Global convergence of a second order augmented Lagrangian method under an error bound condition[*]

R. Andreani[†], G. Haeser[‡], R. W. Prado[§], M. L. Schuverdt[¶], L. D. Secchin[‖]

September 11, 2024

Abstract

This work deals with convergence to points satisfying the weak second-order necessary optimality conditions of a second-order safeguarded augmented Lagrangian method from the literature. To this end, we propose a new second-order sequential optimality condition that is, in a certain way, based on the iterates generated by the algorithm itself. This also allows us to establish the best possible global convergence result for the method studied, from which a companion constraint qualification is derived. Unlike similar results from previous works, the new constraint qualification assures second-order stationarity without the need for constant rank hypotheses. To guarantee this result, we established the convergence of the method under a property slightly stronger than the error bound constraint qualification, which, until now, has not been known to be associated with nonlinear optimization methods.

Keywords: Sequential optimality conditions; Second-order optimization methods; Constraint qualifications; Error bound; Constant rank

1 Introduction

We consider the general nonlinear programming problem (NLP)

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}, \quad \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^p$ and $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^q$ are all twice continuously differentiable. The feasible set of problem (1) is denoted by

$$\Omega \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}, \ \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\}.$$

Due to advances in computing hardware and algorithmic efficiency, there has been an increase in interest in optimization methods that utilize second-order information [26, 29, 34, 1]. It allows

[†]Department of Applied Mathematics, Universidade Estadual de Campinas, Campinas, SP, Brazil. Email: andreani@unicamp.br

[‡]Department of Applied Mathematics, University of São Paulo, São Paulo, SP, Brazil. Email: ghaeser@ime.usp.br

[§]Department of Applied Mathematics, University of São Paulo, São Paulo, SP, Brazil. Email: pradorenanw@gmail.com

[¶]CONICET, Department of Mathematics, FCE, University of La Plata, La Plata, Bs. As., Argentina. Email: schuverd@mate.unlp.edu.ar

[‖]Department of Applied Mathematics, Federal University of Espírito Santo, São Mateus, ES, Brazil. Email: leonardo.secchin@ufes.br

1

achieving high accuracy solutions [20], as it provides additional information about the curvature of the objective and constraints, see, for example, [15] and references therein. This work focuses on the second-order safeguarded augmented Lagrangian method (SALM) proposed in [4], whose first-order counterpart method Algencan [3] is a celebrated algorithm with excellent convergence results [14, 6, 11] and robust numerical performance [22, 21].

Ideally, one would like methods to find a local minimizer. However, practical methods can only guarantee stationary points. In the first-order case, the strength of an algorithm can be measured by the strength of constraint qualifications that ensure the Karush-Kuhn-Tucker (KKT) conditions at its accumulation points. Likewise, the validity of second-order stationary conditions can be considered.

Preferably, as a second-order condition, one wants not only the KKT conditions to hold but also that, for some Lagrange multiplier vector, no direction within the linearized cone is simultaneously a non-ascending direction for the objective function ($\nabla f(\boldsymbol{x})^T \boldsymbol{d} \leq 0$) and of negative curvature for the Hessian of the Lagrangian function (with respect to $\boldsymbol{x}$); that is, one wants that the Hessian of the Lagrangian function is positive semidefinite in the strong critical cone

$$
\mathscr{C}^S(\boldsymbol{x}) = \left\{ \boldsymbol{d} \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \langle \nabla f(\boldsymbol{x}), \boldsymbol{d} \rangle \leq 0, \\ \langle \nabla h_j(\boldsymbol{x}), \boldsymbol{d} \rangle = 0 \text{ for all } j \in \{1, \ldots, q\}, \\ \langle \nabla g_i(\boldsymbol{x}), \boldsymbol{d} \rangle \leq 0 \text{ for all } i \in \mathcal{A}(\boldsymbol{x}^*) \end{array} \right\},
$$

where $\mathcal{A}(\boldsymbol{x}) = \{j \mid g_j(\boldsymbol{x}) = 0\}$ is the set of indexes of active inequality constraints at $\boldsymbol{x}$. Together with the KKT conditions, the resultant condition is called *strong second-order optimality condition* (SSOC). Nevertheless, even testing such a condition in a given and fixed point is hard [33]. For that reason, all commonly applied methods with second-order optimality guarantees deal with the weak critical cone

$$
\mathscr{C}^W(\boldsymbol{x}) = \left\{ \boldsymbol{d} \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \langle \nabla h_j(\boldsymbol{x}), \boldsymbol{d} \rangle = 0 \text{ for all } j \in \{1, \ldots, q\}, \\ \langle \nabla g_i(\boldsymbol{x}), \boldsymbol{d} \rangle = 0 \text{ for all } i \in \mathcal{A}(\boldsymbol{x}) \end{array} \right\}; \tag{2}
$$

for instance, the methods in [15]. As easily seen, this cone is properly included in the strong critical cone when KKT conditions hold. The practical condition associated with the weak critical cone is called *weak second-order conditions* (WSOC); see Definition 2.3.

The convergence of any algorithm to points fulfilling WSOC cannot be ensured even when the *Mangasarian-Fromovitz constraint qualification* (MFCQ) holds; see the counterexample in [17, Section 2.4]. For this reason, until recently, second-order methods had the *linear independence constraint qualification* (LICQ) and strict complementarity as requirements for convergence, conditions that guarantee that $\mathscr{C}^W(\boldsymbol{x}^*) = \mathscr{C}^S(\boldsymbol{x}^*)$ and uniqueness of the multipliers. To guarantee convergence of implementable second-order methods, one may require MFCQ together with some other constant rank-type condition. In [10], this was done assuming MFCQ and that the gradients of all active constraints in a neighborhood of $\boldsymbol{x}^*$ have constant rank – a condition known as *weak constant rank* (WCR), see Definition 2.4. In [5], it was proved that a constraint qualification (CQ) of constant rank-type also fulfills this purpose. Finally, these results were extended in [15] using a second-order sequential optimality condition called *second-order approximate KKT* (AKKT2).

Motivated by [15] and the recent results obtained for the first-order optimality in [14], we identify precisely the strongest global convergence result associated with SALM. This is done by showing that all possible accumulation points reached by the method are fully characterized by a new proposed second-order sequential optimality condition that we call *second-order augmented Lagrangian AKKT* (AL-AKKT2), following the terminology adopted in [14]. As a consequence, we prove that WSOC holds at these points, which include minimizers of (1), under new weak constraint qualifications. In particular, AL-AKKT2 is stronger than AKKT2 and allows us to obtain convergence of SALM in situations where WCR does not hold. It is worth noting that this was not possible before, as AKKT2 requires the validity of WCR to attest WSOC [25, Proposition 5]. This is done with a slightly more stringent condition than the error bound (EB) property, which we call *strong-EB*. Strong-EB is proven weaker than a recently proposed CQ for WSOC in conic programming context, called strong-CRSC [12]. To the best of our knowledge, this is the first time that a second-order method has its global convergence established under an error bound-type condition.

It has been recently discovered that a relaxed version of the *quasi-normality* (QN) CQ is sufficient for the boundedness of the multipliers sequence associated with the first-order SALM [16]. This *relaxed QN* (RQN) condition is a weak CQ, less stringent than QN itself and one of the weakest known CQ of constant rank-type, the *constant rank of the subspace component* (CRSC) [8]. However, the boundedness of dual sequences of SALM ensured by RQN is insufficient to assert WSOC. To address this concern, we show that WSOC holds under RQN as long as WCR or even strong-EB holds. This generalizes all the previous results from the literature regarding second-order algorithms.

The paper is organized as follows. In section 2, we present SALM and the fundamental theory involving WSOC. In section 3, we establish the new sequential optimality condition AL-AKKT2. In section 4, we prove that strong-EB guarantees WSOC under AL-AKKT2 and the boundedness of the AL-AKKT2 multipliers. Furthermore, we prove that WCR is independent of strong-EB and that WCR can also guarantee WSOC under boundedness of the AL-AKKT2 dual sequences. In section 5, we prove that the AL-AKKT2 conditions fully characterize the accumulation points of sequences generated by SALM. Additionally, we define the weakest CQ for WSOC under the AL-AKKT2 condition. Finally, section 6 brings our conclusions and future works.

Notation

- We typically use boldface letters (i.e., $\boldsymbol{v}$) for vectors and vector-valued functions;

- v_i is the $i-$th component of the vector $\boldsymbol{v}$. If $\boldsymbol{v}$ is a vector-valued function, v_i is the function given by its $i-$th component;

- For $\boldsymbol{v} \in \mathbb{R}^n$, $\boldsymbol{v}_+ = (\max\{0, v_1\}, \ldots, \max\{0, v_n\})$;

- $\| \cdot \|$ denotes the Euclidean norm;

- $\mathcal{B}(\boldsymbol{x}, \epsilon)$ is the open ball with center $\boldsymbol{x}$, radius ϵ, and the Euclidean distance;

- $\mathbb{R}_+$ stands for the set of non-negative real numbers;

- $\mathbb{R}_{++}$ denotes the set of strictly positive real numbers;

- Given a natural ℓ, $I_\ell = \{k \in \mathbb{N} \,|\, 1 \leq k \leq \ell\}$;

- $\mathbb{I}$ denotes the identity matrix of appropriate size;

- Given an infinite subset $\mathcal{K} \subset \mathbb{N}$, $\lim_{k \in \mathcal{K}} \boldsymbol{w}^k$ denotes the limit of the subsequence of $\{\boldsymbol{w}^k\}_{k \in \mathbb{N}}$ with indexes $k \in \mathcal{K}$;

- $\text{Sym}(n)$ is the set of the symmetric matrices of order n;

- Given two symmetric matrices A and B of the same order, $A \succeq B$ indicates that $A - B$ is positive semi-definite;

2 Technical background and the augmented Lagrangian method

The algorithm of interest for solving nonlinear programming problems is the (Powell-Hestenes-Rockafellar – PHR) safeguarded augmented Lagrangian method (SALM). This method successively minimizes the PHR-augmented Lagrangian function described below.

Definition 2.1. Fixed a scalar $\rho > 0$, the Powell-Hestenes-Rockafellar (PHR) augmented Lagrangian function $\mathcal{L}_\rho \colon \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q_+$ is given by

$$\mathcal{L}_\rho(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}) \overset{\text{def}}{=} f(\boldsymbol{x}) + \frac{\rho}{2}\left(\left\|\left(\boldsymbol{g}(\boldsymbol{x}) + \frac{\boldsymbol{\mu}}{\rho}\right)_+\right\|^2 + \left\|\left(\boldsymbol{h}(\boldsymbol{x}) + \frac{\boldsymbol{\lambda}}{\rho}\right)\right\|^2 \right)$$

for all $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{\mu} \in \mathbb{R}^p_+$ and $\boldsymbol{\lambda} \in \mathbb{R}^q$.

Specifically, the idea behind SALM is to solve (1) through a sequence of problems

$$\min_{\boldsymbol{x}} \mathcal{L}_{\rho_k}(\boldsymbol{x}, \bar{\boldsymbol{\mu}}^k, \bar{\boldsymbol{\lambda}}^k) \tag{3}$$

where $\{\rho_k\}_{k \in \mathbb{N}}$ is a non-decreasing and positive sequence of real numbers and $\{\bar{\boldsymbol{\lambda}}^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^q$ and $\{\bar{\boldsymbol{\mu}}^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^p_+$ are bounded sequences, which are safeguarded estimates of the Lagrange multipliers for (1). The scalar ρ_k is called the penalty parameter. Under reasonable conditions, local minimizers of (3) are also local minimizers of (1), c.f. [19, Prop. 4.2.3] or [21, Lemma 7.2]. Essentially, SALM is a method that controls the increasing behavior of the sequence of penalty parameters to avoid ill-conditioned subproblems.

Unless (3) is a convex problem, globally minimizing it in practice is tough. Despite the global problem (3) being hard, finding an approximate second-order stationary point is not. Consequently, a natural idea would be to apply to (3) globally convergent algorithms for unconstrained optimization that generate sequences whose accumulation points satisfy a second-order optimality condition. However, the PHR augmented Lagrangian function is not twice differentiable everywhere. To overcome this, we work with an approximation of its Hessian, as stated in the next definition.

Definition 2.2. Considering the functions that appear in problem (1), $\boldsymbol{x} \in \mathbb{R}^n$, $\bar{\boldsymbol{\mu}} \in \mathbb{R}^p_+$, $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^q$, $\eta \in \mathbb{R}_+$ and $\rho \in \mathbb{R}_+$, define the matrix

$$\nabla^2_{\boldsymbol{x}} \mathcal{L}^\eta_\rho(\boldsymbol{x}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}}) \stackrel{\text{def}}{=} \nabla^2_{\boldsymbol{x}} f(\boldsymbol{x}) + \sum_{i=1}^p (\bar{\mu}_i + \rho g_i(\boldsymbol{x}))_+ \nabla^2 g_i(\boldsymbol{x}) + \rho \sum_{\bar{\mu}_i + \rho g_i(\boldsymbol{x}) \geq -\eta} \nabla g_i(\boldsymbol{x}) \nabla g_i(\boldsymbol{x})^T$$

$$+ \sum_{j=1}^q \left(\bar{\lambda}_j + \rho h_j(\boldsymbol{x}) \right) \nabla^2 h_i(\boldsymbol{x}) + \rho \sum_{j=1}^q \nabla h_j(\boldsymbol{x}) \nabla h_j(\boldsymbol{x})^T.$$

By [4, Proposition 1], it is possible to show that $\nabla^2_{\boldsymbol{x}} \mathcal{L}^\eta_\rho(\cdot, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}})$ is positive semidefinite at unconstrained local minima of $\mathcal{L}_\rho(\cdot, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}})$, independently of η. Thus, it is a suitable notion of the second-order derivative of $\mathcal{L}_\rho(\cdot, \boldsymbol{\lambda}, \boldsymbol{\mu})$ w.r.t. $\boldsymbol{x}$, and it can be employed in methods for unconstrained optimization with second-order convergence guarantees. We expect that such methods applied to (3) are able to find $\boldsymbol{x}^k$ fulfilling (4) below. With this, we enunciate the safeguarded augmented Lagrangian method in Algorithm 1.

Remark 1. A small extension of [38, Corollary 1.4.3] can be used to guarantee the existence of points fulfilling (4) under the assumption that the PHR augmented Lagrangian function is bounded below. This can be ensured by assuming that the objective function of (1) has a lower bound. Under such an assumption, any method generating sequences with second-order properties for unconstrained optimization applied to (3) can be used to obtain (4). Andreani et al. proposed an implementation of a SALM meeting the requirements of this section, which they called ALGENCAN-SECOND, see [4, Algorithm 3.1].

From now on, we will refer to the second-order augmented Lagrangian method stated above as SALM. When we want to refer to the first-order method, that is, the one with only the first condition in (4), we explicitly write "first-order SALM". Additionally, any sequence satisfying the steps of SALM for a proper choice of parameters is said to be generated by SALM. Similarly, we might also say that a sequence is generated by the first-order SALM.

When a new method is proposed, it is desirable that the KKT conditions, i.e., for some $\boldsymbol{\mu} \in \mathbb{R}^p_+$, $\boldsymbol{\lambda} \in \mathbb{R}^q$ and $\boldsymbol{x} \in \Omega$,

$$\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \boldsymbol{0} \quad \text{and} \quad \boldsymbol{\mu}^T g(\boldsymbol{x}) = 0,$$

where $\mathcal{L}$ is the usual Lagrangian function associated with (1), hold at its accumulation points under at least one constraint qualification (CQ). The less stringent the CQ is, the more points the method will

---

Algorithm 1 Second-order safeguarded augmented Lagrangian method (SALM)

Step 1 Choose $\lambda_{\min} \leq \lambda_{\max}$, $0 \leq \mu_{\max}$, $\gamma > 1$, $0 < r < 1$, $\{\theta_k\}_{k\in\mathbb{N}}$ and $\{\eta_k\}_{k\in\mathbb{N}}$ with $\theta_k \downarrow 0$ and $\eta_k \downarrow 0$. Let $\bar{\boldsymbol{\lambda}}^1 \in [\lambda_{\min}, \lambda_{\max}]^q$, $\bar{\boldsymbol{\mu}}^1 \in [0, \mu_{\max}]^p$ and $\rho_1 > 0$. Initialize $k \leftarrow 1$.

Step 2 Solve approximately (3) obtaining $\boldsymbol{x}^k$ satisfying

$$\|\nabla \mathcal{L}_{\rho_k}(\boldsymbol{x}^k, \bar{\boldsymbol{\mu}}^k, \bar{\boldsymbol{\lambda}}^k)\| \leq \theta_k \qquad \text{and} \qquad \nabla_{\boldsymbol{x}}^2 \mathcal{L}_{\rho_k}^{\eta_k}(\boldsymbol{x}^k, \bar{\boldsymbol{\mu}}^k, \bar{\boldsymbol{\lambda}}^k) \succeq -\theta_k \mathbb{I}. \qquad (4)$$

Step 3 Calculate $\boldsymbol{\mu}^k \overset{\text{def}}{=} (\bar{\boldsymbol{\mu}}^k + \rho_k \boldsymbol{g}(\boldsymbol{x}^k))_+$ and $\boldsymbol{\lambda}^k \overset{\text{def}}{=} (\bar{\boldsymbol{\lambda}}^k + \rho_k \boldsymbol{h}(\boldsymbol{x}^k))$.

Step 4 For each $i \in I_p$, define

$$V_i^k \overset{\text{def}}{=} \min\left\{g_i(\boldsymbol{x}^k), -\frac{\bar{\mu}_i^k}{\rho_k}\right\}.$$

Step 5 If $k = 1$ or

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

do $\rho_{k+1} = \rho_k$. Otherwise, choose $\rho_{k+1} \geq \gamma \rho_k$. Also, take $\bar{\boldsymbol{\mu}}^{k+1} \in [0, \mu_{\max}]^p$ and $\bar{\boldsymbol{\lambda}}^{k+1} \in [\lambda_{\min}, \lambda_{\max}]^q$. Do $k \leftarrow k + 1$ and return to **Step 2**.

---

be able to find that satisfy the KKT conditions. In [14], it was established the weakest CQ for which it is possible to ensure the KKT conditions at the accumulation points of the first-order SALM, namely, *AL-regularity*. In particular, *AL-regularity* is implied by the *quasi-normality CQ* [19, pp. 337] and CRSC [8], two weak CQs used to guarantee convergence in first-order optimization methods [7, 8, 23].

The most accepted second-order optimality notion linked to the convergence of algorithms is the following:

Definition 2.3. We say that a feasible point $\boldsymbol{x}^*$ satisfies the *weak second-order condition (WSOC)* whenever there are vectors $\boldsymbol{\mu} \in \mathbb{R}_+^p$ and $\boldsymbol{\lambda} \in \mathbb{R}^q$ so that

$$\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \boldsymbol{0}, \quad \boldsymbol{\mu}^T \boldsymbol{g}(\boldsymbol{x}^*) = 0 \quad \text{and} \quad \boldsymbol{d}^T \nabla_{\boldsymbol{x}}^2 \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\mu}, \boldsymbol{\lambda}) \boldsymbol{d} \geq 0 \ \ \forall \boldsymbol{d} \in \mathscr{C}^W(\boldsymbol{x}^*),$$

where $\mathscr{C}^W(\boldsymbol{x}^*)$ is the weak critical cone (2). In this case, $\boldsymbol{x}^*$ is called a *WSOC point*.

In other words, WSOC holds whenever the KKT conditions hold with $(\boldsymbol{\mu}, \boldsymbol{\lambda})$ and the Hessian of the Lagrangian $\mathcal{L}(\cdot, \boldsymbol{\mu}, \boldsymbol{\lambda})$ is positive semidefinite in the subspace orthogonal to the gradients of active constraints at $\boldsymbol{x}^*$.

Since WSOC are stronger than KKT conditions, they are not necessarily valid in all local minimizers. For this, a suitable CQ can be imposed. The convergence analysis of algorithms involving WSOC usually revolves around rigorous CQs, such as MFCQ. However, using second-order sequential optimality conditions (SOC2), it has recently been shown that this can be relaxed, allowing, for instance, the sequences of multipliers generated by the algorithm to be unbounded. Two such conditions were considered: the *second-order approximate-KKT* (AKKT2) [15] condition and the *second-order complementarity-approximate-KKT* (CAKKT2) condition [24]. Optimization methods that employ second-order techniques, like SALM, can ensure both conditions; see [15, 24] for more methods.

When an appropriate property on the constraints holds at a point that satisfies a SOC2, such a point also satisfies WSOC. These properties were usually composed of a CQ together with a constant rank-type property, which we recall below.

Definition 2.4. A feasible point $\boldsymbol{x}^*$ is said to satisfy the *weak constant rank* (WCR) property if there exists a neighborhood V of $\boldsymbol{x}^*$ such that the rank of the space generated by $\{\nabla g_i(\boldsymbol{x})\}_{i \in \mathcal{A}(\boldsymbol{x}^*)} \cup \{\nabla h_j(\boldsymbol{x})\}_{j=1,\ldots,q}$ is constant for all $\boldsymbol{x} \in V$.

Given an arbitrary SOC2, the weakest associated CQ can be built by the techniques used in [15], which, in the case of AKKT2, is the *second-order cone continuity property* (CCP2) [15]. This CQ, together with WCR, guarantees WSOC. A similar discussion is valid for the weakest CQ associated with CAKKT2, namely *CAKKT2-regularity* [24].

The CQs required for convergence of methods like SALM are commonly related to the problem's stability. An instability happens when small changes in the constraints cause significant differences in the optimum values. A well-known property linked to the stability of a problem is the *error bound* [37, §3]. Additionally, methods usually request CQs that are stronger than the error bound. This work uses the validity of a variation of the error bound condition in order to guarantee WSOC. As said in the introduction, this is the first time a condition slightly stronger than the error bound has been used to reach such a strong result.

Definition 2.5. Let Ω be the feasible set of (1). We say that $\boldsymbol{x}^* \in \Omega$ meets the *error bound* (EB) CQ when there are constants $L > 0$ and $\delta > 0$ such that

$$\text{dist}(\boldsymbol{x}, \Omega) \leq L \left(\|\boldsymbol{g}(\boldsymbol{x})_+\| + \|\boldsymbol{h}(\boldsymbol{x})\| \right) \text{ for all } \boldsymbol{x} \in \mathcal{B}(\boldsymbol{x}^*, \delta).$$

We call L a local error bound constant for Ω at $\boldsymbol{x}^*$.

SALM constructs a sequence of approximations of the primal and dual solutions. Even though the primal sequence has an accumulation point, the dual sequence may be unbounded if no extra assumptions about the problem hold. In [16], it has been proposed a weak extension of the quasi-normality CQ called *relaxed quasinormality* that assures boundedness of the dual sequences of the method. We will use this condition in the following sections to obtain WSOC.

Definition 2.6 ([16]). It is said that the *relaxed quasinormality* (RQN) CQ holds at the point $\boldsymbol{x}^*$ whenever there is no sequence $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$ and vector $(\boldsymbol{\mu}, \boldsymbol{\lambda}) \in \mathbb{R}_+^p \times \mathbb{R}^q$ satisfying the following requirements:

1. $\sum_{i=1}^p \mu_i \nabla g_i(\boldsymbol{x}^k) + \sum_{j=1}^q \lambda_j \nabla h_j(\boldsymbol{x}^k) = \boldsymbol{0}$;

2. $(\boldsymbol{\mu}, \boldsymbol{\lambda}) \neq (\boldsymbol{0}, \boldsymbol{0})$;

3. for all $i \in I_p$ and $j \in I_q$,

$$\text{if } \mu_i \neq 0, \text{ then } g_i(\boldsymbol{x}^k) > 0; \quad \text{if } \lambda_j \neq 0, \text{ then } \lambda_j h_j(\boldsymbol{x}^k) > 0;$$

$$\text{if } \mu_i = 0, \text{ then } g_i(\boldsymbol{x}^k)_+ = o(t_k); \quad \text{if } \lambda_j = 0, \text{ then } h_j(\boldsymbol{x}^k) = o(t_k),$$

where $t_k = \min\{\min_{\mu_i > 0} g_i(\boldsymbol{x}^k)_+, \min_{\lambda_j \neq 0} |h_j(\boldsymbol{x}^k)|\}$.

Some CQs, necessarily not weaker than MFCQ, guarantee WSOC at local minimizes. An example is the *relaxed constant rank constraint qualification* (RCRCQ) [15, Proposition 4.12]. This CQ needs the local constant dimension of the subspace generated by many subsets of the gradients of active constraints. This was relaxed in the context of nonlinear conic programming with the *strong-constant rank of the subspace component* (strong-CRSC) [12, Definition 4.1], by reducing the number of subsets needed to obtain WSOC. In the next sections, we will prove, in addition, the global convergence of SALM under strong-CRSC.

Definition 2.7. Considering a feasible point $\boldsymbol{x}^*$ of (1), we define the set

$$\mathcal{J}_-(\boldsymbol{x}^*) = \left\{ i \in \mathcal{A}(\boldsymbol{x}^*) \left| \begin{array}{c} \exists \boldsymbol{\mu} \in \mathbb{R}_+^p, \ \exists \boldsymbol{\lambda} \in \mathbb{R}^q, \\ -\nabla g_i(\boldsymbol{x}^*) = \sum_{i \in \mathcal{A}(\boldsymbol{x}^*)} \mu_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j \in I_q} \lambda_j \nabla h_j(\boldsymbol{x}^*) \end{array} \right. \right\}.$$

Definition 2.8. The strong-CRSC condition is valid at a feasible point $\boldsymbol{x}^*$ of (1) when for every $\mathcal{J}_-(\boldsymbol{x}^*) \subset \mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$, the rank of the subspace generated by the set $\{\nabla g_i(\boldsymbol{x})\}_{i \in \mathcal{A}} \cup \{\nabla h_j(\boldsymbol{x})\}_{j \in I_q}$ remains constant in a neighborhood of $\boldsymbol{x}^*$.

3   A new second-order sequential optimality condition

In the last years, the global convergence of the first-order SALM has been improved by associating it with increasingly strengthened sequential optimality conditions [6, 7, 11, 15, 24, 35]. A recent proposal has provided a definitive answer to such an issue. In [14] (see also [13]), a sequential optimality condition inspired by the augmented Lagrangian method itself was proposed, namely Augmented Lagrangian AKKT (AL-AKKT). This condition provides the best possible global convergence result for the first-order SALM in a suitable sense. Motivated by these ideas, we define the second-order counterpart of AL-AKKT for SALM.

Definition 3.1. We say that a feasible point $\boldsymbol{x}^*$ for problem (1) fulfills the *second-order augmented Lagrangian AKKT* (AL-AKKT2) condition if there are sequences $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, $\{\bar{\boldsymbol{\mu}}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+^p$, $\{\bar{\boldsymbol{\lambda}}^k\} \subset \mathbb{R}^q$, $\{\rho_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_{++}$, $\{\theta_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+$ and $\{\eta_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+$ such that $\lim_{k\to\infty}\boldsymbol{x}^k = \boldsymbol{x}^*$, $\lim_{k\to\infty}\theta_k = \lim_{k\to\infty}\eta_k = 0$, the sequences $\{\bar{\boldsymbol{\mu}}^k\}_{k\in\mathbb{N}}$ and $\{\bar{\boldsymbol{\lambda}}^k\}_{k\in\mathbb{N}}$ are bounded,

$$\|\nabla_{\boldsymbol{x}}\mathcal{L}_{\rho_k}(\boldsymbol{x}^k,\bar{\boldsymbol{\mu}}^k,\bar{\boldsymbol{\lambda}}^k)\| \leq \theta_k, \tag{5}$$

$$\nabla_{\boldsymbol{x}}^2\mathcal{L}_{\rho_k}^{\eta_k}(\boldsymbol{x}^k,\bar{\boldsymbol{\mu}}^k,\bar{\boldsymbol{\lambda}}^k) \succeq -\theta_k\,\mathbb{I}, \text{ for all } k \in \mathbb{N}, \tag{6}$$

and

$$\lim_{k\to\infty}\left\|\min\left\{-g_i(\boldsymbol{x}^k), \frac{\bar{\mu}_i^k}{\rho_k}\right\}\right\| = 0, \text{ for all } i \in I_p. \tag{7}$$

As it is natural, we call $\{\rho_k\}_{k\in\mathbb{N}}$ the penalty parameter sequence and both $\{(\rho_k\boldsymbol{g}(\boldsymbol{x}^k) + \bar{\boldsymbol{\mu}}^k)_+\}_{k\in\mathbb{N}}$ and $\{\rho_k\boldsymbol{h}(\boldsymbol{x}^k) + \bar{\boldsymbol{\lambda}}^k\}_{k\in\mathbb{N}}$ the *AL-AKKT2 dual sequences* or *AL-AKKT2 multipliers*.

Remark 2. The AL-AKKT condition is defined exactly like the AL-AKKT2 condition but without requiring (6).

Theorem 3.2. *The AL-AKKT2 condition is valid in every local minimizer of* (1).

Proof. Follow the proof of [15, Proposition 1] noting that the constructed sequence satisfies the AL-AKKT2 properties. $\qquad\square$

Combining the previous result with the following one has an interesting practical interpretation. Both ensure that the AL-AKKT2 conditions are valid at local minimizers and WSOC points. Conceptually, they mean that the new conditions can be seen as a generalization of WSOC in contexts where CQs cannot be guaranteed.

Proposition 3.3. *If the feasible point $\boldsymbol{x}^*$ is a WSOC point, then $\boldsymbol{x}^*$ is an AL-AKKT2 point. Additionally, a sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ fulfills the SALM steps with $\boldsymbol{x}^*$ being a limit of this sequence.*

Proof. If WSOC holds, there are Lagrange multipliers $\boldsymbol{\mu} \in \mathbb{R}_+^p$ and $\boldsymbol{\lambda} \in \mathbb{R}^q$ such that

$$\nabla_x^2\mathcal{L}(\boldsymbol{x}^*,\boldsymbol{\mu},\boldsymbol{\lambda})^T\boldsymbol{d} \geq 0 \text{ for all } \boldsymbol{d} \in \mathscr{C}^W(\boldsymbol{x}^*).$$

Thus, by [18, Lemma 1.25], there exists $\rho_0 > 0$ such that

$$\nabla_x^2\mathcal{L}(\boldsymbol{x}^*,\boldsymbol{\mu},\boldsymbol{\lambda}) + \rho_0\left(\sum_{i\in\mathcal{A}(\boldsymbol{x}^*)}\nabla g_i(\boldsymbol{x}^*)\nabla g_i(\boldsymbol{x}^*)^T + \sum_{j=1}^q\nabla h_j(\boldsymbol{x}^*)\nabla h_j(\boldsymbol{x}^*)^T\right) \succeq \boldsymbol{0}.$$

Regarding the second part of the statement, if we choose $\mu_{\max}$, $\lambda_{\max}$ and $\lambda_{\min}$ to be big enough in modulus so that $\bar{\boldsymbol{\mu}}^1 = \boldsymbol{\mu} \in [0,\mu_{\max}]^p$ and $\bar{\boldsymbol{\lambda}}^1 = \boldsymbol{\lambda} \in [\lambda_{\min},\lambda_{\max}]^p$, we can start the method with $\boldsymbol{x}^1$ being $\boldsymbol{x}^*$. Thus, the primal sequence produced by the method will be constant, as **Step 2** (with the criterion (4)) is always satisfied with such feasible $\boldsymbol{x}^1$. $\qquad\square$

4   Convergence of SALM under an error bound condition

The work [25] suggests that second-order algorithms generally try to compute a negative-curvature direction from the second-order derivative of the Lagrangian function of Definition 2.2. However, this procedure can fail to identify correctly WSOC points if the sequences associated with the terms $\nabla g_i(\boldsymbol{x}^k)\nabla g_i(\boldsymbol{x}^k)^T$ and $\nabla h_i(\boldsymbol{x}^k)\nabla h_i(\boldsymbol{x}^k)^T$ in $\nabla_{\boldsymbol{x}}^2 \mathcal{L}_{\rho_k}^{\eta}(\boldsymbol{x}^k, \bar{\boldsymbol{\mu}}^k, \bar{\boldsymbol{\lambda}}^k)$ grows in an uncontrolled manner. This problem is present in AKKT2 sequences, making WCR necessary; see [25, Definition 1 and Proposition 1]. On the other hand, in the AL-AKKT2 condition, such terms are closely linked with the penalty parameter sequence $\{\rho_k\}_{k\in\mathbb{N}}$, see Definition 5. Since ρ_k appears in the dual sequences of the AL-AKKT2 condition, its growth is somehow limited by the growth of the dual sequences. This leads us to ask whether it is possible to guarantee convergence to WSOC points without requiring the WCR condition. This section shows it is possible.

Next, we will define the first two concepts required to derive the basic results of this section.

Definition 4.1. For each $\boldsymbol{x}^* \in \mathbb{R}^n$, $\boldsymbol{x} \in \mathbb{R}^n$ and $\mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$, define the vector space

$$\mathscr{C}_{\mathcal{A}}^W(\boldsymbol{x}, \boldsymbol{x}^*) = \left\{ \boldsymbol{d} \in \mathbb{R}^n \,\middle|\, \begin{array}{l} \langle \nabla h_j(\boldsymbol{x}), \boldsymbol{d} \rangle = 0 \text{ for all } j \in I_q \\ \langle \nabla g_i(\boldsymbol{x}), \boldsymbol{d} \rangle = 0 \text{ for all } i \in \mathcal{A} \end{array} \right\}$$

and the set

$$\Omega_{\mathcal{A}} = \{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{g}_{\mathcal{A}}(\boldsymbol{x}) = \boldsymbol{0}, \ \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \},$$

where $\boldsymbol{g}_{\mathcal{A}}(\boldsymbol{x}) = (g_i(\boldsymbol{x}))_{i\in\mathcal{A}}$. Note that $\mathscr{C}_{\mathcal{A}(\boldsymbol{x}^*)}^W(\boldsymbol{x}^*, \boldsymbol{x}^*) = \mathscr{C}^W(\boldsymbol{x}^*)$ (see (2)). Additionally, define the set

$$\mathcal{A}_{=0} = \{ i \in I_p \mid \exists \delta > 0, \forall \boldsymbol{y} \in \mathcal{B}(\boldsymbol{x}^*, \delta) \cap \Omega, \ g_i(\boldsymbol{y}) = 0 \}.$$

Next, we define a slightly stronger condition than EB, which we use to guarantee WSOC at points fulfilling the AL-AKKT2 conditions.

Definition 4.2. We say that the *strong error bound* (strong-EB) condition holds for Ω at the point $\boldsymbol{x}^*$ whenever Ω and $\Omega_{\mathcal{A}}$ fulfill EB at $\boldsymbol{x}^*$ for each $\mathcal{A}$ with $\mathcal{A}_{=0} \subset \mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$, that is, there are $L > 0$ and $\delta > 0$ such that

$$\text{dist}(\boldsymbol{x}, \Omega_{\mathcal{A}}) \leq L\left(\|\boldsymbol{g}_{\mathcal{A}}(\boldsymbol{x})\| + \|\boldsymbol{h}(\boldsymbol{x})\|\right) \quad \text{and} \quad \text{dist}(\boldsymbol{x}, \Omega) \leq L\left(\|\boldsymbol{g}(\boldsymbol{x})_+\| + \|\boldsymbol{h}(\boldsymbol{x})\|\right)$$

for all $\boldsymbol{x} \in \mathcal{B}(\boldsymbol{x}^*, \delta)$ and $\mathcal{A}$ with $\mathcal{A}_{=0} \subset \mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$.

Abadie's CQ holds at $\boldsymbol{x}^*$ when the linearized cone of Ω at $\boldsymbol{x}^*$ and the tangent cone to Ω at $\boldsymbol{x}^*$,

$$T_{\Omega}(\boldsymbol{x}^*) \stackrel{\text{def}}{=} \limsup_{t\downarrow 0} \frac{\Omega - \boldsymbol{x}^*}{t} = \{ \boldsymbol{v} \in \mathbb{R}^n \mid \exists\{\boldsymbol{v}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n, \exists\{t_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+^* \text{ such that}$$

$$\lim_{k\to\infty} \boldsymbol{v}^k = \boldsymbol{v}, \lim_{k\to\infty} t_k = 0 \text{ and } t_k \boldsymbol{v}^k + \boldsymbol{x}^* \in \Omega, \forall k \in \mathbb{N} \}$$

[36, Definition 6.1], coincide. It is known that WSOC can be ensured at local minimizers if Abadie's CQ is valid regarding the set $\Omega_{\mathcal{A}(\boldsymbol{x}^*)}$ [2, Theorem 3.1], which has $\mathscr{C}^W(\boldsymbol{x}^*)$ as the linearized cone at $\boldsymbol{x}^*$. The next technical lemma involves the validity of Abadie's CQ in an intermediate set $\Omega_{\mathcal{A}}$, see Remark 3.

Lemma 4.3. *For a set $\mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$, suppose that*

$$\liminf_{\substack{\boldsymbol{x} \longrightarrow \boldsymbol{x}^* \\ \Omega_{\mathcal{A}}}} T_{\Omega_{\mathcal{A}}}(\boldsymbol{x}) = \mathscr{C}_{\mathcal{A}}^W(\boldsymbol{x}^*, \boldsymbol{x}^*). \tag{8}$$

Then

$$\mathscr{C}_{\mathcal{A}(\boldsymbol{x}^*)}^W(\boldsymbol{x}^*, \boldsymbol{x}^*) \subset \liminf_{\substack{\boldsymbol{x} \longrightarrow \boldsymbol{x}^* \\ \Omega_{\mathcal{A}}}} \mathscr{C}_{\mathcal{A}}^W(\boldsymbol{x}, \boldsymbol{x}^*).$$

Proof. For a set $\mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$, $\boldsymbol{v} \in \mathscr{C}^W_{\mathcal{A}(\boldsymbol{x}^*)}(\boldsymbol{x}^*, \boldsymbol{x}^*)$ and a sequence $\{\boldsymbol{x}^k\} \subset \Omega_{\mathcal{A}}$, as $\mathscr{C}^W_{\mathcal{A}(\boldsymbol{x}^*)}(\boldsymbol{x}^*, \boldsymbol{x}^*) \subset \mathscr{C}^W_{\mathcal{A}}(\boldsymbol{x}^*, \boldsymbol{x}^*)$, by (8), there exists $\boldsymbol{v}^k \to \boldsymbol{v}$ with $\boldsymbol{v}^k \in T_{\Omega_{\mathcal{A}}}(\boldsymbol{x}^k)$ and, by the definition of $T_{\Omega_{\mathcal{A}}}(\boldsymbol{x}^k)$, we have $T_{\Omega_{\mathcal{A}}}(\boldsymbol{x}^k) \subset \mathscr{C}^W_{\mathcal{A}}(\boldsymbol{x}^k, \boldsymbol{x}^*)$, from which $\boldsymbol{v}^k \in \mathscr{C}^W_{\mathcal{A}}(\boldsymbol{x}^k, \boldsymbol{x}^*)$, and the statement follows. $\qquad \square$

Remark 3. Clarke's regularity is a well-known type of regularity in variational analysis and optimization [27, 28, 36]. Considering a closed subset $\mathcal{X} \subset \mathbb{R}^n$, Clarke's regularity is defined in [36, Definition 6.4] by asking for the equality

$$N_{\mathcal{X}}(\boldsymbol{x}^*) = \limsup_{\boldsymbol{x} \xrightarrow{\mathcal{X}} \boldsymbol{x}^*} \hat{N}_{\mathcal{X}}(\boldsymbol{x}^*) = \hat{N}_{\mathcal{X}}(\boldsymbol{x}^*),$$

where $\hat{N}_{\mathcal{X}}(\boldsymbol{x}^*)$ is the regular normal cone and $N_{\mathcal{X}}(\boldsymbol{x}^*)$ is the limiting (Mordukhovich) normal cone. In view of [36, Corollary 6.29], Clarke's regularity can be stated as

$$T_{\mathcal{X}}(\boldsymbol{x}^*) = \limsup_{t \downarrow 0} \frac{\mathcal{X} - \boldsymbol{x}^*}{t} = \liminf_{\boldsymbol{x} \xrightarrow{\mathcal{X}} \boldsymbol{x}^*, \ t \downarrow 0} \frac{\mathcal{X} - \boldsymbol{x}}{t} \stackrel{\text{def}}{=} \hat{T}_{\mathcal{X}}(\boldsymbol{x}^*),$$

where $\hat{T}_{\mathcal{X}}(\boldsymbol{x}^*)$ is the regular tangent cone (see [36, Definition 6.25] or [32, Definition 1.8]). We have $\hat{T}_{\mathcal{X}}(\boldsymbol{x}^*) \subset T_{\mathcal{X}}(\boldsymbol{x}^*)$, and, by [36, Theorem 6.26],

$$\liminf_{\boldsymbol{x} \xrightarrow{\mathcal{X}} \boldsymbol{x}^*} T_{\mathcal{X}}(\boldsymbol{x}) = \hat{T}(\boldsymbol{x}^*).$$

Therefore, condition (8) is equivalent to asking for Abadie's CQ and Clarke's regularity at $\boldsymbol{x}^*$ with respect to the set $\Omega_{\mathcal{A}}$.

Like EB, the boundedness of the dual sequences generated by the method is another important point for the stability of algorithms. RQN condition [16] (Definition 2.6) guarantees the boundedness of multipliers generated by first-order SALM; it is weaker than CRSC [8] and QN, indicating that the boundedness of AL-AKKT multipliers can be obtained under weak and typical CQs. The compelling thing about this is that WSOC can be brought under the boundedness of the approximate Lagrange multipliers and strong-EB, as shown in the next result.

Proposition 4.4. *Suppose that $\boldsymbol{x}^*$ conforms to strong-EB and that there is an AL-AKKT2 sequence converging to $\boldsymbol{x}^*$ with bounded AL-AKKT2 multipliers. Then WSOC holds at $\boldsymbol{x}^*$.*

Proof. Let us consider the AL-AKKT2 sequence $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$ associated with the bounded AL-AKKT2 multipliers $\{(\rho_k \boldsymbol{g}(\boldsymbol{x}) + \bar{\boldsymbol{\mu}})_+\}_{k \in \mathbb{N}}$ and $\{\rho_k \boldsymbol{h}(\boldsymbol{x}) + \bar{\boldsymbol{\lambda}}\}_{k \in \mathbb{N}}$. Then there are vectors $\boldsymbol{\mu} \in \mathbb{R}^p_+$, $\boldsymbol{\lambda} \in \mathbb{R}^q$ and an infinite set of indices $\mathcal{M} \subset \mathbb{N}$ such that

$$\lim_{k \in \mathcal{M}} \left(\bar{\boldsymbol{\mu}}^k + \rho_k \boldsymbol{g}(\boldsymbol{x}^k)\right)_+ = \boldsymbol{\mu} \qquad \text{and} \qquad \lim_{k \in \mathcal{M}} \bar{\boldsymbol{\lambda}}^k + \rho_k \boldsymbol{h}(\boldsymbol{x}^k) = \boldsymbol{\lambda}.$$

Also, the sequences $\{\rho_k \boldsymbol{g}(\boldsymbol{x}^k)_+\}_{k \in \mathbb{N}}$ and $\{\rho_k \boldsymbol{h}(\boldsymbol{x}^k)\}_{k \in \mathbb{N}}$ are bounded, so there is $K > 0$ such that

$$\rho_k \left(\|\boldsymbol{g}(\boldsymbol{x}^k)_+\| + \|\boldsymbol{h}(\boldsymbol{x}^k)\|\right) \leq K, \text{ for all } k \in \mathbb{N}. \tag{9}$$

Now, by a recursive construction, it is possible to obtain an infinite subset $\mathcal{K}$ of $\mathcal{M}$ so that, for all $i \in I_p$,

$$\lim_{k \in \mathcal{K}} \rho_k g_i(\boldsymbol{x}^k) \in \{-\infty\} \cup \mathbb{R}.$$

Consider the set of indices

$$\bar{\mathcal{A}} = \left\{ i \in I_p \ \middle| \ \lim_{k \in \mathcal{K}} \rho_k g_i(\boldsymbol{x}^k) \in \mathbb{R} \right\}. \tag{10}$$

We can, for each $k \in \mathcal{K}$, project $\boldsymbol{x}^k$ onto the feasible set Ω. Let us choose $\boldsymbol{y}^k \in \Omega$ to be one of such projections. We have that $\boldsymbol{y}^k \in \Omega$ and, since the projection function is non-expansive, $\boldsymbol{y}^k \to \boldsymbol{x}^*$. Also, if $i \notin \bar{\mathcal{A}}$, necessarily,

$$\lim_{k \in \mathcal{K}} \rho_k g_i(\boldsymbol{x}^k) = -\infty.$$

9

Observe that, for each $k \in \mathbb{N}$ big enough, $\rho_k |g_i(\boldsymbol{y}^k) - g_i(\boldsymbol{x}^k)| \le L_{g_i} \rho_k \|\boldsymbol{y}^k - \boldsymbol{x}^k\| = L_{g_i} \rho_k \operatorname{dist}(\boldsymbol{x}^k, \Omega) \le L_{g_i} L_\Omega \rho_k \left(\|\boldsymbol{g}(\boldsymbol{x}^k)_+\| + \|\boldsymbol{h}(\boldsymbol{x}^k)\| \right) \le L_{g_i} L_\Omega K$, where L_{g_i} is the local Lipschitz constant of the function g_i and L_Ω is a local error bound constant for Ω at $\boldsymbol{x}^*$ – strong-EB assumes EB, see Definition 4.2. Hence, for each $k \in \mathbb{N}$ big enough,

$$\rho_k g_i(\boldsymbol{y}^k) = \rho_k \left(g_i(\boldsymbol{y}^k) - g_i(\boldsymbol{x}^k) \right) + \rho_k g_i(\boldsymbol{x}^k) \le L_{g_i} L_\Omega K + \rho_k g_i(\boldsymbol{x}^k).$$

Thus, as $\lim_{k \in \mathcal{K}} \rho_k g_i(\boldsymbol{x}^k) = -\infty$, also $\lim_{k \in \mathcal{K}} \rho_k g_i(\boldsymbol{y}^k) = -\infty$. Consequently, since $\lim_{k \in \mathcal{K}} \boldsymbol{y}^k = \boldsymbol{x}^*$ and $\boldsymbol{y}^k \in \Omega$, it cannot be $i \in \mathcal{A}_{=0}$. This all means that $\mathcal{A}_{=0} \subset \bar{\mathcal{A}}$.

Since $\mathcal{A}_{=0} \subset \bar{\mathcal{A}} \subset \mathcal{A}(\boldsymbol{x}^*)$, it is possible to take $\mathcal{A} = \bar{\mathcal{A}}$ and see that, by the definition of the strong-EB condition, see Definition 4.2, EB is valid for the set $\Omega_{\bar{\mathcal{A}}}$ at $\boldsymbol{x}^*$. Consequently, there exists $L > 0$ such that, for all $k \in \mathcal{K}$ big enough,

$$\rho_k \operatorname{dist}(\boldsymbol{x}^k, \Omega_{\bar{\mathcal{A}}}) \le L \rho_k \left(\sum_{i \in \bar{\mathcal{A}}} |g_i(\boldsymbol{x}^k)| + \sum_{j=1}^{q} |h_j(\boldsymbol{x}^k)| \right).$$

Now, considering that all the limits appearing in the definition (10) exist and, consequently, the associated sequences are bounded, the right-hand side of the last equation must be bounded. Consequently, there exists $M > 0$ such that, for all $k \in \mathcal{K}$,

$$\rho_k \operatorname{dist}(\boldsymbol{x}^k, \Omega_{\bar{\mathcal{A}}}) \le M.$$

Now, for each $k \in \mathcal{K}$, we can define the set $\mathcal{I}_k \stackrel{\text{def}}{=} \{i \in I_p \mid \rho_k g_i(\boldsymbol{x}^k) + \bar{\mu}_i^k \ge -\eta_k\}$. In this setting, for each $i \in \mathcal{I}_k$ and $k \in \mathcal{K}$, $\rho_k g_i(\boldsymbol{x}^k) \ge -\eta_k - \bar{\mu}_i^k$. Hence, due to (9), $\rho_k |g_i(\boldsymbol{x}^k)| \le \max\{\eta_k + \bar{\mu}_i^k, K\}$. Consequently, $\mathcal{I}_k \subset \bar{\mathcal{A}}$.

As $\Omega_{\bar{\mathcal{A}}}$ fulfills error bound, by [31, Proposition 1], it holds

$$\hat{T}_{\Omega_{\bar{\mathcal{A}}}}(\boldsymbol{x}^*) = L_{\Omega_{\bar{\mathcal{A}}}}(\boldsymbol{x}^*).$$

Consequently, considering Remark 3, not only $\Omega_{\bar{\mathcal{A}}}$ satisfies Clarke regularity but also Abadie's constraint qualification.

Now, we have all the necessary ingredients to prove the validity of WSOC at $\boldsymbol{x}^*$. Thus, take $\boldsymbol{d} \in \mathscr{C}^W(\boldsymbol{x}^*) = \mathscr{C}^W_{\mathcal{A}(\boldsymbol{x}^*)}(\boldsymbol{x}^*, \boldsymbol{x}^*)$. Hence, considering the sequence AL-AKKT2 $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$, we can, for each $k \in \mathcal{K}$, project $\boldsymbol{x}^k$ onto set $\Omega_{\bar{\mathcal{A}}}$. Let $\bar{\boldsymbol{y}}^k$ be a projection. Observe that $\bar{\boldsymbol{y}}^k \in \Omega_{\bar{\mathcal{A}}}$ and, since the projection function is non-expansive, we have also that $\bar{\boldsymbol{y}}^k \to \boldsymbol{x}^*$. Thus, by Lemma 4.3, we find a sequence $\{\boldsymbol{d}^k\}_{k \in \mathbb{N}}$ converging to $\boldsymbol{d}$ such that $\boldsymbol{d}^k \in \mathscr{C}^W_{\bar{\mathcal{A}}}(\bar{\boldsymbol{y}}^k, \boldsymbol{x}^*)$ for all $k \in \mathbb{N}$. Consequently,

$$\nabla g_i(\bar{\boldsymbol{y}}^k)^T \boldsymbol{d}^k = 0 \text{ and } \nabla h_j(\bar{\boldsymbol{y}}^k)^T \boldsymbol{d}^k = 0, \text{ for all } i \in \bar{\mathcal{A}} \text{ and } j \in I_q.$$

Furthermore, as $\mathcal{I}_k \subset \bar{\mathcal{A}}$, it holds

$$
\begin{aligned}
\rho_k \sum_{\rho_k g_i(\boldsymbol{x}^k) + \bar{\mu}_i^k \ge -\eta_k} \left(\nabla g_i(\boldsymbol{x}^k)^T \boldsymbol{d}^k \right)^2 &= \rho_k \sum_{i \in \mathcal{I}_k} \left(\left(\nabla g_i(\boldsymbol{x}^k) - \nabla g_i(\bar{\boldsymbol{y}}^k) \right)^T \boldsymbol{d}^k \right)^2 \\
&\le \rho_k \sum_{i \in \bar{\mathcal{A}}} \left(\left(\nabla g_i(\boldsymbol{x}^k) - \nabla g_i(\bar{\boldsymbol{y}}^k) \right)^T \boldsymbol{d}^k \right)^2 \\
&\le \rho_k \sum_{i \in \bar{\mathcal{A}}} L_{g_i}^2 \left\| \boldsymbol{x}^k - \bar{\boldsymbol{y}}^k \right\|^2 \left\| \boldsymbol{d}^k \right\|^2 \\
&= \rho_k \operatorname{dist}(\boldsymbol{x}^k, \Omega_{\bar{\mathcal{A}}}) \sum_{i \in \bar{\mathcal{A}}} L_{g_i}^2 \|\boldsymbol{d}^k\|^2 \operatorname{dist}(\boldsymbol{x}^k, \Omega_{\bar{\mathcal{A}}}) \\
&\le M \sum_{i \in \bar{\mathcal{A}}} L_{g_i}^2 \|\boldsymbol{d}^k\|^2 \operatorname{dist}(\boldsymbol{x}^k, \Omega_{\bar{\mathcal{A}}}) \to 0
\end{aligned}
$$

10

and, similarly,

$$\rho_k \sum_{j=1}^{q} \left(\nabla h_j(\boldsymbol{x}^k)^T \boldsymbol{d}^k \right)^2 \leq M \sum_{j=1}^{q} L_{h_j}^2 \|\boldsymbol{d}^k\|^2 \operatorname{dist}(\boldsymbol{x}^k, \Omega_{\bar{\mathcal{A}}}) \to 0.$$

Finally, multiplying (6) by $\boldsymbol{d}^k$ on the right and $(\boldsymbol{d}^k)^T$ on the left, we obtain

$$(\boldsymbol{d}^k)^T \left(\nabla_{\boldsymbol{x}}^2 \mathcal{L}(\boldsymbol{x}^k, \left(\boldsymbol{\mu}^k + \rho_k \boldsymbol{g}(\boldsymbol{x}^k) \right)_+ , \boldsymbol{\lambda}^k + \rho_k \boldsymbol{h}(\boldsymbol{x}^k)) \right) \boldsymbol{d}^k$$

$$+ \rho_k \sum_{\rho_k g_i(\boldsymbol{x}^k) + \bar{\mu}_i^k \geq -\eta_k} \left(\nabla g_i(\boldsymbol{x}^k)^T \boldsymbol{d}^k \right)^2 + \rho_k \sum_{j=1}^{q} \left(\nabla h_j(\boldsymbol{x}^k)^T \boldsymbol{d}^k \right)^2 \geq -\theta_k \|\boldsymbol{d}^k\|^2$$

for each $k \in \mathcal{K}$. Taking the limit in $\mathcal{K}_2$, we get

$$\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{p} \mu_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j=1}^{q} \lambda_j \nabla h_j(\boldsymbol{x}^*) = \boldsymbol{0},$$

$$\boldsymbol{\mu}^T \boldsymbol{g}(\boldsymbol{x}^*) = 0, \quad \boldsymbol{\mu} \geq \boldsymbol{0},$$

$$\boldsymbol{d}^T \left(\nabla^2 f(\boldsymbol{x}^*) + \sum_{i=1}^{p} \mu_i \nabla^2 g_i(\boldsymbol{x}^*) + \sum_{j=1}^{q} \lambda_j \nabla^2 h_j(\boldsymbol{x}^*) \right) \boldsymbol{d} \geq 0,$$

which implies WSOC, since $\boldsymbol{d}$ is a vector that does not depend on $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$. $\qquad\square$

The following result gives a situation where WSOC is guaranteed at points satisfying AL-AKKT2 and strong-EB.

Corollary 4.5. *WSOC hold at AL-AKKT2 points under strong-EB and RQN.*

Proof. The expression of the AL-AKKT2 multipliers is the same as the multipliers generated by the first-order SALM. Consequently, only assuming RQN, we can follow the proof of [16, Theorems 3] and prove that the dual sequences of the AL-AKKT2 are bounded. Thus, strong-EB and RQN are enough to guarantee WSOC at AL-AKKT2 points. $\qquad\square$

The following result directly applies the previous one and shows that strong-CRSC is enough to guarantee WSOC at points satisfying the AL-AKKT2 conditions.

Proposition 4.6. *Strong-CRSC guarantees strong-EB and RQN.*

Proof. First, note that, by hypothesis, CRSC is valid. Thus, Ω not only satisfies EB at $\boldsymbol{x}^*$ but also RQN by [8, Theorem 5.5] and [30, Theorem 2.1]. Additionally, $\mathcal{J}_-(\boldsymbol{x}^*) \subset \mathcal{A}_{=0}$ by [8, Lemma 5.3]. Hence, let us fix $\mathcal{A}$ with $\mathcal{A}_{=0} \subset \mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$ and consider the set $\Omega_{\mathcal{A}}$. It is only necessary to prove the validity of CRSC at $\boldsymbol{x}^*$ applied to $\Omega_{\mathcal{A}}$, since CRSC implies error bound [8, Theorem 5.5]. As $\Omega_{\mathcal{A}}$ has only equality constraints, it is only necessary to prove the constant rank of the gradients of the active constraints in a neighborhood of $\boldsymbol{x}^*$, i.e., the constant rank of $\{\nabla g_j(\boldsymbol{x})\}_{i \in \mathcal{A}} \cup \{\nabla h_j(\boldsymbol{x})\}_{j \in I_q}$ in a neighborhood of $\boldsymbol{x}^*$. Notice that $\mathcal{J}_-(\boldsymbol{x}^*) \subset \mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$. Consequently, by Definition 2.8, strong-CRSC says that the dimension of the subspace generated by $\{\nabla g_j(\boldsymbol{x})\}_{i \in \mathcal{A}} \cup \{\nabla h_j(\boldsymbol{x})\}_{j \in I_q}$ is constant for $\boldsymbol{x}$ in a neighborhood of $\boldsymbol{x}^*$, as we wanted to prove. $\qquad\square$

The next result shows that the quasi-normality CQ for $\Omega_{\mathcal{A}(\boldsymbol{x}^*)}$ is enough to guarantee WSOC at points satisfying the AL-AKKT2 conditions.

Proposition 4.7. *If $\Omega_{\mathcal{A}(\boldsymbol{x}^*)}$ satisfies QN, then Ω satisfies strong-EB and RQN.*

Proof. Quasi-normality applied to the set $\Omega_{\mathcal{A}(\boldsymbol{x}^*)}$ implies the quasi-normality CQ applied to the set $\Omega_{\mathcal{A}}$ in which $\mathcal{A}$ satisfies $\mathcal{A} \subset \mathcal{A}(\boldsymbol{x}^*)$. Additionally, quasi-normality with respect to the feasible set Ω can be derived from the one at $\Omega_{\mathcal{A}(\boldsymbol{x}^*)}$. Consequently, not only quasi-normality and relaxed quasinormality hold, but also EB for $\Omega_{\mathcal{A}}$ and Ω by [30, Theorem 2.1]. Thus, strong-EB and RQN hold. $\qquad\square$

The following example shows that the hypothesis of Proposition 4.7 does not depend on the WCR condition. Thus, although the results previously known in the literature are not broad enough to guarantee WSOC without asking for WCR, the following result shows that this is no longer true.

Example 4.8. Consider the set $\Omega \subset \mathbb{R}^2$ formed by the constraints from [24, Example 1]: $h(x_1, x_2) = x_1 = 0$, $g_1(x_1, x_2) = x_1^3 \leq 0$ and $g_2(x_1, x_2) = x_1 e^{x_2} \leq 0$. Additionally, take the point $\boldsymbol{x}^* = (0, 0)^T$. The set Ω does not meet the WCR condition at the origin, but the set $\Omega_{\mathcal{A}(\boldsymbol{x}^*)}$ fulfills quasi-normality.

Next we show that WCR and strong-EB are in fact independent conditions. In the following example, WCR holds while strong-EB fails.

Example 4.9. Considering the functions $c_1, c_2, c_3 \colon \mathbb{R}^2 \to \mathbb{R}$ such that $c_1(x_1, x_2) = -x_1$, $c_2(x_1, x_2) = x_2 - x_1^2$ and $c_3(x_1, x_2) = -x_2 - x_1^2$ for each $(x_1, x_2)^T \in \mathbb{R}^2$, the set $\Omega = \{(x_1, x_2)^T \in \mathbb{R}^2 \mid c_1(x_1, x_2) \leq 0,\ c_2(x_1, x_2) \leq 0,\ c_3(x_1, x_2) \leq 0\}$ meets the WCR conditions. However, considering the subset $\Omega_{\mathcal{A}}$ with $\mathcal{A} = \{2, 3\}$, we have that $\Omega_{\mathcal{A}} = \{\boldsymbol{0}\}$, also, taking the sequence $\boldsymbol{x}^k = (1/k, 0)^T$, for all $k \in \mathbb{N}$, $\mathrm{dist}(\boldsymbol{x}^k, \Omega_{\mathcal{A}}) = \|\boldsymbol{x}^k\| = 1/k$, but $c_1(\boldsymbol{x}^k)_+ = 0$, $|c_2(\boldsymbol{x}^k)| = (1/k)^2$ and $|c_3(\boldsymbol{x}^k)| = (1/k)^2$. Hence,

$$\lim_{k \to \infty} \frac{\mathrm{dist}(\boldsymbol{x}^k, \Omega_{\mathcal{A}})}{c_1(\boldsymbol{x}^k)_+ + |c_2(\boldsymbol{x}^k)| + |c_3(\boldsymbol{x}^k)|} = \infty.$$

Consequently, the set $\Omega_{\mathcal{A}}$ does not meet the error bound property, and strong-EB is violated.

The previous example shows that some problems are not naturally qualified concerning the local EB-type condition. Hence, presenting convergence results for WSOC points involving the WCR condition is still valid. Next, we use the boundedness of AL-AKKT2 multipliers and WCR conditions to obtain WSOC, as done in [15, 24].

Proposition 4.10. *WCR and boundedness of the AL-AKKT2 multipliers are sufficient for WSOC.*

Proof. By the boundedness of the AL-AKKT2 multipliers there are vectors $\boldsymbol{\mu} \in \mathbb{R}_+^p$ and $\boldsymbol{\lambda} \in \mathbb{R}^q$ such that $\lim_{k \in \mathcal{K}} (\bar{\boldsymbol{\mu}}^k + \rho_k \boldsymbol{g}(\boldsymbol{x}^k))_+ = \boldsymbol{\mu}$ and $\lim_{k \in \mathcal{K}} \bar{\boldsymbol{\lambda}}^k + \rho_k \boldsymbol{h}(\boldsymbol{x}^k) = \boldsymbol{\lambda}$ for some $\mathcal{K} \subset \mathbb{N}$. Now, given $\boldsymbol{d} \in \mathscr{C}^W(\boldsymbol{x}^*)$ we have that the WCR condition guarantees, by [10, Lemma 3.1],

$$\liminf_{\boldsymbol{x} \to \boldsymbol{x}^*} \mathscr{C}_{\mathcal{A}(\boldsymbol{x}^*)}^W(\boldsymbol{x}, \boldsymbol{x}^*) = \mathscr{C}^W(\boldsymbol{x}^*),$$

and thus, considering the AL-AKKT2 sequence $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$, there is a sequence $\{\boldsymbol{d}^k\}_{k \in \mathbb{N}}$ such that $\boldsymbol{d}^k \to \boldsymbol{d}$, $\nabla g_i(\boldsymbol{x}^k)^T \boldsymbol{d}^k = 0$ and $\nabla h_j(\boldsymbol{x}^k)^T \boldsymbol{d}^k = 0$, for all $i \in \mathcal{A}(\boldsymbol{x}^*)$ and $j \in I_q$. Therefore, the result follows similarly to the proof of Proposition 4.4. $\qquad\square$

The following two results give practical situations where WSOC can be guaranteed at points satisfying AL-AKKT2.

Corollary 4.11. *WSOC holds at AL-AKKT2 points under WCR and RQN.*

Proof. The boundedness of the AL-AKKT2 multipliers can be guaranteed under RQN, as discussed in the proof of the Corollary 4.5. Thus, we have, by Proposition 4.10, the validity of WSOC. $\qquad\square$

Corollary 4.12. *Strong-CRSC implies WCR and RQN simultaneously.*

Proof. By Definition 2.8, strong-CRSC ensures that the rank of the Jacobian matrix with only the active constraints is constant. Consequently, WCR holds. Additionally, strong-CRSC implies CRSC, which implies RQN [16, Theorem 2]. $\qquad\square$

5 The strongest possible sequential optimality condition for SALM

In this section, we show that AL-AKKT2 is the strongest second-order sequential optimality condition related to SALM. Then, using AL-AKKT2, we can provide the weakest possible CQ for which SALM guarantees WSOC. Part of the theory presented here could be seen as the counterpart of the theory for first-order SALM discussed in [14].

First, we prove that, as expected by the design of Definition 3.1, SALM's feasible accumulation points meet the AL-AKKT2 condition.

Theorem 5.1. *The AL-AKKT2 condition is valid at every feasible accumulation point of SALM.*

Proof. Let $\boldsymbol{x}^*$ be a feasible accumulation point of SALM. In step k, an unconstrained optimization algorithm is applied to (3) to obtain $\boldsymbol{x}^k$ that fulfills (4). Hence, the first and second-order stationarity conditions related to AL-AKKT2 are valid.

It remains to prove (7), that is, $\lim_{k\to\infty} \|\boldsymbol{V}^k\|_\infty = 0$. If the sequence $\{\rho_k\}_{k\in\mathbb{N}}$ generated by the method is bounded, **Step 5** and $r < 1$ guarantee that $\lim_{k\to\infty} \|\boldsymbol{V}^k\|_\infty = 0$. If $\{\rho_k\}_{k\in\mathbb{N}}$ is unbounded then by the definition of V_i^k we have $|V_i^k - g_i(\boldsymbol{x}^k)_+| \le \bar{\mu}_i^k/\rho_k$ for all $k \in \mathbb{N}$. Consequently, the boundedness of $\{\bar{\mu}_i^k\}_{k\in\mathbb{N}}$ and the feasibility of $\boldsymbol{x}^*$ imply $\lim_{k\to\infty} V_i^k = 0$ for all $i \in I_p$. This concludes the proof. $\square$

When applying a second-order method, the algorithm should obtain accumulation points fulfilling a second-order stationarity condition in at least one good context. In the case of methods that penalize infeasible constraints, such a context is when the sequence of penalty parameters ($\{\rho_k\}_{k\in\mathbb{N}}$) is bounded. The following result shows that it is possible to guarantee WSOC under the boundedness of this sequence.

Proposition 5.2. *WSOC is valid at points fulfilling the AL-AKKT2 conditions whenever its associated sequence $\{\rho_k\}_{k\in\mathbb{N}}$ is bounded.*

Proof. Let us fix the point $\boldsymbol{x}^*$ for which AL-AKKT2 holds, and consider an associated sequence $\{(\rho_k, \boldsymbol{x}^k, \bar{\boldsymbol{\lambda}}^k, \bar{\boldsymbol{\mu}}^k)\}$ as in Definition 3.1. By the hypothesis, the AL-AKKT2 multipliers $\boldsymbol{\lambda}^k = \bar{\boldsymbol{\lambda}}^k + \rho_k \boldsymbol{h}(\boldsymbol{x}^k)$ and $\boldsymbol{\mu}^k = (\bar{\boldsymbol{\mu}}^k + \rho_k \boldsymbol{g}(\boldsymbol{x}^k))_+$ are bounded. Thus, there exists an infinite subset $\mathcal{K}_0 \subset \mathbb{N}$ such that

$$\lim_{k\in\mathcal{K}}(\boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) = (\boldsymbol{\lambda}, \boldsymbol{\mu}) \quad \text{and} \quad \lim_{k\in\mathcal{K}_0} \rho_k = \rho_\infty,$$

with $\boldsymbol{\mu} \ge 0$ and $\rho_\infty > 0$. As $\lim_{k\to\infty}\{-g_i(\boldsymbol{x}^k), \bar{\mu}_i^k/\rho_k\} = 0$ for all i, we have $\lim_{k\to\infty} \bar{\mu}_i^k/\rho_k = 0$ whenever $g_i(\boldsymbol{x}^*) < 0$. Thus, $\mu_i = \lim_{k\to\infty} \rho_k((\bar{\mu}_i^k/\rho_k) + g_i(\boldsymbol{x}^k))_+ = 0$ whenever $g_i(\boldsymbol{x}^*) < 0$, and therefore the KKT complementarity is satisfied. Taking the limit in (5) over $\mathcal{K}$, we conclude that $(\boldsymbol{x}^*, \boldsymbol{\lambda}, \boldsymbol{\mu})$ is a KKT triple.

Now, the second-order stationarity can be seen taking $\boldsymbol{d} \in \mathscr{C}^W(\boldsymbol{x}^*)$, multiplying both sides of equation (6) by $\boldsymbol{d}$ and taking the limit in the resulting expression. Since $\boldsymbol{d}$ does not depend on $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, this ends the proof. $\square$

Next, we show that when the AL-AKKT2 condition is ensured, it is possible to find a sequence accepted by SALM for a particular choice of input parameters. This is consistent with previous results for the first-order SALM discussed in [14]. In other words, AL-AKKT2 provides the strongest global convergence result for SALM, from which we derive the least stringent CQ associated with the method, discussed in the next section.

Theorem 5.3. *Let $\boldsymbol{x}^*$ be a feasible point of problem* (1) *satisfying the AL-AKKT2 conditions, then, for a proper choice of input parameters, there is a convergent sequence to $\boldsymbol{x}^*$ fulfilling the steps of SALM.*

Proof. If WSOC holds at $\boldsymbol{x}^*$, then the statement follows from Proposition 3.3.

Let us assume that WSOC does not hold at $\boldsymbol{x}^*$ and take $\{(\rho_k, \boldsymbol{x}^k, \bar{\boldsymbol{\lambda}}^k, \bar{\boldsymbol{\mu}}^k)\}$ a sequence associated with the AL-AKKT2 conditions. By Proposition 5.2, $\{\rho_k\}_{k\in\mathbb{N}}$ is unbounded. Fixing any $r < 1$ and $\gamma > 1$, there exists a strictly increasing sequence of natural numbers $\{\sigma_k\}_{k\in\mathbb{N}}$ such that

$$\rho_{\sigma_{k+1}} \geq \gamma \rho_{\sigma_k}, \text{ for all } k \in \mathbb{N},$$

$$\max\left\{\|\boldsymbol{V}^{\sigma_k}\|_\infty, \|\boldsymbol{h}(\boldsymbol{x}^{\sigma_k})\|_\infty\right\} \leq r \max\left\{\|\boldsymbol{V}^{\sigma_{k-1}}\|_\infty, \|\boldsymbol{h}(\boldsymbol{x}^{\sigma_{k-1}})\|_\infty\right\}, \text{ for all } k \in \mathbb{N},$$

$$\lim_{k\to\infty} \rho_{\sigma_k} g_i(\boldsymbol{x}^{\sigma_k}) \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}, \text{ for all } i \in I_p,$$

$$\text{and} \quad \lim_{k\in\mathcal{K}_1} \rho_{\sigma_k} h_j(\boldsymbol{x}^{\sigma_k}) \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}, \text{ for all } j \in I_q.$$

Define the new sequence $\{\bar{\sigma}_k\}_{k\in\mathbb{N}}$ by putting

$$\bar{\sigma}_k \overset{\text{def}}{=} \begin{cases} \sigma_{k/2} & , \text{ if } k \text{ is even,} \\ \sigma_{(k+1)/2}, & \text{ if } k \text{ is odd,} \end{cases}$$

for all $k \in \mathbb{N}$. As can be seen, the terms repeat for two iterations before changing their value:

$$\{\bar{\sigma}_1, \bar{\sigma}_2, \bar{\sigma}_3, \bar{\sigma}_4, \bar{\sigma}_5, \bar{\sigma}_6, \cdots\} = \{\sigma_1, \sigma_1, \sigma_2, \sigma_2, \sigma_3, \sigma_3, \cdots\}. \tag{11}$$

Let us take, for all $k \in \mathbb{N}$,

$$\bar{\rho}_k \overset{\text{def}}{=} \rho_{\bar{\sigma}_k}, \quad \boldsymbol{y}^k \overset{\text{def}}{=} \boldsymbol{x}^{\bar{\sigma}_k}, \quad \bar{\boldsymbol{\mu}}^k \overset{\text{def}}{=} \bar{\boldsymbol{\mu}}^{\bar{\sigma}_k}, \quad \bar{\bar{\boldsymbol{\lambda}}}^k \overset{\text{def}}{=} \bar{\boldsymbol{\lambda}}^{\bar{\sigma}_k}, \tag{12}$$

$$\bar{\boldsymbol{V}}^k \overset{\text{def}}{=} \boldsymbol{V}^{\bar{\sigma}_k} = \min\left\{\boldsymbol{c}(\boldsymbol{x}^{\bar{\sigma}_k}), -\frac{\bar{\boldsymbol{\mu}}^{\bar{\sigma}_k}}{\rho_{\bar{\sigma}_k}}\right\}.$$

We proceed by showing that these sequences respect the SALM steps. We analyse separately when k is even or odd.

If k is even, we have $k = 2\ell$ for some ℓ and thus $\bar{\sigma}_k = \bar{\sigma}_{k-1}$ (see (11)). Furthermore,

$$\bar{\sigma}_{k+1} = \bar{\sigma}_{2\ell+1} = \sigma_{((2\ell+1)+1)/2} = \sigma_{\ell+1} \quad \text{and} \quad \bar{\sigma}_k = \bar{\sigma}_{2\ell} = \sigma_{2\ell/2} = \sigma_\ell.$$

Therefore, $\bar{\rho}_{k+1} = \rho_{\bar{\sigma}_{k+1}} = \rho_{\sigma_{\ell+1}} \geq \gamma \rho_{\sigma_\ell} = \gamma \rho_{\bar{\sigma}_k} = \gamma \bar{\rho}_k$ and

$$\begin{aligned} \max\left\{\|\bar{\boldsymbol{V}}^k\|_\infty, \|\boldsymbol{h}(\boldsymbol{y}^k)\|_\infty\right\} &= \max\left\{\|\boldsymbol{V}^{\bar{\sigma}_k}\|_\infty, \|\boldsymbol{h}(\boldsymbol{x}^{\bar{\sigma}_k})\|_\infty\right\} \\ &= \max\left\{\|\boldsymbol{V}^{\bar{\sigma}_{k-1}}\|_\infty, \|\boldsymbol{h}(\boldsymbol{x}^{\bar{\sigma}_{k-1}})\|_\infty\right\} \\ &= \max\left\{\|\bar{\boldsymbol{V}}^{k-1}\|_\infty, \|\boldsymbol{h}(\boldsymbol{y}^{k-1})\|_\infty\right\}, \end{aligned}$$

agreeing with **Step 5** from SALM.

Now, suppose that $k = 2\ell + 1$ for some $\ell \in \mathbb{N}$. Similarly,

$$\bar{\sigma}_{k-1} = \bar{\sigma}_{(2\ell+1)-1} = \bar{\sigma}_{2\ell} = \sigma_{2\ell/2} = \sigma_\ell \quad \text{and}$$

$$\bar{\sigma}_{k+1} = \bar{\sigma}_k = \bar{\sigma}_{(2\ell+1)+1} = \sigma_{((2\ell+1)+1)/2} = \sigma_{\ell+1}$$

and so

$$\begin{aligned} \max\left\{\|\bar{\boldsymbol{V}}^k\|_\infty, \|\boldsymbol{h}(\boldsymbol{y}^k)\|_\infty\right\} &= \max\left\{\|\boldsymbol{V}^{\bar{\sigma}_k}\|_\infty, \|\boldsymbol{h}(\boldsymbol{x}^{\bar{\sigma}_k})\|_\infty\right\} \\ &= \max\left\{\|\boldsymbol{V}^{\sigma_{\ell+1}}\|_\infty, \|\boldsymbol{h}(\boldsymbol{x}^{\sigma_{\ell+1}})\|_\infty\right\} \\ &\leq r \max\left\{\|\boldsymbol{V}^{\sigma_\ell}\|_\infty, \|\boldsymbol{h}(\boldsymbol{x}^{\sigma_\ell})\|_\infty\right\} \\ &= r \max\left\{\|\boldsymbol{V}^{\bar{\sigma}_{k-1}}\|_\infty, \|\boldsymbol{h}(\boldsymbol{x}^{\bar{\sigma}_{k-1}})\|_\infty\right\} \end{aligned}$$

14

$$= r \max \left\{ \|\bar{\boldsymbol{V}}^{k-1}\|_\infty, \|\boldsymbol{h}(\boldsymbol{y}^{k-1})\|_\infty \right\},$$

which corroborates $\bar{\rho}_{k+1} = \bar{\rho}_k$. Thus, **Step 5** is being respected.

The validity of the first- and second-order stationarity stated in equations (5) and (6) can be seen by noticing – considering expression (11) – that the sequences $\{\boldsymbol{y}^k\}$, $\{(\bar{\bar{\boldsymbol{\mu}}}^k + \bar{\rho}_k \boldsymbol{g}(\boldsymbol{y}^k))_+\}_{k \in \mathbb{N}}$, $\{\bar{\bar{\boldsymbol{\lambda}}}^k + \bar{\rho}_k \boldsymbol{h}(\boldsymbol{y}^k)\}_{k \in \mathbb{N}}$ may be regarded as subsequences with repeated terms of the original AL-AKKT2 sequences $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$, $\{(\bar{\boldsymbol{\mu}}^k + \rho_k \boldsymbol{g}(\boldsymbol{x}^k))_+\}_{k \in \mathbb{N}}$ and $\{\bar{\boldsymbol{\lambda}}^k + \rho_k \boldsymbol{h}(\boldsymbol{x}^k)\}_{k \in \mathbb{N}}$. Hence, the defined "subsequences" in (12) must also be AL-AKKT2 sequences.

The ideas regarding SALM's input parameters are straightforward. Here, it is only needed to choose $\mu_{\max}$, $\lambda_{\max}$ and $\lambda_{\min}$ to be big enough in modulus so that the bounded sequences satisfy $\bar{\bar{\boldsymbol{\mu}}}^k \in [0, \mu_{\max}]^p$ and $\bar{\bar{\boldsymbol{\lambda}}}^k \in [\lambda_{\min}, \lambda_{\max}]^q$ for all $k \in \mathbb{N}$. Notice that the sequences $\{\theta_{\bar{\sigma}_k}\}_{k \in \mathbb{N}}$ and $\{\eta_{\bar{\sigma}_k}\}_{k \in \mathbb{N}}$ in SALM can be taken as those associated with AL-AKKT2 (compare SALM's requirements with Definition 3.1). In such a way, $\{\boldsymbol{y}^k\}_{k \in \mathbb{N}}$ respects all the steps of SALM, concluding the proof. $\qquad \square$

The next result is a direct consequence of Theorems 5.1 and 5.3.

Corollary 5.4. *AL-AKKT2 is the strongest sequential optimality condition valid at SALM accumulation points. In particular, it implies AKKT2.*

Remark 4. An analogous result to the above corollary is valid for CAKKT2 instead of AKKT2, provided that the quadratic-like penalty measure of infeasibility associated with problem (1) satisfies the generalized Kurdyka-Lojasiewicz (GKL) inequality [11], a mild hypothesis valid, for example, if the data functions of the problem are subanalytic. GKL is required for SALM to converge to CAKKT/CAKKT2 points, see [11, 24].

5.1   The least stringent CQ associated with SALM

In this section, we present the least stringent CQ associated with AL-AKKT2. In other words, in view of Corollary 5.4 we state the weakest possible CQ for which SALM accumulation points fulfill WSOC. Evidently, the weaker the CQ used, the better the method's theoretical reliability, so ultimately, we are expanding the class of problems to which SALM converges.

As we already mentioned, whenever a SOC2 is proposed, it is possible to construct its weakest companion CQ following the techniques in [9] or [15]. Next, we recall some concepts necessary to build such a CQ for AL-AKKT2.

Consider a multivalued function $\mathcal{F} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $C \subset \mathbb{R}^n$. The *Painlevé–Kuratowski outer/upper limit* [36, p. 152] of $\mathcal{F}(\boldsymbol{x})$ as $\boldsymbol{x} \to \boldsymbol{x}^*$ is the set

$$\limsup_{\boldsymbol{x} \xrightarrow{C} \boldsymbol{x}^*} \mathcal{F}(\boldsymbol{x}) \stackrel{\text{def}}{=} \{\boldsymbol{v} \in \mathbb{R}^n \mid \exists \{\boldsymbol{x}^k\}_{k \in \mathbb{N}}, \exists \{\boldsymbol{v}^k\}_{k \in \mathbb{N}} \text{ such that}$$

$$\lim_{k \to \infty} \boldsymbol{v}^k = \boldsymbol{v}, \lim_{k \to \infty} \boldsymbol{x}^k = \boldsymbol{x}^*, \boldsymbol{x}^k \in C, \forall k \in \mathbb{N}, \text{ and } \boldsymbol{v}^k \in \mathcal{F}(\boldsymbol{x}^k), \forall k \in \mathbb{N}\}.$$

When it holds $\limsup_{\boldsymbol{x} \xrightarrow{C} \boldsymbol{x}^*} \mathcal{F}(\boldsymbol{x}) \subset \mathcal{F}(\boldsymbol{x}^*)$, the multifunction $\mathcal{F}$ is said to be outer semicontinuous in $\boldsymbol{x}^*$ relative to C.

Below, we define the condition we call the weakest CQ for obtaining WSOC points associated with SALM's accumulation points.

Definition 5.5. Define the multivalued function $K^W : \mathbb{R}^n \times \mathbb{R}^p_+ \times \mathbb{R}^q \times \mathbb{R}_{++} \times \mathbb{R}_{++} \rightrightarrows \mathbb{R}^n$ such that

$$K^W(\boldsymbol{x}, \bar{\bar{\boldsymbol{\mu}}}, \bar{\bar{\boldsymbol{\lambda}}}, \rho, \delta) \stackrel{\text{def}}{=}$$

$$
\left\{
\begin{array}{l}
\left(\displaystyle\sum_{i=1}^{p} (\mu_i)_+ \nabla g_i(\boldsymbol{x}) + \sum_{j=1}^{q} \lambda_j \nabla h_j(\boldsymbol{x}), H \right) \in \mathbb{R}^n \times \mathrm{Sym}(n) \mid \\[4mm]
H \preceq \displaystyle\sum_{i=1}^{p} (\mu_i)_+ \nabla^2 g_i(\boldsymbol{x}) + \sum_{j=1}^{q} \lambda_j \nabla^2 h_j(\boldsymbol{x}) \\[4mm]
\quad + \rho \displaystyle\sum_{\mu_i \geq -\delta} \nabla g_i(\boldsymbol{x}) \nabla g_i(\boldsymbol{x})^T + \rho \sum_{j=1}^{q} \nabla h_j(\boldsymbol{x}) \nabla h_j(\boldsymbol{x})^T, \text{ where} \\[4mm]
\boldsymbol{\mu} = \bar{\bar{\boldsymbol{\mu}}} + \rho\, \boldsymbol{g}(\boldsymbol{x}) \qquad \text{and} \qquad \boldsymbol{\lambda} = \bar{\bar{\boldsymbol{\lambda}}} + \rho\, \boldsymbol{h}(\boldsymbol{x})
\end{array}
\right\},
$$

for all $\boldsymbol{x} \in \mathbb{R}^n$, $\bar{\boldsymbol{\mu}} \in \mathbb{R}_+^p$, $\bar{\bar{\boldsymbol{\lambda}}} \in \mathbb{R}^q$, $\rho \in \mathbb{R}_{++}$ and $\delta \in \mathbb{R}_{++}$. We say that a feasible point $\boldsymbol{x}^*$ fulfills the *AL2-regularity CQ* when, for all $\bar{\boldsymbol{\mu}} \in \mathbb{R}_+^p$ and $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^q$, the multivalued function fulfills

$$
\limsup_{(\boldsymbol{x},\bar{\boldsymbol{\mu}},\bar{\bar{\boldsymbol{\lambda}}},\rho,\delta) \to (\boldsymbol{x}^*,\bar{\boldsymbol{\mu}},\bar{\boldsymbol{\lambda}},+\infty,0^+)} K^W(\boldsymbol{x},\bar{\boldsymbol{\mu}},\bar{\bar{\boldsymbol{\lambda}}},\rho,\delta) \subset K^2(\boldsymbol{x}^*),
$$

where

$$
K^2(\boldsymbol{x}^*) \overset{\mathrm{def}}{=} \left\{
\begin{array}{c}
(\sum_{i=1}^{p} \mu_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j=1}^{q} \lambda_j \nabla h_j(\boldsymbol{x}^*), H) \in \mathbb{R}^n \times \mathrm{Sym}(n)| \\
\boldsymbol{\mu} \in \mathbb{R}_+^p, \quad \text{and} \quad \boldsymbol{\lambda} \in \mathbb{R}^q \\
H \preceq \sum_{i=1}^{p} \mu_i \nabla^2 g_i(\boldsymbol{x}^*) + \sum_{j=1}^{q} \lambda_j \nabla^2 h_j(\boldsymbol{x}^*) \text{ in } \mathscr{C}^W(\boldsymbol{x}^*) \\
\text{and} \qquad \boldsymbol{g}(\boldsymbol{x}^*)^T \boldsymbol{\mu} = 0
\end{array}
\right\}.
$$

Remark 5. Considering the previous definition, note that the point $\boldsymbol{x}^*$ satisfies WSOC if and only if $(-\nabla f(\boldsymbol{x}^*), -\nabla^2 f(\boldsymbol{x}^*)) \in K^2(\boldsymbol{x}^*)$.

Guaranteeing that AL2-regularity is sufficient to ensure WSOC under AL-AKKT2 shows that it is, in fact, a CQ since AL-AKKT2 is valid in every minimizer. On the other hand, the reciprocal of this result guarantees that any CQ that guarantees WSOC under the validity of AL-AKKT2 must be stronger than (or equivalent to) AL2-regularity. These two results are fundamental to achieving our objective.

Theorem 5.6. *The point $\boldsymbol{x}^*$ fulfills the AL2-regularity CQ if and only if, for every twice-differentiable objective function f considered in (1), if $\boldsymbol{x}^*$ satisfies AL-AKKT2 with respect to this objective function, it also satisfies WSOC.*

Proof. Suppose that $\boldsymbol{x}^*$ meets with AL2-regularity and that it fulfills AL-AKKT2. Let $\{(\boldsymbol{x}^k, \bar{\boldsymbol{\mu}}^k, \bar{\boldsymbol{\lambda}}^k, \rho_k, \theta_k, \eta_k)\}_{k \in \mathbb{N}}$ be a sequence fulfilling Definition 3.1. We analyze the cases where $\{\rho_k\}_{k \in \mathbb{N}}$ has a bounded subsequence or not separately.

If $\{\rho_k\}_{k \in \mathbb{N}}$ has a bounded subsequence, by taking a subsequence, we may assume that $\{\rho_k\}_{k \in \mathbb{N}}$ is convergent since subsequences of AL-AKKT2 sequences are also AL-AKKT2 sequences. Then, by Proposition 5.2, WSOC is held, and the necessary result is proven. If no subsequence of $\{\rho_k\}_{k \in \mathbb{N}}$ is bounded, necessarily $\lim_{k \to \infty} \rho_k = \infty$. Taking $\delta_k = \eta_k$ for all $k \in \mathbb{N}$, we have that $(-\nabla f(\boldsymbol{x}^k) + \boldsymbol{v}^k, -\nabla^2 f(\boldsymbol{x}^k) - \theta_k \mathbb{I}) \in K^W(\boldsymbol{x}^k, \bar{\boldsymbol{\mu}}^k, \bar{\boldsymbol{\lambda}}^k, \rho_k, \delta_k)$. Thus, as $\{\bar{\boldsymbol{\mu}}^k\}_{k \in \mathbb{N}}$ and $\{\bar{\boldsymbol{\lambda}}^k\}_{k \in \mathbb{N}}$ are bounded, we can assume, without losing generality, that $\{\bar{\boldsymbol{\mu}}^k\}_{k \in \mathbb{N}}$ and $\{\bar{\boldsymbol{\lambda}}^k\}_{k \in \mathbb{N}}$ converges to $\bar{\boldsymbol{\mu}}$ and $\bar{\boldsymbol{\lambda}}$, respectively. Consequently, taking the limit we arrive at

$$
(-\nabla f(\boldsymbol{x}^*), -\nabla^2 f(\boldsymbol{x}^*)) \in \limsup_{(\boldsymbol{x},\bar{\boldsymbol{\mu}},\bar{\bar{\boldsymbol{\lambda}}},\rho,\delta) \to (\boldsymbol{x}^*,\bar{\boldsymbol{\mu}},\bar{\boldsymbol{\lambda}},+\infty,0^+)} K^W(\boldsymbol{x},\bar{\boldsymbol{\mu}},\bar{\bar{\boldsymbol{\lambda}}},\rho,\delta).
$$

Hence, by the AL2-regularity at $\boldsymbol{x}^*$, $(-\nabla f(\boldsymbol{x}^*), -\nabla^2 f(\boldsymbol{x}^*)) \in K^2(\boldsymbol{x}^*)$. That is, $\boldsymbol{x}^*$ is a WSOC point.

Let us prove the reciprocal; that is, we now show that if the point $\boldsymbol{x}^*$ satisfies WSOC whenever it fulfills the AL-AKKT2 condition, then the function K^W satisfies, for a given $\bar{\boldsymbol{\mu}} \in \mathbb{R}_+^p$ and $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^q$,

$$
\limsup_{(\boldsymbol{x},\bar{\boldsymbol{\mu}},\bar{\bar{\boldsymbol{\lambda}}},\rho,\delta) \to (\boldsymbol{x}^*,\bar{\boldsymbol{\mu}},\bar{\boldsymbol{\lambda}},+\infty,0^+)} K^W(\boldsymbol{x},\bar{\boldsymbol{\mu}},\bar{\bar{\boldsymbol{\lambda}}},\rho,\delta) \subset K^2(\boldsymbol{x}^*).
$$

Indeed, let

$$(\boldsymbol{d}, H) \in \limsup_{\left(\boldsymbol{x}, \bar{\bar{\boldsymbol{\mu}}}, \bar{\bar{\boldsymbol{\lambda}}}, \rho, \delta\right) \to \left(\boldsymbol{x}^*, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}}, +\infty, 0^+\right)} K^W(\boldsymbol{x}, \bar{\bar{\boldsymbol{\mu}}}, \bar{\bar{\boldsymbol{\lambda}}}, \rho, \delta).$$

Then, there are sequences $\{\boldsymbol{d}^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$, $\{H^k\}_{k \in \mathbb{N}} \subset \mathrm{Sym}(n)$, $\{\rho_k\}_{k \in \mathbb{N}} \to \infty$ and $\{\delta_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_{++}$ such that $(\boldsymbol{d}^k, H^k) \in K^W(\boldsymbol{x}^k, \bar{\boldsymbol{\mu}}^k, \bar{\boldsymbol{\lambda}}^k, \rho_k, \delta_k)$, $\{\delta_k\}_{k \in \mathbb{N}}$ converges to zero, $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$ converges to $\boldsymbol{x}^*$, $\{(\boldsymbol{d}^k, H^k)\}_{k \in \mathbb{N}}$ converges to $(\boldsymbol{d}, H)$ and $\{\bar{\boldsymbol{\mu}}^k\}_{k \in \mathbb{N}}$ and $\{\bar{\boldsymbol{\lambda}}^k\}_{k \in \mathbb{N}}$ converge to $\bar{\boldsymbol{\mu}}$ and $\bar{\boldsymbol{\lambda}}$, respectively. Therefore, considering the problem

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \boldsymbol{d}^T(\boldsymbol{x} - \boldsymbol{x}^*) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^*)^T H(\boldsymbol{x} - \boldsymbol{x}^*) \quad \text{s.t.} \quad \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}, \quad \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \tag{13}$$

$\boldsymbol{x}^*$ is an AL-AKKT2 point with the primal sequence $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$. Thus, by the hypothesis, $\boldsymbol{x}^*$ conforms with WSOC. And finally, $(\boldsymbol{d}, H) \in K^2(\boldsymbol{x}^*)$. This concludes the proof. $\qquad\square$

The least stringent CQ associated with the AKKT2 conditions is called CCP2, while the one associated with CAKKT2 is called CAKKT2-regularity. Since AL-AKKT2 is stronger than AKKT2 and CAKKT2, it is natural to wonder whether the AL2-regularity is weaker than CCP2 and CAKKT2-regularity. Interestingly, it is, but not only that, AL2-regularity weakens all possible constraint qualifications that can be used to guarantee convergence to WSOC points for SALM's accumulation points. Understanding this matter is straightforward. A proof is given below.

Proposition 5.7. *The least stringent CQ necessary for WSOC points at the SALM accumulation points is AL2-regularity. In particular, CCP2 implies AL2-regularity.*

Proof. Let $\boldsymbol{x}^*$ be a feasible point and assume that a constraint qualification that guarantees WSOC at SALM accumulation points is valid at $\boldsymbol{x}^*$. Let

$$(\boldsymbol{d}, H) \in \limsup_{\left(\boldsymbol{x}, \bar{\bar{\boldsymbol{\mu}}}, \bar{\bar{\boldsymbol{\lambda}}}, \rho, \delta\right) \to \left(\boldsymbol{x}^*, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}}, +\infty, 0^+\right)} K^W(\boldsymbol{x}, \bar{\bar{\boldsymbol{\mu}}}, \bar{\bar{\boldsymbol{\lambda}}}, \rho, \delta).$$

Then, there are sequences $\{\boldsymbol{d}^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$, $\{H^k\}_{k \in \mathbb{N}} \subset \mathrm{Sym}(n)$, $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$, $\{\bar{\boldsymbol{\mu}}^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^p_+$, $\{\bar{\boldsymbol{\lambda}}^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^p_+$, $\{\rho_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_{++}$ and $\{\delta_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_{++}$ such that $(\boldsymbol{d}^k, H^k) \in K^W(\boldsymbol{x}^k, \bar{\boldsymbol{\mu}}^k, \bar{\boldsymbol{\lambda}}^k, \rho_k, \delta_k)$, with $\rho_k \to \infty$, $\{\delta_k\}_{k \in \mathbb{N}}$ converging to zero, $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$ converging to $\boldsymbol{x}^*$, $\{(\boldsymbol{d}^k, H^k)\}_{k \in \mathbb{N}}$ converging to $(\boldsymbol{d}, H)$ and $\{\bar{\boldsymbol{\mu}}^k\}_{k \in \mathbb{N}}$ and $\{\bar{\boldsymbol{\lambda}}^k\}_{k \in \mathbb{N}}$ converging to $\bar{\boldsymbol{\mu}} \in \mathbb{R}^p_+$ and $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^q$, respectively. Therefore, considering the problem (13), $\boldsymbol{x}^*$ is an AL-AKKT2 point with the primal sequence $\{\boldsymbol{x}^k\}_{k \in \mathbb{N}}$. Since AL-AKKT2 guarantees a sequence convergent to $\boldsymbol{x}^*$ generated by SALM applied to (13), and a constraint qualification that guarantees WSOC at SALM accumulation points are valid at the same point, we have the validity of WSOC. Consequently, $(\boldsymbol{d}, H) \in K^2(\boldsymbol{x}^*)$ and the inclusion in Definition 5.5 follows.

Since AL-AKKT2 implies AKKT2, the second statement can be proved using the same reasoning applied to the first part of the proof. $\qquad\square$

Remark 6. It is possible to prove straightforwardly that CAKKT2-regularity implies AL2-regularity under the GKL inequality, as discussed in Remark 4.

Considering the developments in this article, we provide the landscape of CQs associated with WSOC in Figure 1.

6 Final remarks

Safeguarded augmented Lagrangian methods have been widely studied recently due to their broad applicability and general global convergence results. When one can solve the subproblems of an algorithm obtaining second-order stationarity, a second-order stationary accumulation point for the original problem is expected to be found. The first-order variation of this idea and reasonable assumptions
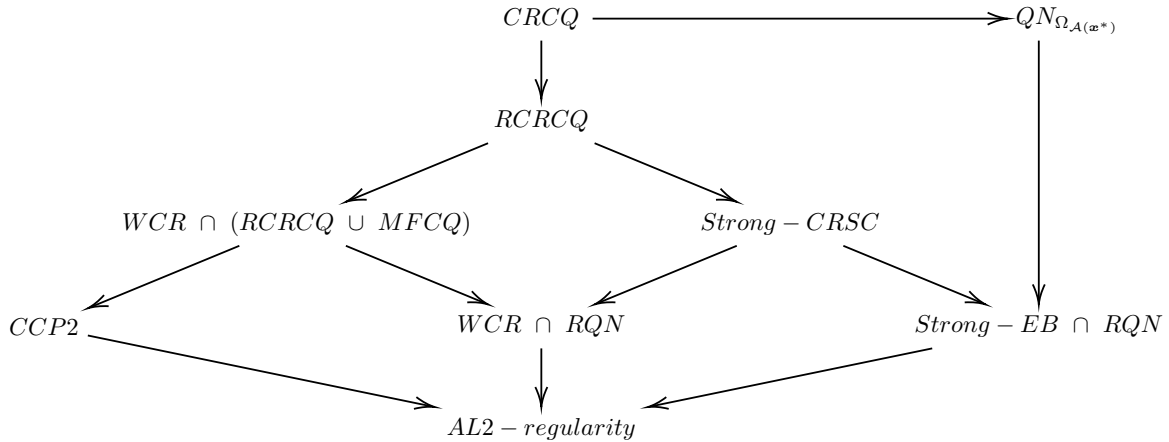
Figure 1: Landscape of the new constraint qualifications for which accumulation points of the safeguarded augmented Lagrangian method are WSOC points. Here, $QN_{\Omega_{\mathcal{A}(\boldsymbol{x}^*)}}$ is the quasi-normality CQ applied to the set $\Omega_{\mathcal{A}(\boldsymbol{x}^*)}$.

that guarantee a similar first-order property are well known, however very few results are known for the second-order case. Previous results suggested that a constant rank property of the gradients of active constraints was needed in order to guarantee the inner continuity of the linearized subspace, inherent to second-order stationarity.

In this paper, however, we propose a strong error bound property, which is enough to guarantee second-order stationarity. The condition corresponds to the traditional error bound constraint qualification for constraint sets formulated with selections of the constraint functions. Our result also requires boundedness of the dual sequences, which can be achieved with the so-called relaxed quasinormality condition. Connection with several other conditions known in the literature are studied; in particular, we show that the recently introduced *strong constant rank of the subspace component* is sufficient for guaranteeing second-order stationarity without any additional requirement.

To formulate our results, we present a new second-order sequential optimality condition, particularly tailored to the algorithm, which allows us to characterize the least restrictive condition that guarantees second-order stationarity, called AL2-regularity. This condition is the least stringent CQ associated with the algorithm in the sense that any CQ that guarantees second-order stationarity at feasible accumulation points reached by the algorithm is necessarily more stringent than AL2-regularity. This suggests a relevant research topic related to our results, which is the characterization of conditions that ensure boundedness of the dual sequences generated by the algorithm. We conjecture that this property can be characterized in terms of error bound type conditions. This will be the subject of future research.

References

[1] Anandkumar, A. and Ge, R. Efficient approaches for escaping higher order saddle points in non-convex optimization. In Feldman, V., Rakhlin, A., and Shamir, O., editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 81–102, Columbia University, New York, New York, USA, 23–26 Jun 2016. PMLR.

[2] Andreani, R., Behling, R., Haeser, G., and Silva, P. J. On second-order optimality conditions in nonlinear optimization. *Optimization Methods and Software*, 32(1):22–38, 2017.

[3] Andreani, R., Birgin, E. G., Martínez, J. M., and Schuverdt, M. L. On Augmented Lagrangian Methods with General Lower-Level Constraints. *SIAM Journal on Optimization*, 18(4):1286–1309, 2008.

[4] Andreani, R., Birgin, E. G., Martínez, J. M., and Schuverdt, M. L. Second-order negative-curvature methods for box-constrained and general constrained optimization. *Computational Optimization and Applications*, 45(2):209–236, Mar 2010.

[5] Andreani, R., Echagüe, C. E., and Schuverdt, M. L. Constant-rank condition and second-order constraint qualification. *Journal of Optimization Theory and Applications*, 146(2):255–266, 2010.

[6] Andreani, R., Fazzio, N. S., Schuverdt, M. L., and Secchin, L. D. A Sequential Optimality Condition Related to the Quasi-normality Constraint Qualification and Its Algorithmic Consequences. *SIAM Journal on Optimization*, 29(1):743–766, 2019.

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

[8] Andreani, R., Haeser, G., Schuverdt, M. L., and Silva, P. J. S. Two New Weak Constraint Qualifications and Applications. *SIAM Journal on Optimization*, 22(3):1109–1135, 2012.

[9] Andreani, R., Martínez, J. M., Ramos, A., and Silva, P. J. S. Strict Constraint Qualifications and Sequential Optimality Conditions for Constrained Optimization. *Mathematics of Operations Research*, 43(3):693–717, 2018.

[10] Andreani, R., Martínez, J. M., and Schuverdt, M. L. On second-order optimality conditions for nonlinear programming. *Optimization*, 56(5-6):529–542, 2007.

[11] Andreani, R., Martínez, J. M., and Svaiter, B. F. A New Sequential Optimality Condition for Constrained Optimization and Algorithmic Consequences. *SIAM Journal on Optimization*, 20(6):3533–3554, 2010.

[12] Andreani, R., Haeser, G., Mito, L., and Cabrera, H. A minimal face constant rank constraint qualification for reducible conic programming. unpublished, 04 2023.

[13] Andreani, R., Haeser, G., Mito, L. M., Ramos, A., and Secchin, L. D. Correction to: On the best achievable quality of limit points of augmented Lagrangian schemes. *Numerical Algorithms*, 90(2):879–880, Jun 2022.

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

[15] Andreani, R., Haeser, G., Ramos, A., and Silva, P. J. S. A second-order sequential optimality condition associated to the convergence of optimization algorithms. *IMA Journal of Numerical Analysis*, 37(4):1902–1929, 01 2017.

[16] Andreani R., Haeser G., Schuverdt M.L. and Secchin L. D. A relaxed quasinormality condition and the boundedness of dual augmented Lagrangian sequences, 2024. *Submitted*.

[17] Arutyunov, A. V. Perturbations of extremal problems with constraints and necessary optimality conditions. *Journal of Soviet Mathematics*, 54(6):1342–1400, May 1991.

[18] Bertsekas, D. P. *Constrained Optimization and Lagrange Multiplier Methods*. Computer science and applied mathematics. Academic Press, London, 1982.

[19] Bertsekas, D. P. *Nonlinear Programming*. Athena scientific optimization and computation series. Athena Scientific, Belmont, Massachussets, 2 edition, 1999.

[20] Birgin, E. G., Gardenghi, J. L., Martínez, J. M., Santos, S. A., and Toint, P. L. Evaluation complexity for nonlinear constrained optimization using unscaled KKT conditions and high-order models. *SIAM Journal on Optimization*, 26(2):951–967, 2016.

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

[22] Birgin, E. G. and Martínez, J. M. Complexity and performance of an Augmented Lagrangian algorithm. *Optimization Methods and Software*, 35(5):885–920, 2020.

[23] Gomes-Ruggiero, M. A., Martínez, J. M., and Santos, S. A. Spectral Projected Gradient Method with Inexact Restoration for Minimization with Nonconvex Constraints. *SIAM Journal on Scientific Computing*, 31(3):1628–1652, 2009.

[24] Haeser, G. A second-order optimality condition with first- and second-order complementarity associated with global convergence of algorithms. *Computational Optimization and Applications*, 70(2):615–639, 2018.

[25] Haeser, G. Some theoretical limitations of second-order algorithms for smooth constrained optimization. *Operations Research Letters*, 46(3):295–299, 2018.

[26] Jin, X.-B., Zhang, X.-Y., Huang, K., and Geng, G.-G. Stochastic conjugate gradient algorithm with variance reduction. *IEEE Transactions on Neural Networks and Learning Systems*, 30(5):1360–1369, 2019.

[27] Jourani, A. and Sene, M. Characterization of the clarke regularity of subanalytic sets. *Proceedings of the American Mathematical Society*, 146:1, 06 2017.

[28] Lewis, A. S., Luke, D. R., and Malick, J. Local Linear Convergence for Alternating and Averaged Nonconvex Projections. *Foundations of Computational Mathematics*, 9(4):485–513, Aug 2009.

[29] Martens, J. *Second-order optimization for neural networks*. Ph.d. thesis, Department of Computer Science, University of Toronto, Toronto, Canada, 03 2016. Accessed: 16 Jul. 2024.

[30] Minchenko, L. and Tarakanov, A. On Error Bounds for Quasinormal Programs. *Journal of Optimization Theory and Applications*, 148(3):571–579, 2011.

[31] Minchenko, L. I. and Sirotko, S. I. On Local Error Bound in Nonlinear Programs. In Olenev, N. N., Evtushenko, Y. G., Jaćimović, M., Khachay, M., and Malkova, V., editors, *Optimization and Applications*, pages 38–49, Cham, 2021. Springer International Publishing.

[32] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation I*, volume 330 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1st edition, 2006.

[33] Murty, K. G. and Kabadi, S. N. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39(2):117–129, Jun 1987.

[34] Nesterov, Y. and Polyak, B. T. Cubic regularization of Newton method and its global performance. *Mathematical Programming*, 108(1):177–205, Aug 2006.

[35] Prado, R. W., Santos, S. A., and Simões, L. E. A. A novel sequential optimality condition for smooth constrained optimization and algorithmic consequences. *Optimization*, 0(0):1–30, 2023.

[36] Rockafellar, R. T. and Wets, R. J.-B. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften*. Springer Science & Business Media, 1 edition, 2009.

[37] Ye, J. J. New uniform parametric error bounds. *Journal of Optimization Theory and Applications*, 98(1):197–219, Jul 1998.

[38] Zalinescu, C. *Convex Analysis in General Vector Spaces*. World Scientific Publishing Company, Singapore, 2002.