New Verifiable Sufficient Conditions for Metric Subregularity of Constraint Systems with Applications to Disjunctive Programs

Matúš Benko

Institute of Computational Mathematics, Johannes Kepler University
Altenberger Str. 69, 4040 Linz, Austria

Michal Červinka
Institute of Economical Studies, Faculty of Social Sciences
Charles University, Opletalova 26
110 00 Prague 1, Czech Republic
and
Institute of Information Theory and Automation
Czech Academy of Sciencess, Pod Vodarenskou vezi 4
180 00 Prague 8, Czech Republic

Tim Hoheisel
Institute of Mathematics and Statistics, McGill University
805 Sherbrooke St West, Room 1114
Montréal, Québec, Canada H3A 0B9

Technical Reports before 1998:

1995

95-1	Hedwig Brandstetter	
	Was ist neu in Fortran 90?	March 1995
95-2	G. Haase, B. Heise, M. Kuhn, U. Langer	
	Adaptive Domain Decomposition Methods for Finite and Boundary Element Equations.	August 1995
95-3	Joachim Schöberl	
	An Automatic Mesh Generator Using Geometric Rules for Two and Three Space Dimensions.	August 1995

1996

96-1	Ferdinand Kickinger	
	Automatic Mesh Generation for 3D Objects.	February 1996
96-2	Mario Goppold, Gundolf Haase, Bodo Heise und Michael Kuhn	
	Preprocessing in BE/FE Domain Decomposition Methods.	February 1996
96-3	Bodo Heise	
	A Mixed Variational Formulation for 3D Magnetostatics and its Finite Element Discretisation.	February 1996
96-4	Bodo Heise und Michael Jung	
	Robust Parallel Newton-Multilevel Methods.	February 1996
96-5	Ferdinand Kickinger	
	Algebraic Multigrid for Discrete Elliptic Second Order Problems.	February 1996
96-6	Bodo Heise	
	A Mixed Variational Formulation for 3D Magnetostatics and its Finite Element Discretisation.	May 1996
96-7	Michael Kuhn	
	Benchmarking for Boundary Element Methods.	June 1996

1997

97-1	Bodo Heise, Michael Kuhn and Ulrich Langer	
	A Mixed Variational Formulation for 3D Magnetostatics in the Space $H(rot) \cap H(div)$	February 1997
97-2	Joachim Schöberl	
	Robust Multigrid Preconditioning for Parameter Dependent Problems I: The Stokes-type Case.	June 1997
97-3	Ferdinand Kickinger, Sergei V. Nepomnyaschikh, Ralf Pfau, Joachim Schöberl	
	Numerical Estimates of Inequalities in $H^{\frac{1}{2}}$.	August 1997
97-4	Joachim Schöberl	
	Programmbeschreibung NAOMI 2D und Algebraic Multigrid.	September 1997

NEW VERIFIABLE SUFFICIENT CONDITIONS FOR METRIC SUBREGULARITY OF CONSTRAINT SYSTEMS WITH APPLICATIONS TO DISJUNCTIVE PROGRAMS

MATÚŠ BENKO*, MICHAL ČERVINKA†,‡, AND TIM HOHEISEL⊥

ABSTRACT. This paper is devoted to the study of the *metric subregularity constraint qualification (MSCQ)* for optimization problems with nonconvex constraints. We propose a unified theory for several prominent sufficient conditions for MSCQ, which is achieved by means of a new constraint qualification that combines the well-established approach via pseudo- and quasi-normality with the recently developed tools of directional variational analysis. When applied to *disjunctive programs* this new constraint qualification unifies Robinson's celebrated result on polyhedral multifunctions and Gfrerer's *second-order sufficient condition for metric subregularity*. Finally, we refine our study by defining the new class of *ortho-disjunctive programs* which comprises prominent problems such as *mathematical problems with complementarity, vanishing or switching constraints*.

1 Introduction

In recent years there has been an increasing interest in optimization problems with inherently nonconvex structures induced by imposing logical or combinatorial conditions on otherwise smooth or convex data [56]. Particularly prominent examples are *mathematical programs with complementarity constraints* (MPCCs), *mathematical programs with vanishing constraints* (MPVCs), *mathematical programs with relaxed cardinality constraints* (MPrCCs), *mathematical programs with relaxed probabilistic constraints* (MPrPCs), as well as the recently introduced *mathematical programs with switching constraints* (MPSCs). For these optimization problems there is a vast array of applications in the natural and social sciences, economics and engineering. Moreover, they are very challenging from both a theoretical and numerical perspective. For the mathematical background and several of these applications we refer the reader to the textbooks [47, 52] for MPCCs as well as to the book [18] on the closely related class of bilevel programs. As for MPVCs we refer to the paper [1] and the thesis [35] and the references therein. For relaxed cardinality constrained problems we point to the papers [11, 12, 13, 16]. For MPrPCs see [2], and for MPSCs see [45, 49].

Thus far, these different types of programs have been studied mainly independently, but using similar techniques to prove analogous results. In this paper, we work in a unified framework for the above problem classes, concentrating on the underlying *disjunctive* structure. Our starting point is a general mathematical program (GMP) given by

$$\min_{x \in \mathbb{R}^n} \ f(x) \quad \text{s.t.} \quad x \in F^{-1}(\Gamma) =: \mathcal{X}, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $F : \mathbb{R}^n \to \mathbb{R}^d$ are continuously differentiable and $\Gamma \subset \mathbb{R}^d$ is closed. We will then progressively specify the structure of Γ during the course of our study, starting with no

*Institute of Computational Mathematics, Johannes Kepler University Linz, A-4040 Linz, Austria, e-mail: benko@numa.uni-linz.ac.at.

†Institute of Economic Studies, Faculty of Social Sciences, Charles University, Opletalova 26, 110 00, Prague 1, Czech Republic, e-mail: michal.cervinka@fsv.cuni.cz

‡Institute of Information Theory and Automation, Czech Academy of Sciences, Pod Vodarenskou vezi 4, 180 00 Prague 8, Czech Republic, e-mail: cervinka@utia.cas.cz.

⊥Institute of Mathematics and Statistics, McGill University, 805 Sherbrooke St West, Room 1114 Montréal, Québec, Canada H3A 0B9.

assumptions (except closedness) in Section 3. In Section 4 we focus on *disjunctive programs* in which Γ is the finite union of polyhedra. Disjunctive programs were already successfully employed in [22, 25] and systematically studied in the thesis [5], and most recently in [4]. In Section 5 we introduce the new class of *ortho-disjunctive programs*. Ortho-disjunctive programs are special disjunctive programs that have another product structure which allows us to address certain issues that cannot be resolved in the more general disjunctive setting. Mathematical programs with complementarity, vanishing and switching constraints are special instances of ortho-disjunctive programs. Our main workhorse throughout, is the recently developed directional approach by Gfrerer and co-authors [23, 24, 25, 26, 27, 6].

The focus of this paper is the study of *constraint qualifications (CQs)* which play a crucial role in the variational analysis of mathematical programs such (1), e.g., when dealing with stationarity and optimality conditions, sensitivity analysis or exact penalization. At the center of our attention is the so-called *metric subregularity constraint qualification* (MSCQ). Known also under other monikers such as *error bound property* or *calmness constraint qualification*, MSCQ is, to the best of our knowledge, the weakest known CQ to ensure the full calculus for (limiting) normal cones and subdifferentials, see [42], which is absolutely essential for a whole variety of issues in optimization. Concretley, the basic necessary optimality condition, based on the regular normal cone, for (1) at $\bar{x} \in \mathcal{X} = F^{-1}(\Gamma)$ reads

$$- \nabla f(\bar{x}) \in \widehat{N}_{\mathcal{X}}(\bar{x}), \tag{2}$$

see e.g. [55, Theorem 6.12]. It is very difficult to efficiently use this condition due to the intractability of the implicitly given set $\mathcal{X}$ and the insufficient calculus of the regular normal cone. Under MSCQ, however, one can work with a more versatile first-order condition, based on the limiting normal cone, namely

$$-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T N_\Gamma(F(\bar{x})),$$

see [42]. This condition is typically referred to as *Mordukhovich (M)-stationarity*, see [22], and it is strictly weaker than (2). Apart from M-stationarity, several other stationarity conditions have been studied in the literature, in particular for MPCCs and MPVCs. Except for so-called *strong stationarity*, however, these standard conditions are even weaker than M-stationarity and thus of limited use from a theoretical perspective, although there is some relevance in an algorithmic setting, see, e.g., [13, 37, 38, 43]. We refer the interested reader to the Gfrerer's newly developed stationarity concepts of $\mathcal{Q}$- and *linearized M-stationarity* with remarkable properties, see [3, 29].

We point out that for programs with disjunctive constraints, M-stationarity of a local minimizer can be also shown under milder generalized Guignard constraint qualification (GGCQ), what was first observed in [22, Theorem 7].

Apart from the area of optimality conditions, MSCQ turns out to be essential also in the second-order variational analysis and closely related areas of stability and sensitivity. For a brief sample of the numerous very recent works, see, e.g., [7, 8, 27, 28] and the references therein. For more details on metric subregularity and related condition we refer to [21, 31, 32, 46, 53, 57].

The main drawback of MSCQ is the difficulty of efficiently verifying that this property holds. There are two main tools for achieving that: The first approach is to consider the stronger property *metric regularity*, closely related to other concepts such as the *Aubin property*, *(generalized) Mangasarian-Fromovitz constraint qualification (GMFCQ)*, *no nonzero abnormal multiplier constraint qualification (NNAMCQ)*. Metric regularity can be efficiently verified in terms of generalized derivatives by the celebrated *Mordukhovich criterion*, see [55, Theorem 9.40] and the bibliographical annotations therein on the evolution of said result. The second approach corresponds to Robinon's famous result on polyhedral multifunctions [54, Proposition 1]. Arguably both approaches have their limitations but, most importantly, there are many situations in which both impose excessively strict assumptions, yet metric subregularity is provably satisfied.

Therefore, a lot of attention has been recently given to conditions that lie between metric regularity and metric subregularity. Two of the most prominent strategies are the following. The first type is obtained by the so-called *pseudo-* and *quasi-normality*, first introduced for nonlinear programming in [9], and later extended to MPCCs in [44, 58] as well as to general programs of the form (1), see [30]. The second one was established and heavily utilized in recent years by

Gfrerer [23, 24, 26] under the name *first/second-order sufficient condition for metric subregularity* (FOSCMS/SOSCMS). FOSCMS resembles the Mordukhovich criterion for metric regularity, but is based on directional counterparts of standard generalized derivatives, making it a less restrictive condition. The main advantage of these conditions is that they are *point-based* which makes it possible to verify them efficiently. Note that these two strategies are, in general, independent and not comparable.

Contributions (with pointers)

The main achievement of the paper is a unification and simplification of existing approaches to sufficient conditions for the metric subregularity constraint qualification for optimization problems with nonconvex constraint structure. More concretely, the contributions of the paper are as follows:

- *New mild CQs for the general program* (1)*:* By successfully synthesizing the directional approach due to Gfrerer with the notion of pseudo- and quasi-normality, we obtain new constraint qualifications, *directional pseudo-/quasi-normality*, that imply MSCQ, and that are (by definition) milder than both pseudo-/quasi-normality and FOSCMS. For the implication that directional quasi-normality implies MSCQ, the very foundation of this paper, see Theorem 3.2 (Theorem 3.6). As a byproduct, we recover new and comparably simpler proofs of the known results that pseudo-/quasi-normality as well as FOSCMS imply MSCQ.
- *New interpretation of pseudo-normality for disjunctive programs:* When considering the general program (1) under the assumption that Γ is the finite union of convex polyhedra, we observe that pseudo-normality can be cast in a simpler way which is, in fact, a proper extension of the definition of pseudo-normality that has already been used for NLPs and MPCCs in the literature. This new definition, however, reveals a significant interpretation of pseudo-normality via certain maximality condition, which is neither visible from the general definition for (1) nor from the specially tailored ones for NLPs and MPCCs, respectively. This in turn yields three striking implications. First, the affine constraint mappings satisfy this condition by default, thus recovering the Robinson's result. Second, considering directional pseudo-normality instead and applying the second-order sufficient conditions for the maximality condition, one obtains Gfrerer's SOSCMS, unifying the concepts of pseudo-normality on the one hand and directional SOSCMS (and Robinson's result) on the other. Third, the analysis of the maximality condition is not restricted to second-order conditions and can even be improved by considering higher-order analysis, ultimately yielding new mild point-based sufficient conditions for MSCQ.

    For the maximality condition, see formula (28) of Corollary 4.6. The equivalence between (28) and pseudo-normality is due to Theorem 3.9, taking into account that Assumption 3.8 is fulfilled by Corollary 4.6. On the other hand, in the directional case one needs to employ Theorem 3.17 to show that (28) *implies* direction pseudo-normality.

    Note also that Corollary 4.6 holds due to Lemma 4.1 and Proposition 4.3. Moreover, for the application of the higher-order conditions as well as neat comparison of various approaches, we refer to Example 4.12. Finally, for the main result, see Theorem 4.11.
- *Quasi-normality and multi-objective optimization:* A similar approach as the one to pseudo-normality can be made to (directional) quasi-normality if one moves from the disjunctive to the even more specialized ortho-disjunctive setting. The corresponding conditions ensuring quasi-normality lead to a surprising connection between quasi-normality and multi-objective optimization, see Corollary 5.4, in particular formula (40). As a result, we obtain *second-order* point-based conditions ensuring MSCQ that are milder than analogous conditions based on pseudo-normality, see Corollary 3.11 for the standard case and Proposition 3.19 for the directional case. In the standard case, these new conditions turn out to be actually strictly milder, see Example 3.15. In the directional case, however, we were not able to determine whether they are in fact strictly milder than SOSCMS.
- *Ortho-disjunctive programs:* As advertised, we propose a new problem class, namely ortho-disjunctive programs which, in addition to the disjunctive structure, exhibits an underlying product structure, see Section 5.2. The ortho-disjunctive programs enable us to resolve certain

issues regarding quasi-normality that cannot be resolved for general disjunctive programs. For the main result, see Theorem 5.5.
- *PQ-normality:* Finally, in the main Section 3 we actually work with the new notion of (directional) PQ-normality, which we invented as a generalization/unification of pseudo- and quasi-normality. Interestingly, PQ-normality turns out to be a very natural extension, suitable in particular for a class of programs studied in Section 5.1.

Organization

The rest of the paper is organized as follows. In Section 2 we introduce some preliminary results and notions from variational analysis as well as the main result regarding constraint qualifications. We also briefly discuss exact penalization and the role of MSCQ. Section 3 contains the most fundamental results of our study dealing with CQs for the general program (1). In Section 4, we study disjunctive programs and obtain full results on pseudo-normality. Section 5 deals with disjunctive programs with additional product structures often present in the problems of interest (MPCCs, MPVCs etc.). In particular, the ortho-disjunctive programs are introduced and complete results on quasi-normality are obtained.

Notation: Most of the notation used is standard: The closed ball in $\mathbb{R}^n$ with center at x and radius r is denoted by $\mathbb{B}_r(x)$ and we use $\mathbb{B} := \mathbb{B}_1(0)$ for the closed unit ball. The extended real line is given by $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$. For $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ its *epigraph* is given by $\operatorname{epi} f := \{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \le \alpha\}$. For a nonempty set $S \subset \mathbb{R}^n$ we define the (Euclidean) *distance function* $\mathrm{d}_S : \mathbb{R}^n \to \mathbb{R}$ through $\mathrm{d}_S(x) := \inf_{y \in S} \|x - y\|$. The *projection mapping* $P_S : \mathbb{R}^n \rightrightarrows S$ associated with S is defined by $P_S(x) := \operatorname{argmin}_{y \in S} \|x - y\|$. For a mapping $f : \mathbb{R}^n \to \mathbb{R}$ we use $\nabla f(\bar{x})$ to denote the *gradient* of f at $\bar{x}$ and $\nabla^2 f(\bar{x})$ to denote its *Hessian* at $\bar{x}$. For a mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ with $m > 1$, however, $\nabla F(\bar{x})$ stands for the *Jacobian* of F at $\bar{x}$. Moreover, given $\lambda \in \mathbb{R}^m$ the scalarized function $\langle \lambda, F \rangle : \mathbb{R}^n \to \mathbb{R}$ is given by $\langle \lambda, F \rangle (x) = \lambda^T F(x)$. Note that for $u \in \mathbb{R}^n$ we have $\nabla \langle \lambda, F \rangle (\bar{x})^T u = \langle \lambda, \nabla F(\bar{x}) u \rangle$ and we often use the latter notation. For a matrix $A \in \mathbb{R}^{m \times n}$, its *range* or *image* is $\operatorname{Im} A := \{A x \mid x \in \mathbb{R}^n\}$. For some vector $v \in \mathbb{R}^n$ we set $\mathbb{R}_+ v := \{tv \mid t \ge 0\}$ and $\mathbb{R}_- v := \{tv \mid t \le 0\}$.

2 Preliminaries

This section is divided into three parts. First, we introduce some basic notions and principles from variational analysis. The second part is devoted to constraint qualifications for the general mathematical program (1), and the last part shows the importance of metric subregularity (MSCQ) for exact penalization.

2.1 Variational analysis

Given a closed set $C \subset \mathbb{R}^n$ and $z \in C$, the *tangent cone* to C at z is defined by

$$T_C(z) := \{d \in \mathbb{R}^n \mid \exists \{d_k\} \to d, \{t_k\} \downarrow 0 : \ z + t_k d_k \in C \ (k \in \mathbb{N})\}.$$

The *regular normal cone* to C at z is given as the polar cone of the tangent cone, i.e.

$$\widehat{N}_C(z) := \{z^* \in \mathbb{R}^n \mid \langle z^*, d \rangle \le 0 \ (d \in T_C(z))\}.$$

The *limiting normal cone* to C at z is given by

$$N_C(z) := \left\{z^* \in \mathbb{R}^n \ \middle| \ \exists \{z_k^*\} \to z^*, \{z_k\} \to z : \ z_k \in C, z_k^* \in \widehat{N}_C(z_k) \ (k \in \mathbb{N})\right\}.$$

If $z \notin C$ we set $T_C(z) := \widehat{N}_C(z) := N_C(z) := \emptyset$. Observe that $\widehat{N}_C(z) \subset N_C(z)$ holds. In case C is a convex set, regular and limiting normal cone coincide with the classical normal cone of convex analysis, i.e.,

$$\widehat{N}_C(z) = N_C(z) = \{z^* \in \mathbb{R}^n \mid \langle z^*, v - z \rangle \le 0 \ (v \in C)\}, \tag{3}$$

and we will use the notation $N_C(z)$ in this case. Finally, given a direction $d \in \mathbb{R}^n$, the *limiting normal cone* to C at z *in direction* d is defined by

$$N_C(z; d) := \left\{ z^* \in \mathbb{R}^n \ \middle|\ \exists \{t_k\} \downarrow 0,\ \{d_k\} \to d,\ \{z_k^*\} \to z^* :\ z_k^* \in \widehat{N}_C(z + t_k d_k)\ (k \in \mathbb{N}) \right\}.$$

Note that, by definition, we have $N_C(z; 0) = N_C(z)$. Furthermore, observe that $N_C(z; d) \subset N_C(z)$ for all $d \in \mathbb{R}^n$ and $N_C(z; d) = \emptyset$ if $d \notin T_C(z)$.

For $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\bar{x}$ such that $f(\bar{x})$ is finite (hence $(\bar{x}, f(\bar{x})) \in \operatorname{epi} f$) the sets

$$\hat{\partial} f(\bar{x}) := \left\{ \xi \in \mathbb{R}^n \ \middle|\ (\xi^*, -1) \in \widehat{N}_{\operatorname{epi} f}\big(\bar{x}, f(\bar{x})\big) \right\}, \quad \partial f(\bar{x}) := \left\{ \xi \in \mathbb{R}^n \ \middle|\ (\xi^*, -1) \in N_{\operatorname{epi} f}\big(\bar{x}, f(\bar{x})\big) \right\}$$

denote the *regular* and *limiting subdifferential* of f at $\bar{x}$, respectively. Observe that, in particular, for the indicator function of a set $C \in \mathbb{R}^n$, given by

$$\delta_C : x \mapsto \begin{cases} 0 & \text{if} \quad x \in C, \\ +\infty & \text{else,} \end{cases}$$

we have $\hat{\partial}\delta_C = N_C$ and $\partial\delta_C = N_C$. The distance function enjoys a rich subdifferential calculus briefly summarized in the next result.

Proposition 2.1 (Subdifferentiation of distance function). *Let $S \subset \mathbb{R}^d$ be closed and $F : \mathbb{R}^n \to \mathbb{R}^d$ continuously differentiable. Then the following hold:*

(i) [55, Example 8.53] $\partial \mathrm{d}_S(y) = \begin{cases} N_S(y) \cap \mathbb{B} & \text{if} \quad y \in S, \\ \frac{y - P_S(y)}{\mathrm{d}_S(y)} & \text{if} \quad y \notin S; \end{cases}$

(ii) [55, Theorem 10.6] $\partial(\mathrm{d}_S \circ F)(x) \subset \nabla F(x)^T \partial \mathrm{d}_S(F(x)).$

We will make use of *Ekeland's variational principle* [20], which we provide for the reader's convenience in the form given in [55, Proposition 1.43].

Proposition 2.2 (Ekeland's variational principle). *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ has closed epigraph* $\operatorname{epi} f$ *with* $\inf_{x \in \mathbb{R}^n} f(x)$ *finite, and let $\bar{x}$ be an ε-minimizer of f for some $\varepsilon > 0$, i.e., $f(\bar{x}) \leq \inf_x f(x) + \varepsilon$. Then for any $\delta > 0$ there exists a point $\tilde{x}$ such that*

$$\|\tilde{x} - \bar{x}\| \leq \varepsilon/\delta, \quad f(\tilde{x}) \leq f(\bar{x}) \quad \text{and} \quad \operatorname{argmin}\{f + \delta \|(\cdot) - \tilde{x}\|\} = \{\tilde{x}\}.$$

2.2 Constraint qualifications

The purpose of this paragraph is to recall several well-established CQs for the general program (1) and to highlight some basic relations between them. We commence with the CQ that is most important to our study.

Definition 2.3 (MSCQ). *Let $\bar{x}$ be feasible for* (1). *We say that the* metric subregularity constraint qualification (MSCQ) *holds at $\bar{x}$ if there exists a neighborhood U of $\bar{x}$ and $\kappa > 0$ such that*

$$\mathrm{d}_{\mathcal{X}}(x) \leq \kappa \mathrm{d}_{\Gamma}(F(x)) \quad (x \in U).$$

Note that MSCQ is exactly metric subregularity in the set-valued sense of the *feasibility mapping* for the constraint system $\mathcal{X} = F^{-1}(\Gamma)$ which is given by $M(x) := F(x) - \Gamma$, see e.g.[26]. For the sake of brevity we do not introduce metric subregularity (and many other concepts) for general multifunctions, but instead restrict our definitions to the particular case of constraint systems. We point out that MSCQ is also known under the monikers *error bound property* or *calmness constraint qualification*.

As mentioned in Introduction, MSCQ plays a crucial role in optimization and variational analysis. However, MSCQ is hard to verify, which is one of the reasons why more attention has been given to the celebrated notion of *metric regularity*, see, e.g., the monographs [19, 41, 50, 51, 55]. Metric regularity is more restrictive than its subregular counterpart, but admits a compact characterization via generalized differentiation known as *Mordukhovich criterion*, see, e.g., [55, Theorem 9.40]. In case of constraint systems, given $\bar{x}$ feasible for (1), metric regularity of $M(x) := F(x) - \Gamma$ around $(\bar{x}, 0)$ holds if and only if there are neighborhoods U of $\bar{x}$ and V of 0 and $\kappa > 0$ such that

$$\mathrm{d}_{M^{-1}(y)}(x) \leq \kappa \mathrm{d}_{M(x)}(y) = \kappa \mathrm{d}_{\Gamma}(F(x) - y) \quad ((x, y) \in U \times V).$$

Since $\mathcal{X} = M^{-1}(0)$, one can easily see that metric subregularity corresponds to metric regularity with $y = 0$, rendering it milder than metric regularity. It is well known that metric regularity of a multifunction is equivalent to the *Aubin property* of the inverse multifunction, see e.g. [55, Theorem 9.43]. In addition, as a constraint qualification, metric regularity frequently appears in different forms and under different names such as *generalized Mangasarian-Fromovitz constraint qualification (GMFCQ)* or *No nonzero abnormal multiplier constraint qualification (NNAMCQ)*. In the rest of the paper, we will mainly stick to the GMFCQ terminology. We say that GMFCQ holds at $\bar{x}$ feasible for (1), if there is no nonzero multiplier $\bar{\lambda} \in N_\Gamma(F(\bar{x}))$ such that

$$\nabla F(\bar{x})^T \bar{\lambda} = 0. \tag{4}$$

We point out that GMFCQ at $\bar{x}$ for (1) is exactly the Mordukhovich criterion applied to the feasibility mapping M, hence characterizing its metric regularity. With a slight abuse of terminology we will henceforth sometimes refer to GMFCQ also as the Mordukhovich criterion and/or metric regularity. It is apparent that the Mordukhovich criterion has a very desirable feature to provide an efficient tool for verification of metric regularity. However, there are still plenty of situations when GMFCQ is not fulfilled but MSCQ is. It is therefore an important and worthwhile endeavor to fill the gap between GMFCQ and MSCQ, ideally with verifiable conditions at that. The next definition lists several such conditions. For these purposes consider the following set of constraint qualifications for (1).

Definition 2.4 (Constraint qualifications)**.** *Let $\bar{x} \in \mathcal{X}$ be feasible for* (1)*. We say that*

(i) pseudo-normality *holds at $\bar{x}$ if there exists no nonzero $\bar{\lambda} \in N_\Gamma(F(\bar{x}))$ such that* (4) *holds and that satisfies the following condition: There exists a sequence $\{(x^k, y^k, \lambda^k) \in \mathbb{R}^n \times \Gamma \times \mathbb{R}^d\} \to (\bar{x}, F(\bar{x}), \bar{\lambda})$ with*

$$\lambda^k \in \widehat{N}_\Gamma(y^k) \quad and \quad \langle \bar{\lambda}, F(x^k) - y^k \rangle > 0 \quad (k \in \mathbb{N});$$

(ii) quasi-normality *holds at $\bar{x}$ if there exists no nonzero $\bar{\lambda} \in N_\Gamma(F(\bar{x}))$ such that* (4) *holds and that satisfies the following condition: There exists a sequence $\{(x^k, y^k, \lambda^k) \in \mathbb{R}^n \times \Gamma \times \mathbb{R}^d\} \to (\bar{x}, F(\bar{x}), \bar{\lambda})$ with*

$$\lambda^k \in \widehat{N}_\Gamma(y^k) \quad and \quad \bar{\lambda}_i(F_i(x^k) - y_i^k) > 0 \quad if \quad \bar{\lambda}_i \neq 0 \quad (k \in \mathbb{N});$$

(iii) first-order sufficient condition for metric subregularity (FOSCMS) *holds at $\bar{x}$ if for every $0 \neq u \in \mathbb{R}^n$ with $\nabla F(\bar{x})u \in T_\Gamma(F(\bar{x}))$ one has*

$$\nabla F(\bar{x})^T \lambda = 0, \ \lambda \in N_\Gamma(F(\bar{x}); \nabla F(\bar{x})u) \implies \lambda = 0;$$

(iv) second-order sufficient condition for metric subregularity (SOSCMS) *holds at $\bar{x}$ if F is twice differentiable at $\bar{x}$, Γ is the union of finitely many convex polyhedra, and for every $0 \neq u \in \mathbb{R}^n$ with $\nabla F(\bar{x})u \in T_\Gamma(F(\bar{x}))$ one has*

$$\nabla F(\bar{x})^T \lambda = 0, \ \lambda \in N_\Gamma(F(\bar{x}); \nabla F(\bar{x})u), \ u^T \nabla^2 \langle \lambda, F \rangle (\bar{x})u \geq 0 \implies \lambda = 0.$$

We point out that asking the (nonexisting) multiplier $\bar{\lambda}$ to be in $N_\Gamma(F(\bar{x}))$ in the definition of pseudo-/quasi-normality is clearly redundant, since

$$\underset{y \xrightarrow{\Gamma} F(\bar{x})}{\operatorname{Lim\,sup}} \widehat{N}_\Gamma(y) = N_\Gamma(F(\bar{x})). \tag{5}$$

Nevertheless, in order to be consistent with the literature and to emphasize the connection to GMFCQ and other CQs, we stick to the original definition. In particular, it is obvious from the definition that GMFCQ implies both pseudo- and hence quasi-normality. The concepts of pseudo- and quasi-normality are well established in the literature. Note that in [30], the condition $\lambda^k \in \widehat{N}_\Gamma(y^k)$ in (i) and (ii) is replaced by $\lambda^k \in N_\Gamma(y^k)$. In order to see that no difference arises, consider the following elementary lemma which follows readily from the definitions of continuity and of the limiting normal cone, respectively.

Lemma 2.5. *Let $\Gamma \subset \mathbb{R}^d$ be closed, $y \in \Gamma, \lambda \in N_\Gamma(y)$ and let $a : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^q$ be continuous. Then for every $\varepsilon > 0$ there exist $\tilde{y} \in \Gamma$ and $\tilde{\lambda} \in \widehat{N}_\Gamma(\tilde{y})$ such that $\left\| a(\tilde{y}, \tilde{\lambda}) - a(y, \lambda) \right\| < \epsilon$.*

Corollary 2.6. *Under assumptions of Definition 2.4 let $\{(x^k, y^k, \lambda^k) \in \mathbb{R}^n \times \Gamma \times \mathbb{R}^d\} \to (\bar{x}, F(\bar{x}), \bar{\lambda})$. Then the following hold:*

(i) *If $\lambda^k \in N_\Gamma(y^k)$ and $\langle \bar{\lambda}, F(x^k) - y^k \rangle > 0$ for all $k \in \mathbb{N}$ then there exists $\{(\tilde{y}^k, \tilde{\lambda}^k)\} \to (F(\bar{x}), \bar{\lambda})$ such that $\tilde{\lambda}^k \in \widehat{N}_\Gamma(\tilde{y}^k)$ and $\langle \bar{\lambda}, F(x^k) - \tilde{y}^k \rangle > 0$ for all $k \in \mathbb{N}$.*

(ii) *If $\lambda^k \in N_\Gamma(y^k)$ and $\bar{\lambda}_i(F_i(x^k) - y_i^k) > 0$ $(i : \bar{\lambda}_i \neq 0)$ for all $k \in \mathbb{N}$ then there exists $\{(\tilde{y}^k, \tilde{\lambda}^k)\} \to (F(\bar{x}), \bar{\lambda})$ such that $\tilde{\lambda}^k \in \widehat{N}_\Gamma(\tilde{y}^k)$ and $\bar{\lambda}_i(F_i(x^k) - \tilde{y}_i^k) > 0$ $(i : \bar{\lambda}_i \neq 0)$ for all $k \in \mathbb{N}$.*

Proof. We only prove part (i); part (ii) can be shown analogously: To this end, define the continuous maps

$$a_k : (y, \lambda) \mapsto (y, \lambda, \langle \bar{\lambda}, F(x^k) - y \rangle) \quad (k \in \mathbb{N}),$$

and set $\epsilon_k := \min\{\frac{1}{k}, \frac{1}{2}\langle \bar{\lambda}, F(x^k) - y^k \rangle\}$. Applying Lemma 2.5 then generates the desired sequences. $\qquad\square$

Corollary 2.6 guarantees that using $\lambda^k \in \widehat{N}_\Gamma(y^k)$ instead of $\lambda^k \in N_\Gamma(y^k)$ in definition of pseudo- and quasi-normality does not play any role. We note that this is also true for the directional versions of these CQs to be established in Definition 3.5.

The obvious drawback of pseudo- and quasi-normality is that they are not point-based and hence it is not easy to check their validity and apply them. Another way of relaxing GMFCQ is provided by FOSCMS and SOSCMS, see Definition 2.4 (iii) and (iv). These conditions are using a directional approach based on techniques developed by Gfrerer and co-authors, see [23, 24, 25, 26, 27, 6]. The great advantage of FOSCMS and SOSCMS is that they are point-based and thus easier to verify.

In order to simplify the notation, given $\bar{x}$ feasible for (1), we define

$$\Lambda^0(\bar{x}; u) := \ker \nabla F(\bar{x})^T \cap N_\Gamma(F(\bar{x}); \nabla F(\bar{x})u) \quad (u \in \mathbb{R}^n) \tag{6}$$

and set

$$\Lambda^0(\bar{x}) := \Lambda^0(\bar{x}; 0) = \ker \nabla F(\bar{x})^T \cap N_\Gamma(F(\bar{x})),$$

i.e., the directional normal cone is replaced by the standard one. With these conventions, GMFCQ at $\bar{x}$ reads

$$\Lambda^0(\bar{x}) = \{0\},$$

while FOSCMS now reads

$$\Lambda^0(\bar{x}; u) = \{0\} \quad (u : \nabla F(\bar{x})u \in T_\Gamma(F(\bar{x}))).$$

The fact that GMFCQ implies FOSCMS is clear from the inclusion

$$N_\Gamma(F(\bar{x}); \nabla F(\bar{x})u) \subset N_\Gamma(F(\bar{x})) \quad (u \in \mathbb{R}^n).$$

The following example shows that this implication is indeed strict. In addition, it also illustrates that MSCQ is strictly weaker than quasi-normality, cf. Theorem 2.8(i).

Example 2.7. *Let $\Gamma := \{y \in \mathbb{R}^2 \mid y_2 \geq |y_1|\} \subset \mathbb{R}^2$, $F : \mathbb{R} \to \mathbb{R}^2$, $F(x) := (x, -x^2)^T$ and set $\bar{x} := 0$. Clearly $\nabla F(\bar{x}) = (1, 0)^T$ and $N_\Gamma(F(\bar{x})) = \{y \in \mathbb{R}^2 \mid y_2 \leq -|y_1|\}$, hence $0 \neq \lambda := (0, -1)^T \in \Lambda^0(\bar{x})$ and the Mordukhovich criterion (GMFCQ) is violated at $\bar{x}$.*

Moreover, setting $x_k := 1/k$, $y_k := F(\bar{x}) = (0, 0)^T$ and $\lambda_k := \lambda = (0, -1)^T$ we obtain $\lambda_2(F_2(x_k) - y_{k,2}) = -1(-1/(k^2)) > 0$, showing that also quasi-normality is violated at $\bar{x}$.

On the other hand, since $N_\Gamma(F(\bar{x}); \nabla F(\bar{x})u) = \emptyset$ for all $u \neq 0$, FOSCMS and hence MSCQ are satisfied at $\bar{x}$.

We point out that the set Γ in Example 2.7 is convex, thus illustrating that even in the convex case one may not be able to verify MSCQ using the non-directional conditions (GMFCQ, pseudo- and quasi-normality), but one may invoke a directional one (here FOSCMS).

Although the directional conditions FOSCMS and SOSCMS are similar in flavor, we point out that FOSCMS is only applicable in the case where Γ has *disjunctive* structure. In this setting, there is yet another condition due to Robinson [54] that ensures MSCQ.

The following proposition summarizes the most important sufficient conditions for MSCQ, other than GMFCQ, which have already been established in the literature and that are important to

our study. We point out, however, that the validity of the cited results will be a simple corollary of our substantially refined analysis in Section 3.

Proposition 2.8 (Sufficient conditions for MSCQ). *Let $\bar{x}$ be feasible for* (1). *Then under either of the following conditions MSCQ holds at $\bar{x}$.*

(i) [30, Theorem 5.2] *quasi-normality (or even pseudo-normality) holds at $\bar{x}$;*
(ii) [26, Corollary 1] *FOSCMS holds at $\bar{x}$;*
(iii) [26, Corollary 1] *SOSCMS holds at $\bar{x}$;*
(iv) [54, Proposition 1] *F is affine and Γ is the union of finitely many convex polyhedra.*

As we can see, apart from GMFCQ, there are currently four important conditions ensuring MSCQ. Two of them are applicable for the general program (1) and are strictly milder than GMFCQ. The other two are restricted to the special structure of disjunctive constraints and hence are in general not comparable with GMFCQ. Interestingly, all four conditions are mutually incomparable and were obtained by different approaches. The only available comparison is for the disjunctive constraints, where FOSCMS clearly implies SOSCMS.

In Section 3 we introduce a new constraint qualification that synthesizes the directional approach with quasi-normality. We show that this new condition implies MSCQ. In addition, we prove that it unifies not only FOSCMS and quasi-normality, but, quite surprisingly, also SOSCMS and condition (iv) in Theorem 2.8, to which we will refer as *Robinson's result*. Thus, we have found a new CQ that is milder than all four well-established sufficient conditions for MSCQ. In particular cases of disjunctive programs, see Section 4, and ortho-disjunctive programs, see Section 5, we even obtain new point-based conditions to verify our new CQ, which slightly improve SOSCMS.

2.3 Exact penalization under MSCQ

Let us briefly discuss the role of MSCQ in the context of exact penalization. Note that the general mathematical program (1) is equivalent to the unconstrained (but extended real-valued) problem

$$\min\ f(x) + \delta_\Gamma(F(x)). \tag{7}$$

A natural approximation for (7) (and hence (1)) is given by minimization of the following penalty function

$$P_\alpha := f + \alpha \mathrm{d}_\Gamma \circ F \quad (\alpha > 0), \tag{8}$$

which is a classical technique employed to tackle program (1), see, e.g., [14, 15, 30, 36, 40, 44, 48, 39].

The crucial issue is the exactness of the penalty function, which holds true under MSCQ as is stated in the following theorem. The proof essentially coincides with the proof of Theorem [44, Theorem 4.5], but we provide it for the sake of completeness and also to realize that the special structure underlying in [44] is not needed at all.

Note also that for general programs the following result was first established in [36, Proposition 3.5] and it was based on results by Burke [14, Theorem 1.1] and Clarke [17, Proposition 6.4.3].

Theorem 2.9. *Let $\bar{x}$ be a local minimizer of* (1) *such that MSCQ holds at $\bar{x}$. Then the penalty function P_α from* (8) *is exact at $\bar{x}$, i.e., $\bar{x}$ is a local minimizer of P_α for all $\alpha > 0$ sufficiently large. In particular, $\bar{x}$ is an M-stationary point of* (1).

Proof. By MSCQ at $\bar{x}$ there exist $\delta, \kappa > 0$ such that

$$\mathrm{d}_{\mathcal{X}}(x) \le \kappa \mathrm{d}_\Gamma(F(x)) \quad (x \in \mathbb{B}_\delta(\bar{x})).$$

As $\bar{x}$ is a local minimizer of f over $\mathcal{X}$, we can choose $0 < \varepsilon < \delta/2$ such that $\bar{x} \in \mathrm{argmin}_{x \in \mathbb{B}_{2\varepsilon}(\bar{x}) \cap \mathcal{X}} f(x)$. Since f is locally Lipschitz, by compactness, f is L-Lipschitz on $\mathbb{B}_{2\varepsilon}(\bar{x})$ for some $L > 0$. Now let $x \in \mathbb{B}_\varepsilon(\bar{x})$. In particular, we find $y \in P_{\mathcal{X}}(x) \subset \mathcal{X} \cap \mathbb{B}_{2\varepsilon}(\bar{x})$, hence it follows that

$$f(\bar{x}) \le f(y) \le f(x) + L\|y - x\| = f(x) + L\mathrm{d}_{\mathcal{X}}(x) \le f(x) + \kappa L \mathrm{d}_\Gamma(F(x)).$$

This shows that $\bar{x}$ is a local minimizer of $f + \kappa L \mathrm{d}_\Gamma \circ F$ and the exactness of P_α follows. Moreover, applying a nonsmooth Fermat's rule (cf. [55, Theorem 10.1]), invoking [55, Exercise 10.10] and Proposition 2.1 yields

$$0 \in \partial(f + \kappa L \mathrm{d}_\Gamma \circ F)(\bar{x}) = \nabla f(\bar{x}) + \nabla F(\bar{x})^T N_\Gamma(F(\bar{x})),$$

which gives the M-stationarity at $\bar{x}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3 New constraint qualifications for GMP

In this section we are primarily concerned with constraint qualifications for the general mathematical program (1). In particular, we establish directional counterparts of pseudo- and quasi-normality from Definition 2.4, and introduce a new CQ called PQ-normality that unifies pseudo- and quasi-normality. We then show that all new CQs imply MSCQ. In particular, by means of directional quasi-normality, we recover Proposition 2.8 statements (i) and (ii). In Section 3.2 and Section 3.3 we study these CQs under some additional structural assumptions in standard and directional form, respectively, and we establish various sufficient conditions for them. In particular, when applied to the disjunctive constraints in Section 4, these sufficient conditions recover Gfrerer's SOSCMS as well as Robinson's result about polyhedral multifunctions, see Proposition 2.8 statements (iii) and (iv), respectively.

3.1 Directional constraint qualifications and PQ-normality

FOSCMS can be viewed as a directional counterpart of GMFCQ, see Definition 2.4. This naturally raises the question whether one can define directional counterparts of pseudo- and quasi-normality, and such that they fit naturally in the existing tapestry of CQs. Our study shows that this is indeed possible. We start with the following observation, where we invoke definitions of $\Lambda^0(\bar{x}; u)$ and $\Lambda^0(\bar{x})$.

Lemma 3.1. *Let $\bar{x}$ be feasible for (1) such that MSCQ is violated at $\bar{x}$. Then there exist sequences $\{x^k \notin \mathcal{X}\} \to \bar{x}$ and $\{\xi^k \in \partial(\mathrm{d}_\Gamma \circ F)(x^k)\} \to 0$ as well as $u \in \mathbb{R}^n \setminus \{0\}$ with $\|u\| = 1$ such that*

$$\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to u, \quad \frac{y^k - F(\bar{x})}{\|x^k - \bar{x}\|} \to \nabla F(\bar{x})u \quad (y_k \in P_\Gamma(F(x^k))) \quad and \quad \nabla F(\bar{x})u \in T_\Gamma(F(\bar{x})). \quad (9)$$

Proof. Violation of MSCQ at $\bar{x}$ readily yields a sequence $\{\tilde{x}^k\} \to \bar{x}$ with $\mathrm{d}_\mathcal{X}(\tilde{x}^k) > k\mathrm{d}_\Gamma(F(\tilde{x}^k))$. We put $\varepsilon_k := \mathrm{d}_\Gamma(F(\tilde{x}^k))$ and find that $\tilde{x}^k$ is an ε_k-minimizer of $\mathrm{d}_\Gamma \circ F$ for all $k \in \mathbb{N}$. Hence by Ekeland's variational principle (Proposition 2.2) with $\delta = \frac{1}{k}$ ($k \in \mathbb{N}$), there exists a sequence $\{x^k\}$ such that $x^k = \operatorname{argmin}\{\mathrm{d}_\Gamma \circ F + \frac{1}{k}\|(\cdot) - x^k\|\}$ and $\|x^k - \tilde{x}^k\| \leq k\varepsilon_k < \mathrm{d}_\mathcal{X}(\tilde{x}^k)$ for all $k \in \mathbb{N}$. This implies $\{x^k \notin \mathcal{X}\} \to \bar{x}$ as well as $0 \in \partial(\mathrm{d}_\Gamma \circ F)(x^k) + \frac{1}{k}\mathbb{B}$ for all $k \in \mathbb{N}$ by applying a nonsmooth Fermat's rule (cf. [55, Theorem 10.1]) and invoking a sum rule for locally Lipschitz functions (cf. [55, Exercise 10.10]). In particular, there exists a sequence $\{\xi^k \in \partial(\mathrm{d}_\Gamma \circ F)(x^k)\} \to 0$. As $x^k \neq \bar{x}$, w.l.o.g. we may assume that $\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to u$ with $\|u\| = 1$. Now let $y^k \in P_\Gamma(F(x^k))$ for all $k \in \mathbb{N}$. Then

$$\left\| \frac{y^k - F(\bar{x})}{\|x^k - \bar{x}\|} - \nabla F(\bar{x})u \right\| \leq \frac{\|y^k - F(x^k)\|}{\|x^k - \bar{x}\|} + \left\| \frac{F(x^k) - F(\bar{x})}{\|x^k - \bar{x}\|} - \nabla F(\bar{x})u \right\| \quad (k \in \mathbb{N}). \quad (10)$$

As x^k minimizes $\mathrm{d}_\Gamma \circ F + \frac{1}{k}\|(\cdot) - x^k\|$ for all $k \in \mathbb{N}$, we find that $\mathrm{d}_\Gamma(F(x^k)) \leq 1/k\|\bar{x} - x^k\|$. Hence we infer that the first term on the right in (10) satisfies

$$\frac{\|y^k - F(x^k)\|}{\|x^k - \bar{x}\|} = \frac{\mathrm{d}_\Gamma(F(x^k))}{\|x^k - \bar{x}\|} \leq \frac{1}{k} \to 0.$$

The second term on the right in (10) goes to zero by differentiability of F and we conclude from (10) that $\frac{y^k - F(\bar{x})}{\|x^k - \bar{x}\|} \to \nabla F(\bar{x})u$. Finally, as $y^k \in \Gamma$ for all $k \in \mathbb{N}$, we have $\nabla F(\bar{x})u \in T_\Gamma(F(\bar{x}))$. $\quad\square$

Theorem 3.2. *Let $\bar{x}$ be feasible for (1) and assume that the following holds: For every $u \in \mathbb{R}^n$ with $\|u\| = 1$ and $\nabla F(\bar{x})u \in T_\Gamma(F(\bar{x}))$ there does not exist a nonzero $\bar{\lambda} \in \Lambda^0(\bar{x}; u)$ that satisfies*

the following condition: There exists a sequence $\{(x^k, y^k, \lambda^k) \in \mathbb{R}^n \times \Gamma \times \mathbb{R}^d\} \to (\bar{x}, F(\bar{x}), \bar{\lambda})$ *such that for all* $k \in \mathbb{N}$ *we have*

$$(x^k - \bar{x})/\|x^k - \bar{x}\| \to u, \qquad (y^k - F(\bar{x}))/\|x^k - \bar{x}\| \to \nabla F(\bar{x})u,$$

$$\lambda^k \in \widehat{N}_\Gamma(y^k), \qquad \bar{\lambda}_i(F_i(x^k) - y_i^k) > 0 \quad (\bar{\lambda}_i \neq 0).$$

Then MSCQ is fulfilled at $\bar{x}$.

Proof. Assume that MSCQ is not satisfied at $\bar{x}$. Consider sequences $\{x^k \notin \mathcal{X}\} \to \bar{x}$, $\{\xi^k \in \partial (\mathrm{d}_\Gamma \circ F)(x^k)\} \to 0$ and $u \in \mathbb{R}^n$ with $\|u\| = 1$ provided by Lemma 3.1. Recall that

$$\partial(\mathrm{d}_\Gamma \circ F)(x) \subset \nabla F(x)^T \partial \mathrm{d}_\Gamma(F(x)) \quad (x \in \mathbb{R}^n),$$

see Proposition 2.1 (ii). Moreover, by Proposition 2.1 (i), it holds that

$$\partial \mathrm{d}_\Gamma(F(x^k)) = \frac{F(x^k) - P_\Gamma(F(x^k))}{\mathrm{d}_\Gamma(F(x^k))} \quad (k \in \mathbb{N}),$$

since $x^k \notin \mathcal{X}$ $(k \in \mathbb{N})$. Consequently, there exists $\{y^k \in P_\Gamma(F(x^k))\}$ such that with

$$\lambda^k := \frac{F(x^k) - y^k}{\mathrm{d}_\Gamma(F(x^k))} \tag{11}$$

we have

$$\xi^k = \nabla F(x^k)^T \lambda^k \quad \text{and} \quad \|\lambda^k\| = 1 \quad (k \in \mathbb{N}). \tag{12}$$

Moreover, by the definition of λ^k in (11) and the fact that $y^k \in P_\Gamma(F(x^k))$ $(k \in \mathbb{N})$, [55, Example 6.16] implies that

$$\lambda^k \in \widehat{N}_\Gamma(y^k) \quad (k \in \mathbb{N}).$$

Since $\{\lambda^k\}$ is bounded, we may assume w.l.o.g. that $\lambda^k \to \bar{\lambda}$ for some $\bar{\lambda} \neq 0$. Then from (11) we infer that $y^k \to F(\bar{x})$. Hence, passing to the limit in (12) we obtain

$$0 = \nabla F(\bar{x})^T \bar{\lambda} \quad \text{and} \quad \bar{\lambda} \neq 0.$$

Now, if $\bar{\lambda}_i > 0$ then w.l.o.g. $F_i(x^k) - y_i^k = \mathrm{d}_\Gamma(F(x^k))\lambda_i^k > 0$ and hence $\bar{\lambda}_i(F_i(x^k) - y_i^k) > 0$. Analogously, we argue for $\bar{\lambda}_i < 0$. Altogether, we find that

$$\bar{\lambda}_i(F_i(x^k) - y_i^k) > 0 \quad \text{if} \quad \bar{\lambda}_i \neq 0 \quad (k \in \mathbb{N}).$$

Finally, Lemma 3.1 yields that $(y^k - F(\bar{x}))/\|x^k - \bar{x}\| \to \nabla F(\bar{x})u$, showing $\bar{\lambda} \in N_\Gamma(F(\bar{x}), \nabla F(\bar{x})u)$, which establishes a contradiction. □

Remark 3.3. *The proof of Theorem 3.2, via Lemma 3.1, is based on Ekeland's variational principle and the rich subdifferential calculus for the distance function, see Proposition 2.1. We would like to emphasize here, that this approach provides, as a by-product, a new proof of Proposition 2.8(i), i.e., the fact that (standard, nondirectional) quasi-normality implies MSCQ. To realize that, consider first the statement of Lemma 3.1 without the directional part* (9). *In addition, note that the proof of Theorem 3.2, with some minor modifications, readily yields the auxiliary result that quasi-normality at* $\bar{x} \in \mathcal{X}$ *implies the existence of* $\delta, c > 0$ *such that for all* $x \in B_\delta(\bar{x}) \setminus \mathcal{X}$ *and all* $\xi \in \partial (\mathrm{d}_\Gamma \circ F)(x)$ *we have* $\|\xi\| \geq \frac{1}{c}$. *Combining these two observation then establishes the fact that quasi-normality implies MSCQ.*

A similar technique of proof via the above mentioned auxiliary result was already used in [44, Lemma 4.4], where Γ *models the cartesian product of the complementarity manifold. However, the authors did not observe that this result holds for general closed sets* Γ *and they also did not exploit the rich subdifferential calculus for the distance function which is the workhorse in our proof.*

Theorem 3.2 will be our guiding principle for establishing new directional constraint qualifications. Instead of directly extracting directional versions of quasi- and pseudo-normality from Theorem 3.2, we introduce the notion of PQ-normality which serves as a bridge between directional pseudo- and quasi-normality, which turn out to be simply a special cases of PQ-normality. We strongly emphasize that introducing PQ-normality does not merely serve the academic purpose of unifying the two concepts. In fact, it has important consequences for the class of programs in Section 5 where the set Γ possesses an underlying product structure in addition to its disjunctive nature.

Prior to the definition of new CQs, we introduce additional notation. For $z \in \mathbb{R}^d$ we denote by z_i $(i \in I := \{1, \ldots, d\})$ its scalar components. More generally, suppose that $\mathbb{R}^d$ is expressed via factors as $\mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_l}$ and introduce the d multi-indices $\delta := (d_1, \ldots, d_l) \in \mathbb{N}^l$ with $|\delta| := d_1 + \ldots + d_l = d$. Note that there is one-to-one correspondence between such multi-indices and factorizations of $\mathbb{R}^d$. The components of some $z \in \mathbb{R}^d$ we denote as z_ν for $\nu \in I_\delta$, where I_δ is some (abstract) index set of l elements. Note that we do not identify I_δ with $\{1, \ldots, l\}$ in order to avoid ambiguity of notation, e.g., $z_1 \subset \mathbb{R}$ stands only for the first, scalar, component of z. Moreover, we use a Greek letter to indicate the vector components z_ν of z and a Latin letter to indicate the scalar components z_i.

Given a multi-index δ fix $\nu \in I_\delta$. The component z_ν, vector in general, can also be written via its scalar components, i.e., there exists an index set, denoted by I^ν, such that $z_\nu = (z_i)_{i \in I^\nu}$. Note that $\cup_{\nu \in I_\delta} I^\nu = I$. Finally, given two multi-indices δ, δ' with $|\delta| = |\delta'| = d$, we say that δ' is a *refinement* of δ and write $\delta' \subset \delta$, provided for every $\nu \in I_\delta$ there exists an index set $I^\nu_{\delta'}$ such that

$$z_\nu = (z_{\nu'})_{\nu' \in I^\nu_{\delta'}} \quad \text{and} \quad I_{\delta'} = \cup_{\nu \in I_\delta} I^\nu_{\delta'}.$$

Note that the special multi-indices $\delta^P := d \in \mathbb{N}^1$ and $\delta^Q := (1, \ldots, 1) \in \mathbb{N}^d$ are in fact maximal and minimal in the sense that for any multi-index $\delta \in \mathbb{N}^l$ with $|\delta| = d$ one has $\delta^Q \subset \delta \subset \delta^P$.

The following example illustrates the use of notation introduced above.

Example 3.4. *Let $n = 7$, $I := \{1, \ldots, 7\}$ and consider a multi-index $\delta := (1, 4, 2)$ corresponding to the factorization $\mathbb{R}^7 = \mathbb{R} \times \mathbb{R}^4 \times \mathbb{R}^2$. Consider also an element $z = (z_1, \ldots, z_7) \in \mathbb{R}^7$. Since $|\delta| = 3$, we may set, e.g., $I_\delta = \{a, b, c\}$ yielding $z_a = z_1$, $z_b = (z_2, z_3, z_4, z_5)$ and $z_c = (z_6, z_7)$. Clearly, we have $I^a = \{1\}$, $I^b = \{2, 3, 4, 5\}$ and $I^c = \{6, 7\}$.*

Moreover, the multi-index $\delta' := (1, (3, 1), (1, 1))$ is a refinement of δ, since we may set

$$I^a_{\delta'} := \{a\}, \ I^b_{\delta'} := \{b_1, b_2\} \ and \ I^c_{\delta'} := \{c_1, c_2\}$$

to obtain

$$z_a = z_1, \ z_{b_1} = (z_2, z_3, z_4), \ z_{b_2} = z_5, \ z_{c_1} = z_6, \ z_{c_2} = z_7,$$

$$and \quad I_{\delta'} = I^a_{\delta'} \cup I^b_{\delta'} \cup I^c_{\delta'} = \{a, b_1, b_2, c_1, c_2\}, \quad z_a = z_a, \ z_b = (z_{b_1}, z_{b_2}), \ z_c = (z_{c_1}, z_{c_2}).$$

We now proceed with the definition of PQ-normality which embeds quasi- and pseudo-normality as extremal cases in a whole family of constraint qualifications.

Definition 3.5 (PQ-normality). *Let $\bar{x} \in \mathcal{X}$ be feasible for (1), consider $u \in \mathbb{R}^n$ with $\|u\| = 1$, and let $\delta \in \mathbb{N}^l$ be a multi-index such that $|\delta| = d$. We say that*

(i) *PQ-normality w.r.t. δ holds at $\bar{x}$, if <u>there exists no nonzero</u> $\bar{\lambda} \in \Lambda^0(\bar{x})$ that satisfies the following condition: There exists a sequence $\{(x^k, y^k, \lambda^k) \in \mathbb{R}^n \times \Gamma \times \mathbb{R}^d\} \to (\bar{x}, F(\bar{x}), \bar{\lambda})$ with $\lambda^k \in \widehat{N}_\Gamma(y^k)$ and*

$$\langle \bar{\lambda}_\nu, F_\nu(x^k) - y^k_\nu \rangle > 0 \ \ for \ \ \nu \in I_\delta(\bar{\lambda}) := \{\nu \in I_\delta \,|\, \bar{\lambda}_\nu \neq 0\} \quad (k \in \mathbb{N}). \tag{13}$$

(ii) *PQ-normality w.r.t. δ in direction u holds at $\bar{x}$, if <u>there exists no nonzero</u> $\bar{\lambda} \in \Lambda^0(\bar{x}; u)$ that satisfies the following condition: There exists a sequence $\{(x^k, y^k, \lambda^k) \in \mathbb{R}^n \times \Gamma \times \mathbb{R}^d\} \to (\bar{x}, F(\bar{x}), \bar{\lambda})$ with $\lambda^k \in \widehat{N}_\Gamma(y^k)$, (13) and*

$$(x^k - \bar{x})/\|x^k - \bar{x}\| \to u, \quad (y^k - F(\bar{x}))/\|x^k - \bar{x}\| \to \nabla F(\bar{x})u. \tag{14}$$

We say that directional PQ-normality w.r.t. δ holds at $\bar{x}$, if PQ-normality w.r.t. δ in direction u holds at $\bar{x}$ for all $u \in \mathbb{R}^n$ with $\|u\| = 1$. In particular, we refer to PQ-normality w.r.t. δ^P (in direction u) as pseudo-normality (in direction u), while PQ-normality w.r.t. δ^Q we call quasi-normality.

As advertised above, PQ-normality w.r.t. δ contains pseudo- and quasi-normality as extreme cases by setting $\delta := \delta^P$ and $\delta := \delta^Q$, respectively. It is clear from the definition that PQ-normality w.r.t. δ implies PQ-normality w.r.t. δ' provided $\delta' \subset \delta$. In particular, since $\delta^Q \subset \delta \subset \delta^P$ for all $\delta \in \mathbb{N}^l$ with $|\delta| = d$, we conclude that pseudo-normality implies PQ-normality w.r.t. any δ and this further implies quasi-normality. Naturally, all of the above comments remain true for the corresponding directional CQs.

For the sake of completeness, we reformulate Theorem 3.2 in terms of directional PQ-normality.

Theorem 3.6. *Let $\bar{x}$ be feasible for* (1) *and let the directional PQ-normality w.r.t. any $\delta \in \mathbb{N}^l$, in particular directional pseudo- or quasi-normality, hold at $\bar{x}$. Then MSCQ is fulfilled at $\bar{x}$. In particular, if $\bar{x}$ is also a local minimizer of* (1), *the penalty function P_α from* (8) *is exact at $\bar{x}$ and $\bar{x}$ is M-stationary for* (1).

Proof. The statement follows from Theorems 3.2 and 2.9.   □

We point out that directional quasi-normality is strictly weaker than both FOSCMS (clear from the definition of the respective CQs) as well as quasi-normality, see Example 2.7. Hence it constitutes, to the best of our knowledge, one of the weakest conditions to imply MSCQ for the general optimization problem (1), but which can be efficiently verified in some very important cases as shown in Section 3.2 and Section 3.3 below. Note that as a by-product we thus established a new and rather simple proof of Proposition 2.8 statement (ii).

3.2 Simplified CQs and second-order sufficient conditions: The standard case

For some important instances of the general program (1), the concepts of pseudo- and quasi-normality were introduced without the undesirable additional sequence $\{y^k\}$, see [9] for standard NLPs and [44] for MPCCs. In this section and Section 3.3, we address the question as to when this is possible for much more general instances of (1) and for the whole PQ-normality family (thus containing pseudo- and quasi-normality as special cases).

For clarity of exposition, we split our analysis into the standard (non-directional) and the directional case.

We begin our study of the non-directional case by the following straightforward result, which follows readily from Lemma 2.5 using similar arguments as in the proof of Corollary 2.6.

Lemma 3.7. *Let $\bar{x}$ be feasible for* (1). *If PQ-normality w.r.t. $\delta \in \mathbb{N}^l$ holds at $\bar{x}$ then there exists no nonzero $\bar{\lambda} \in \Lambda^0(\bar{x})$ that satisfies the following condition: There exists a sequence $\{x^k\} \to \bar{x}$ with*

$$\left\langle \bar{\lambda}_\nu, F_\nu(x^k) - F_\nu(\bar{x}) \right\rangle > 0 \ \ for \ \ \nu \in I_\delta(\bar{\lambda}) \quad (k \in \mathbb{N}). \tag{15}$$

Note that in case of MPCCs, by the geometry of the feasible set and the resulting normal cones, one always has $\left\langle \bar{\lambda}, F(\bar{x}) \right\rangle = 0$. Thus the conditions used in [44] simplify to $\left\langle \bar{\lambda}, F(x^k) \right\rangle > 0$ and $\bar{\lambda}_i F_i(x^k) > 0$ if $\bar{\lambda}_i \neq 0$, respectively. However, in the general setting of problem (1), as well as in the case of general disjunctive constraints, we cannot make this simplification. In order to obtain the reverse implication, however, we have to impose some additional assumptions on the constraints of (1).

Assumption 3.8. *Let $\delta \in \mathbb{N}^l$ be a multi-index and let $\bar{x}$ be feasible for* (1). *Assume that for every $\bar{\lambda} \in \Lambda^0(\bar{x})$ and every sequence $\{(y^k, \lambda^k) \in \Gamma \times \mathbb{R}^d\} \to (F(\bar{x}), \bar{\lambda})$ with $\lambda^k \in \widehat{N}_\Gamma(y^k)$, there exists a subsequence $K \subset \mathbb{N}$ such that*

$$\left\langle \bar{\lambda}_\nu, y_\nu^k - F_\nu(\bar{x}) \right\rangle \geq 0 \quad (\nu \in I_\delta, \ k \in K). \tag{16}$$

Theorem 3.9 (Simplified PQ-normality under Ass. 3.8)**.** *Let $\bar{x}$ be feasible for* (1) *and $\delta \in \mathbb{N}^l$ such that Assumption 3.8 holds. Then PQ-normality w.r.t. $\delta \in \mathbb{N}^l$ at $\bar{x}$ is equivalent to the following simplified PQ-normality w.r.t. $\delta \in \mathbb{N}^l$ at $\bar{x}$, i.e.:*

There exists no nonzero $\bar{\lambda} \in \Lambda^0(\bar{x})$ such that there exists a sequence $\{x^k\} \to \bar{x}$ fulfilling (15).

Proof. The fact that PQ-normality implies the simplified PQ-normality follows from Lemma 3.7.

In turn, if PQ-normality w.r.t. δ is violated, there exist $\bar{\lambda} \in \Lambda^0(\bar{x}) \setminus \{0\}$ and $\{(x^k, y^k, \lambda^k) \in \mathbb{R}^n \times \Gamma \times \mathbb{R}^d\} \to (\bar{x}, F(\bar{x}), \bar{\lambda})$ with $\lambda^k \in \widehat{N}_\Gamma(y^k)$ and $\left\langle \bar{\lambda}_\nu, F_\nu(x^k) - y_\nu^k \right\rangle > 0$ for some $\nu \in I_\delta(\bar{\lambda})$. Relabeling $\{x^k\}$ by only using the indices $k \in K$ and then summing up the above expression with (16) for all $k \in K$ shows that the simplified PQ-normality is then violated as well.   □

As the above theorem shows, under Assumption 3.8, the simplified PQ-normality (without the sequence $\{y^k\}$) is equivalent to PQ-normality, hence sufficient for MSCQ. Without Assumption 3.8

this is, in general, false, see Example 3.13. In the following sections, however, we establish various classes of optimization problems which automatically satisfy Assumption 3.8 for any multi-index δ, including δ^P and δ^Q, at every feasible point.

As we will now show, Theorem 3.9 also reveals a striking connection between PQ-normality and vector optimization. This, in turn, paves the way to a variety of sufficient conditions for PQ-normality, hence also for MSCQ.

Let us recall some standard terminology. Given $\varphi : \mathbb{R}^n \to \mathbb{R}^q$, a point $\bar{x}$ is called a *local weak efficient solution* of the unconstrained vector optimization problem $\max_{x \in \mathbb{R}^n} \varphi(x)$ if there exists a neighborhood U of $\bar{x}$ such that no $x \in U$ satisfies $\varphi_j(x) > \varphi_j(\bar{x})$ for all $j = 1, \ldots, q$. Given $\delta = (d_1, \ldots, d_l) \in \mathbb{N}^l$ and $\lambda = (\lambda_\nu)_{\nu \in I_\delta} \in \mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_l} = \mathbb{R}^d$, we define the function

$$\varphi^\lambda : \mathbb{R}^n \to \mathbb{R}^{|I_\delta(\lambda)|}, \quad \varphi^\lambda(x) := (\langle \lambda_\nu, F_\nu \rangle (x))_{\nu \in I_\delta(\lambda)}. \tag{17}$$

The next result then follows directly from the definitions of local weak efficient solutions and simplified PQ-normality established in Theorem 3.9.

Theorem 3.10. *Let $\bar{x}$ be feasible for (1) and let Assumption 3.8 for some $\delta \in \mathbb{N}^l$ be fulfilled. Then PQ-normality w.r.t. δ holds at $\bar{x}$ if and only if for every $\bar{\lambda} \in \Lambda^0(\bar{x})$, the vector $\bar{x}$ is a local weak efficient solution of the unconstrained vector optimization problem $\max_{x \in \mathbb{R}^n} \varphi^{\bar{\lambda}}(x)$ for $\varphi^{\bar{\lambda}}$ given by (17).*

Proof. If there exists $\bar{\lambda} \in \Lambda^0(\bar{x})$ such that $\bar{x}$ is not a local weak efficient solution of $\max_{x \in \mathbb{R}^n} \varphi^{\bar{\lambda}}(x)$, then $\bar{\lambda} \neq 0$ and there exists $x^k \to \bar{x}$ together with some $\nu \in I_\delta(\bar{\lambda})$ such that $\langle \bar{\lambda}_\nu, F_\nu(x^k) \rangle > \langle \bar{\lambda}_\nu, F_\nu(\bar{x}) \rangle$ for all $k \in \mathbb{N}$. This shows that PQ-normality w.r.t. δ is violated due to Theorem 3.9.

In turn, if pseudo-normality is violated, there exists $\bar{\lambda} \in \Lambda^0(\bar{x}) \setminus \{0\}$ and a sequence $x^k \to \bar{x}$ such that $\langle \bar{\lambda}_\nu, F_\nu(x^k) - F_\nu(\bar{x}) \rangle > 0$ for all $\nu \in I_\delta(\bar{\lambda})$ and all $k \in \mathbb{N}$, which shows that $\bar{x}$ is not a local weak efficient solution of $\max_{x \in \mathbb{R}^n} \varphi^{\bar{\lambda}}(x)$. $\qquad\square$

This simple observation obviously has several strong consequences. In particular, it allows one to use the standard sufficient conditions for a local weak efficient solution to obtain the following point-based second-order sufficient condition for PQ-normality w.r.t. δ (SOSCPQN(δ)).

Corollary 3.11. *Let $\bar{x}$ be feasible for (1) with F twice differentiable at $\bar{x}$ and let Assumption 3.8 for some $\delta \in \mathbb{N}^l$ be fulfilled. Then PQ-normality w.r.t. δ, in particular MSCQ, holds at $\bar{x}$ if the following SOSCPQN(δ) is fulfilled: For every $\bar{\lambda} \in \Lambda^0(\bar{x}) \setminus \{0\}$, every $u \in \mathbb{R}^n \setminus \{0\}$ with $\langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x}) u \rangle = 0$ for all $\nu \in I_\delta(\bar{\lambda})$ and every w with $\langle w, u \rangle = 0$ one has*

$$\min_{\nu \in I_\delta(\bar{\lambda})} \left(\langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x}) w \rangle + u^T \nabla^2 \langle \bar{\lambda}_\nu, F_\nu \rangle (\bar{x}) u \right) < 0. \tag{18}$$

Proof. Consider $\bar{\lambda} \in \Lambda^0(\bar{x}) \setminus \{0\}$ and $\varphi^{\bar{\lambda}}$ given by (17) and let $z \in \mathbb{R}^n$ be arbitrary. Then

$$\sum_{\nu \in I_\delta(\bar{\lambda})} \nabla \varphi_\nu^{\bar{\lambda}}(\bar{x}) z = \sum_{\nu \in I_\delta(\bar{\lambda})} \langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x}) z \rangle = \langle \bar{\lambda}, \nabla F(\bar{x}) z \rangle = 0, \tag{19}$$

since $\bar{\lambda} \in \Lambda^0(\bar{x})$. Hence, every u with $\nabla \varphi_\nu^{\bar{\lambda}}(\bar{x}) u \leq 0$ for all $\nu \in I_\delta(\bar{\lambda})$ in fact fulfills $\nabla \varphi_\nu^{\bar{\lambda}}(\bar{x}) u = \langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x}) u \rangle = 0$ for all $\nu \in I_\delta(\bar{\lambda})$. The proof thus follows from [10, Theorem 4] and Theorem 3.10. $\qquad\square$

Remark 3.12. *Note that $\bar{\lambda} \in \Lambda^0(\bar{x})$ implies the validity of first-order necessary conditions for local efficient solution, $\min_{\nu \in I_\delta(\bar{\lambda})} \langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x}) w \rangle \leq 0$ for all $w \in \mathbb{R}^n$, as can be seen from (19).*

We refer to SOSCPQN(δ^P) and SOSCPQN(δ^Q) as second-order sufficient condition for pseudo-/quasi-normality (SOSCPN and SOSCQN).

Naturally, one can also consider higher-order sufficient conditions. We indeed do so in Section 4, where we focus on pseudo-normality. Note that pseudo-normality is connected to standard maximality since φ^λ is a scalar function in that case.

The following example shows that SOSCPN on its own, in particular without Assumption 3.8 for δ^P, does not guarantee pseudo-normality and not even MSCQ.

Example 3.13. *Consider* $\Gamma \subset \mathbb{R}^2$ *given by* $\Gamma := \{y \in \mathbb{R}^2 \,|\, y_2 \geq |y_1|^{3/2}\}$ *and* $F : \mathbb{R} \to \mathbb{R}^2$ *defined by* $F(x) := (x, x^2)^T$ *and let* $\bar{x} := 0$. *Clearly* $\nabla F(\bar{x}) = (1, 0)^T$ *and* $\Lambda^0(\bar{x}) = \mathbb{R}_+(0, -1)^T$. *Moreover,* $\nabla^2 F(\bar{x}) = (0, 2)$ *and thus for every* $\lambda \in \Lambda^0(\bar{x}) \setminus \{0\}$ *and every* $u \in \mathbb{R} \setminus \{0\}$ *we have* $u^T \nabla^2 \langle \lambda, F \rangle (\bar{x}) u = -2\alpha u^2 < 0$, *showing that SOSCPN holds at* $\bar{x}$. *On the other hand, for a sequence* $x_k \to 0$ *we obtain* $\mathrm{d}_{F^{-1}(\Gamma)}(x_k) = |x_k|$, *while*

$$\mathrm{d}_\Gamma(F(x_k)) \leq \left\| (x_k, x_k^2) - (x_k, |x_k|^{3/2}) \right\| \leq |x_k|^{3/2},$$

showing the violation of MSCQ and consequently the violation pseudo-normality as well.

We point out that the set Γ in Example 3.13 equals epi $|\cdot|^{3/2}$ and is therefore convex, yet SOSCPN still does not imply MSCQ.

Theorem 3.14. *Let* $\bar{x}$ *be feasible for* (1) *with* F *twice differentiable at* $\bar{x}$. *Given two multi-indices* $\delta \in \mathbb{N}^l, \delta' \in \mathbb{N}^{l'}$ *with* $\delta' \subset \delta$, *assume that Assumption 3.8 for* δ' *is fulfilled. Then SOSCPQN(δ) implies SOSCPQN(δ'). In particular, we have SOSCPN $\Rightarrow$ SOSCPQN(δ) $\Rightarrow$ SOSCQN, provided Assumption 3.8 is fulfilled for* δ^Q.

Proof. Assuming that SOSCPQN(δ) holds, given $0 \neq \bar{\lambda} \in \Lambda^0(\bar{x})$, $0 \neq u \in \mathbb{R}^n$ and arbitrary $w \in \mathbb{R}^n$, consider $\nu \in I_\delta(\bar{\lambda})$ with

$$\langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x})w \rangle + u^T \nabla^2 \langle \bar{\lambda}_\nu, F_\nu \rangle (\bar{x})u < 0.$$

Since $\delta' \subset \delta$, it yields the existence of index set $I_{\delta'}^\nu \subset I_{\delta'}$ such that $z_\nu = (z_{\nu'})_{\nu' \in I_{\delta'}^\nu}$. We obtain

$$\sum_{\nu' \in I_{\delta'}^\nu} \left(\langle \bar{\lambda}_{\nu'}, \nabla F_{\nu'}(\bar{x})w \rangle + u^T \nabla^2 \langle \bar{\lambda}_{\nu'}, F_{\nu'} \rangle (\bar{x})u \right) = \langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x})w \rangle + u^T \nabla^2 \langle \bar{\lambda}_\nu, F_\nu \rangle (\bar{x})u < 0.$$

This yields, however, that SOSCPQN(δ') is fulfilled.

The second statement now follows from the obvious relation $\delta^Q \subset \delta \subset \delta^P$ valid for any δ. $\qquad\square$

The following example shows that SOSCQN is, in fact, strictly milder than SOSCPN. Since both SOSCQN as well as SOSCPN are special cases of SOSCQPN in turns out that, in general, SOSCQN is strictly milder than SOSCQPN, and this is further strictly milder than SOSCPN. Moreover, the example demonstrates that one can effectively verify MSCQ by means of SOSCQN even when pseudo-normality is not fulfilled.

Example 3.15. *Let* $\Gamma := \Gamma_1 \times \Gamma_2 \subset \mathbb{R}^2$ *for two convex polyhedral sets* $\Gamma_1 = \Gamma_2 := \mathbb{R}_-$ *and let* $F := (F_1, F_2)^T : \mathbb{R} \to \mathbb{R}^2$ *for* $F_1(x) := -x$ *and* $F_2(x) := x + x^2$ *and let* $\bar{x} := 0$. *In particular, Assumption 3.8 for* δ^Q *is fulfilled by Corollary 5.4. Clearly,* $\nabla F_1(\bar{x}) = -1, \nabla F_2(\bar{x}) = 1$ *and hence* $\Lambda^0(\bar{x}) = \mathbb{R}_+(1, 1)^T$.

SOSCQN is fulfilled since for any $\lambda = (\lambda_1, \lambda_2) = \alpha(1, 1)^T$ *for some* $\alpha > 0$ *and for* $u = \pm 1$ *one has* $|\lambda_i \nabla F_i(\bar{x})u| = \alpha \neq 0$, $i = 1, 2$. *In particular, quasi-normality and MSCQ follows.*

On the other hand, let $\bar{\lambda} := (1, 1)^T$ *and consider a sequence* $x_k \downarrow 0$. *We obtain*

$$\langle \bar{\lambda}, F(x_k) - F(\bar{x}) \rangle = -x_k + x_k + x_k^2 > 0,$$

showing the violation of pseudo-normality.

The next example shows that in general, i.e., without Assumption 3.8 for δ^Q, the simplified form of quasi-normality from Lemma 3.7 does not imply metric subregularity even if we consider a convex polyhedral set Γ.

Example 3.16. *Let* $\Gamma \subset \mathbb{R}^2$ *be convex polyhedral set given by* $\Gamma := \{y \in \mathbb{R}^2 \,|\, y_2 \geq y_1\}$ *and* $F : \mathbb{R} \to \mathbb{R}^2$ *given by* $F(x) := (x, \sin x)^T$ *and let* $\bar{x} := 0$. *Clearly* $\nabla F(\bar{x}) = (1, 1)^T$ *and we find that* $\Lambda^0(\bar{x}) = \mathbb{R}_+(1, -1)^T$. *For every* $\lambda = (\lambda_1, \lambda_2) = \alpha(1, -1)^T$ *for some* $\alpha > 0$ *and every* $x \in \mathbb{R}$ *close to* $\bar{x}$ *we have* $\lambda_1(F_1(x) - F_1(\bar{x})) = \alpha x < 0$ *if* $x < 0$ *and* $\lambda_2(F_2(x) - F_1(\bar{x})) = -\alpha \sin x \leq 0$ *if* $x \geq 0$, *showing that the simplified form of quasi-normality holds at* $\bar{x}$. *On the other hand, for a sequence* $x_k \downarrow 0$ *we obtain* $\mathrm{d}_{F^{-1}(\Gamma)}(x_k) = |x_k|$, *while*

$$\mathrm{d}_\Gamma(F(x_k)) \leq \|(x_k, \sin x_k) - (x_k, x_k)\| = o(|x_k|),$$

showing the violation of MSCQ.

3.3 Simplified CQs and second-order sufficient conditions: The directional case

In this subsection, we consider the directional case, where the situation is slightly different.

Theorem 3.17. *Let $\bar{x}$ be feasible for* (1) *and consider $u \in \mathbb{R}^n$ with $\|u\| = 1$. Then under Assumption 3.8 for $\delta \in \mathbb{N}^l$, PQ-normality w.r.t. δ at $\bar{x}$ in direction u follows provided: there exists no nonzero $\bar{\lambda} \in \Lambda^0(\bar{x}; u)$ such that there exists a sequence $x^k \to \bar{x}$ with $(x^k - \bar{x})/\|x^k - \bar{x}\| \to u$ fulfilling* (15).

Proof. The proof follows by the same arguments as used in the proof of Theorem 3.9. □

In contrast to the standard case, the following example shows that the reverse implication in the above theorem is not true in general.

Example 3.18. *Consider $\Gamma \subset \mathbb{R}^2$ given by $\Gamma := \{y \in \mathbb{R}^2 \,|\, y_2 \le y_1^2\}$ and $F : \mathbb{R} \to \mathbb{R}^2$ defined by $F(x) := (x, x^4)^T$ and let $\bar{x} := 0$ and $u := 1$. Clearly $\nabla F(\bar{x}) = (1, 0)^T$ and $\Lambda^0(\bar{x}; 1) = \Lambda^0(\bar{x}) = \mathbb{R}_+(0, 1)^T$. Set $\bar{\lambda} := (0, 1)^T$ and note that any sequence $x_k \downarrow 0$ fulfills $(x_k - \bar{x})/\|x_k - \bar{x}\| \to u$ as well as* (15) *for δ^P, since $\langle \bar{\lambda}, F(x_k) - F(\bar{x}) \rangle = x_k^4 > 0$.*

On the other hand, for arbitrary sequence $y_k = (y_{k,1}, y_{k,2})^T \to F(\bar{x}) = (0, 0)^T$ with $N_\Gamma(y_k) \ne \emptyset$ we have $y_k = (y_{k,1}, y_{k,1}^2)^T$. Hence, for any $\lambda \in \mathbb{R}_+(0, 1)^T$ one has $\langle \lambda, y_k - F(\bar{x}) \rangle = \lambda_2 y_{k,1}^2 \ge 0$, showing that Assumption 3.8 for δ^P is fulfilled. Moreover $(y_{k,1}/x_k, y_{k,1}^2/x_k)^T = (y_k - F(\bar{x}))/\|x_k - \bar{x}\| \to \nabla F(\bar{x})u = (1, 0)^T$ yields $y_{k,1}/x_k \to 1$. Then, however, we obtain

$$\langle \lambda, F(x_k) - y_k \rangle = \lambda_2(x_k^4 - y_{k,1}^2) = \lambda_2 x_k^2(x_k^2 - y_{k,1}^2/x_k^2) \le 0,$$

showing that pseudo-normality at $\bar{x}$ in direction u is fulfilled.

Nevertheless, the previous theorem still allows us to use sufficient conditions. Consider the following second-order sufficient condition for directional PQ-normality w.r.t. δ (SOSCdirPQN(δ)).

Proposition 3.19. *Let $\bar{x}$ be feasible for* (1) *with F twice differentiable at $\bar{x}$ and let Assumption 3.8 for some $\delta \in \mathbb{N}^l$ be fulfilled. Then directional PQ-normality w.r.t. δ, in particular MSCQ, holds at $\bar{x}$ if the following SOSCdirPQN(δ) is fulfilled: For every $u \in \mathbb{R}^n$ with $\|u\| = 1$, every $0 \ne \bar{\lambda} \in \Lambda^0(\bar{x}; u)$ with $\langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x})u \rangle = 0$, for all $\nu \in I_\delta(\bar{\lambda})$ and every w with $\langle w, u \rangle = 0$ condition* (18) *is fulfilled.*

Proof. Assume that directional PQ-normality w.r.t. δ is violated. Theorem 3.17 yields the existence of $u \in \mathbb{R}^n$, $0 \ne \bar{\lambda} \in \Lambda^0(\bar{x}; u)$ and a sequence $x^k \to \bar{x}$ with $(x^k - \bar{x})/\|x^k - \bar{x}\| \to u$ such that $\varphi_\nu^{\bar{\lambda}}(x_k) - \varphi_\nu^{\bar{\lambda}}(\bar{x}) > 0$ for all $\nu \in I_\delta(\bar{\lambda})$ with $\varphi^{\bar{\lambda}}$ as in (17). Hence, by passing to a subsequence if necessary, we can assume that $(\varphi(x_k) - \varphi(\bar{x}))/\|\varphi(x_k) - \varphi(\bar{x})\| \to p$ with $p \ge 0$ and $\|p\| = 1$, where for simplification we dropped the upper index $\bar{\lambda}$ from φ.

By Taylor expansion, we have

$$\frac{\|\varphi(x_k) - \varphi(\bar{x})\|}{\|x_k - \bar{x}\|^2} \frac{\varphi(x_k) - \varphi(\bar{x})}{\|\varphi(x_k) - \varphi(\bar{x})\|} = \nabla\varphi(\bar{x})\frac{(x_k - \bar{x})}{\|x_k - \bar{x}\|^2} + \frac{(x_k - \bar{x})^T}{\|x_k - \bar{x}\|}\nabla^2\varphi(\bar{x})\frac{(x_k - \bar{x})}{\|x_k - \bar{x}\|} + o(1). \quad (20)$$

If there exists a subsequence K such that $\|\varphi(x_k) - \varphi(\bar{x})\| / \|x_k - \bar{x}\|^2 \to \infty$, we conclude from (20) that

$$\frac{\varphi(x_k) - \varphi(\bar{x})}{\|\varphi(x_k) - \varphi(\bar{x})\|} = \nabla\varphi(\bar{x})\frac{(x_k - \bar{x})}{\|\varphi(x_k) - \varphi(\bar{x})\|} + q_k,$$

where $q_k \to 0$ for $k \in K$. Passing to a subsequence if necessary, and taking into account that $\nabla\varphi(\bar{x})(x_k - \bar{x})/\|\varphi(x_k) - \varphi(\bar{x})\| \in \mathrm{Im}(\nabla\varphi(\bar{x}))$ with $\mathrm{Im}(\nabla\varphi(\bar{x}))$ being a closed set, we conclude that $p \in \mathrm{Im}(\nabla\varphi(\bar{x}))$, i.e., there exists $z \in \mathbb{R}^n$ with $\nabla\varphi(\bar{x})z = (\langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x})z \rangle)_{\nu \in I_\delta(\bar{\lambda})} = p$. This is, however, a contradiction with $\|p\| = 1$, since by $p \ge 0$ and (19) we obtain that $p = 0$.

Consequently, $\|\varphi(x_k) - \varphi(\bar{x})\| / \|x_k - \bar{x}\|^2$ remains bounded and by passing to a subsequence K if necessary we assume that $\|\varphi(x_k) - \varphi(\bar{x})\| / \|x_k - \bar{x}\|^2 \to \alpha \ge 0$. By similar arguments as before, (20) now yields the existence of w such that

$$\alpha p = \nabla\varphi(\bar{x})w + u^T\nabla^2\varphi(\bar{x})u.$$

Moreover, we can clearly take w with $\langle w, u \rangle = 0$ since $\mathbb{R}^n$ is the direct sum of the span of u and its orthogonal complement, and $\nabla \varphi(\bar{x})u = 0$ by (20). The assumed SSOSCdirPQN(δ) (18) now yields the existence of $\nu \in I_\delta(\bar{\lambda})$ with $p_\nu < 0$, a contradiction. This completes the proof.   □

In the definition of SOSCdirPQN(δ) we explicitly assume $\langle \bar{\lambda}_\nu, \nabla F_\nu(\bar{x})u \rangle = 0$ for all $\nu \in I_\delta(\bar{\lambda})$ in order to make it clear that SOSCdirPQN(δ) is indeed milder than SOSCPQN(δ). If fact we can omit it from the assumption since it actually follows from $\bar{\lambda} \in \Lambda^0(\bar{x}; u)$.

Naturally, the higher-order approach can be utilized here as well. As before, we will refer to SOSCdirPQN(δ^P) and SOSCdirPQN(δ^Q) as second-order sufficient condition for directional pseudo/quasi-normality (SOSCdirPN and SOSCdirQN).

The following directional counterpart of Theorem 3.14 follows by the same arguments.

Theorem 3.20. *Let $\bar{x}$ be feasible for (1) with F twice differentiable at $\bar{x}$. Given two multiindexes $\delta \in \mathbb{N}^l, \delta' \in \mathbb{N}^{l'}$ with $\delta' \subset \delta$, assume that Assumption 3.8 for δ' is fulfilled. Then SOSCdirPQN(δ) implies SOSCdirPQN(δ'). In particular, we have SOSCMS $\Rightarrow$ SOSCdirPQN(δ) $\Rightarrow$ SOSCdirQN, provided Assumption 3.8 is fulfilled for δ^Q.*

We point out here that, unlike in the non-directional case, we could not find an example to show that the above implications can be indeed strict.

3.4 Summary

We now summarize our findings of this section: We introduced several new constraint qualifications for the general program (1) by considering directional versions of the well-established pseudo- and quasi-normality, respectively. In addition, we introduced the new concept of PQ-normality, together with its directional counterpart, that unifies the two standard CQs. In our study we obtained novel, improved results for the metric subregularity constraint qualification (MSCQ) and we established intriguing connections among the well-known CQs and the new ones.

In the following diagram, we summarize the relations between the various constraint qualifications weaker than metric regularity (GMFCQ) that imply metric subregularity. The point-based conditions are naturally of primary interest and are hence emphasized in double-framed boxes. Note that pseudo- and quasi-normality are included as special cases of PQ-normality for δ^P and δ^Q.
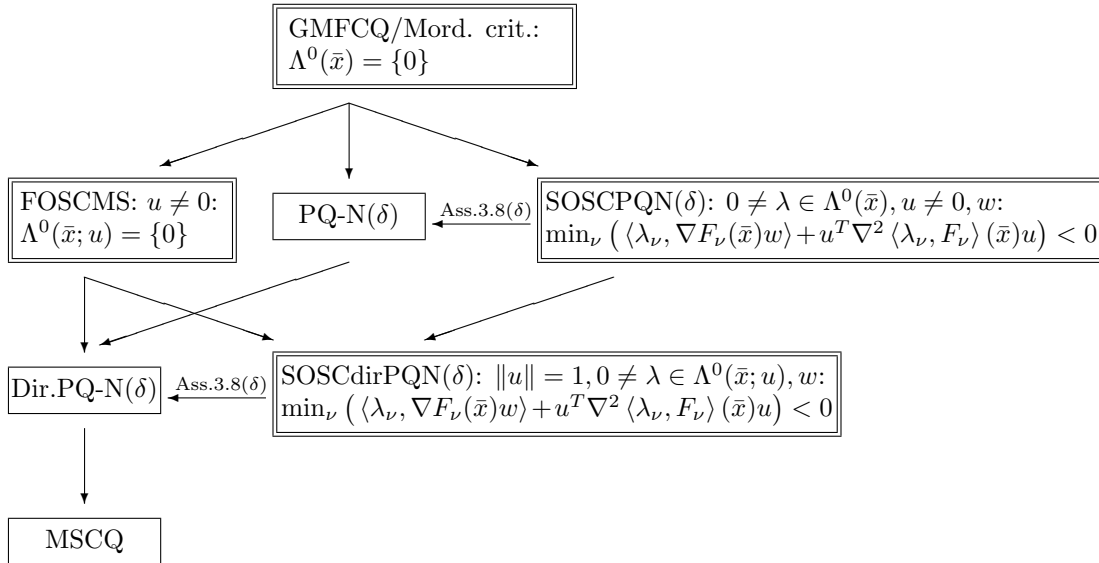

FIGURE 1. Constraint qualifications for GMP (1).

A few comments on the above diagram are in order.

- The downward left arrows indicate a refinement from standard CQ to its directional version, taking into account that one can view FOSCMS as a directional version of the Mordukhovich criterion, and their validity follows directly from definition. It seems that it is always favorable to switch to a directional version as it introduces no disadvantages. In fact, in Example 2.7 one obtains that all directional CQs are fulfilled (since the strongest FOSCMS is fulfilled) but at the same time none of their non-directional counterparts are (since the mildest - quasi-normality - is violated).
- On the other hand, the downward right arrows indicate refinement to second-order sufficient conditions and are also easily obtained.
- For the horizontal arrows pointing left on needs an additional Assumption 3.8 in order to make these conditions applicable, i.e., to eventually show that they imply metric subregularity. They follow from Theorems 3.9 and 3.17, respectively, with the exception of the implication from SOSCdirPQN(δ) to directional PQ-normality w.r.t. δ for $\delta \neq \delta^P$, where some additional effort was needed to prove Corollary 3.19.
- The downward arrows denote the switch to some milder condition. The arrow pointing to metric subregularity is fundamental for our work and highly nontrivial. This implication was established in Theorem 3.2 (Theorem 3.6).

4 Programs with disjunctive constraints

In this section we study a special case of problem (1) in which the set Γ is *disjunctive*, which means that it can be written as a union of finitely many polyhedra, i.e.,

$$\Gamma = \bigcup_{\ell=1}^{N} \Gamma^\ell \quad \text{with} \quad \Gamma^\ell \subset \mathbb{R}^d \text{ convex polyhedral},\tag{21}$$

where we refer the reader to Section 4.1 for a definition of convex polyhedral set. Subsequently, we call problem (1) with Γ disjunctive (in the sense of (21)) as a *(mathematical) program with disjunctive constraints* or simply a *disjunctive program*. For the readers convenience we recall here that disjunctive programs were studied in several papers [22, 25, 4] and in the thesis [5]. The most prominent examples of disjunctive programs are provided by the aforementioned classes of MPCCs, MPVCs, MPrCCs, MPrPCs and MPSCs, see Section 1. Dropping standard constraints for brevity, all of these programs exhibit the general form

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{s.t.} \quad (G_i(x), H_i(x)) \in \widetilde{\Gamma} \ (i \in V),\tag{22}$$

where $f, G_i, H_i : \mathbb{R} \to \mathbb{R}$ are continuously differentiable, V is a finite index set and $\widetilde{\Gamma}$ is given by

(a) (complementarity constraints)

$$\widetilde{\Gamma} := \Gamma_{\mathrm{CC}} := \{(a,b) \mid ab = 0, \ a,b \geq 0\} = (\mathbb{R}_+ \times \{0\}) \cup (\{0\} \times \mathbb{R}_+);$$

(b) (vanishing constraints)

$$\widetilde{\Gamma} := \Gamma_{\mathrm{VC}} := \{(a,b) \mid ab \leq 0, \ b \geq 0\} = (\mathbb{R}_- \times \mathbb{R}_+) \cup (\mathbb{R}_+ \times \{0\});$$

(c) (relaxed cardinality constraints)

$$\widetilde{\Gamma} := \Gamma_{\mathrm{rCC}} := \{(a,b) \mid ab = 0, \ b \in [0,1]\} = (\mathbb{R} \times \{0\}) \cup (\{0\} \times [0,1]);$$

(d) (relaxed probabilistic constraints)

$$\widetilde{\Gamma} := \Gamma_{\mathrm{rPC}} := \{(a,b) \mid ab \leq 0, \ b \in [0,1]\} = (\mathbb{R}_- \times [0,1]) \cup (\mathbb{R}_+ \times \{0\});$$

(e) (switching constraints)

$$\widetilde{\Gamma} := \Gamma_{\mathrm{SC}} := \{(a,b) \mid ab = 0\} = (\mathbb{R} \times \{0\}) \cup (\{0\} \times \mathbb{R}),$$

Clearly, Γ_{CC}, Γ_{VC}, Γ_{rCC}, Γ_{rPC} and Γ_{SC} are disjunctive, rendering the resulting optimization problem a disjunctive program. We point out that there is generally not a unique way to write the disjunctive sets in (a)-(e) as a union of convex polyhedral sets. For instance, Γ_{VC} can be alternatively written as $\Gamma_{\mathrm{VC}} = (\mathbb{R}_- \times \mathbb{R}_+) \cup (\mathbb{R} \times \{0\})$.

The main finding of this section is to show that the crucial Assumption 3.8 is automatically fulfilled for disjunctive programs. In addition, we also prove that directional pseudo-normality does not only imply, but is in fact equivalent to its simplified form from Theorem 3.17, which suggests that our sufficient conditions are not too restrictive. Recall that Example 3.18 shows that, in general, the simplified form is strictly stronger. For these purposes, we commence our study with a preliminary section on the variational geometry of convex polyhedral sets and how these extend to a more general setting.

4.1 Key properties of convex polyhedral sets

Recall that a set is said to be *convex polyhedral* (or a *convex polyhedron*) if it is the intersection of finitely many closed half-spaces. In particular, for a convex polyhedron $P \subset \mathbb{R}^s$ there exist $p \in \mathbb{N}$ and $a_j \in \mathbb{R}^s$, $\beta_j \in \mathbb{R}$ $(j = 1, \ldots, p)$ such that

$$P = \{y \mid \langle a_j, y \rangle \leq \beta_j \ (j = 1, \ldots, p)\}.$$

Clearly, every convex polyhedron is closed. Due to convexity of P, the regular and limiting normal cone to P coincide with the classical normal cone of convex analysis, see (3). Given $y \in P$, we have

$$N_P(y) = \left\{ \sum_{j \in J(y)} \lambda_j a_j \mid \lambda_j \geq 0 \right\},$$

where $J(y) := \{j \in \{1, \ldots, p\} \mid \langle a_j, y \rangle = \beta_j\}$, i.e., the normal cone of P at y is the convex cone generated by $\{a_j \mid j \in J(y)\}$, see e.g. [34, p. 67]. Therefore, there is only a finite number of different normal cones induced by a convex polyhedral set, in fact, this number is bounded by $p!$ (as there can be at most $p!$ active sets in $\{1, \ldots, p\}$).

We will make use of the essential properties of convex polyhedra. The first one is the well-known *exactness of tangent approximation*, see [55, Exercise 6.47]: Given a convex polyhedron P, for any $\bar{y} \in P$ there exists a neighborhood U of $\bar{y}$ such that

$$P \cap U = (\bar{y} + T_P(\bar{y})) \cap U. \tag{23}$$

In particular, taking into account [55, Exercise 6.44], one has

$$N_P(\bar{y}) = N_{T_P(\bar{y})}(0).$$

The second property is closely related to Assumption 3.8 as stated in the following lemma, part (ii).

Lemma 4.1. *Let $P \subset \mathbb{R}^s$ be closed and convex, let $\{y^k \in P\} \to \bar{y}$ and $\{\lambda^k \in N_P(y^k)\} \to \bar{\lambda}$. Then there exists a subsequence $K \subset \mathbb{N}$ such that the following hold:*

  (i) We have $\langle \bar{\lambda}, y^k - \bar{y} \rangle \leq 0$ for all $k \in K$;
  (ii) Moreover, if P is polyhedral then $\langle \bar{\lambda}, y^k - \bar{y} \rangle = 0$ for all $k \in K$.

Proof. (i) Taking the limit in $\lambda^k \in N_P(y^k)$ yields $\bar{\lambda} \in N_P(\bar{y})$. In particular, as $y^k \in P$ we get $\langle \bar{\lambda}, y^k - \bar{y} \rangle \leq 0$ $(k \in \mathbb{N})$.

(ii) Recall from the discussion above, that for a convex polyhedral set there are only finitely many different normal cones. Hence, there exists a subsequence $K \subset \mathbb{N}$ such that $N_P(y^k) \equiv \mathcal{N}$ for all $k \in K$ and some closed convex cone $\mathcal{N}$. Consequently, from $\lambda^k \in N_P(y^k)$ we obtain $\bar{\lambda} \in \mathcal{N} = N_P(y^k)$ and hence $\langle \bar{\lambda}, y^k - \bar{y} \rangle \geq 0$ due to convexity of P and $\bar{y} \in P$. □

The above lemma immediately yields that Assumption 3.8 for the multi-index $\delta^P := d$ is fulfilled at every feasible point for program (1) with convex polyhedral Γ, regardless of the constraint mapping F. However, since we are not primarily interested in this convex polyhedral setting, we now state the desirable properties from (23) and Lemma 4.1 (ii) in a general form. To this end, given an arbitrary closed set $C \subset \mathbb{R}^d$ and $\bar{y} \in C$, consider the following condition:

$$\exists\, U(\bar{y}): \quad C \cap U(\bar{y}) = (\bar{y} + T_C(\bar{y})) \cap U(\bar{y}), \tag{P1}$$

where $U(\bar{y})$ denotes a neighborhood of $\bar{y}$. Moreover, given also a multi-index $\delta \in \mathbb{N}^l$ with $|\delta| = d$ and $\bar{\lambda} \in \mathbb{R}^d$, consider the condition:

$$\forall \{y^k \in C\} \to \bar{y}, \ \{\lambda^k \in \widehat{N}_C(y^k)\} \to \bar{\lambda}, \ \exists \, K \subset \mathbb{N} : \ \langle \bar{\lambda}_\nu, y^k_\nu - \bar{y}_\nu \rangle = 0 \ (\nu \in I_\delta, \ k \in K), \qquad \text{(P2)}$$

where K is a subsequence of $\mathbb{N}$. Note that (P2) is automatically fulfilled if $\bar{\lambda} \notin N_C(\bar{y})$. We will repeatedly refer to these conditions in the subsequent study and hence we formulated it for an arbitrary multi-index δ. Clearly, if $\bar{x}$ is feasible for (1) and Γ satisfies (P2) for δ, $\bar{y} = F(\bar{x})$ and every multiplier $\bar{\lambda} \in N_\Gamma(F(\bar{x}))$, then Assumption 3.8 for δ is fulfilled at $\bar{x}$.

Motivated by the disjunctive setting in (21), for the remainder of our study we deal with sets generated by unions and, in addition, Cartesian products of convex polyhedra (see the product setting in Section 5). Hence, we now examine properties (P1) and (P2) under these set operations on convex polyhedra.

Consider first a collection of closed sets $C^i \subset \mathbb{R}^d$ for $i = 1, \ldots, q$ and set $C := \bigcup_{i=1}^q C^i$. We start with some elementary observations about tangent and normal cones. To this end, for $y \in C$, let us denote $\mathcal{I}(y) := \{ i \in \{1, \ldots, q\} \mid y \in C^i \}$ and observe that, by the definition of the tangent cone, we have

$$T_C(y) = \bigcup_{i \in \mathcal{I}(y)} T_{C^i}(y), \qquad (24)$$

hence, by polarization

$$\widehat{N}_C(y) = \bigcap_{i \in \mathcal{I}(y)} \widehat{N}_{C^i}(y). \qquad (25)$$

This yields the following elementary estimate which could also be derived from the more general result [6, Proposition 3.1].

Lemma 4.2 (Elementary estimate for normal cone of union). *Let $C = \bigcup_{i=1}^q C^i$ with $C^i \subset \mathbb{R}^d$ $(i = 1, \ldots, q)$ closed and let $y \in C$. Then*

$$N_C(y) \subset \bigcup_{i \in \mathcal{I}(y)} N_{C^i}(y).$$

Proof. Let $\lambda \in N_C(y)$. Then there exists $y^k \in C$ with $y^k \to y$ and $\lambda^k \in \widehat{N}_C(y^k)$ with $\lambda^k \to \lambda$. By (25), for all $k \in \mathbb{N}$ and $i \in \mathcal{I}(y^k)$ we have $\lambda^k \in \widehat{N}_{C^i}(y^k)$. By closedness of the C^i, we have $\mathcal{I}(y^k) \subset \mathcal{I}(y)$ for all $k \in \mathbb{N}$ sufficiently large. Hence by finiteness of $\mathcal{I}(y)$ we can assume that there exists $j \in \mathcal{I}(y)$ and a subsequence $K \subset \mathbb{N}$ such that

$$\lambda^k \in \widehat{N}_{C^j}(y^k) \quad (k \in K),$$

and we conclude $\lambda \in N_{C^j}(y)$. $\qquad \square$

On the other hand, consider now $C = \prod_{i=1}^r C_i$, where $C_i \subset \mathbb{R}^{d_i}$ is closed for $i = 1, \ldots, r$ and let $y = (y_1, \ldots, y_r) \in C$. By [55, Proposition 6.41], we have

$$\widehat{N}_C(y) = \prod_{i=1}^r \widehat{N}_{C_i}(y_i) \ \text{ and } \ N_C(y) = \prod_{i=1}^r N_{C_i}(y_i). \qquad (26)$$

Note that for the tangent cones, [55, Proposition 6.41] in general yields only the inclusion $T_C(y) \subset \prod_{i=1}^r T_{C_i}(y_i)$. It can be easily seen, however, that

$$T_C(y) = \prod_{i=1}^r T_{C_i}(y_i) \qquad (27)$$

holds, provided C_i satisfies (P1) at $\bar{y}_i$ for all $i = 1, \ldots, r$. Indeed, for $v = (v_i) \in \prod_{i=1}^r T_{C_i}(y_i)$ we readily obtain from (P1) for every $i = 1, \ldots, r$ the existence of $\alpha_i > 0$ such that $y_i + \alpha v_i \in C_i$ holds for all $\alpha \leq \alpha_i$. Taking $\bar{\alpha} := \min \alpha_i$ yields $y + \alpha v \in C$ for all $\alpha \leq \bar{\alpha}$ and $v \in T_C(y)$ follows.

Next we show that conditions (P1) and (P2) are preserved under unions and products, where (P2) is preserved under products with the obvious modifications of multi-index, point and multiplier.

Proposition 4.3. *Let $C = \bigcup_{i=1}^{q} C^i$ with $C^i \subset \mathbb{R}^d$ ($i = 1, \ldots, q$) closed and let $\bar{y} \in C$.*
- *(i) If C^i satisfies (P1) at $\bar{y}$ for all $i \in \mathcal{I}(\bar{y})$, then C also satisfies (P1) at $\bar{y}$.*
- *(ii) If C^i satisfies (P2) for some multi-index δ, the point $\bar{y}$ and some $\bar{\lambda}$ for all $i \in \mathcal{I}(\bar{y})$, then C also satisfies (P2) for δ, $\bar{y}$ and $\bar{\lambda}$.*

Proof. Denoting $U^i(\bar{y})$ for $i \in \mathcal{I}(\bar{y})$ the neighborhoods given by the assumption (i) and taking into account (24), the first statement follows easily by setting $U(\bar{y}) := \bigcap_{i \in \mathcal{I}(\bar{y})} U^i(\bar{y}) \cap \widetilde{U}(\bar{y})$, where $\widetilde{U}(\bar{y})$ is a neighborhood of $\bar{y}$ such that $C \cap \widetilde{U}(\bar{y}) = \bigcup_{i \in \mathcal{I}(\bar{y})} C^i \cap \widetilde{U}(\bar{y})$. Clearly, the existence of $\widetilde{U}(\bar{y})$ is guaranteed by the closedness of C^i ($i \notin \mathcal{I}(\bar{y})$).

In order to prove (ii), consider sequences $\{y^k \in C\} \to \bar{y}$ and $\{\lambda^k \in \widehat{N}_C(y^k)\} \to \bar{\lambda}$. By the proof of Lemma 4.2 we obtain the existence of $j \in \mathcal{I}(\bar{y})$ and a subsequence $\tilde{K} \subset \mathbb{N}$ such that

$$\lambda^k \in \widehat{N}_{C^j}(y^k) \quad (k \in \tilde{K}).$$

The assumption now yields the existence of a subsequence $K \subset \tilde{K}$ such that $\langle \bar{\lambda}_\nu, y_\nu^k - \bar{y}_\nu \rangle = 0$ for $\nu \in I_\delta$ and $k \in K$. □

Recall that if $\bar{\lambda} \notin N_{C^i}(\bar{y})$ for some $i \in \mathcal{I}(\bar{y})$, then C^i automatically satisfies (P2).

Proposition 4.4. *Let $C = \prod_{i=1}^{r} C_i$ with $C_i \subset \mathbb{R}^{d_i}$ ($i = 1, \ldots, r$) closed and $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_r) \in C$.*
- *(i) If C_i satisfies (P1) at $\bar{y}_i$ for all $i = 1, \ldots, r$, then C satisfies (P1) at $\bar{y}$.*
- *(ii) If C_i satisfies (P2) for multi-index δ_i with $|\delta_i| = d_i$, the point $\bar{y}_i$ and $\bar{\lambda}_i$ for all $i = 1, \ldots, r$, then C satisfies (P2) for $\delta = (\delta_1, \ldots, \delta_r)$, $\bar{y}$ and $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_r)$.*

Proof. Denoting by $U_i(\bar{y}_i)$ ($i = 1, \ldots, r$) the neighborhoods given by the assumption in (i), the first statement follows by simply setting $U(\bar{y}) := \prod_{i=1}^{r} U_i(\bar{y}_i)$ and applying (27).

In order to prove (ii), consider sequences $\{y^k \in C\} \to \bar{y}$ and $\{\lambda^k \in \widehat{N}_C(y^k)\} \to \bar{\lambda}$. By (26), we have $\lambda_i^k \in \widehat{N}_{C_i}(y_i^k)$ for every $i = 1, \ldots, r$ and $k \in \mathbb{N}$. By assumption, there exists a subsequence $K_1 \subset \mathbb{N}$ with $\langle \bar{\lambda}_{1,\nu_1}, y_{1,\nu_1}^k - \bar{y}_{1,\nu_1} \rangle = 0$ ($\nu_1 \in I_{\delta_1}, k \in K_1$). Consequently, by assumption, there exists a subsequence $K_2 \subset K_1$ such that $\langle \bar{\lambda}_{2,\nu_2}, y_{2,\nu_2}^k - \bar{y}_{2,\nu_2} \rangle = 0$ ($\nu_2 \in I_{\delta_2}, k \in K_2$). Repeating this argument $r - 1$ times, we find that there exists a subsequence $K(= K_r)$ such that $\langle \bar{\lambda}_{i,\nu_i}, y_{i,\nu_i}^k - \bar{y}_{i,\nu_i} \rangle = 0$ ($\nu_i \in I_{\delta_i}, k \in K$) for all $i = 1, \ldots, r$. This proves the statement. □

We conclude this subsection by showing that not only the program (1) with Γ fulfilling properties (P1) and (P2) automatically satisfies the crucial Assumption 3.8, but, moreover, directional PQ-normality is equivalent to its simplified counterpart in this case. We point out that this result is the very foundation for all remaining results of the paper.

Proposition 4.5. *Let $\bar{x}$ be feasible for (1) with Γ closed and satisfying (P1) at $\bar{y} = F(\bar{x})$ as well as (P2) for some multi-index δ, the point $\bar{y} = F(\bar{x})$ and every multiplier $\bar{\lambda} \in N_\Gamma(F(\bar{x}))$. Then Assumption 3.8 for δ is fulfilled at $\bar{x}$ and, moreover, (directional) PQ-normality w.r.t. δ at $\bar{x}$ is equivalent to its simplified form (15) from Theorem 3.9 (Theorem 3.17).*

Proof. Assumption 3.8 for δ at $\bar{x} \in \mathcal{X}$ follows from (P2) for Γ with δ at $\bar{y} = F(\bar{x}) \in \Gamma$. Hence, the statement for the nondirectional version follows from Theorem 3.9. Similarly, the implication from the directional simplified form to directional PQ-normality follows from Theorem 3.17.

It remains to show that PQ-normality w.r.t. δ in direction u implies its simplified form. We do this by contraposition, so let us assume that there exists $\bar{\lambda} \in \Lambda^0(\bar{x}; u) \setminus \{0\}$ and $x^k \to \bar{x}$ such that $(x^k - \bar{x})/\|x^k - \bar{x}\| \to u$ and

$$\langle \bar{\lambda}_\nu, F_\nu(x^k) - F_\nu(\bar{x}) \rangle > 0 \quad \text{for} \ \ \nu \in I_\delta(\bar{\lambda}), \ (k \in \mathbb{N}).$$

By the definition of the directional normal cone, there exists $\{t_k\} \downarrow 0$ and $\{w^k\} \to \nabla F(\bar{x})u$ as well as $\{\lambda^k \in \widehat{N}_\Gamma(F(\bar{x}) + t_k w^k)\} \to \bar{\lambda}$. Taking into account (P1) together with [55, Exercise 6.44] we obtain

$$
\begin{aligned}
\lambda^k &\in& \widehat{N}_\Gamma(F(\bar{x}) + t_k w^k) = \widehat{N}_{F(\bar{x}) + T_\Gamma(F(\bar{x}))}(F(\bar{x}) + t_k w^k) \subset \widehat{N}_{T_\Gamma(F(\bar{x}))}(t_k w^k) \\
&=& \widehat{N}_{T_\Gamma(F(\bar{x}))}(\alpha w^k) = \widehat{N}_{\Gamma - F(\bar{x})}(F(\bar{x}) + \alpha w^k - F(\bar{x})) \subset \widehat{N}_\Gamma(F(\bar{x}) + \alpha w^k)
\end{aligned}
$$

for any $\alpha > 0$ sufficiently small. Hence by setting $y^k := F(\bar{x}) + \|x^k - \bar{x}\| w^k$ we conclude $\lambda^k \in \widehat{N}_\Gamma(y^k)$. Moreover, (P2) for δ yields that, by passing to a subsequence if necessary, we may take y^k such that $\langle \bar{\lambda}_\nu, y_\nu^k - \bar{y}_\nu \rangle = 0$, for all $\nu \in I_\delta$ and $k \in \mathbb{N}$. Consequently, we obtain

$$\langle \bar{\lambda}_\nu, F_\nu(x^k) - y_\nu^k \rangle = \langle \bar{\lambda}_\nu, F_\nu(x^k) - F_\nu(\bar{x}) \rangle > 0.$$

Finally, $(y^k - F(\bar{x}))/\|x^k - \bar{x}\| = w^k \to \nabla F(\bar{x})u$, showing the violation of PQ-normality w.r.t. δ in direction u and the proof is complete. $\qquad\square$

4.2 Pseudo-normality for disjunctive programs

The desired results for the disjunctive setting (21) can be viewed as a simple corollary of our analysis in Section 4.1. Indeed, Proposition 4.3 yields that a disjunctive set Γ satisfies properties (P1) and (P2). In particular, due to (P1), the endeavor of computing the normal cone to disjunctive Γ at some point can be reduced to computing the normal cone to a union of finitely many polyhedral cones at zero, i.e.,

$$N_\Gamma(\bar{y}) = N_{\bigcup_{\ell=1}^N T_{\Gamma^\ell}(\bar{y})}(0) = N_{T_\Gamma(\bar{y})}(0),$$

see [33, p. 59]. More importantly, we can derive the following corollary from Proposition 4.5.

Corollary 4.6. *Let Γ be disjunctive in the sense of* (21). *Then Γ satisfies* (P1) *at every point $\bar{y} \in \Gamma$ as well as* (P2) *for the multi-index $\delta^P := d$ at every point $\bar{y}$ and every $\bar{\lambda}$. In particular, Assumption 3.8 for δ^P is fulfilled at every feasible point $\bar{x}$ for disjunctive programs. Moreover, (directional) pseudo-normality at $\bar{x}$ is equivalent to its simplified form: (for any $u \in \mathbb{R}^n$ with $\|u\| = 1$) there exists no nonzero $\bar{\lambda} \in \Lambda^0(\bar{x})$ $(\bar{\lambda} \in \Lambda^0(\bar{x}; u))$ such that there exists a sequence $x^k \to \bar{x}$ (with $(x^k - \bar{x})/\|x^k - \bar{x}\| \to u$) fulfilling*

$$\langle \bar{\lambda}, F(x^k) - F(\bar{x}) \rangle > 0 \quad (k \in \mathbb{N}). \tag{28}$$

We strongly emphasize that Corollary 4.6 clearly yields that the various definitions of pseudo-normality used in the literature stem from the same concept. In the general setting (1), pseudo-normality contains the additional sequence $\{y^k\}$, but in the special cases of disjunctive programs it reduces to the simplified version without $\{y^k\}$.

Corollary 4.6 also allows us to use all the sufficient conditions for pseudo-normality, hence also for MSCQ, studied in Section 3. These conditions now take on simpler forms since the vector optimization techniques reduce to standard optimization in the disjunctive setting. This can be seen from (28), which yields that pseudo-normality of $\bar{x}$ is equivalent to $\bar{x}$ being a local maximizer of $\langle \bar{\lambda}, F(x) \rangle$ for all $\bar{\lambda} \in \Lambda^0(\bar{x})$, cf. Theorem 3.10. In particular, the second-order sufficient conditions from Corollary 3.11 and Theorem 3.19 read as follows.

Corollary 4.7. *Let $\bar{x}$ be feasible for* (1) *with Γ disjunctive and F twice differentiable at $\bar{x}$. Consider the following two conditions:*

   (i) *second-order sufficient condition for pseudo-normality (SOSCPN): For every $0 \neq \bar{\lambda} \in \Lambda^0(\bar{x})$ and every $0 \neq u \in \mathbb{R}^n$ one has*

$$u^T \nabla^2 \langle \bar{\lambda}, F \rangle (\bar{x}) u < 0; \tag{29}$$

   (ii) *second-order sufficient condition for directional pseudo-normality (SOSCdirPN): For every $u \in \mathbb{R}^n$ with $\|u\| = 1$ and every $\neq \bar{\lambda} \in \Lambda^0(\bar{x}; u)$ one has* (29).

Then condition (i) (condition (ii)) implies (directional) pseudo-normality at $\bar{x}$. In particular, either of the two conditions implies MSCQ at $\bar{x}$.

Clearly, affine F can never fulfill the strict inequality of SOSCPN. The required maximality of $\bar{x}$ expressed in (28) can be secured nonetheless.

Corollary 4.8. *Let $\bar{x}$ be feasible for* (1) *with Γ disjunctive. If F is affine then pseudo-normality, and consequently also MSCQ, holds at $\bar{x}$.*

Proof. For F affine we have $F(x) = F(\bar{x}) + \nabla F(\bar{x})(x - \bar{x})$ for all $x \in \mathbb{R}^n$. Hence, taking into account $\bar{\lambda} \in \Lambda^0(\bar{x})$ we find that

$$\langle \bar{\lambda}, F(x) - F(\bar{x}) \rangle = \langle \nabla F(\bar{x})^T \bar{\lambda}, x - \bar{x} \rangle = 0,$$

showing that $\bar{x}$ is a local maximizer of $\langle \bar{\lambda}, F(x) \rangle$ and pseudo-normality thus follows.   □

We point out the that the sufficiency of SOSCdirPN for MSCQ established in Corollary 4.7 corresponds to the sufficiency of Gfrerer's SOSCMS for MSCQ (Proposition 2.8 (iii)). In turn, Corollary 4.8 corresponds to Robinson's result (Proposition 2.8 (iv)). Hence, by employing the notion of (directional) pseudo-normality and its sufficiency for MSCQ, we found new proofs for these highly important results. Moreover, the notion of directional quasi-normality, our weakest CQ that implies MSCQ by Theorem 3.6, unifies all sufficient conditions for MSCQ from Proposition 2.8.

4.3 Higher-order conditions

Our approach enables us to extend the above results by means of higher-order analysis. To this end, we rely once more on the notion of multi-indices. First, we introduce the following standard notation: Given $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we set

$$|\alpha| := \alpha_1 + \ldots + \alpha_n, \quad \alpha! := \alpha_1! \ldots \alpha_n!, \quad x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}.$$

Given a function $g : \mathbb{R}^n \to \mathbb{R}$, m-times differentiable at $\bar{x}$, and $\alpha \in \mathbb{N}^n$ with $|\alpha| \le m$ we set

$$D^\alpha g(\bar{x}) = \frac{\partial^{|\alpha|} g(\bar{x})}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}.$$

Note that $q! \sum_{|\alpha|=q} \frac{D^\alpha g(\bar{x})}{\alpha!} w^\alpha$ for some $q \le m$, $w \in \mathbb{R}^n$ corresponds to value $A(w, \ldots, w)$ of the multilinear mapping A of q arguments that represents the q-th derivative.

Corollary 4.9. *Let $\bar{x}$ be feasible for a disjunctive program with F being m-times differentiable at $\bar{x}$. Consider the following two conditions:*

(i) for every $0 \ne \bar{\lambda} \in \Lambda^0(\bar{x})$, $q < m$, $w \in \mathbb{R}^n$ and all $0 \ne u \in \mathbb{R}^n$ one has

$$\sum_{|\alpha|=q} \frac{D^\alpha \langle \bar{\lambda}, F \rangle (\bar{x})}{\alpha!} w^\alpha \le 0 \quad \text{and} \quad \sum_{|\alpha|=m} \frac{D^\alpha \langle \bar{\lambda}, F \rangle (\bar{x})}{\alpha!} u^\alpha < 0; \tag{30}$$

(ii) for every $u \in \mathbb{R}^n$ with $\|u\| = 1$, $0 \ne \bar{\lambda} \in \Lambda^0(\bar{x}; u)$, $q < m$ and all $w \in \mathbb{R}^n$ one has (30).

Then condition (i) (condition (ii)) implies (directional) pseudo-normality at $\bar{x}$. In particular, either of the two conditions implies MSCQ at $\bar{x}$.

Proof. Both statements follows from the same arguments, namely, given $0 \ne \bar{\lambda} \in \Lambda^0(\bar{x})$ and $q < m$ and setting $u_k := (x^k - \bar{x}) / \|x^k - \bar{x}\|$, Taylor expansion together with (30) yield

$$\begin{aligned} \langle \bar{\lambda}, F(x^k) - F(\bar{x}) \rangle &= \sum_{1 \le |\alpha| \le m} \frac{D^\alpha \langle \bar{\lambda}, F \rangle (\bar{x})}{\alpha!} (x^k - \bar{x})^\alpha + o(\|x^k - \bar{x}\|^m) \\ &\le \|x^k - \bar{x}\|^m \Big(\sum_{|\alpha|=m} \frac{D^\alpha \langle \bar{\lambda}, F \rangle (\bar{x})}{\alpha!} u_k^\alpha + o(1) \Big) < 0. \end{aligned}$$

□

Similarly as in the case of affine F, the strict inequality of the above higher-order sufficient conditions does not have to be fulfilled, as long as F has polynomial structure, i.e., for every $i = 1, \ldots, d$, and every x, we have

$$F_i(x) = \sum_{|\alpha| \le m} c_{i,\alpha} x^\alpha \tag{31}$$

for some $m \in \mathbb{N}$, denoting the degree of F, and $c_{i,\alpha} \in \mathbb{R}$. We point out that one actually has $c_{i,\alpha} = D^\alpha F_i(0)/\alpha!$ and (31) can be equivalently rewritten as

$$F_i(x) = \sum_{|\alpha| \le m} \frac{D^\alpha F_i(\bar{x})}{\alpha!} (x - \bar{x})^\alpha \tag{32}$$

for arbitrary $\bar{x} \in \mathbb{R}^n$.

Corollary 4.10. *Let $\bar{x}$ be feasible for a disjunctive program with F being polynomial of degree m, i.e., given by (31). Consider the following two conditions:*

(i) for every $0 \neq \bar{\lambda} \in \Lambda^0(\bar{x})$, $q \leq m$, and for all $w \in \mathbb{R}^n$ one has

$$\sum_{|\alpha|=q} \frac{D^\alpha \langle \bar{\lambda}, F \rangle (\bar{x})}{\alpha!} w^\alpha \leq 0; \tag{33}$$

(ii) for every $u \in \mathbb{R}^n$ with $\|u\| = 1$, $0 \neq \bar{\lambda} \in \Lambda^0(\bar{x}; u)$, $q \leq m$, and for all $w \in \mathbb{R}^n$ one has (33).

Then condition (i) (condition (ii)) implies (directional) pseudo-normality at $\bar{x}$. In particular, either of the two conditions implies MSCQ at $\bar{x}$.

Proof. Denoting $c_\alpha := (c_{1,\alpha}, \ldots, c_{d,\alpha})$ and taking into account (32), for any $\bar{\lambda} \neq 0$, one has

$$\langle \bar{\lambda}, F(x) \rangle = \sum_{|\alpha| \leq m} \langle \bar{\lambda}, c_\alpha \rangle x^\alpha = \sum_{|\alpha| \leq m} \frac{D^\alpha \langle \bar{\lambda}, F \rangle (\bar{x})}{\alpha!} (x - \bar{x})^\alpha$$

for every x. Hence, given $0 \neq \bar{\lambda} \in \Lambda^0(\bar{x})$ and $q \leq m$, both statements follows from (33) since

$$\langle \bar{\lambda}, F(x) - F(\bar{x}) \rangle = \sum_{1 \leq |\alpha| \leq m} \frac{D^\alpha \langle \bar{\lambda}, F \rangle (\bar{x})}{\alpha!} (x - \bar{x})^\alpha \leq 0.$$

$\square$

Of course the above higher-order conditions are sufficient for pseudo-normality and MSCQ also for general programs (1) fulfilling Assumption 3.8 for δ^P.

4.4 Summary and example for the disjunctive case

For the sake of completeness, we summarize the sufficient conditions for pseudo-normality and MSCQ in the disjunctive setting in the following theorem. Recall that the penalty function P_α given by (8) for Γ disjunctive in the sense of (21) reads

$$P_\alpha = f + \alpha \min_{\ell=1,\ldots,N} d_{\Gamma^\ell} \circ F \quad (\alpha > 0). \tag{34}$$

Theorem 4.11 (Sufficient conditions for pseudo-normality, MSCQ and exact penalization). *Consider (1) with Γ disjunctive in the sense of (21) and a feasible point $\bar{x}$. Then any of the conditions from Proposition 4.5 and Corollaries 4.7, 4.8, 4.9 and 4.10 implies (directional) pseudo-normality and MSCQ at $\bar{x}$. In particular, each of them implies exactness of the penalty function P_α from (34) and M-stationarity of $\bar{x}$, provided $\bar{x}$ is a local minimizer of the disjunctive program.*

The following example with parameters nicely demonstrates the applicability of our conditions.

Example 4.12. *Let $\Gamma \subset \mathbb{R}^3$ be given by $\Gamma := \mathbb{R} \times \{y \in \mathbb{R}^2 \,|\, y_2 \leq -|y_1|\}$, $F : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $F(x) := (x_1, x_2, ax_1^2 + bx_1^4 + cx_2^2)^T$ for some parameters $a, b, c \in \mathbb{R}$ and let $\bar{x} := (0, 0)$. Clearly,*

$$\nabla F(\bar{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \nabla^2 \langle \lambda, F \rangle (\bar{x}) = \begin{pmatrix} 2a & 0 \\ 0 & 2c \end{pmatrix}$$

and for any $\lambda \in \Lambda^0(\bar{x}) = \Lambda^0(\bar{x}; (\pm 1, 0)) = \mathbb{R}_+(0, 0, 1)^T$. Note also that $\Lambda^0(\bar{x}; u) = \emptyset$ for all directions $u \neq (\pm 1, 0)$ with $\|u\| = 1$ since $T_\Gamma(F(\bar{x})) = \Gamma$ and $\nabla F(\bar{x})u = (u_1, u_2, 0)$.

It is easy to see that sequence $\{x_k := (1/k, 0)\}$ shows violation of MSCQ if either $a > 0$ or $a = 0$ and $b > 0$. We will show that in all other cases MSCQ holds. In particular, we observe that the fulfillment of MSCQ does not depend on c. Nevertheless, in order to see which sufficient conditions can be used to verify MSCQ we split the analysis into 3 cases depending on c. In the following two tables corresponding to $c = 0$ and $c < 0$ we depict the mildest sufficient conditions ensuring the MSCQ for given parameters.

	$a = 0$	$a > 0$	$a < 0$
$b = 0$	Robinson SC	-	Polynomial SC
$b > 0$	-	-	SOSCMS
$b < 0$	4th-OSC	-	4th-OSC

(A) $c = 0$

	$a = 0$	$a > 0$	$a < 0$
$b = 0$	Polynomial SC	-	SOSCPN
$b > 0$	-	-	SOSCPN
$b < 0$	4th-OSC	-	SOSCPN

(B) $c < 0$

TABLE 1. *Robinson SC* refers to Robinson's result (Theorem 2.8 (iv)), *Polynomial SC* refers to the sufficient condition for polynomial F (Corollary 4.10 (i)) and 4th-*OSC* stands for the fourth-order sufficient condition based on Corollary 4.9 (i).

The most challenging and interesting is the case $c > 0$ that can not be handled by non-directional sufficient conditions. The conditions specified in the following table are the mildest sufficient condition ensuring MSCQ for $c > 0$, but are in fact sufficient for arbitrary $c \in \mathbb{R}$.

	$a = 0$	$a > 0$	$a < 0$
$b = 0$	Dir. Polynomial SC	-	SOSCMS
$b > 0$	-	-	SOSCMS
$b < 0$	Dir. $4^{th} - OSC$	-	SOSCMS

(C) $c < 0$

TABLE 2. *Dir. Polynomial SC* refers to the directional version of the sufficient condition for polynomial F (Corollary 4.10 (ii)) and *Dir. 4th-OSC* stands for the directional version of the fourth-order sufficient condition based on Corollary 4.9 (ii).

The power of our new sufficient conditions is nicely demonstrated for $a = 0$, when Gfrerer's SOSCMS can never be used. At the same time, for $a = b = 0$ we can also see the limitation of Robinson's result that can not be applied if $c \neq 0$.

5 Disjunctive programs with product structure

Revisiting the prototypical disjunctive programs from (22) (a)-(e), we observe that there are two additional product structures that are worth exploring: First, we see that these programs, with the constraints given by $(G_i, H_i) \in \widetilde{\Gamma}$ $(i \in V)$, fit the general framework (1) by setting $F(x) := (G_i(x), H_i(x))_{i \in V}$ and

$$\Gamma = \tilde{\Gamma} \times \ldots \times \tilde{\Gamma} = \tilde{\Gamma}^{|V|}.$$

Hence, in Section 5.1 we investigate this product structure in a generalized fashion. This leads to a natural justification for PQ-normality and, in particular, we establish an analysis of PQ-normality similar to the analysis of pseudo-normality from the previous section.

On the other hand, we also realize that the factors Γ_ν $(\nu \in \{CC, VC, rCC, rPC, SC\})$ in (22) (a)-(e) are unions of products of closed intervals. This motivates our study of *ortho-disjunctive programs* in Section 5.2 where we expand our analysis of quasi-normality for this class of problems. In particular, we recover several known results for MPCCs and MPVCs and obtain new corresponding results for MPSCs, MPrCCs and MPrPCs based on the standard (nondirectional) approach. Our directional results are new for all of the considered special classes of programs.

5.1 Cartesian products of disjunctive sets

As Example 3.16 demonstrates, the simplified form of quasi-normality is not sufficient for metric subregularity even in case the set Γ under consideration is convex polyhedral. This rules out the usage of sufficient conditions SOSCQN and SOSCdirQN to verify MSCQ. However, under an additional product structure, which is present in the myriad of applications that we have in the

back of our mind, this issue can be overcome. Indeed, we will later show that in this special setting (directional) quasi-normality coincides with its simplified form. In this first part, however, let us begin with the detailed study of PQ-normality. To this end, consider an instance of (1), where

$$\Gamma = \prod_{\nu \in I_\delta} \Gamma_\nu, \quad \Gamma_\nu = \bigcup_{\ell=1}^{N_\nu} \Gamma_\nu^\ell, \quad \Gamma_\nu^\ell \text{ convex polyhedral}, \tag{35}$$

for some multi-index $\delta \in \mathbb{N}^l$, i.e., Γ is the cartesian product of disjunctive sets. We emphasize that Γ given by (35) is still a disjunctive set in the sense of (21). Indeed, denoting $\mathcal{J} := \prod_{\nu \in I_\delta} \{1, \ldots, N_\nu\}$, for every $\boldsymbol{\ell} \in \mathcal{J}$ the set $\Gamma^{\boldsymbol{\ell}} := \prod_{\nu \in I_\delta} \Gamma_\nu^{\boldsymbol{\ell}_\nu}$ is convex polyhedral and $\Gamma = \bigcup_{\boldsymbol{\ell} \in \mathcal{J}} \Gamma^{\boldsymbol{\ell}}$. In particular, this means that Γ fulfills (P1) at every $\bar{y} \in \Gamma$.

Regardless, it turns out to be advantageous to exploit the underlying product structure of Γ rather than just treating Γ as a disjunctive set. One of the reasons is that we deal with the unions of only N_ν sets, which is often a small number ($N_\nu = 2$ for all ν for MPCCs, MPVCs, etc.), instead of dealing with the union of $|\mathcal{J}| = \prod_{\nu \in I_\delta} N_\nu$ sets. We point out that the concept of $\mathcal{Q}$-stationarity from [3, 4], mentioned in Section 1, takes advantage of this very observation.

Note that in this setting we partition F according to the disjunctive factors of Γ_ν, i.e.,

$$F(x) \in \Gamma \iff F_\nu(x) \in \Gamma_\nu \quad (\nu \in I_\delta).$$

We point out that, combining the formulas for unions (24) and (25) as well as Lemma 4.2 with the formulas for products (27) and (26), one can derive explicit formulas for tangent and normal cones of Γ given by (35). For our analysis, however, this is not needed and we istead directly apply Propositions 4.3 and 4.4 to conclude that such Γ again satisfies properties (P1) and (P2). As a result, Proposition 4.5 yields the following corollary.

Corollary 5.1. *Let Γ be given by (35) for some multi-index δ. Then Γ satisfies (P1) at every point $\bar{y} \in \Gamma$ as well as (P2) for δ at every point $\bar{y}$ and every $\bar{\lambda}$. In particular, Assumption 3.8 for δ is fulfilled at every feasible point $\bar{x}$ of (1) with such Γ. Moreover, (directional) PQ-normality w.r.t. δ at $\bar{x}$ is equivalent to its simplified form (15) from Theorem 3.9 (3.17).*

For Γ given by (35) the penalty function P_α from (8) based on the l_1-norm reads

$$P_\alpha = f + \alpha \sum_{\nu \in I_\delta} \mathrm{d}_{\Gamma_\nu} \circ F_\nu = f + \alpha \sum_{\nu \in I_\delta} \min_{\ell=1,\ldots,N_\nu} \mathrm{d}_{\Gamma_\nu^\ell} \circ F_\nu \quad (\alpha > 0), \tag{36}$$

Naturally, other l_p-norms can be used as well in definition of the penalty function and the following results remains true by equivalence of all norms in finite dimension.

In the following theorem, we sum up the obtained results.

Theorem 5.2 (Sufficient conditions for PQ-normality, MSCQ and exact penalization)**.** *Consider (1) with Γ given by (35) and a feasible point $\bar{x}$. Then each of the following conditions implies (directional) PQ-normality w.r.t. δ and MSCQ at $\bar{x}$: the maximality condition from Theorem 3.10, SOSCPQN(δ) from Corollary 3.11 (and SOSCdirPQN(δ) from Theorem 3.19). In particular, each of them implies exactness of the penalty function P_α from (36) and M-stationarity of $\bar{x}$, provided $\bar{x}$ is a local minimizer of (1).*

Since quasi-normality is a special case of PQ-normality, applying the above results to the multi-index $\delta^Q := (1, \ldots, 1) \in \mathbb{N}^d$, we get results regarding quasi-normality. In such case, however, Γ reduces to a product of unions of closed intervals. Hence, by this approach we may study quasi-normality for, e.g., NLPs, but the prominent examples from Section 4 do not fit such setting. This issue is addressed in the next subsection.

5.2 Ortho-disjunctive constraints

As advertized at the beginning of Section 5, another inspection of the sets Γ_{CC}, Γ_{VC}, Γ_{SC}, Γ_{rCC} and Γ_{rPC} from (22) (a)-(e) reveals yet another product structure "inside" the union. In a very

general form, given a multi-index δ, this can be cast by sets Γ of the form

$$\Gamma = \bigcup_{\ell=1}^{N} \Gamma^\ell, \ \Gamma^\ell = \prod_{\mu \in I_\delta} \Gamma^\ell_\mu, \ \Gamma^\ell_\mu \text{ convex polyhedral.} \tag{37}$$

As before, we skip writing down the explicit formulas for tangent and normal cones, since all the hard work has already been done in Section 4.1 and we just collect the results. Indeed, Propositions 4.3 and 4.4 again yield that Γ given by (37) satisfies properties (P1) and (P2) and hence Proposition 4.5 gives us the following result.

Corollary 5.3. *Let δ be a multi-index. Then Γ given by (37) satisfies (P1) at every point $\bar{y} \in \Gamma$ as well as (P2) for δ at every $\bar{y} \in \Gamma$ and every $\bar{\lambda}$. In particular, for program (1) with Γ given by (37), Assumption 3.8 for δ is fulfilled at every feasible point $\bar{x}$ and, moreover, (directional) PQ-normality w.r.t. δ at $\bar{x}$ is equivalent to its simplified form (15) from Theorem 3.9 (3.17).*

Naturally, one can proceed as in the previous subsection and consider products of sets of the type (37), combining the two approaches of "outer" and "inner" products. Here we focus on such analysis in the special case when factors the Γ^ℓ_μ are one-dimensional. On the one hand, this eases the notational burden tremendously, while on the other it still covers all the instances in (22) (a)-(e). Moreover, it allows for a refined study of quasi-normality. For these purposes, given a multi-index δ, we consider sets of the form

$$\Gamma = \prod_{\nu \in I_\delta} \Gamma_\nu, \quad \Gamma_\nu = \bigcup_{\ell=1}^{N_\nu} \Gamma^\ell_\nu, \tag{38}$$

where each set Γ^ℓ_ν ($\ell = 1, \ldots, N_\nu$, $\nu \in I_\delta$) is a product of closed convex subsets of $\bar{\mathbb{R}}$, i.e., closed intervals

$$\Gamma^\ell_\nu = \prod_{i \in I^\nu} [a^\ell_i, b^\ell_i], \tag{39}$$

where $a^\ell_i, b^\ell_i \in \bar{\mathbb{R}}$, $a^\ell_i \leq b^\ell_i$ and $a^\ell_i < \infty, b^\ell_i > -\infty$. Given $y = (y_\nu)_{\nu \in I_\delta} \in \Gamma$ we denote $\mathcal{I}^\nu(y_\nu) := \{\ell \in \{1, \ldots, N_\nu\} \mid y_\nu \in \Gamma^\ell_\nu\}$.

We call the sets Γ_ν ($\nu \in I_\delta$) defined by (38)-(39) *ortho-disjunctive*. Moreover, we refer to programs (1) with ortho-disjunctive Γ as *mathematical programs with ortho-disjunctive constraints* or briefly *ortho-disjunctive programs*.

In Section 5.1 we saw that a product of disjunctive sets is also a disjunctive set. Here one can proceed similarly to show that a product of ortho-disjunctive sets remains ortho-disjunctive, since a product of sets Γ^ℓ_ν given by (39) can again be written as a product of intervals.

For Γ_{CC}, Γ_{VC}, Γ_{SC}, Γ_{rCC} and Γ_{rPC} we have $|I^\nu| = 2$ and Γ_ν is the same for every ν.

Note that for a closed interval $[a, b]$ and $c \in [a, b]$ we have

$$N_{[a,b]}(c) = \begin{cases} 0 & \text{if } c \in (a, b), \\ \mathbb{R}_- & \text{if } c = a, \\ \mathbb{R}_+ & \text{if } c = b. \end{cases}$$

Hence, in this setting, the normal cones as well as tangent cones obviously possess very nice descriptions that can be fruitfully exploited in a different context.

Proceeding as before, applying Propositions 4.4, 4.3 and, again, Proposition 4.4, we conclude that an ortho-disjunctive Γ satisfies properties (P1) and (P2). Alternatively, one can also start with Corollary 5.3 to deduce that each factor Γ_ν fulfils (P1) and (P2) (with multi-index $(1, \ldots, 1) \in \mathbb{N}^{|I^\nu|}$) and then apply only Proposition 4.4 once. We emphasize that it results in Γ satisfying (P2) with multi-index $\delta^Q := (1, \ldots, 1)$, which is clearly very crucial. By means of Proposition 4.5 we thus obtain the following corollary.

Corollary 5.4. *Set Γ given by (38)-(39) for any multi-index δ satisfies (P1) at every point $\bar{y} \in \Gamma$ as well as (P2) for multi-index $\delta^Q := (1, \ldots, 1)$ at every $\bar{y}$ and every $\bar{\lambda}$. In particular, for program (1) with Γ given by (38)-(39), Assumption 3.8 for δ^Q is fulfilled at every feasible point $\bar{x}$ and, moreover, the (directional) quasi-normality at $\bar{x}$ is equivalent to its simplified form: (for any*

$u \in \mathbb{R}^n \setminus \{0\}$) *there exists no nonzero* $\bar{\lambda} \in \Lambda^0(\bar{x})$ ($\bar{\lambda} \in \Lambda^0(\bar{x}; u)$) *such that there exists a sequence* $x^k \to \bar{x}$ *(such that* $(x^k - \bar{x})/\|x^k - \bar{x}\| \to u$) *fulfilling*

$$\bar{\lambda}_i \left(F_i(x^k) - F_i(\bar{x}) \right) > 0 \quad \text{if} \quad \bar{\lambda}_i \neq 0, \ (k \in \mathbb{N}). \tag{40}$$

Analogous to the case of pseudo-normality, cf. the comments below Corollary 4.6, we have now clarified that, in fact, there is only one concept of quasi-normality which, in general, contains the additional sequence $\{y^k\}$, but in special cases simplifies into the known versions without $\{y^k\}$. Moreover, the above proposition provides the definition of quasi-normality for all other ortho-disjunctive programs.

Before we state the main result of this subsection that parallels Theorem 5.2 for PQ-normality, we write down explicitly the conditions from Theorems 3.10 and 3.19 and Corollary 3.11 for multi-index δ^Q corresponding to quasi-normality:

Given $\lambda = (\lambda_i)_{i \in I}$ with $I = \{1, \ldots, d\}$, φ^λ from (17) reads as

$$\varphi^\lambda(x) = (\lambda_i F_i(x))_{i \in I(\lambda)}, \tag{41}$$

where $I(\lambda) := I_{\delta^Q}(\lambda) = \{i = 1, \ldots, d \,|\, \lambda_i \neq 0\}$. Moreover, assuming that F is twice differentiable at $\bar{x}$, the second-order sufficient conditions from Corollary 3.11 and Theorem 3.19, respectively, read as follows:

- *Second-order sufficient condition for quasi-normality (SOSCQN):* For every $\bar{\lambda} \in \Lambda^0(\bar{x})$, every $u \in \mathbb{R}^n$ with $\|u\| = 1$ and $\nabla F_i(\bar{x})u = 0$ for all $i \in I(\bar{\lambda})$ and every $w \in \mathbb{R}^n$ with $\langle w, u \rangle = 0$ one has

$$\min_{i \in I(\bar{\lambda})} \left(\bar{\lambda}_i \nabla F_i(\bar{x})w + u^T \nabla^2 (\bar{\lambda}_i F_i)(\bar{x})u \right) < 0; \tag{42}$$

- *Second-order sufficient condition for directional quasi-normality (SOSCdirQN):* For every $u \in \mathbb{R}^n$ with $\|u\| = 1$ and $\nabla F(\bar{x})u \in T_\Gamma(F(\bar{x}))$, every $\bar{\lambda} \in \Lambda^0(\bar{x}; u)$ with $\nabla F_i(\bar{x})u = 0$ for all $i \in I(\bar{\lambda})$ and every w with $\langle w, u \rangle = 0$ one has (42).

Moreover, for a closed interval $[a, b]$ and $c \in \mathbb{R}$ we have

$$\mathrm{d}_{[a,b]}(c) = (c - a)^- + (c - b)^+,$$

where $(q)^- := -\min\{q, 0\}$ and $(q)^+ := \max\{q, 0\}$ denotes the negative and the positive part of any number $q \in \mathbb{R}$, respectively, extended to $\bar{\mathbb{R}}$ by the natural convention $(\infty)^- = (-\infty)^+ = 0$. Thus, the penalty function now reads as

$$P_\alpha = f + \alpha \sum_{\nu \in I_\delta} \min_{\ell = 1, \ldots, N_\nu} \mathrm{d}_{\Gamma_\nu^\ell} \circ F_\nu \tag{43}$$

$$= \begin{cases} f + \alpha \sum_{\nu \in I_\delta} \min_{\ell \in \mathcal{I}^\nu} \sum_{i \in I^\nu} \left((F_i(\cdot) - a_i^\ell)^- + (F_i(\cdot) - b_i^\ell)^+ \right) & (l_1\text{-norm}), \\ f + \alpha \sum_{\nu \in I_\delta} \min_{\ell \in \mathcal{I}^\nu} \max_{i \in I^\nu} \left((F_i(\cdot) - a_i^\ell)^- + (F_i(\cdot) - b_i^\ell)^+ \right) & (l_\infty\text{-norm}), \end{cases} \quad (\alpha > 0),$$

where for the "outer" product we stick to the l_1-norm, resulting in the "outer" sum, while for the "inner" product we consider both the l_1- as well as the l_∞-norm.

Theorem 5.5 (Sufficient conditions for quasi-normality, MSCQ and exact penalization). *Consider an ortho-disjunctive program, i.e., (1) with Γ given by (38)-(39) and a feasible point $\bar{x}$. Then each of the following conditions implies (directional) quasi-normality and MSCQ at $\bar{x}$: $\bar{x}$ being local weak efficient solution of the problem $\max_{x \in \mathbb{R}^n} \varphi^{\bar{\lambda}}(x)$ for every $\bar{\lambda} \in \Lambda^0(\bar{x})$, SOSCQN (and SOSCdirQN). In particular, each of them implies the exactness of penalty function P_α from (43) and M-stationarity of $\bar{x}$, provided $\bar{x}$ is a local minimizer of (1).*

Let us briefly comment on the importance of the previous theorem (together with Corollary 5.4). First, consider only the statement that the (simplified form of) quasi-normality (40) implies MSCQ and hence M-stationarity and exactness of the penalty function at local minimizers. For MPCCs, we thus recover the following results: [44, Theorem 3.3] (quasi-normality implies M-stationarity), [44, Lemma 4.3 and 4.4] (pseudo-normality implies MSCQ), [44, Theorem 4.5 and Corollary 4.6] (pseudo-normality implies exactness of l_1 and l_∞ penalty function), as well as [58, Theorem 3.1] (quasi-normality implies MSCQ). Similarly, for MPVCs we recover and improve [39, Theorem 3.1]

(pseudo-normality implies exactness of the penalty function) and the fact that quasi-normality implies M-stationarity, which is not stated in the paper, but follows directly from [39, Theorem 2.1 and Definition 2.3]. Moreover, to the best of our knowledge, pseudo- and quasi-normality were not yet introduced for MPSCs, MPrCCs and MPrPCs and all our results are hence new when applied to these problem classes.

Second, we also provide several verifiable sufficient conditions for quasi-normality, such as SOSCQN, but also conditions sufficient for pseudo- and PQ-normality (higher-order conditions, polynomiality of F), which make our results practically applicable.

Finally, we open a path for a refined analysis using directional quasi-normality as well as all the corresponding sufficient conditions (SOSCdirQN etc.).

In order to illuminate and compare our results with the literature, we conclude this section by writing it down explicitly for MPCCs, while the same exercise could be executed for MPVCs, MP-SCs, MPrCCs and MPrPCs. Recall that, dropping standard equality and inequality constraints, an MPCC is given as

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad G_i(x), H_i(x) \geq 0, \ G_i(x)H_i(x) = 0, \ i \in V.$$

The constraints of MPCCs fit the general setting $F(x) \in \Gamma$ with $F(x) := (G_i(x), H_i(x))_{i \in V}$, and $\Gamma := \Gamma_{\mathrm{CC}}^{|V|}$, where $\Gamma_{\mathrm{CC}} = (\mathbb{R}_+ \times \{0\}) \cup (\{0\} \times \mathbb{R}_+)$ is clearly ortho-disjunctive in the sense of (38)-(39). We point out that the standard approach to MPCCs is to consider $\Gamma := -\Gamma_{\mathrm{CC}}^{|V|}$ and $F(x) := (-G_i(x), -H_i(x))_{i \in V}$ in order to work with nonnegative signs of certain multipliers, while in our case we obtain the opposite sign restrictions.

A simple computation yields that for $(G, H) \in \Gamma_{\mathrm{CC}}$ we have

$$N_{\Gamma_{\mathrm{CC}}}(G, H) = \begin{cases} \{0\} \times \mathbb{R} & \text{if } G > 0 = H, \\ \mathbb{R} \times \{0\} & \text{if } G = 0 < H, \\ (\mathbb{R}_- \times \mathbb{R}_-) \cup (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}) & \text{if } G = 0 = H, \end{cases}$$

Hence, denoting

$$\begin{aligned}
I^{+0}(\bar{x}) &:= \{i \in V \mid G_i(\bar{x}) > 0 = H_i(\bar{x})\}, \\
I^{0+}(\bar{x}) &:= \{i \in V \mid G_i(\bar{x}) = 0 < H_i(\bar{x})\}, \\
I^{00}(\bar{x}) &:= \{i \in V \mid G_i(\bar{x}) = 0 = H_i(\bar{x})\}
\end{aligned}$$

for some feasible point $\bar{x}$, we conclude that $\lambda = (\lambda_i^G, \lambda_i^H)_{i \in V} \in \mathbb{R}^{2|V|} \in N_{\Gamma_{\mathrm{CC}}^{|V|}}(F(\bar{x}))$ if and only if

$$\lambda_i^G = 0, \ i \in I^{+0}(\bar{x}), \ \lambda_i^H = 0, \ i \in I^{0+}(\bar{x}) \text{ and } \lambda_i^G, \lambda_i^H \leq 0 \text{ or } \lambda_i^G \lambda_i^H = 0, \ i \in I^{00}(\bar{x}). \tag{44}$$

Consequently, $\bar{x}$ satisfies quasi-normality provided there is no nonzero $\bar{\lambda} = (\bar{\lambda}_i^G, \bar{\lambda}_i^H)_{i \in V}$ fulfilling

$$0 = \sum_{i \in V} \left(\bar{\lambda}_i^G \nabla G_i(\bar{x})^T + \bar{\lambda}_i^H \nabla H_i(\bar{x})^T \right)$$

together with (44) such that there exists a sequence $x^k \to \bar{x}$ with

$$\bar{\lambda}_i^G G_i(x^k) > 0 \text{ if } \bar{\lambda}_i^G \neq 0 \text{ and } \bar{\lambda}_i^H H_i(x^k) > 0 \text{ if } \bar{\lambda}_i^H \neq 0, \ (k \in \mathbb{N}).$$

On the other hand, $\bar{x}$ satisfies M-stationarity provided there exists $\bar{\lambda} = (\bar{\lambda}_i^G, \bar{\lambda}_i^H)_{i \in V}$ such that

$$0 = \nabla f(\bar{x}) + \sum_{i \in V} \left(\bar{\lambda}_i^G \nabla G_i(\bar{x})^T + \bar{\lambda}_i^H \nabla H_i(\bar{x})^T \right)$$

and (44).

Moreover, using the l_∞-norm, for $G, H \in \mathbb{R}^2$ we have $\mathrm{d}_{\Gamma_{\mathrm{CC}}}(G, H) = |\min\{G, H\}|$ and this agrees with the corresponding expression from (43), which reads as $\min\{\max\{(G)^-, |H|\}, \max\{|G|, (H)^-\}\}$. Consequently, we obtain

$$P_\alpha(x) = \sum_{i \in V} |\min\{G_i(x), H_i(x)\}|.$$

Conclusion

Building on newly developed directional techniques from variational analysis, this paper contains a complex and self-contained study of the metric subregularity constraint qualification (MSCQ) for broad classes of nonconvex optimization problems including, most importantly, disjunctive programs. Our findings reveal a common denominator of several prominent sufficient conditions for MSCQ occuring in the literature. Thus, our study unifies these powerful and seemingly independent approaches and provides a new essential insight. Moreover, it offers a wider spectrum of sufficient conditions for MSCQ, including point-based ones, and consequently not only unifies, but improves existing sufficient conditions, thus helping to close the gap between the metric regularity and metric *sub*regularity constraint qualification. Furthermore, by introducing the new class of ortho-disjunctive programs we established the appropriate framework for a unified study of several nonconvex optimization problems such as mathematical programs with complementarity, vanishing or switching constraints. These ortho-disjunctive programs hence provide an intriguing area for future research.

Acknowledgments

The authors would like to dedicate this paper to Helmut Gfrerer in honor of his 60th birthday.

References

[1] Achtziger, W., Kanzow, C.: Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications. Math. Program. **114**, 69–99 (2008)

[2] Adam, L., Branda, M.: Nonlinear Chance Constrained Problems: Optimality Conditions, Regularization and Solvers. J. Optim. Theory Appl. **170**(2), 419–436 (2016)

[3] Benko, M., Gfrerer, H.: On estimating the regular normal cone to constraint systems and stationary conditions. Optimization **66**, 61–92 (2017)

[4] Benko, M., Gfrerer, H.: New verifiable stationary concepts for a class of mathematical programs with disjunctive constraints. Optimization **67**, 1–23 (2018)

[5] Benko, M.: Numerical methods for mathematical programs with disjunctive constraints. PhD-thesis, Univ. Linz (2016).

[6] Benko, M., Gfrerer, H., Outrata, J. V.: Calculus for directional limiting normal cones and subdifferentials. Set-Valued Var. Anal., first online (14 August 2018), https://doi.org/10.1007/s11228-018-0492-5

[7] Benko, M., Gfrerer, H., Mordukhovich, B. S.: Characterizations of tilt-stable minimizers in second-order cone programming. submitted to SIAM J. Optim., 28 p. (2018), arXiv: 1809.03607.

[8] Benko, M., Gfrerer, H., Outrata, J. V., Stability analysis for parameterized variational systems with implicit constraints. submitted to SVVA, 26 p. (2018), arXiv: 1902.07464.

[9] Bertsekas, D., Ozdaglar, A. E.: Pseudonormality and a Lagrange multiplier theory for constrained optimization. J. Optim. Theory Appl. **114**, 287–343 (2002)

[10] Bigi, G.: On sufficient second order optimality conditions in multiobjective optimization. Math. Meth. Oper. Res. **63**, 77–85 (2006)

[11] Branda, M., Bucher, M., Červinka, M., Schwartz, A.: Convergence of a Scholtes-type regularization method for cardinality-constrained optimization problems with an application in sparse robust portfolio optimization. Comput. Optim. Appl. **70**(2), 503–530 (2018)

[12] Bucher, M., Schwartz, A.: Second-Order Optimality Conditions and Improved Convergence Results for Regularization Methods for Cardinality-Constrained Optimization Problems. J. Optim. Theory Appl. **178**(2), 383–410 (2018)

[13] Burdakov, O. P., Kanzow, C., Schwartz, A.: Mathematical Programs with Cardinality Constraints: Reformulation by Complementarity-Type Conditions and a Regularization Method. SIAM J. Optim. **26**(1), 397–425 (2016)

[14] Burke, J. V.: Calmness and exact penalization. SIAM J. Control and Optim. **29**, 493–497 (1991)

[15] Burke, J. V.: An exact penalization viewpoint of constrained optimization. SIAM J. Control Optim. **29**, 968–998 (1991)

[16] Červinka, M., Kanzow, C., Schwartz, A.: Constraint qualifications and optimality conditions for optimization problems with cardinality constraints. Math. Program. **160**, 353–377 (2016)

[17] Clarke, F. H.: Optimization and Nonsmooth Analysis. John Wiley & Sons, New York (1983)

[18] Dempe, S.: Foundations of Bilevel Programming. Nonconvex Optimization and Its Applications **61**, Kluwer Academic Publishers, Dordrecht, The Netherlands (2002)

[19] Dontchev, A. L., Rockafellar, R. T.: Implicit Functions and Solution Mappings. Springer, Heidelberg (2014)

[20] Ekeland, I.: On the variational principle. J. Math. Anal. Appl. **47**, 324–353 (1974)

[21] Fabian, M. , Henrion, R., Kruger, A. Y., Outrata, J. V.: Error bounds: necessary and sufficient conditions. Set-Valued Var. Anal. **18**, 121–149 (2010)

[22] Flegel, M. L., Kanzow, C., Outrata, J. V.: Optimality Conditions for Disjunctive Programs with Application to Mathematical Programs with Equilibrium Constraints. Set-Valued Anal. **15**, 139–162 (2007)

[23] Gfrerer, H.: On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs. Set-Valued Var. Anal. **21**, 151–176 (2013)

[24] Gfrerer, H.: On directional metric subregularity and second-order optimality conditions for a class of nonsmooth mathematical programs. SIAM J. Optim. **23**, 632–665 (2013)

[25] Gfrerer, H.: Optimality conditions for disjunctive programs based on generalized differentiation with application to mathematical programs with equilibrium constraints. SIAM J. Optim. **24**, 898–931 (2014)

[26] Gfrerer, H., Klatte, D.: Lipschitz and Hölder stability of optimization problems and generalized equations. Math. Program. **158**, 35–75 (2016)

[27] Gfrerer, H., Outrata, J. V.: On Lipschitzian properties of implicit multifunction. SIAM J. Optim. **26**, 2160–2189 (2016)

[28] Gfrerer, H., Mordukhovich, B. S.: Second-order variational analysis of parametric constraint and variational systems. SIAM J. Optim. **29**, 423–453 (2019)

[29] Gfrerer, H.: Linearized M-stationarity Conditions for General Optimization Problems. Set-Valued Var. Anal., first online (8 August 2018), `https://doi.org/10.1007/s11228-018-0491-6`

[30] Guo, L., Ye, J. J., Zhang, J.: Mathematical Programs with Geometric Constraints in Banach Spaces: Enhanced Optimality, Exact Penalty, and Sensitivity. SIAM J. Optim. **23**, 2295–2319 (2013)

[31] Henrion, R., Jourani, A., Outrata, J. V.: On the calmness of a class of multifunctions. SIAM J. Optim. **13**, 603–618 (2002)

[32] Henrion, R., Outrata, J. V.: Calmness of constraint systems with applications. Math. Program. **104** 437–464 (2005)

[33] Henrion, R., Outrata, J. V.: On calculating the normal cone to a finite union of convex polyhedra. Optimization **57**(1), 57–78 (2008)

[34] Hiriart-Urrruty, J.-B., Lemaréchal, C.: Fundamentals of Convex Analysis. Grundlehren Text Editions, Springer, Berlin, Heidelberg (2001)

[35] Hoheisel, T.: Mathematical programs with vanishing constraints. PhD-thesis, Julius–Maximilians–Universität Würzburg (2009)

[36] Hoheisel, T., Kanzow, C., Outrata, J. V.: Exact penalty results for mathematical programs with vanishing constraints. Nonlinear Anal. **72**, 2514–2526 (2010)

[37] Hoheisel, T., Kanzow, C., Schwartz, A.: Mathematical programs with vanishing constraints: A new regularization approach with strong convergence properties. Optimization **61**(6), 619-636 (2012)

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

[39] Hu, Q., Zhang, H., Chen, Y., Tang, M.: An Improved Exact Penalty Result for Mathematical Programs with Vanishing Constraints. Journal of Adv. in Appl. Math. **3**, 43–49 (2018)

[40] Hu, X. M., Ralph, D.: Convergence of a penalty method for mathematical programming with complementarity constraints. J. Optim. Theory Appl. **123**, 365–390 (2004)

[41] Ioffe, A. D.: Variational Analysis of Regular Mappings. Springer Monographs in Mathematics, Springer, Cham (2017)

[42] Ioffe, A. D., Outrata, J. V.: On metric and calmness qualification conditions in subdifferential calculus. Set-Valued Anal. **16**, 199–227 (2018)

[43] Izmailov, A. F., Solodov, M. V.: Mathematical programs with vanishing constraints: Optimality conditions, sensitivity, and a relaxation method. J. Optim. Theory Appl. **142**, 501–532 (2009)

[44] Kanzow, C., Schwartz, A.: Mathematical Programs with Equilibrium Constraints: Enhanced Fritz John-conditions, New Constraint Qualifications, and Improved Exact Penalty Results. SIAM J. Optim. **20**, 2730–2753 (2010)

[45] Kanzow. C, Mehlitz, P., and Steck, D.: Relaxation schemes for mathematical programs with switching constraints. arxiv.org/abs/1809.02388 (2018)

[46] Kruger, A. Y.: Error bounds and metric subregularity. Optimization **64**, 49–79 (2015)

[47] Luo, Z. Q., Pang, J. S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, New York, Melbourne (1996)

[48] Luo, Z. Q., Pang, J. S., Ralph, D., Wu, S. Q.: Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints. Math. Program. **75**, 19–76 (1996)

[49] Mehlitz, P.: Stationarity conditions and constraint qualifications for mathematical programs with switching constraints. Math. Program., first online (6 March 2019), `https://doi.org/10.1007/s10107-019-01380-5`

[50] Mordukhovich, B. S.: Variational analysis and generalized differentiation, I: Basic theory. Springer, Berlin, Heidelberg (2006)

[51] Mordukhovich, B. S.: Variational analysis and applications. Springer, Cham (2018)

[52] Outrata, J. V., Kočvara, M., Zowe, J.: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands (1998)

[53] Penot, J. P.: Error bounds, calmness and their applications in nonsmooth analysis. In Nonlinear Analysis and Optimization II: Optimization, A. Leizarowitz, B. S. Mordukhovich, I. Shafrir, and A. J. Zaslavski, eds., vol. 514 of Contemp. Math., AMS, Providence, RI, 225–248 (2010)

[54] Robinson, S. M.: Some continuity properties of polyhedral multifunctions. Math. Program. Studies **14**, 206–214 (1981)

[55] Rockafellar, R. T., Wets, R. J.-B.: Variational Analysis. A Series of Comprehensive Studies in Mathematics Vol. **317**, Springer, Berlin, Heidelberg (1998)

[56] Scholtes, S.: Nonconvex structures in nonlinear programming. Oper. Res. **52**, 368-383 (2004)

[57] Wu, Z., Ye, J. J.: First-order and second-order conditions for error bounds. SIAM J. Optim. **14**(3), 621–645 (2003)

[58] Ye, J. J., Zhang, J.: Enhanced Karush-Kuhn-Tucker conditions for mathematical programs with equilibrium constraints. J. Optim. Theory Appl. **164**, 777–794 (2014)

Latest Reports in this series

2009 - 2017

[..]

2018

2019