\min} = 1/\alpha_{\max}$, $L_{\max} = 1/\alpha_{\min}$ and $L_t^0 = 1/\alpha_0$, this because L mimics the inverse of the spectral steplength α in the regular SPG. For both methods, we set the maximum number of iterations that to 10,000 and the following additional stopping criteria: $\|\text{opt}_k\|_\infty \leq \varepsilon_k$, lack of progress due to too small steplength, and lack of progress in reduction of L during 3,000 consecutive iterations.

Some MacMPEC problems presented an error while parsing the `nl` file, and thus were discarded. Each of the remaining 127 problems is executed maintaining complementarities in the subproblem (Algorithm 1) and penalizing it (Standard AL). The last case is achieved by explicitly rewriting the complementarities as regular constraints using the `AmplMPECModel` command from the `AmplNLReader.jl` package.

Table 1 brings the results for the problems solved by at least one method and that the final objective values f_1 and f_{std} in both do not differ significantly, in the sense that $|f_1 - f_{\text{std}}|/\max\{1, |f_1|, |f_{\text{std}}|\} \leq 10^{-2}$. In this manner, we capture problems where both algorithms are comparable. Column "opt/com" contains the values $\max\{\|\text{opt}_k\|_\infty, \|V^k\|_\infty\}$, i.e., the optimality and complementarity violations. Italic values mean that the prescribed tolerance was not reached. Column "max dual" brings the maximum sup-norm of the Lagrange multipliers of penalized constraints achieved during all the minimization process (as ρ never decreases, it is usually the sup-norm of the last multiplier). Smaller multipliers values are marked in bold for problems where tolerances are met. We observe that Algorithm 1 works with smaller multipliers more often. This is coherent with the fact that in this method we do not deal directly with the multipliers of complementarities, which tend to explode as illustrated by Example 1.

Table 1: Problems solved by at least one method.

	Algorithm 1				Standard AL			
Problem	st	infeas	opt/com	max dual	st	infeas	opt/com	max dual
`bard3`	0	7.9e−06	2.0e−06	1.3e+00	0	6.4e−06	1.7e−06	1.3e+00
`bilevel3`	0	1.7e−06	8.9e−07	**1.7e+00**	2	*8.9e−01*	*1.1e+40*	3.4e+20
`design-cent-31`	0	1.4e−13	6.5e−07	**4.2e−01**	0	7.4e−12	4.3e−06	1.9e+00
`desilva`	0	9.9e−06	4.9e−06	5.0e−01	0	5.7e−06	9.2e−06	5.0e−01
`ex9.2.1`	0	3.3e−06	1.1e−06	**3.3e+00**	0	4.4e−06	1.5e−06	2.7e+02
`ex9.2.4`	0	1.4e−06	1.4e−06	1.0e+00	0	5.3e−06	5.3e−06	1.0e+00
`ex9.2.6`	0	9.0e−06	8.3e−06	5.0e−01	0	9.2e−06	8.4e−06	5.0e−01
`ex9.2.7`	0	3.3e−06	1.1e−06	**3.3e+00**	0	4.4e−06	1.5e−06	2.7e+02
`flp2`	0	9.2e−06	3.9e−06	**6.6e−03**	0	0.0e+00	1.1e−10	2.7e+02
`flp4-2`	2	*2.9e−01*	*5.0e+18*	1.4e+17	0	0.0e+00	5.5e−10	**0.0e+00**
`flp4-4`	2	*2.6e−02*	*1.5e+17*	4.2e+15	0	0.0e+00	4.6e−10	**0.0e+00**
`gauvin`	0	8.8e−06	1.1e−06	**1.0e+00**	0	0.0e+00	7.7e−08	3.0e+02
`jr1`	0	2.6e−06	2.6e−06	**1.0e+00**	0	0.0e+00	5.7e−06	2.0e+00
`jr2`	0	9.1e−06	9.1e−06	1.0e+00	0	0.0e+00	5.7e−06	**1.1e−05**
`kth2`	0	2.0e−06	6.8e−06	8.1e−06	0	0.0e+00	0.0e+00	0.0e+00
`nash1`	0	5.2e−08	5.9e−07	**2.1e+00**	0	3.3e−06	1.7e−06	3.9e+00
`nash1a`	0	3.7e−07	6.8e−07	7.2e−05	0	4.3e−06	1.6e−06	2.4e−06
`nash1b`	0	1.5e−07	5.9e−07	2.6e−03	0	3.0e−06	1.3e−06	3.2e−04
`nash1e`	0	7.0e−06	2.7e−06	8.1e−04	0	9.2e−06	3.6e−06	6.8e−04
`outrata31`	0	2.0e−06	2.0e−06	**1.2e+00**	3	0.0e+00	*2.7e+00*	1.1e+01
`qpec1`	0	1.9e−08	3.0e−06	8.6e−07	0	0.0e+00	1.5e−16	0.0e+00
`scale1`	0	3.1e−06	3.1e−06	2.0e+00	0	0.0e+00	1.0e−06	**4.0e−04**
`scale3`	0	3.8e−06	7.7e−06	1.5e−03	0	0.0e+00	2.8e−06	6.8e−04
`scholtes5`	0	2.8e−06	2.8e−06	**1.0e+00**	0	0.0e+00	3.3e−06	8.9e+00

5 Final remarks

In this work, we analyzed AL methods for MPCCs under a framework that preserves the complementarity structure in the subproblems, specifically the one proposed in [28]. Motivated by recent developments in nonlinear programming [3], we introduced two new tailored CQs for MPCCs, namely, specialized MPCC-quasi-normality (sMPCC-QN, Definition 3) and its relaxed counterpart (sMPCC-RQN, Definition 6). From our analysis, we prove that both conditions, as well as the MPCC-QN from [33], are sufficient to ensure boundedness of the multiplier sequence in such AL methods, an open question until then. This extends to the MPCC context the weakest assumption guaranteeing this property [3]. We also demonstrated that these new CQs imply a novel error bound property (sMPCC-EB, Definition 8), which is equivalent to the one defined in [24], and naturally aligned with algorithms that enforce feasibility with respect to complementarity.

The theoretical advances were complemented by numerical experiments, which reinforce that preserving complementarity constraints in the AL subproblems, as done in [28], rather than penalize them is beneficial to the numerical performance. This raises the need to develop robust inner solvers that directly deal with complementarity constraints. In fact, the (first order) projected gradient-type method described in [28], used in our implementation, is not able to achieve high accuracy in ill-conditioned problems. One possibility is to design Newtonian strategies, as the one employed in the current implementation of ALGENCAN [18] solver (`https://www.ime.usp.br/~egbirgin/tango`). In this package, box-constrained subproblems are solved by GENCAN [15], an active-set method that employs Newton's method within the constraint faces. Also, acceleration strategies [17] may be promising. This is a future topic of research. Another open question is to determine whether sMPCC-RQN is equivalent to sMPCC-EB and whether it is the weakest assumption ensuring bounded multipliers in Algorithm 1. Although this was established recently for standard nonlinear programming [3], the adaptations are not straightforward.

References

[1] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt. On augmented Lagrangian methods with general lower-level constraints. *SIAM Journal on Optimization*, 18(4):1286–1309, 2007.

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

[3] R. Andreani, G. Haeser, R. W. Prado, and L. D. Secchin. Primal-dual global convergence of an augmented Lagrangian method under the error bound condition. Technical report, 2025.

[4] R. Andreani, G. Haeser, M. L. Schuverdt, and L. D. Secchin. A relaxed quasinormality condition and the boundedness of dual augmented Lagrangian sequences. *To appear in SIAM Journal on Optimization*, 2025.

[5] R. Andreani, G. Haeser, M. L. Schuverdt, L. D. Secchin, and P. J. S. Silva. On scaled stopping criteria for a safeguarded augmented Lagrangian method with theoretical guarantees. *Mathematical Programming Computation*, 14(1):121–146, 2022.

[6] R. Andreani, G. Haeser, L. D. Secchin, and P. J. S. Silva. New sequential optimality conditions for mathematical programs with complementarity constraints and algorithmic consequences. *SIAM Journal on Optimization*, 29(4):3201–3230, 2019.

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

[8] R. Andreani, L. D. Secchin, and P. J. S. Silva. Convergence properties of a second order augmented Lagrangian method for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 28(3):2574–2600, 2018.

[9] M. Anitescu. Global convergence of an elastic mode approach for a class of mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 16(1):120–145, 2005.

[10] M. Anitescu. On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 15(4):1203–1236, 2005.

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

[12] M. Benko and H. Gfrerer. An SQP method for mathematical programs with complementarity constraints with strong convergence properties. *Kybernetika*, 52(2):169–208, 2016.

[13] D. P. Bertsekas. *Nonlinear Programming: 3nd Edition.* Athena Scientific, 2016.

[14] D. P. Bertsekas and A. E. Ozdaglar. Pseudonormality and a Lagrange multiplier theory for constrained optimization. *Journal of Optimization Theory and Applications*, 114(2):287–343, 2002.

[15] E. G. Birgin and J. M. Martínez. Large-scale active-set box-constrained optimization method with spectral projected gradients. *Computational Optimization and Applications*, 23(1):101–125, 2002.

[16] E. G. Birgin and J. M. Martínez. Augmented Lagrangian method with nonmonotone penalty parameters for constrained optimization. *Computational Optimization and Applications*, 51(3):941–965, 2012.

[17] E. G. Birgin and J. M. Martínez. Improving ultimate convergence of an augmented Lagrangian method. *Optimization Methods and Software*, 23(2):177–195, 2008.

[18] E. G. Birgin and J. M. Martínez. *Practical Augmented Lagrangian Methods for Constrained Optimization.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2014.

[19] E. G. Birgin, J. M. Martínez, and M. Raydan. Nonmonotone spectral projected gradient methods on convex sets. *SIAM Journal on Optimization*, 10(4):1196–1211, 2000.

[20] E. G. Birgin, J. M. Martínez, and M. Raydan. Algorithm 813: SPG-software for convex-constrained optimization. *ACM Transactions on Mathematical Software*, 27(3):340–349, 2001.

[21] E. G. Birgin, J. M. Martínez, and M. Raydan. Inexact spectral projected gradient methods on convex sets. *IMA Journal of Numerical Analysis*, 23(4):539–559, 2003.

[22] E. G. Birgin, J. M. Martínez, and M. Raydan. Spectral projected gradient methods: review and perspectives. *Journal of Statistical Software, Articles*, 60(3):1–21, 2014.

[23] G. A. Carrizo, N. S. Fazzio, M. D. Sánchez, and M. L. Schuverdt. Scaled-PAKKT sequential optimality condition for multiobjective problems and its application to an augmented Lagrangian method. *Computational Optimization and Applications*, 89(3):769–803, 2024.

[24] N. H. Chieu and G. M. Lee. A relaxed constant positive linear dependence constraint qualification for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 158(1):11–32, 2013.

[25] N. S. Fazzio, M. D. Sánchez, and M. L. Schuverdt. Sequential optimality conditions for optimization problems with additional abstract set constraints. *Revista de la Unión Matemática Argentina*, 67(1):257–279, 2024.

[26] M. L. Flegel and C. Kanzow. Abadie-type constraint qualification for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 124(3):595–614, 2005.

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

[28] L. Guo and Z. Deng. A new augmented Lagrangian method for MPCCs - theoretical and numerical comparison with existing augmented Lagrangian methods. *Mathematics of Operations Research*, 47(2):1229–1246, 2022.

[29] L. Guo and G. Li. Approximation methods for a class of non-Lipschitz mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 202(3):1421–1445, July 2024.

[30] L. Guo and G.-H. Lin. Notes on some constraint qualifications for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 156(3):600–616, 2013.

[31] A. F. Izmailov, M. V. Solodov, and E. I. Uskov. Global convergence of augmented Lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints. *SIAM Journal on Optimization*, 22(4):1579–1606, 2012.

[32] H. Jiang and D. Ralph. Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. *SIAM Journal on Optimization*, 10(3):779–808, 2000.

[33] C. Kanzow and A. Schwartz. Mathematical programs with equilibrium constraints: enhanced Fritz John-conditions, new constraint qualifications, and improved exact penalty results. *SIAM Journal on Optimization*, 20(5):2730–2753, 2010.

[34] C. Kanzow and A. Schwartz. The price of inexactness: convergence properties of relaxation methods for mathematical programs with complementarity constraints revisited. *Mathematics of Operations Research*, 40(2):253–275, 2015.

[35] X. Liu and J. Sun. Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints. *Mathematical Programming*, 101(1):231–261, 2004.

[36] H. Z. Luo, X. L. Sun, and Y. F. Xu. Convergence properties of modified and partially-augmented Lagrangian methods for mathematical programs with complementarity constraints. *Journal of Optimization Theory and Applications*, 145(3):489–506, 2010.

[37] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints.* Cambridge University Press, United Kingdom, 1996.

[38] L. Minchenko and A. Tarakanov. On error bounds for quasinormal programs. *Journal of Optimization Theory and Applications*, 148(3):571–579, 2011.

[39] M. T. T. Monteiro and H. S. Rodrigues. Combining the regularization strategy and the SQP to solve MPCC — a MATLAB implementation. *Journal of Computational and Applied Mathematics*, 235(18):5348–5356, 2011.

[40] B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation, vol 1. Basic Theory.* Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, New York, 2006.

[41] J. V. Outrata. Optimality conditions for a class of mathematical programs with equilibrium constraints: strongly regular case. *Kybernetika*, 35(2):177–193, 1999.

[42] A. U. Raghunathan and L. T. Biegler. An interior point method for mathematical programs with complementarity constraints (MPCCs). *SIAM Journal on Optimization*, 15(3):720–750, 2005.

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

[44] K. Wang and L. T. Biegler. MPCC strategies for nonsmooth nonlinear programs. *Optimization and Engineering*, 24(3):1883–1929, Sept. 2022.