On constraint qualifications for lower-level sets and an augmented Lagrangian method*

Roberto Andreani† Gabriel Haeser‡ Mariana da Rosa* Daiana O. Santos§

December 18, 2025

Abstract

In this paper we consider an augmented Lagrangian method with general lower-level constraints, that is, where some of the constraints are penalized while others are kept as subproblem constraints. Motivated by some recent results on optimization problems on manifolds, we present a general theory of global convergence when a feasible approximate KKT point is found for the subproblems at each iteration. In particular, we formulate new constant rank constraint qualifications that do not require a constant rank assumption in a full dimensional neighborhood of the point of interest. We also formulate an appropriate quasinormality and relaxed-quasinormality conditions which guarantee boundedness of the dual sequences generated by the algorithm. These assumptions apply, in particular, to the current ALGENCAN implementation that keeps box constraints within the subproblems.

Key words: Nonlinear optimization, augmented Lagrangian methods, constraint qualifications, global convergence.

AMS subject classifications: 90C30, 65K05.

1 Introduction

In nonlinear optimization, the first-order optimality conditions by Karush/Kuhn-Tucker (KKT) are a primordial tool for solving an optimization problem and, in particular, for guiding an iterative procedure towards a solution. Notably, approximate Lagrange multipliers (dual solutions) provide relevant information that are exploited in an augmented Lagrangian iteration for speeding up convergence.

In order for the KKT conditions to hold at a local minimizer for any objective function, a *constraint qualification* (CQ) must be imposed on the analytical description of the feasible set. The weakest condition for this purpose is termed Guignard's CQ and it requires the equality of the polars of Bouligand's tangent cone and the linearized feasible set. We are interested in more stringent CQs (termed *strict* CQs in [18]) which can be used to guarantee global convergence of numerical algorithms. In particular, even when a sequence of approximate Lagrange multipliers is unbounded, under a strict

*This work was supported by FAPESP grants 2022/06745-7 and 2023/08706-1, and CNPq grants 306988/2021-6, 302000/2022-4, and 407147/2023-3.

†Department of Applied Mathematics, Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Rua Sérgio Buarque de Holanda, 651, Campinas, SP, Brazil. Email: andreani@ime.unicamp.br, marianadarosa13@gmail.com

‡Department of Applied Mathematics, Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão, 1010, Cidade Universitária, 05508-090, São Paulo, SP, Brazil. e-mail: ghaeser@ime.usp.br

§Department of Actuarial Science, Paulista School of Politics, Economics, and Business, Federal University of São Paulo, Rua General Newton Estilac Leal, 932, Quitaúna, Osasco, SP, Brazil. e-mail: daiana.santos@unifesp.br

CQ one can conclude that the limit primal point satisfies the KKT conditions. We will focus on two important strict CQs namely the *constant rank of the subspace component* (CRSC [16]) and the *quasinormality condition* (QN [30, 19]) together with the recently introduced *relaxed quasinormality condition* (RQN [14]). Condition QN (and RQN) is associated with boundedness of dual augmented Lagrangian sequences [6, 26] which allows identifying a true Lagrange multiplier while also employing a loose stopping criterion for the subproblems [22]. Condition CRSC is the weakest of a family of constant rank conditions ([31, 36, 15, 34]) which have been defined mostly associated with global convergence of algorithms. Notably, [31] introduced a constant rank condition in order to compute the derivative of the value function while condition CRSC can be used to reformulate the feasible set by identifying inequality constraints that can only be satisfied as equalities in the feasible set [16]. Constant rank conditions have been extended to the context of conic programming in [11, 9, 8, 12, 10], where, in particular, a nonlinear facial reduction property and a strong second-order necessary optimality condition are proved under CRSC (see [10]). In addition, QN has also been extended to the conic programming setting [39] and for mathematical programming with complementarity constraints (MPCCs) [29]. Conditions CRSC and QN were recently extended to the more general context of Riemannian manifolds in [2].

We are interested in safeguarded augmented Lagrangian methods as described in [24]. In an augmented Lagrangian method, an approximate unconstrained stationary point of an augmented Lagrangian function is sought at each iteration and an approximate Lagrange multiplier is computed while a bounded/safeguarded variation of it is fed back to the augmented Lagrangian function for the next subproblem. In practice, however, in the ALGENCAN implementation, box constraints are not penalized and are kept as subproblem constraints. This is done in order to guarantee that a solution for the subproblem always exists while also allowing that an active set method, which uses a spectral projected gradient step, is used to solve the subproblems. In particular, the iterates computed always satisfy exactly the box constraints.

An alternative approach keeps all equality constraints as explicit subproblem constraints, penalizing only bound constraints similarly to an interior point strategy [21]. Extensive numerical results are presented for the convex quadratic case in [27]. When the constraints define a differentiable manifold, one may keep feasibility of these constraints in the subproblems by means of an unconstrained minimization algorithm for manifold optimization, which is associated with better numerical performance [3]. Finally, recent analyses [5] indicate improved numerical stability of the dual sequence in augmented Lagrangian algorithms applied to MPCCs when simple complementarity constraints are kept explicitly in the subproblems, rather than penalized.

The augmented Lagrangian implementations mentioned previously are part of so-called augmented Lagrangian methods with general lower-level constraints, that is, when any set of constraints can be kept as subproblem constraints. Sequential optimality conditions of first- [1] and second-order [23] have been defined in order to study global convergence of this general variant, but no associated constraint qualification have been defined for this purpose. Sequential optimality conditions for abstract lower-level constraints are considered in [28].

Motivated by the theoretical and numerical improvements observed when the lower-level set is preserved in the subproblems (as in box-constraints, manifold-constraints, and MPCCs formulations with simple complementarity constraints kept explicit), we introduce lower-level versions of the CRSC and QN/RQN conditions. More precisely, we propose "lower-level" CRSC-type and quasinormality-type conditions that allow us to prove global convergence and dual stability of augmented Lagrangian methods with a general lower-level constraint under weak assumptions. In particular, these assumptions require conditions only on constrained neighborhoods of the point of interest, rather than on a full neighborhood of the ambient space. Notably, these conditions are not constraint qualifications in general; nevertheless, they are sufficient to obtain global convergence and boundedness of dual sequences for the algorithm.

We also investigate when these lower-level conditions actually become constraint qualifications. This part of the analysis is motivated by the lower-level constraint qualifications defined in [3], where the feasible set includes a manifold constraint, i.e., a set of equality constraints that remains linearly independent near the point of interest, and by the condition proposed in [4], in which a manifold is identified in the constraint set by requiring a CRSC property on the lower-level set. Moreover, [4] introduces a "constrained" version of the QN/RQN conditions for the MPCC setting, which is independent of the classical MPCC-QN condition available in the literature [33], and which ensures Mordukhovich-stationarity [35] (a notion weaker than KKT stationarity, but widely accepted in this context). Our results generalize these results under a version of QN/RQN considering a general lower-level set.

In what follows, we provide a detailed development of our results. In Section 2, we recall basic definitions and tools that will be used throughout the paper. In Section 3, we propose a constant rank type condition adapted to the lower-level setting and analyze its consequences. Section 4 introduces the notions of lower-level quasinormality and relaxed quasinormality conditions, discussing their main properties and relations with existing conditions in the literature. An illustrative analysis of the improvements obtained with the augmented Lagrangian method, as well as the fact that no constraint qualification is required, is presented in Section 5. Finally, concluding remarks and perspectives for future research are presented in Section 6.

Notations. Given vectors $a, b \in \mathbb{R}^\ell$ we define the componentwise maximum and minimum as $\max\{a, b\}$ and $\min\{a, b\}$, respectively. The non-negative orthant of $\mathbb{R}^\ell$ is denoted by $\mathbb{R}^\ell_+$ while $\mathbb{R}^\ell_- := -\mathbb{R}^\ell_+$. We use $\|\cdot\|$ to denote the Euclidean norm in any $\mathbb{R}^\ell$. The gradient of a differentiable function $f \colon \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ is denoted by $\nabla f(x)$.

2 Preliminaries

Let us consider the general optimization problem

$$\text{Minimize } f(x) \quad \text{subject to} \quad x \in \Omega_L \cap \Omega_U, \tag{NLP}$$

where $\Omega_L := \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$ defines the lower-level set of constraints and $\Omega_U := \{x \in \mathbb{R}^n \mid H(x) = 0, G(x) \leq 0\}$ defines the upper-level set of constraints. Here, $f \colon \mathbb{R}^n \to \mathbb{R}, h \colon \mathbb{R}^n \to \mathbb{R}^m, g \colon \mathbb{R}^n \to \mathbb{R}^p, H \colon \mathbb{R}^n \to \mathbb{R}^{\bar{m}}, G \colon \mathbb{R}^n \to \mathbb{R}^{\bar{p}}$ are continuously differentiable functions. We will abuse the notation and refer to Ω_L and Ω_U in order to refer to the functions h, g and H, G, respectively. It is to be understood that the lower-level set of constraints will be non-relaxable constraints, that is, one may think that the objective function and/or the constraint functions defining the upper level are not defined for $x \notin \Omega_L$. Therefore, an algorithm for solving the problem must necessarily maintain feasibility within its iterates with respect to Ω_L. This is the case, for instance, in the ALGENCAN software [24] where Ω_L correspond to a compact box where the iterates lie or in the MANOPT software for manifold optimization [25] where Ω_L corresponds to a manifold where all iterates must lie.

We will consider the general lower-level safeguarded augmented Lagrangian method [24] which is defined as follows. Given a penalty parameter $\rho > 0$ and safeguarded Lagrange multipliers $v \in \mathbb{R}^{\bar{m}}$, and $u \in \mathbb{R}^{\bar{p}}$, we define the augmented Lagrangian function with respect to the upper-level constraints as

$$x \mapsto L_\rho(x; v, u) := f(x) + \frac{\rho}{2}\left(\left\|H(x) + \frac{v}{\rho}\right\|^2 + \left\|\max\left\{0, G(x) + \frac{u}{\rho}\right\}\right\|^2\right),$$

which will be approximately minimized in Ω_L at each iteration, followed by a judicious update of the parameters ρ, v, and u. The detailed algorithm is described as follows.

Algorithm 1: General lower-level safeguarded augmented Lagrangian method

Step 0. Given non-empty compact sets $S_{\bar{m}} \subseteq \mathbb{R}^{\bar{m}}$ and $S_{\bar{p}} \subseteq \mathbb{R}^{\bar{p}}$, choose initial safeguarded multipliers $\bar{v}^1 \in S_{\bar{m}}$ and $\bar{u}^1 \in S_{\bar{p}}$. Choose a sequence of penalty parameters $\{\rho_k\}_{k=1}^{\infty} \to +\infty$ and a sequence of subproblem tolerances $\{\varepsilon_k\}_{k=1}^{\infty} \to 0^+$. Set $k := 1$.

Step 1. Compute $x^k \in \Omega_L$, $\lambda^k \in \mathbb{R}^m$, and $\mu^k \in \mathbb{R}^p_+$ such that

$$\|\nabla L_{\rho_k}(x^k; \bar{v}^k, \bar{u}^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k)\| \leq \varepsilon_k,$$
$$\|\min\{-g(x^k), \mu^k\}\| \leq \varepsilon_k.$$

Step 2. Define approximate Lagrange multipliers $v^k := \bar{v}^k + \rho_k H(x^k)$, $u^k := \max\{0, \bar{u}^k + \rho_k G(x^k)\}$ and safeguarded variants $\bar{v}^{k+1} \in S_{\bar{m}}$ and $\bar{u}^{k+1} \in S_{\bar{p}}$, respectively. Set $k := k + 1$ and go to Step 1.

Remark 1. *In a practical implementation, the penalty parameter sequence $\{\rho_k\}_{k=1}^{\infty}$ is computed during the execution of the algorithm by measuring the progress in terms of feasibility and complementarity of the upper-level constraints, where the penalty parameter is increased only when sufficient progress is not obtained. For simplicity of the analysis, we fixed a sequence with $\rho_k \to +\infty$. Nevertheless, our results also extend to the practical variant of the algorithm where the penalty parameter sequence may remain bounded; see [24].*

Remark 2. *When Ω_L is bounded, the problem of minimizing $L_{\rho_k}(x; \bar{v}^k, \bar{u}^k)$ subject to $x \in \Omega_L$ admits a solution. When Ω_L satisfies a sufficient interior property [7], a point satisfying the conditions of Step 1 can be computed with an arbitrary precision $\varepsilon_k > 0$, even when a solution does not admit Lagrange multipliers. Clearly, a KKT point of this problem also satisfies these conditions.*

A classical global convergence result of this algorithm can be stated under the *constant rank of the subspace component* (CRSC) condition, introduced in [16]. Its definition for an arbitrary feasible set Ω_L is given as follows.

Definition 2.1 ([16])**.** *We say that a point $\bar{x} \in \Omega_L$ satisfies the* constant rank of the subspace component *(CRSC) condition for the constraint set Ω_L when the gradients*

$$\nabla h_i(x), i = 1, \ldots, m, \quad \nabla g_j(x), j \in J_-$$

have constant rank for x in a neighborhood of $\bar{x}$. The index set J_- is defined as

$$J_- := \left\{ j \in \{1, \ldots, p\} \mid g_j(\bar{x}) = 0 \text{ and } -\nabla g_j(\bar{x}) \in \mathcal{K}^L(\bar{x}, \bar{x}) \right\},$$

where

$$\mathcal{K}^L(x, \bar{x}) := \left\{ \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j: g_j(\bar{x})=0} \mu_j \nabla g_j(x) \mid \lambda_i \in \mathbb{R}, \mu_j \geq 0 \right\}.$$

The perturbed KKT cone $\mathcal{K}^L(x, \bar{x})$ was defined in [17] and the reason for defining it for $x \neq \bar{x}$ will be clear later on when we discuss a cone-continuity property CQ (CCP, [17]). Notice that the KKT conditions with respect to the problem of minimizing $\tilde{f}(x)$ subject to $x \in \Omega_L$ at a feasible point $\bar{x}$ can be simply stated as $-\nabla \tilde{f}(\bar{x}) \in \mathcal{K}^L(\bar{x}, \bar{x})$. Under CRSC, the KKT cone coincides with the polar of the classical tangent cone, which implies that CRSC is a constraint qualification. See [16]. A strong second-order necessary optimality condition can also be proved under a variation of CRSC [10]. The set J_- is the subset of active inequality constraints such that its gradients behave as equality constraints in the definition of $\mathcal{K}^L(\bar{x}, \bar{x})$. In fact, it was proved in [16] that under CRSC it holds that for $j \in J_-$, it is always the case that $g_j(x) = 0$ for all $x \in \Omega_L$ in a neighborhood of $\bar{x}$, that is, the inequalities in J_- are indeed locally equality constraints in disguise. It is shown in [16] that CRSC implies a local error bound property. We are interested in global convergence properties of Algorithm 1 under CRSC. The following result has been stated in [1] under a stronger CQ, which was later updated to CRSC in [16]. See also [15, Corollary 2 and 3].

Theorem 2.1. *Let $\bar{x}$ be a limit point of a sequence $\{x^k\}_{k=1}^{\infty}$ generated by Algorithm 1. If $\bar{x}$ satisfies CRSC for the constraint set Ω_L, then $\bar{x}$ satisfies the KKT conditions for the problem*

$$\textit{Minimize } \|H(x)\|^2 + \|\max\{0, G(x)\}\|^2, \textit{ subject to } h(x) = 0, g(x) \leq 0.$$

And, if $\bar{x} \in \Omega_L \cap \Omega_U$ and satisfies CRSC for the constraint set $\Omega_L \cap \Omega_U$, then $\bar{x}$ satisfies the KKT conditions for problem (NLP).

In [6, 26], a stronger result is obtained under the quasinormality (QN) condition, which is independent of CRSC, for the case of unconstrained subproblems when the limit point is feasible. We proceed with the definition of QN for a general feasible set Ω_L.

Definition 2.2 ([30, 19])**.** *A point $\bar{x} \in \Omega_L$ satisfies* quasinormality *(QN) with respect to Ω_L when there is no $0 \neq (\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ with*

$$\sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j=1}^{p} \mu_j \nabla g_j(\bar{x}) = 0,$$

together with a sequence $x^k \to \bar{x}$ such that for all k, $\lambda_i h_i(x^k) > 0$ when $i \in I_{\neq} := \{i \mid \lambda_i \neq 0\}$ and $g_j(x^k) > 0$ when $j \in J_{\neq} := \{j \mid \mu_j > 0\}$.

This condition can be thought as a refinement of the classical Mangasarian-Fromovitz's condition in order to bound the sequence of dual approximations generated by the algorithm in the case of unconstrained subproblems. Namely, when the dual sequence is unbounded, the corresponding primal sequence $\{x^k\}$ must necessarily violate the constraints with the same sign as the sign of the corresponding dual component. Condition QN simply requires that these sequences do not exist. The statement of the theorem as in [26, 6] is as follows

Theorem 2.2. *Let $\bar{x}$ be a limit point of a sequence $\{x^k\}_{k=1}^{\infty}$ generated by Algorithm 1 with $\Omega_L = \mathbb{R}^n$ and $\{x^k\}_{k \in K} \to \bar{x}$ for some infinite set $K \subseteq \mathbb{N}$. Then $\bar{x}$ is a stationary point for the problem*

$$\textit{Minimize } \|H(x)\|^2 + \|\max\{0, G(x)\}\|^2.$$

And, if $\bar{x} \in \Omega_U$ and satisfies QN with respect to Ω_U, then the dual sequences $\{v^k\}_{k \in K}$ and $\{u^k\}_{k \in K}$ are bounded. In particular, any limit point of these sequences are Lagrange multipliers associated with $\bar{x}$ with respect to problem (NLP).

We end this preliminary presentation with the formal definition of a recently introduced relaxation of QN, termed *relaxed-quasinormality.*

Definition 2.3 ([14]). *A feasible point* $\bar{x} \in \Omega_L$ *satisfies the* relaxed-quasinormality *(RQN) condition with respect to* Ω_L *when there is no* $0 \neq (\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ *with*

$$\sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j=1}^{p} \mu_j \nabla g_j(\bar{x}) = 0,$$

together with a sequence $x^k \to \bar{x}$ *such that for all* k*,*

- $\lambda_i h_i(x^k) > 0$ *when* $i \in I_{\neq} := \{i \mid \lambda_i \neq 0\}$ *and* $g_j(x^k) > 0$ *when* $j \in J_{\neq} := \{j \mid \mu_j > 0\}$*;*
- $|h_i(x^k)| = o(w(x^k))$ *for every* $i \notin I_{\neq}$ *and* $g_j(x^k)_+ = o(w(x^k))$ *for every* $j \notin J_{\neq}$*, where*

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

While QN is independent of CRSC, the relaxed form RQN is *implied* by CRSC. More importantly, RQN was shown to be *equivalent*, under mild assumptions (see [13, Theorem 3.3]), to an error bound condition, which was shown to be the weakest condition that guarantees boundedness of the sequence of dual multipliers generated by the safeguarded augmented Lagrangian method with unconstrained subproblems, that is, the weakest assumption that can be used to replace QN in Theorem 2.2.

Next, we will propose generalizations of CRSC and QN/RQN that are more suitable to the global convergence analysis of the algorithm with lower-level constraints. These conditions are motivated by a previous work on manifold-constrained optimization [2]. There, the ambient space $\mathbb{R}^n$ was replaced by a general manifold. Thus, following an intrinsic approach it makes little sense to consider points outside of the manifold, therefore, constant rank conditions are formulated naturally considering a neighborhood of the point of interest restricted to the manifold. In [3], a manifold embedded in $\mathbb{R}^n$ is interpreted as a lower-level constraint, and generalizations of classical constant rank conditions that consider only neighborhoods constrained to the manifold by using the theory developed in [2] was defined. Later, in [4], a more general condition is proposed based on the local identification of a submanifold, equivalently yielding a lower-level set satisfying CRSC and containing the original feasible set. These generalized conditions are weaker than CRSC while still ensuring useful properties such as global convergence of Algorithm 1 and a local error bound property. It turns out that a much more general approach is possible, as we formulate next.

3 A lower-level constant rank constraint qualification

We start with the definition of Lower-CRSC. This is an extension of the definitions introduced in [3, 4], where the lower-level constraint set is defined by linearly independent equality constraints or by some constraints satisfying the CRSC condition.

Definition 3.1. *Given* $\bar{x} \in \Omega_L \cap \Omega_U$*, we say that Lower-CRSC holds at* $\bar{x}$ *for the constraint set* $\Omega_L \cap \Omega_U$ *with lower-level set* Ω_L *when the gradients*

$$\nabla H_i(x), i = 1, \ldots, \bar{m}, \quad \nabla G_j(x), j \in J_-^U, \quad \nabla h_i(x), i = 1, \ldots, m, \quad \nabla g_j(x), j \in J_-^L$$

have constant rank for all $x \in \Omega_L$ *in a neighborhood of* $\bar{x}$*, where*

$$J_-^U := \{j \in \{1, \ldots, \bar{p}\} \mid G_j(\bar{x}) = 0 \textit{ and } -\nabla G_j(\bar{x}) \in \mathcal{K}(\bar{x}, \bar{x})\},$$

$$J_-^L := \{j \in \{1, \ldots, p\} \mid g_j(\bar{x}) = 0 \textit{ and } -\nabla g_j(\bar{x}) \in \mathcal{K}(\bar{x}, \bar{x})\},$$

and

$$\mathcal{K}(x,\bar{x}) := \left\{ \sum_{i=1}^{\bar{m}} v_i \nabla H_i(x) + \sum_{j:G_j(\bar{x})=0} u_j \nabla G_j(x) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j:g_j(\bar{x})=0} \mu_j \nabla g_j(x) \mid v_i, \lambda_i \in \mathbb{R}; u_j, \mu_j \in \mathbb{R}_+ \right\}.$$

The difference of Lower-CRSC with CRSC with respect to $\Omega_L \cap \Omega_U$ is that the constant rank condition is assumed for *feasible points with respect to* Ω_L. Hence, Lower-CRSC is strictly weaker than CRSC. However, without additional assumptions about the lower-level set Ω_L, this condition alone does not qualify as a genuine constraint qualification, as the next example illustrates.

Example 1. *For $x := (x_1, x_2) \in \mathbb{R}^2$, consider the optimization problem of minimizing $f(x) := x_1$ subject to $x \in \Omega_L \cap \Omega_U$, where $\Omega_U := \{x \mid G_1(x) := x_1 \leq 0\}$ and $\Omega_L := \{x \mid g_1(x) := -x_1^3 + x_2 \leq 0, g_2(x) := -x_2 \leq 0, g_3(x) := x_1 \leq 0\}$. The point $\bar{x} := (0,0)$ is a global minimizer that satisfies Lower-CRSC, since any subset of gradients have constant rank in the set $\Omega_L = \{(0,0)\}$. Nevertheless, $\bar{x}$ is not a KKT point.*

This example shows that further assumptions on Ω_L are needed. In fact, we will show that it is enough to assume Guignard's CQ on Ω_L in a neighborhood of $\bar{x}$ to ensure that Lower-CRSC is a CQ, which is weaker than the assumptions in [3, 4]. Now let us show that Lower-CRSC is strictly weaker than CRSC. This example is particularly relevant to the ALGENCAN implementation due to the use of box-constraints in the lower-level.

Example 2. *Consider the feasible set given by $\Omega_U := \{x \mid H_1(x) := x_2 x_1^2 = 0\}$ and $\Omega_L := \{x \mid g_1(x) := -x_1 \leq 0, g_2(x) := x_1 \leq 0, g_3(x) := x_2 - 1 \leq 0\}$ at the feasible point $\bar{x} := (0,1)$. The gradients are given by $\nabla H_1(x) = (2x_1x_2, x_1^2)$, $\nabla g_1(x) = (-1,0)$, $\nabla g_2(x) = (1,0)$, and $\nabla g_3(x) = (0,1)$ for all x. Condition CRSC fails as the rank of the gradients $\nabla H_1(x)$, $\nabla g_1(x)$, $\nabla g_2(x)$ is equal to 1 at $x := \bar{x}$ and equal to 2 for x in any small neighborhood of $\bar{x}$ with $x_1 \neq 0$. In contrast, Lower-CRSC holds as for all $x \in \Omega_L$ one has $x_1 = 0$, and thus the rank of the same set of gradients is equal to 1 for all $x \in \Omega_L$.*

In order to simplify our proofs, inspired by [17], we will make use of the following weaker condition.

Definition 3.2. *Given $\bar{x} \in \Omega_L \cap \Omega_U$, we say that the Lower Cone Continuity Property (Lower-CCP) holds at $\bar{x}$ for the constraint set $\Omega_L \cap \Omega_U$ with lower-level set Ω_L when the mapping $x \mapsto \mathcal{K}(x,\bar{x})$ has the following property: every $w \in \mathbb{R}^n$ such that there exists a sequence $\{x^k\} \subset \Omega_L$ with $x^k \to \bar{x}$ and $w^k \in \mathcal{K}(x^k,\bar{x})$ with $w^k \to w$ is such that $w \in \mathcal{K}(\bar{x},\bar{x})$.*

This definition is simply the usual lower-continuity in Ω_L of the point-to-set mapping $x \mapsto K(x,\bar{x})$. When the lower-level constraint set satisfies some CQ for points in Ω_L nearby a local minimizer $\bar{x}$ of (NLP), then Lower-CCP ensures the KKT conditions hold. That is, Lower-CCP is a constraint qualification, as we show next.

Theorem 3.1. *If $\bar{x}$ is a local minimizer of* (NLP) *that satisfies Lower-CCP for the constraints $\Omega_L \cap \Omega_U$ with lower-level Ω_L, and Ω_L satisfies Guignard's CQ at all points in Ω_L in a neighborhood of $\bar{x}$, then $\bar{x}$ is a KKT point of problem* (NLP).

Proof. From a standard external penalty approach (see, for instance, the proof of [28, Theorem 3.3]), there exists a sequence $x^k \to \bar{x}$ such that, for all k, x^k is a local minimizer of $f(x) + \frac{1}{2}\|x - \bar{x}\|^2 + \frac{k}{2}(\|H(x)\|^2 + \|\max\{0, G(x)\}\|^2)$, subject to $x \in \Omega_L$. For k large enough, since x^k satisfies Guignard's CQ, there exist Lagrange multipliers $(\lambda^k, \mu^k) \in \mathbb{R}^m \times \mathbb{R}^p_+$ with $\min\{-g(x^k), \mu^k\} = 0$ such that

$$\nabla f(x^k) + \sum_{i=1}^{\bar{m}} v_i^k \nabla H_i(x^k) + \sum_{j=1}^{\bar{p}} u_j^k \nabla G_j(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) = (\bar{x} - x^k),$$

where $v^k := kH(x^k), u^k := k \max\{0, G(x^k)\}$. Moreover, by the definition of $\mathcal{K}(x, \bar{x})$ in Definition 3.1, we may write

$$w^k := \left[-\nabla f(x^k) - \sum_{j: G_j(\bar{x})<0} u_j^k \nabla G_j(x^k) - \sum_{j: g_j(\bar{x})<0} \mu_j^k \nabla g_j(x^k) + (\bar{x} - x^k) \right] \in \mathcal{K}(x^k, \bar{x}).$$

Notice that $w^k \to -\nabla f(\bar{x})$. Indeed, since $\min\{-g(x^k), \mu^k\} = 0$, it must be that $\mu_j^k \to 0$ whenever $g_j(\bar{x}) < 0$. Also, $u_j^k = 0$ whenever $G_j(x^k) < 0$, which is the case for sufficiently large k when $G_j(\bar{x}) < 0$. By Lower-CCP we conclude that $-\nabla f(\bar{x}) \in \mathcal{K}(\bar{x}, \bar{x})$, which proves the desired result. □

Theorem 3.2. *Let $\bar{x} \in \Omega_L \cap \Omega_U$ satisfy Lower-CRSC for $\Omega_L \cap \Omega_U$ with lower-level Ω_L, then $\bar{x}$ satisfies Lower-CCP. In particular, Lower-CRSC for $\Omega_L \cap \Omega_U$ with lower-level Ω_L satisfying Guignard's CQ at all points in Ω_L in a neighborhood of $\bar{x}$ is a constraint qualification.*

Proof. Let $w \in \mathbb{R}^n$ be such that there exist $x^k \in \Omega_L$ and $w^k \in \mathcal{K}(x^k, \bar{x})$ for all k such that $x^k \to \bar{x}$ and $w^k \to w$. Let us show that $w \in \mathcal{K}(\bar{x}, \bar{x})$. By the definition of $\mathcal{K}(x^k, \bar{x})$, for all k, there exist $(v^k, u^k, \lambda^k, \mu^k) \in \mathbb{R}^{\bar{m}} \times \mathbb{R}_+^{\bar{p}} \times \mathbb{R}^m \times \mathbb{R}_+^p$ such that

$$w^k = \sum_{i=1}^{\bar{m}} v_i^k \nabla H_i(x^k) + \sum_{j: G_j(\bar{x})=0} u_j^k \nabla G_j(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j: g_j(\bar{x})=0} \mu_j^k \nabla g_j(x^k).$$

Using the sets J_-^U and J_-^L given in Definition 3.1, and defining $J_+^U := \{j \mid G_j(\bar{x}) = 0\} \backslash J_-^U$ and $J_+^L := \{j \mid g_j(\bar{x}) = 0\} \backslash J_-^L$, we can write w^k as

$$\begin{aligned} w^k = & \sum_{i=1}^{\bar{m}} v_i^k \nabla H_i(x^k) + \sum_{j \in J_-^U} u_j^k \nabla G_j(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j \in J_-^L} \mu_j^k \nabla g_j(x^k) \\ & + \sum_{j \in J_+^U} u_j^k \nabla G_j(x^k) + \sum_{j \in J_+^L} \mu_j^k \nabla g_j(x^k). \end{aligned}$$

Now, let $I^U \subseteq \{1, \ldots, \bar{m}\}$, $J^U \subseteq J_-^U$, $I^L \subseteq \{1, \ldots, m\}$, and $J^L \subseteq J_-^L$ be such that

$$\nabla H_i(x),\ i \in I^U; \quad \nabla G_j(x),\ j \in J^U; \quad \nabla h_i(x),\ i \in I^L; \quad \nabla g_j(x),\ j \in J^L \tag{1}$$

form a basis of the subspace generated by

$$\nabla H_i(x),\ i \in \{1, \ldots, \bar{m}\}; \quad \nabla G_j(x),\ j \in J_-^U; \quad \nabla h_i(x),\ i \in \{1, \ldots, m\}; \quad \nabla g_j(x),\ j \in J_-^L$$

at $x = \bar{x}$. Since the dimension of this subspace is constant for all $x \in \Omega_L$ in a neighborhood of $\bar{x}$, and the vectors in (1) at $x = x^k$ are linearly independent for all sufficiently large k, the vectors in (1) also form a basis of this subspace when evaluated at $x = x^k$. Therefore, for some new scalars $\tilde{v}_i^k \in \mathbb{R}$, $\tilde{u}_j^k \in \mathbb{R}$, $\tilde{\lambda}_i^k \in \mathbb{R}$, $\tilde{\mu}_j^k \in \mathbb{R}$, we can rewrite w^k as

$$\begin{aligned} w^k = & \sum_{i \in I^U} \tilde{v}_i^k \nabla H_i(x^k) + \sum_{j \in J^U} \tilde{u}_j^k \nabla G_j(x^k) + \sum_{i \in I^L} \tilde{\lambda}_i^k \nabla h_i(x^k) + \sum_{j \in J^L} \tilde{\mu}_j^k \nabla g_j(x^k) \\ & + \sum_{j \in J_+^U} u_j^k \nabla G_j(x^k) + \sum_{j \in J_+^L} \mu_j^k \nabla g_j(x^k). \end{aligned}$$

For each k, let M_k be the maximum of the absolute values of all the scalars in the sum above, and assume that the sequence $\{M_k\}$ is unbounded. Taking a subsequence if necessary, assume $M_k \to +\infty$. Writing $\frac{w^k}{M_k}$ and taking the limit in an appropriate subsequence such that the corresponding scalars converge, say,

$$\frac{\tilde{v}_i^k}{M_k} \to \tilde{v}_i,\, i \in I^U; \quad \frac{\tilde{u}_j^k}{M_k} \to \tilde{u}_j \geq 0,\, j \in J^U; \quad \frac{\tilde{\lambda}_i^k}{M_k} \to \tilde{\lambda}_i^L,\, i \in I^L; \quad \frac{\tilde{\mu}_j^k}{M_k} \to \tilde{\mu}_j,\, j \in J^L$$
$$\frac{u_j^k}{M_k} \to u_j,\, j \in J_+^U; \quad \frac{\mu_j^k}{M_k} \to \mu_j,\, j \in J_+^L$$

where the limits are not all zero, we arrive at

$$0 = \sum_{i \in I^U} \tilde{v}_i \nabla H_i(\bar{x}) + \sum_{j \in J^U} \tilde{u}_j \nabla G_j(\bar{x}) + \sum_{i \in I^L} \tilde{\lambda}_i \nabla h_i(\bar{x}) + \sum_{j \in J^L} \tilde{\mu}_j \nabla g_j(\bar{x}) + \sum_{j \in J_+^U} u_j \nabla G_j(\bar{x}) + \sum_{j \in J_+^L} \mu_j \nabla g_j(\bar{x}).$$

Since the scalars associated with the inequality constraints are non-negative, by the definition of the index sets J_+^U and J_+^L, we have $u_j = 0$, $j \in J_+^U$ and $\mu_j = 0$, $j \in J_+^L$. The remaining equation contradicts the linear independence of the vectors. Therefore, the sequence $\{M_k\}$ is bounded and we can take the limit in w^k for a suitable subsequence such that the scalars converge, say,

$$w = \sum_{i \in I^U} \tilde{v}_i^* \nabla H_i(\bar{x}) + \sum_{j \in J^U} \tilde{u}_j^* \nabla G_j(\bar{x}) + \sum_{i \in I^L} \tilde{\lambda}_i^* \nabla h_i(\bar{x}) + \sum_{j \in J^L} \tilde{\mu}_j^* \nabla g_j(\bar{x}) + \sum_{j \in J_+^U} u_j^* \nabla G_j(\bar{x}) + \sum_{j \in J_+^L} \mu_j^* \nabla g_j(\bar{x}),$$

where $w^k \to w$, $\tilde{v}_i^k \to \tilde{v}_i^*$, $\tilde{u}_j^k \to \tilde{v}_j^*$, $\tilde{\lambda}_i^k \to \tilde{\lambda}_i^*$, $\tilde{\mu}_j^k \to \tilde{\mu}_j^*$, $u_j^k \to u_j^* \geq 0$, and $\mu_j^k \to \mu_j^* \geq 0$. Now, since $J^U \subseteq J_-^U$ and $J^L \subseteq J_-^L$, when some scalar $\tilde{u}_j^* < 0$ or $\tilde{\mu}_j^* < 0$, one can rewrite the sum with new scalars to conclude that $w \in \mathcal{K}(\bar{x}, \bar{x})$, which establishes the first statement. Finally, assuming that Ω_L satisfies Guignard's CQ at all points in Ω_L in a neighborhood of a local minimizer $\bar{x}$ and using that Lower-CRSC implies Lower-CCP, we conclude from Theorem 3.1 that $\bar{x}$ is a KKT point, which completes the proof. □

Although Lower-CRSC and Lower-CCP are not constraint qualifications in general (see Example 1), they are sufficient for a global convergence result as long as the subproblems are approximately solved by finding an exact feasible point for the lower-level constraints. Such a feasible point for the lower-level constraint is always obtained, for instance, in the ALGENCAN implementation, where the lower-level set corresponds to box constraints or in the MANOPT implementation, where the iterates lie exactly on a Riemannian manifold as the lower-level constraints.

Theorem 3.3. *Let $\bar{x}$ be a limit point of a sequence $\{x^k\}_{k=1}^{\infty}$ generated by Algorithm 1. If $\bar{x}$ satisfies Lower-CCP for the constraint set Ω_L with lower-level set Ω_L, then $\bar{x}$ satisfies the KKT conditions for the problem*

$$\textit{Minimize } F(x) := \|H(x)\|^2 + \|\max\{0, G(x)\}\|^2, \textit{ subject to } h(x) = 0, g(x) \leq 0.$$

Additionally, if $\bar{x} \in \Omega_L \cap \Omega_U$ and satisfies Lower-CCP for the constraint set $\Omega_L \cap \Omega_U$ with lower-level set Ω_L, then $\bar{x}$ satisfies the KKT conditions for problem (NLP).

Proof. Let $K \subseteq \mathbb{N}$ be such that $\{x^k\}_{k \in K} \to \bar{x}$. By Steps 1 and 2 of the algorithm and computing the derivative of $L_{\rho_k}(x; \bar{v}^k, \bar{u}^k)$ evaluated at $x^k \in \Omega_L$, we have

$$z^k := \nabla f(x^k) + \sum_{i=1}^{\bar{m}} v_i^k \nabla H_i(x^k) + \sum_{j=1}^{\bar{p}} u_j^k \nabla G_j(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) \to 0,$$

where $v_i^k = \bar{v}_i^k + \rho_k H_i(x^k)$, $u_j^k = \max\{0, \bar{u}_j^k + \rho_k G_j(x^k)\}$, and $\|\min\{-g(x^k), \mu^k\}\| \to 0$. Since Ω_L is closed, we have $\bar{x} \in \Omega_L$. If $g_j(\bar{x}) < 0$, then it must be $\mu_j^k \overset{k\in K}{\to} 0$ since otherwise this would violate $\|\min\{-g(x^k), \mu^k\}\| \to 0$. Therefore, we can write

$$\begin{aligned} w^k &:= z^k - \nabla f(x^k) - \sum_{i=1}^{\bar{m}} v_i^k \nabla H_i(x^k) - \sum_{j=1}^{\bar{p}} u_j^k \nabla G_j(x^k) - \sum_{j:g_j(\bar{x})<0} \mu_j^k \nabla g_j(x^k) \\ &= \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j:g_j(\bar{x})=0} \mu_j^k \nabla g_j(x^k). \end{aligned}$$

Then, it is easy to see that

$$\frac{w^k}{\rho_k} \overset{k\in K}{\to} -\sum_{i=1}^{\bar{m}} H_i(\bar{x}) \nabla H_i(\bar{x}) - \sum_{j=1}^{\bar{p}} \max\{0, G_j(\bar{x})\} \nabla G_j(\bar{x}) = -\nabla F(\bar{x}),$$

with $\frac{w^k}{\rho_k} \in \mathcal{K}^L(x^k, \bar{x})$. By Lower-CCP we conclude that $-\nabla F(\bar{x}) \in \mathcal{K}^L(\bar{x}, \bar{x})$, which proves the first result.

Now, suppose $\bar{x} \in \Omega_L \cap \Omega_U$ and notice that if $G_j(\bar{x}) < 0$ then, since $\rho_k \to +\infty$ and $\{\bar{u}_j^k\}$ is bounded, we have that $u_j^k = 0$ for sufficiently large $k \in K$. Therefore, we can define

$$\tilde{w}^k := z^k - \nabla f(x^k) - \sum_{j:G_j(\bar{x})<0} u_j^k \nabla G_j(x^k) - \sum_{j:g_j(\bar{x})<0} \mu_j^k \nabla g_j(x^k) \in \mathcal{K}(x^k, \bar{x})$$

with $\tilde{w}^k \overset{k\in K}{\to} -\nabla f(\bar{x})$. By Lower-CCP we conclude that $-\nabla f(\bar{x}) \in \mathcal{K}(\bar{x}, \bar{x})$, and the result follows. □

It is important to mention that the conditions in the last theorem ensure convergence to a KKT point without requiring a constraint qualification. Let us emphasize that augmented Lagrangian methods that penalize all constraints do not provide this type of theoretical guarantee. We illustrate this with the next simple example.

Example 3. *Consider the problem of minimizing $f(x) := x_1 - x_2$ subject to $x \in \Omega_L \cap \Omega_U$ where*

$$\Omega_U := \{x \mid G_1(x) := x_1 \le 0, G_2(x) := -x_2 \le 0\} \text{ and } \Omega_L := \{x \mid h_1(x) := x_1^2 = 0, g_1(x) := x_2 \le 0\}.$$

It is easy to see that $\bar{x} := (0,0)$ is the global minimizer of this problem, but it is not a KKT point and, thus, no CQ is satisfied at $\bar{x}$. Since $\nabla h_1(x) = (0,0)$ for any $x \in \Omega_L = \{0\} \times \mathbb{R}_-$ and the remaining gradients are constant, it follows that $\bar{x}$ satisfies Lower-CRSC for the constraint set $\Omega_L \cap \Omega_U$ with lower-level set Ω_L and Lower-CCP for Ω_L with lower-level set Ω_L.

For this objective function, Algorithm 1 fails to generate a sequence satisfying Step 1, because the step requires computing approximate Lagrange multipliers such that

$$\left[(1,0) + (1,0)u_1^k + (0,-1)u_2^k + (0,0)\lambda^k + (0,1)\mu^k\right] \to 0$$

with $u_1^k \ge 0, u_2^k \ge 0, \mu^k \ge 0$, which clearly does not exist. This behavior is consistent with the theory developed.

On the other hand, let us consider minimizing $\tilde{f}(x) := -x_2$ subject to $x \in \Omega_L \cap \Omega_U$. The point $\bar{x}$ is also the global minimizer of this problem, which is a KKT point. Now, we affirm that $\bar{x}$ can be *reached by Algorithm 1. In fact, consider the sequence* $x^k := (0, -1/\rho_k^2) \in \Omega_L$, $\rho_k > 0$. *For each* k, *take* $\bar{u}_1^k := 0, \bar{u}_2^k \geq 0$ *and* $\lambda^k := 0, \mu^k := 1 + \bar{u}_2^k$. *We have*

$$\nabla L_{\rho_k}(x^k, u^k) + \nabla h_1(x^k)\lambda^k + \nabla g_1(x^k)\mu^k = \begin{bmatrix} 0 \\ -1 \end{bmatrix} + (1/\rho_k + \bar{u}_2^k) \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \mu^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ -1 - (1/\rho_k + \bar{u}_2^k) + \mu^k \end{bmatrix} = \begin{bmatrix} 0 \\ -(1/\rho_k + \bar{u}_2^k) + \bar{u}_2^k \end{bmatrix}$$

which converges to zero, when $\rho_k \to \infty$. *Moreover, in this case, the approximate Lagrange multiplier sequence turns out to be bounded as* $u^k = \max\{0, \bar{u}^k + \rho^k G(x^k)\} = (0, \bar{u}_2^k + 1/\rho_k)$, *yielding true Lagrange multipliers at its limit points, as Theorem 3.3 predicts, even though the problem does not satisfy a constraint qualification.*

4 Lower-level quasinormality and relaxed-quasinormality conditions

We now introduce the Lower-QN condition. This notion extends the definition given in [3] for manifold-constrained optimization in $\mathbb{R}^n$ to the broader framework of a general lower-level set.

Definition 4.1. *A point* $\bar{x} \in \Omega_L \cap \Omega_U$ *satisfies Lower-QN with respect to* $\Omega_L \cap \Omega_U$ *with lower-level set* Ω_L *when there is no* $(v, u, \lambda, \mu) \in \mathbb{R}^{\bar{m}} \times \mathbb{R}_+^{\bar{p}} \times \mathbb{R}^m \times \mathbb{R}_+^p$ *such that the following conditions hold*

1. $\sum_{i=1}^{\bar{m}} v_i \nabla H_i(\bar{x}) + \sum_{j=1}^{\bar{p}} u_j \nabla G_j(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j: g_j(\bar{x})=0} \mu_j \nabla g_j(\bar{x}) = 0;$
2. $(v, u) \neq 0;$
3. *There is a sequence* $\{x^k\} \subset \Omega_L$ *with* $x^k \to \bar{x}$ *such that for all* k,
   (a) $v_i H_i(x^k) > 0$ *for* $i \in I_{\neq} := \{i \mid v_i \neq 0\}$ *and* $G_j(x^k) > 0$ *for* $j \in J_{\neq} := \{j \mid u_j > 0\}$.

Compared to the standard QN condition for $\Omega_L \cap \Omega_U$ (using Definition 2.2 for $\Omega_L \cap \Omega_U$), the feasibility of the sequence with respect to Ω_L replaces the usual sign requirements involving (λ, μ) and the sequence $\{x^k\}$. Also, the requirement $(v, u, \lambda, \mu) \neq 0$ is replaced by $(v, u) \neq 0$. In contrast to Lower-QN introduced in [3], when Ω_L consists only of equality constraints with linearly independent gradients, where it is proved that QN implies Lower-QN, in this more general setting QN and Lower-QN are independent conditions, as shown by the following examples.

Example 4 (QN does not imply Lower-QN)**.** *For* $x = (x_1, x_2) \in \mathbb{R}^2$, *consider the upper- and lower-level sets*

$$\Omega_U := \{x \mid G_1(x) := x_1 \leq 0\} \quad \text{and} \quad \Omega_L := \{x \mid g_1(x) := -x_1 \leq 0,\ g_2(x) := -x_2 \leq 0\}.$$

The feasible point $\bar{x} := (0, 0)$ *satisfies QN, since the condition*

$$u \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \mu_1 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \mu_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 0$$

implies $u = \mu_1 \geq 0$ *and* $\mu_2 = 0$. *Moreover, there does not exist a sequence* $x^k \to \bar{x}$ *such that* $G_1(x^k) := x_1^k > 0$ *and* $g_1(x^k) := -x_1^k > 0$. *On the other hand, if we take* $u = 1$, $\mu_1 = 1$, $\mu_2 = 0$, *and consider the sequence* $\{x^k := (1/k, 0)\} \subset \Omega_L$, *then* $G_1(x^k) > 0$. *This shows that Lower-QN fails at* $\bar{x}$.

Example 5 (Lower-QN does not imply QN). *For $x = (x_1, x_2) \in \mathbb{R}^2$, consider the sets*

$$\Omega_U := \{x \mid H_1(x) := x^2 = 0\} \quad \textit{and} \quad \Omega_L := \{x \mid h_1(x) := x = 0\}.$$

Condition QN is not satisfied at the feasible point $\bar{x} := 0$, since $1 \cdot \nabla H_1(\bar{x}) + 0 \cdot \nabla h_1(\bar{x}) = 0$ and for the sequence $x^k := 1/k$, we have $H_1(x^k) > 0$. On the other hand, Lower-QN is satisfied because $v\nabla H_1(\bar{x}) + \lambda \nabla h_1(\bar{x}) = 0$ implies $\lambda = 0$. If $v \neq 0$, the only sequence $x^k \in \Omega_L$ is $x^k = 0$, which violates item 3 (a) in the Lower-QN definition. Therefore, Lower-QN holds at $\bar{x}$.

Assuming that the Lower-CCP condition is satisfied at $\bar{x} \in \Omega_L$ with lower-level Ω_L, we can verify that Lower-QN is a constraint qualification. To this end, we establish a more general result regarding necessary conditions for a point to be a minimizer of (NLP), namely a Fritz–John type condition. This result is a refinement of [20, Proposition 2.1] for the case where Ω_L is not *regular* (see [38]), providing a more precise relation. Indeed, in [20, Proposition 2.1], a similar relation is given, but in terms of the limiting normal cone rather than $\mathcal{K}^L(\bar{x}, \bar{x})$. Note that with our assumptions, $\mathcal{K}^L(\bar{x}, \bar{x})$ coincides with the normal cone, which differs from the limiting normal cone in the absence of regularity.

Theorem 4.1. *Let $\bar{x}$ be a local minimizer of problem* (NLP) *that satisfies Lower-CCP for the constraint set Ω_L with lower-level set Ω_L. Assume further that Ω_L satisfies Guignard's CQ at all points in Ω_L in a neighborhood of $\bar{x}$. Then there exists $(\mu_0, v, u, \lambda, \mu) \in \mathbb{R}_+ \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{p}}_+ \times \mathbb{R}^m \times \mathbb{R}^p_+$ satisfying the following conditions:*

1. $\mu_0 \nabla f(\bar{x}) + \sum_{i=1}^{\bar{m}} v_i \nabla H_i(\bar{x}) + \sum_{j=1}^{\bar{p}} u_j \nabla G_j(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j : g_j(\bar{x}) = 0} \mu_j \nabla g_j(\bar{x}) = 0$;
2. $(\mu_0, v, u) \neq 0$;
3. *Let $I_{\neq} := \{i \mid v_i \neq 0\}$ and $J_{\neq} := \{j \mid u_j > 0\}$. If $I_{\neq} \cup J_{\neq} \neq \emptyset$, there exists a sequence $\{x^k\} \subset \Omega_L$ that converges to $\bar{x}$ and is such that, for all k,*
   - *(a)* $f(x^k) < f(\bar{x})$, $v_i H_i(x^k) > 0$ *for all* $i \in I_{\neq}$, *and* $G_j(x^k) > 0$ *for all* $j \in J_{\neq}$;
   - *(b)* $|H_i(x^k)| = o(w(x^k))$ *for* $i \notin I_{\neq}$ *and* $G_j(x^k)_+ = o(w(x^k))$ *for* $j \notin J_{\neq}$, *where*

$$w(x^k) := \min \left\{ \min_{i \in I_{\neq}} |H_i(x^k)|,\ \min_{j \in J_{\neq}} G_j(x^k)_+ \right\}. \tag{2}$$

Proof. The argument essentially follows [20, Proposition 2.1] with additional observations that in our setting $\mathcal{K}^L(\bar{x}, \bar{x})^\circ = T_{\Omega_L}(\bar{x})^\circ$, the polar of the tangent cone to Ω_L at $\bar{x}$. We recall the main steps for completeness. We can obtain sequences $\{x^k\} \subset \Omega_L$, $\{v^k := kH(x^k)\} \subset \mathbb{R}^{\bar{m}}$, and $\{u^k := k \max\{0, G(x^k)\}\} \subset \mathbb{R}^{\bar{p}}_+$ such that

$$-\left[\mu_0^k \nabla f(x^k) + \sum_{i=1}^{\bar{m}} v_i^k \nabla H_i(x^k) + \sum_{j=1}^{\bar{p}} u_j^k \nabla G_j(x^k) \right] \in T_{\Omega_L}(x^k)^\circ, \qquad \|(\mu_0^k, v^k, u^k)\| = 1,$$

which converge to some $(\mu_0, v, u) \neq 0$. Since $g_j(x^k) = 0$ implies that $g_j(\bar{x}) = 0$ for k in some subset $\mathcal{K}$ and sufficiently large, we obtain by Guignard's CQ that $T_{\Omega_L}(x^k)^\circ \subset \mathcal{K}^L(x^k, \bar{x})$. Therefore, by the definition of $\mathcal{K}^L(x^k, \bar{x})$, there are multipliers $(\lambda^k, \mu^k) \in \mathbb{R}^m \times \mathbb{R}^p$ such that

$$w^k := \left[\sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j : g_j(\bar{x}) = 0} \mu_j^k \nabla g_j(x^k) \right] \in \mathcal{K}^L(x^k, \bar{x}) \tag{3}$$

and $\mu_0^k \nabla f(x^k) + \sum_{i=1}^{\bar{m}} v_i^k \nabla H_i(x^k) + \sum_{j=1}^{\bar{p}} u_j^k \nabla G_j(x^k) = -w^k$. Thus w^k converges to some w. By the Lower-CCP condition at $\bar{x} \in \Omega_L$, we obtain $w \in \mathcal{K}^L(\bar{x}, \bar{x})$. Hence, there is $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p_+$ such that item 1 holds. The other items follow directly from construction of the sequences $\{x^k\}$, $\{v^k\}$, and $\{u^k\}$. □

The previous result motivates the introduction of a lower-level condition, which generalizes Lower-QN, which we call *Lower relaxed-quasinormality* (Lower-RQN) condition.

Definition 4.2. *A point $\bar{x} \in \Omega_L \cap \Omega_U$ satisfies Lower-RQN with respect to $\Omega_L \cap \Omega_U$ with lower-level set Ω_L when there is no $(u, v, \lambda, \mu) \in \mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{p}}_+ \times \mathbb{R}^m \times \mathbb{R}^p_+$ such that the following conditions hold*

1. $\displaystyle\sum_{i=1}^{\bar{m}} v_i \nabla H_i(\bar{x}) + \sum_{j=1}^{\bar{p}} u_j \nabla G_j(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j: g_j(\bar{x})=0} \mu_j \nabla g_j(\bar{x}) = 0;$
2. $(v, u) \neq 0$;
3. *There is a sequence $\{x^k\} \subset \Omega_L$ with $x^k \to \bar{x}$ such that for all k,*
   (a) $v_i H_i(x^k) > 0$ *for* $i \in I_{\neq} := \{i \mid v_i \neq 0\}$ *and* $G_j(x^k) > 0$ *for* $j \in J_{\neq} := \{j \mid u_j > 0\}$.
   (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* $j \notin J_{\neq}$, *where*

$$w(x^k) := \min\left\{\min_{i \in I_{\neq}} |H_i(x^k)|, \ \min_{j \in J_{\neq}} G_j(x^k)_+\right\}.$$

It follows directly from Theorem 4.1 that both Lower-RQN and Lower-QN are constraint qualifications whenever the lower-level set Ω_L is regular enough.

Corollary 4.1. *Let $\bar{x} \in \Omega_L \cap \Omega_U$ be a local minimizer of* (NLP) *that satisfies Lower-CCP for the constraint set Ω_L with lower-level set Ω_L. Assume further that Ω_L satisfies Guignard's CQ at all points in Ω_L in a neighborhood of $\bar{x}$. If $\bar{x}$ satisfies Lower-RQN (or Lower-QN) with respect to $\Omega_L \cap \Omega_U$ with lower-level set Ω_L, then $\bar{x}$ is a KKT point of* (NLP)*.*

Proof. From Theorem 4.1, there exists a vector $(\mu_0, v, u, \lambda, \mu) \in \mathbb{R}_+ \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{p}}_+ \times \mathbb{R}^m \times \mathbb{R}^p_+$ satisfying items 1–3. If $\mu_0 \neq 0$, we can divide the expression in item 1 by μ_0. Using condition 3(a), it follows that $\mu_j = 0$ whenever $G_j(\bar{x}) < 0$, and therefore $\bar{x}$ satisfies the KKT conditions. On the other hand, if $\mu_0 = 0$, then the Lower-RQN condition is not valid at $\bar{x}$, which contradicts our assumption. □

A notion of Lower-RQN was first introduced in [5], where the lower-level set is defined by simple complementarity constraints,

$$x_i y_i = 0, \quad x_i \geq 0, \quad y_i \geq 0, \quad i = 1, \ldots, n.$$

In that context, the authors show that their condition is sufficient to ensure global convergence and boundedness of approximate Lagrange multipliers generated by the augmented Lagrangian method studied there, a method that admits M-stationary multipliers. In our framework, we extend their result to a general lower-level set satisfying Lower-CCP.

Theorem 4.2. *Let $\bar{x}$ be a limit point of a sequence $\{x^k\}_{k=1}^{\infty}$ generated by Algorithm 1 with $\{x^k\}_{k \in K} \to \bar{x}$ for some $K \subseteq \mathbb{N}$. If $\bar{x} \in \Omega_L \cap \Omega_U$ and satisfies Lower-RQN (or Lower-QN) for the constraint set $\Omega_L \cap \Omega_U$ with lower-level set Ω_L and $\bar{x}$ satisfies Lower-CCP for the constraint set Ω_L with lower-level set Ω_L, then the dual sequences $\{v^k\}_{k \in K}$ and $\{u^k\}_{k \in K}$ are bounded and any limit points of $\{v^k\}_{k \in K}$ and $\{u^k\}_{k \in K}$ are Lagrange multipliers associated with $\bar{x}$.*

Proof. Consider z^k and $\tilde{w}^k$ as defined in the proof of Theorem 3.3. We have $\tilde{w}^k \to -\nabla f(\bar{x})$ with

$$\tilde{w}^k := \sum_{i=1}^{\bar{m}} v_i^k \nabla H_i(x^k) + \sum_{j:G_j(\bar{x})=0} u_j^k \nabla G_j(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j:g_j(\bar{x})=0} \mu_j^k \nabla g_j(x^k). \quad (4)$$

Let $M_k := \max\{|v_i^k|, i = 1, \dots, \bar{m}; |u_j^k|, j : G_j(\bar{x}) = 0\}$ and suppose that $\{M_k\}_{k \in K}$ is unbounded. Consider a subsequence $K_1 \subseteq K$ such that

$$M_k \overset{k\in K_1}{\to} +\infty, \qquad \frac{v_i^k}{M_k} \overset{k\in K_1}{\to} \tilde{v}_i \in \mathbb{R},\ i = 1, \dots, \bar{m}, \qquad \frac{u_j^k}{M_k} \overset{k\in K_1}{\to} \tilde{u}_j \geq 0,\ j : G_j(\bar{x}) = 0,$$

where the limits are not all zero. Using that $\mathcal{K}^L(x^k, \bar{x})$ is a cone, we obtain

$$\left(\frac{\tilde{w}^k}{M_k} - \sum_{i=1}^{\bar{m}} \frac{v_i^k}{M_k} \nabla H_i(x^k) - \sum_{j:G_j(\bar{x})=0} \frac{u_j^k}{M_k} \nabla G_j(x^k) \right) \in \mathcal{K}^L(x^k, \bar{x}).$$

Taking the limit when $k \in K_1$ and $k \to \infty$, we conclude by Lower-CCP that

$$-\sum_{i=1}^{\bar{m}} \tilde{v}_i \nabla H_i(\bar{x}) - \sum_{j:G_j(\bar{x})=0} \tilde{u}_j \nabla G_j(\bar{x}) \in \mathcal{K}^L(\bar{x}, \bar{x}).$$

From Step 2 of Algorithm 1,

$$v^k = \bar{v}^k + \rho_k H(x^k) \quad \text{and} \quad u^k = \max\{0, \bar{u}^k + \rho_k G(x^k)\},$$

moreover, $\rho_k \to +\infty$, $\{\bar{u}^k\}$ and $\{\bar{v}^k\}$ are bounded, therefore we have $v_i H_i(x^k) > 0$, $i \in I_{\neq}$ and $G_j(x^k) > 0$, $j \in J_{\neq}$. Consequently, we obtain a sequence $\{x^k\}$ that contradicts Lower-QN. We now show that this sequence also violates Lower-RQN.

Assume that there exists some index l such that $v_l = 0$. Take $i \in I_{\neq}$ and consider the sequence $\{A_k\}$ defined by

$$A_k := \frac{|\bar{v}_l^k + \rho_k H_l(x^k)|}{|\rho_k H_i(x^k)|} = \left| \frac{\bar{v}_l^k}{\rho_k H_i(x^k)} + \frac{H_l(x^k)}{H_i(x^k)} \right|,$$

which is well-defined in view of item 1 of the Lower-RQN definition.

If $A_k \geq \varepsilon > 0$ for all $k \in K_1$, we have $|\bar{v}_l^k + \rho_k H_l(x^k)| \geq \varepsilon |\rho_k H_i(x^k)|$, and therefore

$$0 < \varepsilon |\tilde{v}_i| = \lim_{k \in K} \varepsilon \frac{|\bar{v}_i^k + \rho_k H_i(x^k)|}{M_k} = \lim_{k \in K} \varepsilon \frac{|\rho_k H_i(x^k)|}{M_k} \leq \lim_{k \in K} \frac{|\bar{v}_l^k + \rho_k H_l(x^k)|}{M_k} = |\tilde{v}_l| = 0,$$

which is a contradiction. Thus, $\liminf_{k \in K_1} A_k = 0$. Consequently, passing to a subsequence if necessary, we obtain

$$\lim_{k \in K_1} \frac{|H_l(x^k)|}{|H_i(x^k)|} = \lim_{k \in K_1} A_k = 0.$$

An analogous argument applies when l corresponds to a zero component of (v, u) and i to a nonzero component of the other vector, i.e., when $v_l = 0$ with $u_i > 0$, or $u_l = 0$ with either $v_i \neq 0$ or $u_i > 0$. By applying the same reasoning successively and passing to subsequences if necessary, we obtain item 2 of the Lower-RQN definition.

Therefore, $\{M_k\}$ is bounded and we can take a subsequence such that $v_i^k \to \tilde{v}_i \in \mathbb{R}$ for $i = 1, \ldots, \bar{m}$ and $u_j^k \to \tilde{u}_j \geq 0$ for j with $G_j(\bar{x}) = 0$, and take the limit in the expression of $\tilde{w}^k$ in (4). Using Lower-CCP at $\bar{x} \in \Omega_L$, we conclude that

$$-\nabla f(\bar{x}) - \sum_{i=1}^{\bar{m}} \tilde{v}_i \nabla H_i(\bar{x}) - \sum_{j:G_j(\bar{x})=0} \tilde{u}_j \nabla G_j(\bar{x}) \in \mathcal{K}^L(\bar{x}, \bar{x}),$$

which completes the proof. □

Remark 3. *The feasibility of $\bar{x}$ with respect to the upper-level Ω_U is not essential in the proof above. Hence, instead of assuming that the limit point $\bar{x}$ is feasible with respect to the upper-level Ω_U, one can instead extend the definition of Lower-RQN to a possible infeasible point with respect to the upper level with exactly the same formulation. By considering the index set $\{j : G_j(\bar{x}) \geq 0\}$ in the definition of $\tilde{w}^k$ in the proof above, one can obtain boundedness of the dual sequences in the same fashion, which in turn implies feasibility of $\bar{x}$ with respect to the upper-level Ω_U due to the update formula for v^k and u^k in Step 2 of the algorithm. See [14].*

Finally, we note that, as an alternative, the Lower-CCP condition with respect to the lower-level set Ω_L can be replaced by the assumption that Ω_L is regular at $\bar{x}$, in the sense that the limiting normal cone coincides with the classical normal cone at $\bar{x}$, and that it also satisfies a constraint qualification condition. In this case, [20, Proposition 2.1] (similar to Theorem 4.1) guarantees that the Lower-QN and Lower-RQN conditions are valid CQs. In the same spirit, one could replace Step 1 of Algorithm 1 by

$$\text{“compute } x^k \in \Omega_L \text{ such that } -\nabla L_{\rho_k}(x^k; \bar{v}^k, \bar{\mu}^k) \in N_{\Omega_L}(x^k)\text{”}$$

where $N_{\Omega_L}(x^k)$ denotes the limiting normal cone of Ω_L at x^k. In this way, we obtain a result analogous to Theorem 4.2 for this modified method, now relying on regularity rather than continuity of the mapping $x \mapsto \mathcal{K}^L(x, \bar{x})$.

5 Illustrations and further discussion of the algorithm

The convergence results established in Sections 3–4 show that the proposed lower-level constraint qualifications (Lower-CRSC and Lower-QN/Lower-RQN), combined with the continuity properties of the lower-level KKT cone, provide global convergence of Algorithm 1 under mild assumptions. In this section, we complement these theoretical developments with examples and qualitative insights illustrating why the use of lower-level constraints, rather than penalizing all constraints, leads to a markedly different behavior of the augmented Lagrangian method. We highlight scenarios in which the algorithm enjoys stable dual iterates, even in the absence of constraint qualifications, and contrast this behavior with fully penalized augmented Lagrangian schemes.

5.1 When the assumptions of the main theorems do not represent a CQ

We have established global convergence of Algorithm 1 to KKT points under either one of the following conditions:

- Theorem 3.3 - Lower-CRSC at $\bar{x}$ for the constraint set $\Omega_L \cap \Omega_U$, together with Lower-CCP at $\bar{x}$ for the constraint set Ω_L;
- Theorem 4.2 - Lower-RQN at $\bar{x}$ for the constraint set $\Omega_L \cap \Omega_U$, together with Lower-CCP at $\bar{x}$ for the constraint set Ω_L.

However, neither of these assumptions is a constraint qualification (and thus, they do not guarantee that every local minimizer is a KKT point). To see this, consider the problem of minimizing $f(x) := x$ subject to $x \in \Omega_L \cap \Omega_U$ where

$$\Omega_U := \{x \mid G_1(x) := x \leq 0\} \text{ and } \Omega_L := \{x \mid h_1(x) := x^2 = 0\}.$$

The feasible set is $\{0\}$, and the unique minimizer is $\bar{x} := 0$. Since Ω_L consists only of the point $\bar{x}$, both above conditions are trivially satisfied at $\bar{x}$. Nevertheless, $\bar{x}$ is not a KKT point.

Theorems 3.3 and 4.2 assert something subtler: if Algorithm 1 generates a sequence $\{x^k\}$ with a limit point $\bar{x}$ that satisfies the corresponding assumptions, then this limit point must be a KKT point of the original problem. In particular, in the example above, the algorithm cannot generate a sequence $\{x^k\}$ such that $\|\nabla f(x^k) + \nabla G_1(x^k)v^k + \nabla h_1(x^k)\lambda^k\| \to 0$ with $x^k \to \bar{x}$, since $\bar{x}$ is not a KKT point. Therefore the subproblems must be ill-defined. See also Example 3.

5.2 Illustration: qualitative difference between lower-level and full penalization

We now illustrate how the assumptions of the main theorems lead to a qualitative difference between Algorithm 1 and the method that penalizes *all* constraints.

Example 6. *Consider the feasible set $\Omega_L \cap \Omega_U$, where*

$$\Omega_U := \{x \mid G_1(x) := -x_1 \leq 0\} \quad \textit{and} \quad \Omega_L := \{x \mid g_1(x) := -x_1 \leq 0,\ g_2(x) := x_1^3 \exp(x_2^2) \leq 0\}.$$

We consider the feasible point $\bar{x} := (0,0)$. It is easy to see that $\Omega_L = \{0\} \times \mathbb{R}$ and therefore

$$\nabla g_1(x) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \qquad \nabla g_2(x) = \begin{bmatrix} 3x_1^2 \exp(x_2^2) \\ 2x_2 x_1^3 \exp(x_2^2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \textit{for all } x \in \Omega_L.$$

Consequently, for every $x \in \Omega_L$ near $\bar{x}$ the cone generated by the lower-level constraints is

$$\mathcal{K}^L(x, \bar{x}) = \{\mu_1 \nabla g_1(x) + \mu_2 \nabla g_2(x) \mid \mu_1, \mu_2 \geq 0\} = \mathbb{R}_- \times \{0\},$$

and this coincides with $\mathcal{K}^L(\bar{x}, \bar{x})$. Hence Lower-CCP for Ω_L with lower-level set Ω_L is satisfied at $\bar{x}$. In order to check Lower-QN, we consider items 1-2 of Definition 4.1 to form the system

$$u_1 \nabla G_1(\bar{x}) + \mu_1 \nabla g_1(\bar{x}) + \mu_2 \nabla g_2(\bar{x}) = 0$$

with $(u_1, \mu_1, \mu_2) \geq 0$ and $u_1 \neq 0$. However, since $\nabla G_1(\bar{x}) = (-1, 0)$ the only possibility is $u_1 = \mu_1 = 0$. This shows that Lower-QN holds. We have therefore verified that, at $\bar{x}$, the assumptions of Theorem 4.2 are satisfied: $\bar{x}$ satisfies Lower-QN for $\Omega_L \cap \Omega_U$ together with Lower-CCP for the constraint set Ω_L. Consequently, if a sequence $\{x^k\}$ is generated by Algorithm 1, any limit point will be a KKT point of problem (NLP) *and the dual sequence generated by the method via Step 2, will be bounded.*

Suppose now we penalize all constraints. That is, we consider the same problem but with a different partition $\Omega_U := \{x \mid G_1(x) \leq 0, g_1(x) \leq 0, g_2(x) \leq 0\}$ and $\Omega_L := \mathbb{R}^2$. It is easy to see that $\bar{x} := (0,0)$ is a global minimizer of $\tilde{f}(x) := -x_1$ subject to $x \in \Omega_L \cap \Omega_U$ and that the KKT conditions are not valid. Hence, no constraint qualification holds at $\bar{x}$, and, in particular, no classical global convergence result is available (RQN fails). Consider now $f(x) := -x_1^2 + x_2^2$ and the application of Algorithm 1 with all constraints penalized. The point $\bar{x} := (0,0) \in \Omega_L \cap \Omega_U$ is a KKT point and a global minimizer. Consider $\rho_k := \frac{2}{3}k^2$ and $x^k := \big(1/\sqrt{k}, 0\big) \to (0,0)$. Along the sequence $\{x^k\}$,

$$G_1(x^k) = -1/\sqrt{k} < 0, \qquad g_1(x^k) = -1/\sqrt{k} < 0, \qquad g_2(x^k) = (1/\sqrt{k})^3 > 0.$$

Thus, the first-order stationary condition (Step 1 of Algorithm 1) is satisfied exactly as

$$\begin{bmatrix} -2/\sqrt{k} \\ 0 \end{bmatrix} + u_1^k \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \mu_1^k \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \mu_2^k \begin{bmatrix} 3/k \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where $u_1^k := \max\{0, \bar{u}_1^k + \rho_k G_1(x^k)\}$ *and* $\mu_1^k := \max\{0, \bar{\mu}_1^k + \rho_k g_1(x^k)\}$ *for some bounded sequences* $\{\bar{u}_1^k\}$ *and* $\{\bar{\mu}_1^k\}$*, which implies* $u_1^k \to 0$ *and* $\mu_1^k \to 0$*. Moreover,* $\mu_2^k := \max\{0, \bar{\mu}_2^k + \rho_k g_2(x^k)\} = \bar{\mu}_2^k + \frac{2}{3}\sqrt{k}$ *for some bounded sequence* $\{\bar{\mu}_1^k\}$*. Hence* $\{\mu_2^k\}$ *is unbounded, and the method fails to produce a corresponding KKT multiplier.*

Of course, the examples presented in this section serve an illustrative purpose. Nevertheless, our results show that global convergence can be achieved even in the absence of any constraint qualification. Numerical stability of the multipliers is also preserved in this setting without a constraint qualification. This stability arises precisely because we avoid penalizing all constraints in our framework. This phenomenon is consistent with the improved numerical performance observed in the literature, see for instance [3, 5, 32].

6 Conclusions

In recent years, several new constraint qualifications have appeared with the goal of providing sharper global convergence results for safeguarded augmented Lagrangian methods that penalize all constraints. The Constant Rank of the Subspace Component condition (CRSC, [16]) plays a major role in this analysis, as it is the weakest condition within a family of constant rank conditions. Moreover, the relaxed-quasinormality condition (RQN, [14]), which is equivalent to the error bound property [13], has been shown to be the weakest condition guaranteeing boundedness of the dual augmented Lagrangian sequences [13].

These conditions can also be used to analyze the version of the algorithm that keeps some of the constraints within the subproblems; however, only recently has it been observed that a tailored analysis can be carried out in order to obtain weaker conditions in the case of constrained subproblems. There are at least two settings in which such an analysis has been conducted. The first corresponds to subproblem constraints defining a manifold [3], as in the MANOPT implementation [25]. The second concerns simple complementarity constraints [5], as in the implementation proposed in [29]. In [3], tailored CRSC and quasinormality (QN) conditions were introduced, while in [5], tailored QN and RQN conditions were defined.

In this paper, we define tailored CRSC, QN, and RQN conditions for an arbitrary choice of constrained subproblems. This is particularly relevant in the box-constrained case, since the ALGENCAN implementation [24] considers box-constrained subproblems. Our developments take into account the fact that all these implementations compute approximate solutions to the subproblems, but always from within the subproblems' feasible set. That is, the sequence generated by the algorithm remains feasible with respect to the constrained subproblems.

Somewhat surprisingly, the conditions required for global convergence and stability of the algorithm with constrained subproblems are not necessarily constraint qualifications. More precisely, when an objective function associated with the KKT conditions is used, the algorithm behaves as expected, whereas when the KKT conditions fail for a given objective function, the corresponding subproblems become ill-defined.

Finally, we expect that further studies should be carried out in the context of constrained subproblems, in particular concerning the equivalence between RQN and the error bound property, as well as the minimality of RQN with respect to boundedness of dual augmented Lagrangian sequences.

Extensions to conic programming are also expected due to recent studies of CRSC [10] and QN/RQN [39] in this context as well as extensions to infinite dimensional spaces [37].

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, 2008.

[2] R. Andreani, K. R. Couto, O. P. Ferreira, and G. Haeser. Constraint qualifications and strong global convergence properties of an augmented lagrangian method on riemannian manifolds. *SIAM Journal on Optimization*, 34(2):1799–1825, 2024.

[3] R. Andreani, K. R. Couto, O. P. Ferreira, G. Haeser, and L. F. Prudente. Global convergence of an augmented Lagrangian method for nonlinear programming via Riemannian optimization. *To appear in SIAM Journal on Optimization*, 2025.

[4] R. Andreani, M. da Rosa, and L. D. Secchin. A new constant-rank-type condition related to mfcq and local error bounds. Technical report, 2025.

[5] R. Andreani, M. da Rosa, and L. D. Secchin. On the boundedness of multipliers in augmented lagrangian methods for mathematical programs with complementarity constraints. Technical report, Optimization Online, Aug 2025.

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

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

[8] R. Andreani, G. Haeser, L. M. Mito, C. H. Ramírez, and T. P. Silveira. Global convergence of algorithms under constant rank conditions for nonlinear second-order cone programming. *Journal of Optimization Theory and Applications*, 195(1):42–78, 2022.

[9] R. Andreani, G. Haeser, L. M. Mito, and H. Ramírez. Sequential constant rank constraint qualifications for nonlinear semidefinite programming with algorithmic applications. *Set-Valued and Variational Analysis*, 31(1), 2023.

[10] R. Andreani, G. Haeser, L. M. Mito, and H. Ramírez. A minimal face constant rank constraint qualification for reducible conic programming. *Mathematical Programming*, 2025. DOI: 10.1007/s10107-025-02237-w.

[11] R. Andreani, G. Haeser, L. M. Mito, H. Ramírez, D. O. Santos, and T. P. Silveira. Naive constant rank-type constraint qualifications for multifold second-order cone programming and semidefinite programming. *Optimization Letters*, 16(2):589–610, 2021.

[12] R. Andreani, G. Haeser, L. M. Mito, H. Ramírez, and T. P. Silveira. First- and second-order optimality conditions for second-order cone and semidefinite programming under a constant rank condition. *Mathematical Programming*, 202(1–2):473–513, 2023.

[13] 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.

[14] R. Andreani, G. Haeser, M. L. Schuverdt, and L. D. Secchin. A relaxed quasinormality condition and the boundedness of dual augmented lagrangian sequences. *SIAM Journal on Optimization*, 35(4):2474–2489, 2025.

[15] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Mathematical Programming*, 135(1–2):255–273, 2011.

[16] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. Two new weak constraint qualifications and applications. *SIAM Journal on Optimization*, 22(3):1109–1135, 2012.

[17] 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.

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

[19] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 2nd edition, 1999.

[20] 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.

[21] E. G. Birgin, L. F. Bueno, and J. M. Martínez. Sequential equality-constrained optimization for nonlinear programming. *Computational Optimization and Applications*, 65(3):699–721, 2016.

[22] E. G. Birgin, G. Haeser, and J. M. Martínez. Safeguarded augmented lagrangian algorithms with scaled stopping criterion for the subproblems. *Computational Optimization and Applications*, 91:491–509, 2025.

[23] E. G. Birgin, G. Haeser, and A. Ramos. Augmented lagrangians with constrained subproblems and convergence to second-order stationary points. *Computational Optimization and Applications*, 69(1):51–75, 2017.

[24] E. G. Birgin and J. M. Martínez. *Practical augmented Lagrangian methods for constrained optimization*, volume 10 of *Fundamentals of Algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2014.

[25] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research*, 15(42):1455–1459, 2014.

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

[27] L. F. Bueno, G. Haeser, and L.-R. Santos. Towards an efficient augmented lagrangian method for convex quadratic programming. *Computational Optimization and Applications*, 76(3):767–800, 2019.

[28] 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*, page 257–279, 2024.

[29] 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.

[30] M. R. Hestenes. *Optimization Theory: The Finite Dimensional Case.* John Wiley & Sons, New York, NY, 1975.

[31] R. Janin. Directional derivative of the marginal function in nonlinear programming. *Math. Program. Stud.*, 21:110–126, 1984.

[32] X. Jia, C. Kanzow, P. Mehlitz, and G. Wachsmuth. An augmented Lagrangian method for optimization problems with structured geometric constraints. *Mathematical Programming*, 199(1–2):1365–1415, 2023.

[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] L. Minchenko and S. Stakhovski. On relaxed constant rank regularity condition in mathematical programming. *Optimization*, 60(4):429–440, 2011.

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

[36] L. Qi and Z. Wei. On the constant positive linear dependence condition and its application to sqp methods. *SIAM Journal on Optimization*, 10(4):963–981, 2000.

[37] A. Ramos. New Constraint Qualifications based on the decomposition of the cone defined by the Karush-Kuhn-Tucker conditions. *Journal of Optimization Theory and Applications*, 208(1):43, 2026.

[38] R. T. Rockafellar and R. J. B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, 1998.

[39] D. O. Santos. Quasinormality and pseudonormality for nonlinear semidefinite programming. Preprint, 2025.