New verifiable stationarity concepts for a class of mathematical programs with disjunctive constraints

Matus Benko, Helmut Gfrerer*

Abstract

In this paper we consider a sufficiently broad class of nonlinear mathematical programs with disjunctive constraints, which, e.g., include mathematical programs with complemetarity/vanishing constraints. We present an extension of the concept of $\mathcal{Q}$-stationarity as introduced in the recent paper [2]. $\mathcal{Q}$-stationarity can be easily combined with the well-known notion of M-stationarity to obtain the stronger property of so-called $\mathcal{Q}_M$-stationarity. We show how the property of $\mathcal{Q}_M$-stationarity (and thus also of M-stationarity) can be efficiently verified for the considered problem class by computing $\mathcal{Q}$-stationary solutions of a certain quadratic program. We consider further the situation that the point which is to be tested for $\mathcal{Q}_M$-stationarity, is not known exactly, but is approximated by some convergent sequence, as it is usually the case when applying some numerical method.

Key words. Mathematical programs with disjunctive constraints, B-stationarity, M-stationarity, $\mathcal{Q}_M$-stationarity.
AMS subject classification. 49J53 90C33 90C46

1 Introduction

In this paper we consider the following *mathematical program with disjunctive constraints* (MPDC)

$$\min_{x \in \mathbb{R}^n} \quad f(x) \tag{1}$$

$$\text{subject to} \quad F_i(x) \in D_i := \bigcup_{j=1}^{K_i} D_i^j, \ i = 1, \ldots, m_D,$$

where the mappings $f : \mathbb{R}^n \to \mathbb{R}$ and $F_i : \mathbb{R}^n \to \mathbb{R}^{l_i}$, $i = 1, \ldots, m_D$ are assumed to be continuously differentiable and $D_i^j \subset \mathbb{R}^{l_i}$, $j = 1, \ldots, K_i$, $i = 1, \ldots, m_D$ are convex polyhedral sets.

Denoting $m := \sum_{i=1}^{m_D} l_i$,

$$F := (F_1, \ldots, F_{m_D}) : \mathbb{R}^n \to \mathbb{R}^m, \ D := \prod_{i=1}^{m_D} D_i \tag{2}$$

---

*Institute of Computational Mathematics, Johannes Kepler University Linz, A-4040 Linz, Austria, benko@numa.uni-linz.ac.at, helmut.gfrerer@jku.at

we can rewrite the MPDC (1) in the form

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad F(x) \in D. \tag{3}$$

It is easy to see that D can also be written as the union of $\prod_{i=1}^{m_D} K_i$ convex polyhedral sets by

$$D = \bigcup_{\nu \in \mathcal{J}} D(\nu) \quad \text{with} \quad \mathcal{J} := \prod_{i=1}^{m_D} \{1, \ldots, K_i\}, \ D(\nu) := \prod_{i=1}^{m_D} D_i^{\nu_i}. \tag{4}$$

As an example for MPDC consider a *mathematical program with complementarity constraints* (MPCC) given by

$$\min_{x \in \mathbb{R}^n} \quad f(x) \tag{5}$$
$$\text{subject to} \quad g_i(x) \leq 0, \ i = 1, \ldots m_I,$$
$$h_i(x) = 0, \ i = 1, \ldots m_E,$$
$$G_i(x) \geq 0, \ H_i(x) \geq 0, G_i(x)H_i(x) = 0, \ i = 1, \ldots m_C$$

with $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m_I$, $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m_E$, $G_i, H_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m_C$. This problem fits into our setting (1) with $m_D = m_C + 1$,

$$F_1 = (g_1, \ldots, g_{m_I}, h_1 \ldots, h_{m_E})^T, \ D_1^1 = \mathbb{R}_-^{m_I} \times \{0\}^{m_E}, \ l_1 = m_I + m_E, \ K_1 = 1$$
$$F_{i+1} = (-G_i, -H_i)^T, \ D_{i+1}^1 = \{0\} \times \mathbb{R}_-, \ D_{i+1}^2 = \mathbb{R}_- \times \{0\}, \ l_{i+1} = K_{i+1} = 2, \ i = 1, \ldots, m_C.$$

MPCC is known to be a difficult optimization problem, because, due to the complementarity constraints $G_i(x) \geq 0$, $H_i(x) \geq 0$, $G_i(x)H_i(x) = 0$, many of the standard constraint qualifications of nonlinear programming are violated at any feasible point. Hence it is likely that the usual Karush-Kuhn-Tucker conditions fail to hold at a local minimizer and various first-order optimality conditions such as Abadie (A-), Bouligand (B-), Clarke (C-), Mordukhovich (M-) and Strong (S-) stationarity conditions have been studied in the literature [6, 9, 17, 19, 20, 24, 25, 26, 27].

Another prominent example is the *mathematical program with vanishing constraints* (MPVC)

$$\min_{x \in \mathbb{R}^n} \quad f(x) \tag{6}$$
$$\text{subject to} \quad g_i(x) \leq 0, \ i = 1, \ldots m_I,$$
$$h_i(x) = 0, \ i = 1, \ldots m_E,$$
$$H_i(x) \geq 0, \ G_i(x)H_i(x) \leq 0, \ i = 1, \ldots m_V$$

with $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m_I$, $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m_E$, $G_i, H_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m_V$. Again, the problem MPVC can be written in the form (1) with $m_D = m_V + 1$, F_1, D_1^1 as in the case of MPCC and

$$F_{i+1} = (-H_i, G_i)^T, \ D_{i+1}^1 = \{0\} \times \mathbb{R}, \ D_{i+1}^2 = \mathbb{R}_-^2, \ l_{i+1} = K_{i+1} = 2, \ i = 1, \ldots, m_V.$$

Similar as in the case of MPCC, many of the standard constraint qualifications of nonlinear programming can be violated at a local solution of (6) and a lot of stationarity concepts have been introduced. For a comprehensive overview for MPVC we refer to [15] and the references therein.

However, when we do not formulate MPCC or MPVC as a nonlinear program but as a disjunctive program MPDC, then first-order optimality conditions can be formulated which are valid under weak constraint qualifications. We know that a local minimizer is always B-stationary, which geometrically means that no feasible descent direction exists, or, in a dual formulation, that the negative gradient of the objective belongs to the regular normal cone of the feasible region, cf. [23, Theorem 6.12]. The difficult task is now to estimate this regular normal cone. For this regular normal cone always a lower inclusion is available, which yields so-called S-stationarity conditions. For an upper estimate one can use the limiting normal cone which results in the so-called M-stationarity conditions. The notions of S-stationarity and M-stationarity have been introduced in [7] for general programs (3). S-stationarity always implies B-stationarity, but it requires some strong qualification condition on the constraints. On the other hand, M-stationarity requires only some weak constraint qualification but it does not preclude the existence of feasible descent directions. Further, it is not known in general how to efficiently verify the M-stationarity conditions, since the description of the limiting normal cone involves some combinatorial structure which is not known to be resolved without enumeration techniques. These difficulties in verifying M-stationarity have also some impact for numerical solution procedures. E.g., for many algorithms for MPCC it cannot be guaranteed that a limit point is M-stationary, cf. [18].

In the recent paper [2] we derived another upper estimate for the regular normal cone yielding so-called $\mathcal{Q}$-stationarity conditions. $\mathcal{Q}$-stationarity can be easily combined with M-stationarity to obtain so-called $\mathcal{Q}_M$ stationarity which is stronger than M-stationarity. For the disjunctive formulations of the problems MPCC and MPVC the $\mathcal{Q}$- and $\mathcal{Q}_M$-stationarity conditions have been worked out in detail in [2]. In this paper we extend this approach to the general problem MPDC. We show that under a qualification condition which ensures S-stationarity of local minimizers, $\mathcal{Q}$-stationarity and S-stationarity are equivalent. Further we prove that under some weak constraint qualification every local minimizer of MPDC is a $\mathcal{Q}_M$-stationary solution and we provide an efficient algorithm for verifying $\mathcal{Q}_M$-stationarity of some feasible point. More exactly, this algorithm either proves the existence of some feasible descent direction, i.e. the point is not B-stationary, or it computes multipliers fulfilling the $\mathcal{Q}_M$-stationarity condition. To this end we consider quadratic programs with disjunctive constraints (QPDC), i.e., the objective function f in MPDC is a convex quadratic function and the mappings F_i, $i = 1, \ldots, m_D$ are linear. We propose a basic algorithm for QPDC, which either returns a $\mathcal{Q}$-stationary point or proves that the problem is unbounded. Further we show that M-stationarity for MPDC is related with $\mathcal{Q}$-stationarity of some QPDC and the combination of the two parts yields the algorithm for verifying $\mathcal{Q}_M$-stationarity.

Our approach is well suited to the MPDC (1) when all the numbers K_i, $i = 1, \ldots, m_D$ are small or of moderate size. Our disjunctive structure is not induced by integral variables like, e.g., in [16]. It is also not related to the approach of considering the convex hull of a family of convex sets like in [1, 4].

The outline of the paper is as follows. In Section 2 we recall some basic definitions from variational analysis and discuss various stationarity concepts. In Section 3 we introduce the concepts of $\mathcal{Q}$- and $\mathcal{Q}_M$-stationarity for general optimization problems. These concepts are worked out in more detail for MPDC in Section 4. In Section 5 we consider quadratic programs with disjunctive linear constraints. We present a basic algorithm for solving such problems, which either return a $\mathcal{Q}$-stationary solution or prove that the problem is not bounded below. In the next section we demonstrate how this basic algorithm can be applied to a certain quadratic program with disjunctive linear constraints in order to verify M-stationarity or

$\mathcal{Q}_M$-staionarity of a point or to compute a descent direction. In the last Section 7 we present some results for numerical methods for solving MPDC which prevent convergence to non M-stationary and non $\mathcal{Q}_M$-stationary points.

Our notation is fairly standard. In Euclidean space $\mathbb{R}^n$ we denote by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ the Euclidean norm and scalar product, respectively, whereas we denote by $\|u\|_\infty := \max\{|u_i|\,|\,i = 1,\ldots,n\}$ the maximum norm. The closed ball around some point x with radius r is denoted by $\mathcal{B}(x,r)$. Given some cone $Q \subset \mathbb{R}^n$, we denote by $Q^\circ := \{q^* \in \mathbb{R}^n \,|\, \langle q^*, q\rangle \leq 0 \forall q \in Q\}$ its polar cone. By $\mathrm{d}(x,A) := \inf\{\|x - y\|\,|\,y \in A\}$ we refer to the usual distance of some point x to a set A. We denote by 0^+C the recession cone of a convex set C.

2 Preliminaries

For the reader's convenience we start with several notions from variational analysis. Given a set $\Omega \subset \mathbb{R}^n$ and a point $\bar{z} \in \Omega$, the cone

$$T_\Omega(\bar{z}) = \{w \,|\, \exists w_k \to w, t_k \downarrow 0 \text{ with } \bar{z} + t_k w_k \in \Omega\}$$

is called the (Bouligand/Severi) *tangent/contingent cone* to Ω at $\bar{z}$. The (Fréchet) *regular normal cone* to Ω at $\bar{z} \in \Omega$ can be equivalently defined either by

$$\widehat{N}_\Omega(\bar{z}) := \left\{v^* \in \mathbb{R}^d \,\Big|\, \limsup_{z \xrightarrow{\Omega} \bar{z}} \frac{\langle v^*, z - \bar{z}\rangle}{\|z - \bar{z}\|} \leq 0\right\},$$

where $z \xrightarrow{\Omega} \bar{z}$ means that $z \to \bar{z}$ with $z \in \Omega$, or as the dual/polar to the contingent cone, i.e., by

$$\widehat{N}_\Omega(\bar{z}) := T_\Omega(\bar{z})^\circ.$$

For convenience, we put $\hat{N}_\Omega(\bar{z}) := \emptyset$ for $\bar{z} \notin \Omega$. Further, the (Mordukhovich) *limiting/basic normal cone* to Ω at $\bar{z} \in \Omega$ is given by

$$N_\Omega(\bar{z}) := \left\{w^* \in \mathbb{R}^d \,|\, \exists z_k \to \bar{z}, w_k^* \to w^* \text{ with } w_k^* \in \widehat{N}_\Omega(z_k) \text{ for all } k\right\}.$$

If Ω is convex, then both the regular and the limiting normal cones coincide with the normal cone in the sense of convex analysis. Therefore we will use in this case the notation N_Ω.

Consider now the general mathematical program

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad F(x) \in D \tag{7}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $F : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable and $D \subset \mathbb{R}^m$ is a closed set. Let

$$\Omega := \{x \in \mathbb{R}^n \,|\, F(x) \in D\} \tag{8}$$

denote the feasible region of the program (7). Then a necessary condition for a point $\bar{x} \in \Omega$ being locally optimal is

$$\langle \nabla f(\bar{x}), u\rangle \geq 0 \ \forall u \in T_\Omega(\bar{x}), \tag{9}$$

which is the same as

$$-\nabla f(\bar{x}) \in \widehat{N}_\Omega(\bar{x}), \tag{10}$$

cf. [23, Theorem 6.12]. The main task of applying this first-order optimality condition now is the computation of the regular normal cone $\widehat{N}_\Omega(\bar{x})$ which is very difficult for nonconvex D.

We always have the inclusion

$$\nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x})) \subset \widehat{N}_\Omega(\bar{x}), \tag{11}$$

but equality will hold in (11) for nonconvex sets D only under comparatively strong conditions, e.g. when $\nabla F(\bar{x})$ is surjective, see [23, Exercise 6.7]. The following weaker sufficient condition for equality in (11) uses the notion of metric subregularity.

Definition 1. *A multifunction $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is called* metrically subregular *at a point $(\bar{x}, \bar{y})$ of its graph* $\mathrm{gph}\,\Psi$ *with modulus $\kappa > 0$, if there is a neighborhood U of $\bar{x}$ such that*

$$\mathrm{d}(x, \Psi^{-1}(\bar{y})) \leq \kappa \mathrm{d}(\bar{y}, \Psi(x)) \; \forall x \in U.$$

Theorem 1 ([12, Theorem 4])**.** *Let Ω be given by (8) and $\bar{x} \in \Omega$. If the multifunction $x \rightrightarrows F(x) - D$ is metrically subregular at $(\bar{x}, 0)$ and if there exists a subspace $L \subset R^m$ such that*

$$T_D(F(\bar{x})) + L \subset T_D(F(\bar{x})) \tag{12}$$

and

$$\nabla F(\bar{x})\mathbb{R}^n + L = \mathbb{R}^m, \tag{13}$$

then

$$\widehat{N}_\Omega(\bar{x}) = \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x})).$$

In order to state an upper estimate for the regular normal cone $\widehat{N}_\Omega(\bar{x})$ we need some constraint qualification.

Definition 2 ([7, Definition 6])**.** *Let Ω be given by (8) and let $\bar{x} \in \Omega$.*

1. *We say that the* generalized Abadie constraint qualification *(GACQ) holds at $\bar{x}$ if*

$$T_\Omega(\bar{x}) = T_\Omega^{\mathrm{lin}}(\bar{x}), \tag{14}$$

   where $T_\Omega^{\mathrm{lin}}(\bar{x}) := \{u \in \mathbb{R}^n \,|\, \nabla F(\bar{x})u \in T_D(F(\bar{x}))\}$ denotes the linearized cone.

2. *We say that the* generalized Guignard constraint qualification *(GGCQ) holds at $\bar{x}$ if*

$$(T_\Omega(\bar{x}))^\circ = (T_\Omega^{\mathrm{lin}}(\bar{x}))^\circ. \tag{15}$$

Obviously GGCQ is weaker than GACQ, but GACQ is easier to verify by using some advanced tools of variational analysis. E.g., if the mapping $x \rightrightarrows F(x) - D$ is metrically subregular at $(\bar{x}, 0)$ then GACQ is fulfilled at $\bar{x}$, cf. [14, Proposition 1]. Tools for verifying metric subregularity of constraint systems can be found e.g. in [11].

Proposition 1 ([2, Proposition 3])**.** *Let Ω be given by (8), let $\bar{x} \in \Omega$ and assume that GGCQ is fulfilled, while the mapping $u \rightrightarrows \nabla F(\bar{x})u - T_D(F(\bar{x}))$ is metrically subregular at $(0, 0)$. Then*

$$\widehat{N}_\Omega(\bar{x}) \subset \nabla F(\bar{x})^T N_{T_D(F(\bar{x}))}(0) \subset \nabla F(\bar{x})^T N_D(F(\bar{x})). \tag{16}$$

Note that we always have $N_{T_D(F(\bar{x}))}(0) \subset N_D(F(\bar{x}))$, see [23, Proposition 6.27]. However, if D is the union of finitely many convex polyhedral sets, then equality

$$N_{T_D(F(\bar{x}))}(0) = N_D(F(\bar{x})) \tag{17}$$

holds. This is due to the fact that by the assumption on D there is some neighborhood V of 0 such that $(D - F(\bar{x})) \cap V = T_D(F(\bar{x})) \cap V$.

Let us mention that metric subregularity of the constraint mapping $x \rightrightarrows F(x) - D$ at $(\bar{x}, 0)$ does not only imply GACQ and consequently GGCQ, but also metric subregularity of the mapping $u \rightrightarrows \nabla F(\bar{x})u - T_D(F(\bar{x}))$ at $(0, 0)$ with the same modulus, see [10, Proposition 2.1].

The concept of metric subregularity has the drawback that, in general, it is not stable under small perturbations. It is well known that the stronger property of metric regularity is robust.

Definition 3. *A multifunction* $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *is called* metrically regular *near a point* $(\bar{x}, \bar{y})$ *of its graph* $\mathrm{gph}\,\Psi$ *with modulus* $\kappa > 0$, *if there are neighborhoods* U *of* $\bar{x}$ *and* V *of* $\bar{y}$ *such that*

$$\mathrm{d}(x, \Psi^{-1}(y)) \leq \kappa \mathrm{d}(y, \Psi(x)) \ \forall (x, y) \in U \times V.$$

The infimum of the moduli κ *for which the property of metric regularity holds is denoted by*

$$\mathrm{reg}\,\Psi(\bar{x}, \bar{y}).$$

In the following proposition we gather some well-known properties of metric regularity:

Proposition 2. *Let* $\bar{x} \in F^{-1}(D)$ *where* $F : \mathbb{R}^n \to \mathbb{R}^m$ *is continuously differentiable and* D *is the union of finitely many convex polyhedral sets and consider the multifunctions* $x \rightrightarrows \Psi(x) := F(x) - D$ *and* $u \rightrightarrows D\Psi(\bar{x})(u) := \nabla F(\bar{x})u - T_D(F(\bar{x}))$. *Then*

$$\mathrm{reg}\,\Psi(\bar{x}, \bar{y}) = \mathrm{reg}\,D\Psi(\bar{x})(0, 0) = \max\left\{\frac{1}{\|\nabla F(\bar{x})^T\lambda\|} \ \Big|\ \lambda \in N_D(F(\bar{x})) = N_{T_D(F(\bar{x}))}(0),\ \|\lambda\| = 1\right\}.$$

Moreover for every $\kappa > \mathrm{reg}\,\Psi(\bar{x}, \bar{y})$ *there is a neighborhood* W *of* $\bar{x}$ *such that for all* $x \in W$ *the mapping* $u \rightrightarrows \nabla F(x)u - T_D(F(\bar{x}))$ *is metrically regular near* $(0, 0)$ *with modulus* κ,

$$\|\lambda\| \leq \kappa\|\nabla F(x)^T\lambda\| \ \forall \lambda \in N_D(F(\bar{x})) = N_{T_D(F(\bar{x}))}(0) \tag{18}$$

and

$$\mathrm{d}(u, \nabla F(x)^{-1}T_D(F(\bar{x}))) \leq \kappa \mathrm{d}(\nabla F(x)u, T_D(F(\bar{x}))) \ \forall u \in \mathbb{R}^n.$$

Proof. The statement follows from [23, Exercise 9.44] together with the facts that by our assumption on D condition (17) holds and that $T_D(F(\bar{x}))$ is a cone. $\square$

We now recall some well known stationarity concepts based on the considerations above.

Definition 4. *Let* $\bar{x}$ *be feasible for the program* (7).

(i) *We say that* $\bar{x}$ *is* B-stationary, *if* (9) *or, equivalently,* (10) *hold.*

(ii) *We say that* $\bar{x}$ *is* S-stationary, *if*

$$-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x})).$$

(iii) We say that $\bar{x}$ is M-stationary, if

$$-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T N_D(F(\bar{x})).$$

S- and M-stationarity have been introduced in [7] as a generalization of these notions for MPCC. Using the inclusion (5) it immediately follows, that S-stationarity implies B-stationarity. However the reverse implication only holds true under some additional condition on the constraints, e.g. under the assumptions of Theorem 1. A B-stationary point is M-stationary under the assumptions of Proposition 1. However, the inclusion $\widehat{N}_\Omega(\bar{x}) \subset \nabla F(\bar{x})^T N_D(F(\bar{x}))$ can be strict, implying that a M-stationary point $\bar{x}$ needs not to be B-stationary. Hence M-stationarity does eventually not preclude the existence of feasible descent directions, i.e. directions $u \in T_\Omega(\bar{x})$ with $\langle \nabla f(\bar{x}), u \rangle < 0$.

3 On $\mathcal{Q}$- and $\mathcal{Q}_M$-stationarity

In this section we consider an extension of the concept of $\mathcal{Q}$-stationarity as introduced in the recent paper [2]. $\mathcal{Q}$-stationarity is based on the following simple observation.

Consider the general program (7), assume that GGCQ holds at the point $\bar{x} \in \Omega$ and assume that we are given K convex cones $Q_i \subset T_D(F(\bar{x}))$, $i = 1, \ldots, K$. Then for each $i = 1, \ldots, K$ we obviously have $T_\Omega^{\text{lin}}(\bar{x}) = \nabla F(\bar{x})^{-1} T_D(F(\bar{x})) \supset \nabla F(\bar{x})^{-1} Q_i$ implying

$$\widehat{N}_\Omega(\bar{x}) = (T_\Omega^{\text{lin}}(\bar{x}))^\circ \subset (F(\bar{x})^{-1} Q_i)^\circ.$$

If we further assume that $(F(\bar{x})^{-1} Q_i)^\circ = \nabla F(\bar{x})^T Q_i^\circ$ and by taking into account, that by [2, Lemma 1] we have

$$(\nabla F(\bar{x})^T S_1) \cap (\nabla F(\bar{x})^T S_2) = \nabla F(\bar{x})^T (S_1 \cap (\ker \nabla F(\bar{x})^T + S_2))$$

for arbitrary sets $S_1, S_2 \subset \mathbb{R}^m$, we obtain

$$
\begin{aligned}
\widehat{N}_\Omega(\bar{x}) \;&\subset\; \bigcap_{i=1}^{K} \nabla F(\bar{x})^T Q_i^\circ = \nabla F(\bar{x})^T \big(Q_1^\circ \cap (\ker \nabla F(\bar{x})^T + Q_2^\circ)\big) \cap \bigcap_{i=3}^{K} \nabla F(\bar{x})^T Q_i^\circ \\
&=\; \nabla F(\bar{x})^T \big(Q_1^\circ \cap (\ker \nabla F(\bar{x})^T + Q_2^\circ) \cap (\ker \nabla F(\bar{x})^T + Q_3^\circ)\big) \cap \bigcap_{i=4}^{K} \nabla F(\bar{x})^T Q_i^\circ = \ldots \\
&=\; \nabla F(\bar{x})^T \big(Q_1^\circ \cap \bigcap_{i=2}^{K} (\ker \nabla F(\bar{x})^T + Q_i^\circ)\big).
\end{aligned}
$$

Here we use the convention that for for sets $S_1, \ldots, S_K \subset \mathbb{R}^m$ we set $\bigcap_{i=l}^{K} S_i = \mathbb{R}^m$ for $l > K$. It is an easy consequence of (11), that equality holds in this inclusion, provided $\nabla F(\bar{x})^T \big(Q_1^\circ \cap \bigcap_{i=2}^{K} (\ker \nabla F(\bar{x})^T + Q_i^\circ)\big) \subset \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x}))$. Hence we have shown the following theorem.

Theorem 2. *Assume that GGCQ holds at $\bar{x} \in \Omega$ and assume that $Q_1, \ldots, Q_K$ are convex cones contained in $T_D(F(\bar{x}))$. If*

$$(\nabla F(\bar{x})^{-1} Q_i)^\circ = \nabla F(\bar{x})^T Q_i^\circ, \; i = 1, \ldots, K, \tag{19}$$

then

$$\widehat{N}_\Omega(\bar{x}) \subset \nabla F(\bar{x})^T \big(Q_1^\circ \cap \bigcap_{i=2}^K (\ker \nabla F(\bar{x})^T + Q_i^\circ)\big) = \bigcap_{i=1}^K \nabla F(\bar{x})^T Q_i^\circ. \qquad (20)$$

Further, if

$$\nabla F(\bar{x})^T \big(Q_1^\circ \cap \bigcap_{i=2}^K (\ker \nabla F(\bar{x})^T + Q_i^\circ)\big) \subset \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x})), \qquad (21)$$

then equality holds in (20) *and* $\widehat{N}_\Omega(\bar{x}) = \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x}))$.

Remark 1. *Condition* (19) *is e.g. fulfilled, if for each* $i = 1, \dots, K$ *either there is a direction* u_i *with* $\nabla F(\bar{x}) u_i \in \mathrm{ri}\, Q_i$ *or* Q_i *is a convex polyhedral set, cf. [2, Proposition 1].*

The proper choice of $Q_1, \dots, Q_K$ is crucial in order that (20) provides a good estimate for the regular normal cone. It is obvious that we want to choose the cones Q_i, $i = 1, \dots, K$ as large as possible in order that the inclusion (20) is tight. Further it is reasonable that a good choice of $Q_1, \dots, Q_K$ fulfills

$$\bigcap_{i=1}^K Q_i^\circ = \widehat{N}_D(F(\bar{x})) \qquad (22)$$

because then equation (21) holds whenever $\nabla F(\bar{x})$ has full rank. We now show that (21) holds not only under this full rank condition but also under some weaker assumption.

Theorem 3. *Assume that GGCQ holds at* $\bar{x} \in \Omega$ *and assume that we are given convex cones* $Q_1, \dots, Q_K \subset T_D(F(\bar{x}))$ *fulfilling* (19), (22) *and*

$$\ker \nabla F(\bar{x})^T \cap (Q_1^\circ - Q_i^\circ) = \{0\}, i = 2, \dots, K. \qquad (23)$$

Then

$$\widehat{N}_\Omega(\bar{x}) = \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x})) = \nabla F(\bar{x})^T \big(Q_1^\circ \cap \bigcap_{i=2}^K (\ker \nabla F(\bar{x})^T + Q_i^\circ)\big).$$

In particular, (23) *holds if there is a subspace*

$$L \subset \bigcap_{i=1}^K \big(Q_i \cap (-Q_i)\big) \qquad (24)$$

such that (13) *holds.*

Proof. The statement follows from Theorem 2 if we can show that (21) holds. Consider $x^* \in \nabla F(\bar{x})^T \big(Q_1^\circ \cap \bigcap_{i=2}^K (\ker \nabla F(\bar{x})^T + Q_i^\circ)\big)$. Then there are elements $\lambda^i \in Q_i^\circ$, $i = 1, \dots, K$ and $\mu^i \in \ker \nabla F(\bar{x})^T$ such that $\lambda^1 = \mu^i + \lambda^i$, $i = 2, \dots, K$ and $x^* = \nabla F(\bar{x})^T \lambda^1$. We conclude $\mu^i = \lambda^1 - \lambda^i \in Q_1^\circ - Q_i^\circ$, implying $\mu^i \in \ker \nabla F(\bar{x})^T \cap (Q_1^\circ - Q_i^\circ) = \{0\}$ and thus

$$\lambda^1 = \lambda^2 = \dots = \lambda^K \in \bigcap_{i=1}^K Q_i^\circ = \widehat{N}_D(F(\bar{x})).$$

Hence $x^* \in \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x}))$ and (21) is verified. In order to show the last assertion note that from (24) we conclude $L \subset Q_i$ and consequently $Q_i^\circ \subset L^\circ = L^\perp$. Thus $Q_1^\circ - Q_i^\circ \subset L^\perp - L^\perp = L^\perp$, $i = 2, \dots, K$. Since

$$\ker \nabla F(\bar{x})^T \cap L^\perp = \big((\ker \nabla F(\bar{x})^T)^\perp + L\big)^\perp = (\nabla F(\bar{x}) \mathbb{R}^n + L)^\perp = \{0\},$$

it follows that (23) holds. $\qquad \square$

Corollary 1. *Assume that GGCQ holds at $\bar{x} \in \Omega$ and assume that we are given convex cones $Q_1, \ldots, Q_K \subset T_D(F(\bar{x}))$ fulfilling (19) and (22). Further assume that there is some subspace L fulfilling (12) and (13). Then the sets*

$$\tilde{Q}_i := Q_i + L, \; i = 1, \ldots, K$$

are convex cones contained in $T_D(F(\bar{x}))$,

$$(\nabla F(\bar{x})^{-1}\tilde{Q}_i)^\circ = \nabla F(\bar{x})^T \tilde{Q}_i^\circ, i = 1, \ldots, K \tag{25}$$

and

$$\widehat{N}_\Omega(\bar{x}) = \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x})) = \nabla F(\bar{x})^T \Big(\tilde{Q}_1^\circ \cap \bigcap_{i=2}^{K} (\ker \nabla F(\bar{x})^T + \tilde{Q}_i^\circ)\Big).$$

Proof. Firstly observe that $\tilde{Q}_i = Q_i + L \subset T_D(F(\bar{x})) + L \subset T_D(F(\bar{x}))$ by (12). Next consider $z \in \operatorname{ri} \tilde{Q}_i$. By (13) there exists $u \in \mathbb{R}^n$ and $l \in L$ such that $\nabla F(\bar{x})u + l = z$. Because of $-l \in L \subset \tilde{Q}_i$ we have $z - 2l \in \tilde{Q}_i$ and thus $\nabla F(\bar{x})u = z - l = \frac{1}{2}z + \frac{1}{2}(z - 2l) \in \operatorname{ri} \tilde{Q}_i$ by [22, Theorem 6.1] implying (25) by taking into account Remark 1. Further, from $Q_i \subset \tilde{Q}_i \subset T_D(F(\bar{x}))$ it follows that

$$\widehat{N}_D(F(\bar{x})) = (T_D(F(\bar{x})))^\circ \subset \bigcap_{i=1}^{K} \tilde{Q}_i^\circ \subset \bigcap_{i=1}^{K} Q_i^\circ = \widehat{N}_D(F(\bar{x})).$$

Finally note that $L \subset \tilde{Q}_i \cap (-\tilde{Q}_i), \; i = 1, \ldots, K$ and the assertion follows from Theorem 3. $\square$

The following definition is motivated by Theorem 2.

Definition 5. *Let $\bar{x}$ be feasible for the program (7) and let $Q_1, \ldots, Q_K$ be convex cones contained in $T_D(F(\bar{x}))$ fulfilling (19).*

(i) *We say that $\bar{x}$ is $\mathcal{Q}$-stationary with respect to $Q_1, \ldots, Q_K$, if*

$$-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T \Big(Q_1^\circ \cap \bigcap_{i=2}^{K} (\ker \nabla F(\bar{x})^T + Q_i^\circ)\Big).$$

(ii) *We say that $\bar{x}$ is $\mathcal{Q}_M$-stationary with respect to $Q_1, \ldots, Q_K$, if*

$$-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T \Big(N_D(F(\bar{x})) \cap Q_1^\circ \cap \bigcap_{i=2}^{K} (\ker \nabla F(\bar{x})^T + Q_i^\circ)\Big).$$

Note that this definition is an extension of the definition of $\mathcal{Q}$- and $\mathcal{Q}_M$-stationarity in [2], where only the case $K = 2$ was considered.

The following corollary is an immediate consequence of the definitions and Theorem 2.

Corollary 2. *Assume that GGCQ is fulfilled at the point $\bar{x}$ feasible for (7). Further assume that we are given convex cones $Q_1, \ldots, Q_K \subset T_D(F(\bar{x}))$ fulfilling (19). If $\bar{x}$ is B-stationary, then $\bar{x}$ is $\mathcal{Q}$-stationary with respect to $Q_1, \ldots, Q_K$. Conversely, if $\bar{x}$ is $\mathcal{Q}$-stationary with respect to $Q_1, \ldots, Q_K$ and (21) is fulfilled, then $\bar{x}$ is S-stationary and consequently B-stationary.*

We know that under the assumptions of Proposition 1 every B-stationary point $\bar{x}$ for the problem (7) is both M-stationary and $\mathcal{Q}$-stationary with respect to every collection of cones $Q_1, \ldots, Q_K \subset T_D(F(\bar{x}))$ fulfilling (19), i.e.

$$
\begin{aligned}
-\nabla f(\bar{x}) \quad &\in \quad \nabla F(\bar{x})^T N_D(F(\bar{x})) \cap \nabla F(\bar{x})^T \big(Q_1^\circ \cap \bigcap_{i=2}^{K} (\ker \nabla F(\bar{x})^T + Q_i^\circ) \big) \\
&= \quad \nabla F(\bar{x})^T \Big(\big(\ker \nabla F(\bar{x})^T + N_D(F(\bar{x})) \big) \cap Q_1^\circ \cap \bigcap_{i=2}^{K} (\ker \nabla F(\bar{x})^T + Q_i^\circ) \Big).
\end{aligned}
$$

Comparing this relation with the definition of $\mathcal{Q}_M$-stationarity we see that $\mathcal{Q}_M$-stationarity with respect to $Q_1, \ldots, Q_K$ is stronger than the simultaneous fulfillment of M-stationarity and $\mathcal{Q}$-stationarity with respect to $Q_1, \ldots, Q_K$. We refer to [2, Example 2] for an example which shows that $\mathcal{Q}_M$-stationarity is strictly stronger than M-stationarity. However, to ensure $\mathcal{Q}_M$-stationarity of a B-stationary point $\bar{x}$, some additional assumption has to be fulfilled.

Lemma 1. *Let $\bar{x}$ be B-stationary for the program (7) and assume that the assumptions of Proposition 1 are fulfilled at $\bar{x}$. Further assume that for every $\lambda \in N_{T_D(F(\bar{x}))}(0)$ there exists a convex cone $Q_\lambda \subset T_D(F(\bar{x}))$ containing λ and satisfying $(\nabla F(\bar{x})^{-1} Q_\lambda)^\circ = \nabla F(\bar{x})^T Q_\lambda^\circ$. Then there exists a convex cone $Q_1 \subset T_D(F(\bar{x}))$ fulfilling $(\nabla F(\bar{x})^{-1} Q_1)^\circ = \nabla F(\bar{x})^T Q_1^\circ$ such that for every collection $Q_2, \ldots, Q_K \subset T_D(F(\bar{x}))$ fulfilling (19) the point $\bar{x}$ is $\mathcal{Q}_M$ stationary with respect to $Q_1, \ldots, Q_K$.*

Proof. From the definition of B-stationarity and (16) we deduce the existence of $\lambda \in N_{T_D(F(\bar{x}))}(0)$ fulfilling $-\nabla f(\bar{x}) = \nabla F(\bar{x})^T \lambda$. By taking $Q_1 = Q_\lambda$ we obviously have $\lambda \in N_{T_D(F(\bar{x}))}(0) \cap Q_1^\circ \subset N_D(F(\bar{x})) \cap Q_1^\circ$ implying $-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T (N_D(F(\bar{x})) \cap Q_1^\circ)$. Now consider cones $Q_2, \ldots, Q_K \subset T_D(F(\bar{x}))$ fulfilling (19). Similar to the derivation of Theorem 2 we obtain

$$
\begin{aligned}
-\nabla f(\bar{x}) \quad &\in \quad \nabla F(\bar{x})^T (N_D(F(\bar{x})) \cap Q_1^\circ) \cap \bigcap_{i=2}^{K} \nabla F(\bar{x})^T Q_i^\circ \\
&= \quad \nabla F(\bar{x})^T \big(N_D(F(\bar{x})) \cap Q_1^\circ \cap \bigcap_{i=2}^{K} (\ker \nabla F(\bar{x})^T + Q_i^\circ) \big)
\end{aligned}
$$

and the lemma is proved. $\qquad \square$

Lemma 2. *Let $\bar{x}$ be feasible for (7) and assume that $T_D(F(\bar{x}))$ is the union of finitely many closed convex cones $C_1, \ldots, C_p$. Then for every $\lambda \in N_{T_D(F(\bar{x}))}(0)$ there is some $\bar{i} \in \{1, \ldots, p\}$ satisfying $\lambda \in C_{\bar{i}}^\circ$.*

Proof. Consider $\lambda \in N_{T_D(F(\bar{x}))}(0)$. By the definition of the limiting normal cone there are sequences $t_k \xrightarrow{T_D(F(\bar{x}))} 0$ and $\lambda_k \to \lambda$ with

$$
\lambda_k \in \widehat{N}_{T_D(F(\bar{x}))}(t_k) = \big(\bigcup_{i:t_k \in C_i} T_{C_i}(t_k) \big)^\circ = \bigcap_{i:t_k \in C_i} (T_{C_i}(t_k))^\circ = \bigcap_{i:t_k \in C_i} N_{C_i}(t_k).
$$

By passing to a subsequence if necessary we can assume that there is an index $\bar{i}$ such that $t_k \in C_{\bar{i}}$ for all k and we obtain $\lambda_k \in N_{C_{\bar{i}}}(t_k) = \{c^* \in C_{\bar{i}}^\circ \mid \langle c^*, t_k \rangle = 0\} \subset C_{\bar{i}}^\circ$. Since the polar cone $C_{\bar{i}}^\circ$ is closed, we deduce $\lambda \in C_{\bar{i}}^\circ$. $\qquad \square$

If $T_D(F(\bar{x}))$ is the union of finitely many convex polyhedral cones $C_1,\ldots,C_p$, then the mapping $u \rightrightarrows \nabla F(\bar{x})u - T_D(F(\bar{x}))$ is a polyhedral multifunction and thus metrically subregular at $(0,0)$ by Robinson's result [21]. Further we know that for any convex polyhedral cone Q we have $(\nabla F(\bar{x})^{-1}Q)^\circ = \nabla F(\bar{x})^T Q^\circ$. Hence we obtain the following corollary.

Corollary 3. *Assume that $\bar{x}$ is B-stationary for the program* (7)*, that GGCQ is fulfilled at $\bar{x}$ and that $T_D(F(\bar{x}))$ is the union of finitely many convex polyhedral cones. Then there is a convex polyhedral cone $Q_1 \subset T_D(F(\bar{x}))$ such that for every collection $Q_2,\ldots,Q_K$ of convex polyhedral cones contained in $T_D(F(\bar{x}))$ the point $\bar{x}$ is $\mathcal{Q}_M$-stationary with respect to $Q_1,\ldots,Q_K$.*

4 Application to MPDC

It is clear that $\mathcal{Q}$-stationarity is not a very strong optimality condition for every choice of $Q_1,\ldots,Q_K \subset T_D(F(\bar{x}))$. As mentioned above the fulfillment of (22) is desirable. For the general problem (7) it can be impossible to choose the cones $Q_1,\ldots,Q_K$ such that (22) holds. If $T_D(F(\bar{x}))$ is the union of finitely many convex cones $C_1,\ldots,C_p$ then we obviously have

$$\widehat{N}_D(F(\bar{x})) = \bigcap_{i=1}^{p} C_i^\circ.$$

However, to consider $\mathcal{Q}$-stationarity with respect to $C_1,\ldots,C_p$ is in general not a feasible approach because p is often very large. We will now work out that the concepts of $\mathcal{Q}$- and $\mathcal{Q}_M$-stationarity are tailored for the MPDC (1). In what follows let D and F be given by (2).

Given a point $y = (y_1,\ldots,y_{m_D}) \in D$, we denote by

$$\mathcal{A}_i(y) := \{j \in \{1,\ldots,K_i\} \,|\, y_i \in D_i^j\}, \ i = 1,\ldots,m_D$$

the indices of sets D_i^j which contain y_i. Further we choose for each $i = 1,\ldots,m_D$ some index set $\mathcal{J}_i(y) \subset \mathcal{A}_i(y)$ such that

$$T_{D_i}(y_i) = \bigcup_{j \in \mathcal{J}_i(y)} T_{D_i^j}(y_i). \tag{26}$$

Obviously the choice $\mathcal{J}_i(y) = \mathcal{A}_i(y)$ is feasible but for practical reasons it is better to choose $\mathcal{J}_i(y)$ smaller if possible. E.g., if $T_{D_i^{j_2}}(y_i) \subset T_{D_i^{j_1}}(y_i)$ holds for some indices $j_1, j_2 \in \mathcal{A}_i(y)$, then we will not include j_2 in $\mathcal{J}_i(y)$. Such a situation can occur e.g. in case of MPVC when $(-H_i(\bar{x}), G_i(\bar{x})) = (0,a)$ with $a < 0$.

Now consider

$$\nu \in \mathcal{J}(y) := \prod_{i=1}^{m_D} \mathcal{J}_i(y).$$

Since for every $i = 1,\ldots,m_D$ the set D_i is the union of finitely many convex polyhedral sets, for every tangent direction $t \in T_{D_i}(y_i)$ we have $y_i + \alpha t \in D_i$ for all $\alpha > 0$ sufficiently small. Hence we can apply [13, Proposition 1] to obtain

$$T_{D(\nu)}(y) = \prod_{i=1}^{m_D} T_{D_i^{\nu_i}}(y_i), \ \nu \in \mathcal{J}(y)$$

with $D(\nu)$ given by (4), and

$$T_D(y) = \prod_{i=1}^{m_D} T_{D_i}(y_i) = \prod_{i=1}^{m_D} \Big(\bigcup_{j \in \mathcal{J}_i(y)} T_{D_i^j}(y_i) \Big) = \bigcup_{\nu \in \mathcal{J}(y)} T_{D(\nu)}(y). \tag{27}$$

We will apply this setting in particular to points $y = F(\bar{x})$ with $\bar{x}$ feasible for MPDC.

Lemma 3. *Let $\bar{x}$ be feasible for the MPDC (1) and assume that we are given K elements $\nu^1, \ldots, \nu^K \in \mathcal{J}(F(\bar{x}))$ such that*

$$\{\nu_i^1, \ldots, \nu_i^K\} = \mathcal{J}_i(F(\bar{x})), \ i = 1, \ldots, m_D. \tag{28}$$

Then for each $l = 1, \ldots, K$ the cone $Q_l := T_{D(\nu^l)}(F(\bar{x}))$ is a convex polyhedral cone contained in $T_D(F(\bar{x}))$, $\big(\nabla F(\bar{x})^{-1} Q_l \big)^\circ = \nabla F(\bar{x})^T Q_l^\circ$, and

$$\bigcap_{l=1}^{K} Q_l^\circ = \widehat{N}_D(F(\bar{x})).$$

Proof. Obviously for every $l = 1, \ldots, K$ the cone Q_l is convex and polyhedral because it is the product of convex polyhedral cones. This implies $\big(\nabla F(\bar{x})^{-1} Q_l \big)^\circ = \nabla F(\bar{x})^T Q_l^\circ$ and $Q_l \subset T_D(F(\bar{x}))$ follows from (27). By taking into account (27) the last assertion follows from

$$
\begin{aligned}
\widehat{N}_D(F(\bar{x})) &= \big(T_D(F(\bar{x})) \big)^\circ = \prod_{i=1}^{m_D} \Big(\bigcup_{j \in \mathcal{J}_i(F(\bar{x}))} T_{D_i^j}(F_i(\bar{x})) \Big)^\circ = \prod_{i=1}^{m_D} \Big(\bigcup_{l=1}^{K} T_{D_i^{\nu_i^l}}(F_i(\bar{x})) \Big)^\circ \\
&= \prod_{i=1}^{m_D} \Big(\bigcap_{l=1}^{K} (T_{D_i^{\nu_i^l}}(F_i(\bar{x})))^\circ \Big) = \bigcap_{l=1}^{K} \Big(\prod_{i=1}^{m_D} (T_{D_i^{\nu_i^l}}(F_i(\bar{x})))^\circ \Big) = \bigcap_{l=1}^{K} \Big(\prod_{i=1}^{m_D} T_{D_i^{\nu_i^l}}(F_i(\bar{x})) \Big)^\circ \\
&= \bigcap_{l=1}^{K} Q_l^\circ.
\end{aligned}
$$

$\square$

Definition 6. *Let $\bar{x}$ be feasible for the MPDC (1) and let index sets $\mathcal{J}_i(F(\bar{x})) \subset \mathcal{A}_i(\bar{x})$, $i = 1, \ldots, m_D$ fulfilling (26) be given. Further we denote by $\mathcal{Q}(\bar{x})$ the collection of all elements $(\nu^1, \ldots, \nu^K)$ with $\nu^l \in \mathcal{J}(F(\bar{x})) = \prod_{i=1}^{m_D} \mathcal{J}_i(F(\bar{x}))$, $l = 1, \ldots, K$ such that (28) holds.*

1. *We say that $\bar{x}$ is $\mathcal{Q}$-stationary ($\mathcal{Q}_M$-stationary) for (1) with respect to $(\nu^1, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$, if $\bar{x}$ is $\mathcal{Q}$-stationary ($\mathcal{Q}_M$-stationary) with respect to $Q_1, \ldots, Q_K$ in the sense of Definition 5 with $Q_l := T_{D(\nu^l)}(F(\bar{x}))$, $l = 1, \ldots, K$.*

2. *We say that $\bar{x}$ is $\mathcal{Q}$-stationary ($\mathcal{Q}_M$-stationary) for (1) if $\bar{x}$ is $\mathcal{Q}$-stationary ($\mathcal{Q}_M$-stationary) for (1) with respect to some $(\nu^1, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$.*

Definition 6 is an extension of the definition of $\mathcal{Q}$- and $\mathcal{Q}_M$-stationarity made for MPCC and MPVC in [2]. Note that the number K appearing in the definition of $\mathcal{Q}(\bar{x})$ is not fixed. Denoting $K_{\min}(\bar{x})$ the minimal number K such that $(\nu^1, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$, we obviously have

$$K_{\min}(\bar{x}) = \max_{i=1,\ldots,m_D} |\mathcal{J}_i(F(\bar{x}))| \le \max_{i=1,\ldots,m_D} K_i.$$

We see from (27) that the tangent cone $T_D(F(\bar{x}))$ is the union of the $|\mathcal{J}(F(\bar{x}))| = \prod_{i=1}^{m_D} |\mathcal{J}_i(F(\bar{x}))|$ convex polyhedral cones $T_{D(\nu)}(y)$. Hence the minimal number $K_{\min}(\bar{x})$ is much smaller than the number of components of the tangent cone, except when all or nearly all sets $\mathcal{J}_i(F(\bar{x}))$ have cardinality 1. Further it is clear that for every $\nu^1 \in \mathcal{J}(F(\bar{x}))$ and every $K \geq K_{\min}(\bar{x})$ we can find $\nu^2, \ldots, \nu^K \in \mathcal{J}(F(\bar{x}))$ such that $(\nu^1, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$.

We allow K to be greater than $K_{\min}(\bar{x})$ for numerical reasons primarily. Recall that for testing $\mathcal{Q}$-stationarity with respect to $(\nu^1, \ldots, \nu^K)$, we have to check for all $l = 1, \ldots, K$ whether $-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T Q_l^\circ$, or equivalently, that $u = 0$ is a solution of the linear optimization program

$$\min \langle \nabla f(\bar{x}), u \rangle \quad \text{subject to} \quad \nabla F(\bar{x}) u \in Q_l$$

with $Q_l = T_{D(\nu^l)}(F(\bar{x}))$. Theoretically the treatment of degenerated linear constraints is not a big problem but the numerical practice tells us the contrary. In [3] we have implemented an algorithm for solving MPVC based on $\mathcal{Q}$-stationarity and the degeneracy of the linear constraints was the reason when the algorithm crashed. The possibility of choosing $K > K_{\min}(\bar{x})$ gives us more flexibility to avoid linear programs with degenerated constraints.

The following theorem follows from Corollaries 2, 3, Theorem 3 and the considerations above.

Theorem 4. *Let $\bar{x}$ be feasible for the MPDC (1) and assume that GGCQ is fulfilled at $\bar{x}$.*

(i) *If $\bar{x}$ is B-stationary then $\bar{x}$ is $\mathcal{Q}$-stationary with respect to every element $(\nu^1, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$ and there exists some $\bar{\nu}^1 \in \mathcal{J}(F(\bar{x}))$ such that $\bar{x}$ is $\mathcal{Q}_M$-stationary with respect to every $(\bar{\nu}^1, \nu^2, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$.*

(ii) *Conversely, if $\bar{x}$ is $\mathcal{Q}$-stationary with respect to some $(\nu^1, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$ and*

$$\nabla F(\bar{x})^T \left(Q_1^\circ \cap \bigcap_{l=2}^K (\ker \nabla F(\bar{x})^T + Q_l^\circ) \right) \subset \nabla F(\bar{x})^T \widehat{N}_D(F(\bar{x})), \tag{29}$$

where $Q_l := T_{D(\nu^l)} F(\bar{x})$, $l = 1, \ldots, K$, then $\bar{x}$ is S-stationary and consequently B-stationary. In particular, (29) is fulfilled if

$$\ker \nabla F(\bar{x})^T \cap \left(Q_1^\circ - Q_l^\circ \right) = \{0\}, \; l = 2, \ldots, K. \tag{30}$$

5 On quadratic programs with disjunctive constraints

In this section we consider the special case of quadratic programs with disjunctive constraints (QPDC)

$$\min_{x \in \mathbb{R}^n} \quad q(x) := \frac{1}{2} x^T B x + d^T x \tag{31}$$

$$\text{subject to} \quad A_i x \in D_i := \bigcup_{j=1}^{K_i} D_i^j, \; i = 1, \ldots, m_D,$$

where B is a positive semidefinite $n \times n$ matrix, $d \in \mathbb{R}^n$, A_i, $i = 1, \ldots, m_D$ are $l_i \times n$ matrices and $D_i^j \subset \mathbb{R}^{l_i}$, $i = 1, \ldots, m_D$, $j = 1, \ldots, K_j$ are convex polyhedral sets, i.e., QPDC is a special

case of MPDC with $f(x) = q(x)$ and $F_i(x) = A_i x$, $i = 1, \ldots, m_D$. In what follows we denote by A the $m \times n$ matrix

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_{m_D} \end{pmatrix},$$

where $m := \sum_{i=1}^{m_D} l_i$.

We start our analysis with the following preparatory lemma.

Lemma 4. *Assume that the convex quadratic program*

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T B x + d^T x \text{ subject to } Ax \in C \tag{32}$$

is feasible, where B is some symmetric positive semidefinite $n \times n$ matrix, $d \in \mathbb{R}^n$, A is a $m \times n$ matrix and $C \subset \mathbb{R}^m$ is a convex polyhedral set. Then either there exists a direction w satisfying

$$Bw = 0, \ Aw \in 0^+ C, \ d^T w < 0, \tag{33}$$

or the program (32) has a global solution $\bar{x}$.

Proof. Assume that for every w with $Bw = 0$, $Aw \in 0^+ C$ we have $d^T w \geq 0$, i.e. 0 is a global solution of the program

$$\min d^T w \quad \text{subject to} \quad w \in S := \left\{ w \,\middle|\, \begin{pmatrix} B \\ A \end{pmatrix} w \in \{0\}^n \times 0^+ C \right\}.$$

Since C is a convex polyhedral set, its recession cone $0^+ C$ is a convex polyhedral cone and so is $\{0\}^n \times 0^+ C$ as well. Hence

$$\widehat{N}_S(0) = S^\circ = (B^T \vdots A^T)(\{0\}^n \times 0^+ C)^\circ = B^T \mathbb{R}^n + A^T (0^+ C)^\circ$$

and from the first-order optimality condition $-d \in \widehat{N}_S(0)$ we derive the existence of multipliers $\mu_B \in \mathbb{R}^n$ and $\mu_C \in (0^+ C)^\circ$ such that

$$-d = B^T \mu_B + A^T \mu_C.$$

The convex polyhedral set C is the sum of the convex hull Σ of its extreme points and its recession cone. Hence for every x feasible for (32) there is some $c_1 \in \Sigma$ and some $c_2 \in 0^+ C$ such that $Ax = c_1 + c_2$ and, by taking into account $\mu_C^T c_2 \leq 0$, we obtain

$$\begin{aligned}
\frac{1}{2} x^T B x + d^T x &= \frac{1}{2} x^T B x - \mu_B^T B x - \mu_C^T A x \\
&= \frac{1}{2}(x - \mu_B)^T B (x - \mu_B) - \frac{1}{2} \mu_B^T B \mu_B - \mu_C^T c_1 - \mu_C^T c_2 \\
&\geq -\frac{1}{2} \mu_B^T B \mu_B - \mu_C^T c_1.
\end{aligned} \tag{34}$$

The set Σ is compact and we conclude that the objective of (32) is bounded below on the feasible domain $A^{-1} C$ by $-\frac{1}{2} \mu_B^T B \mu_B - \max_{c_1 \in \Sigma} \mu_C^T c_1$. Thus

$$\alpha := \inf\{ \frac{1}{2} x^T B x + d^T x \,|\, Ax \in C \}$$

is finite and there remains to show that the infimum is attained. Consider some sequence $x_k \in A^{-1}C$ with $\lim_{k\to\infty} \frac{1}{2} x_k^T B x_k + d^T x_k = \alpha$. We conclude from (34) that $(x_k - \mu_B)^T B(x_k - \mu_B)$ is bounded which in turn implies that the sequence $B^{1/2} x_k$ is bounded. Hence the sequence $x_k^T B x_k = \|B^{1/2} x_k\|^2$ is bounded as well and we can conclude also the boundedness of $d^T x_k$. By passing to a subsequence we can assume that the sequence $(B^{1/2} x_k, d^T x_k)$ converges to some (z, β) and it follows that $\alpha = \frac{1}{2}\|z\|^2 + \beta$. Since C is a convex polyhedral set, it follows by applying [22, Theorem 19.3] twice, that the sets $A^{-1}C$ and $\{(B^{1/2}u, d^T u) \,|\, u \in A^{-1}C\}$ are convex and polyhedral. Since convex polyhedral sets are closed, it follows that $(z, \beta) \in \{(B^{1/2}u, d^T u) \,|\, u \in A^{-1}C\}$. Thus there is some $\bar{x} \in A^{-1}C$ with $(z, \beta) = (B^{1/2}\bar{x}, d^T\bar{x})$ and $\frac{1}{2}\bar{x}^T B \bar{x} + d^T \bar{x} = \alpha$ follows. This shows that $\bar{x}$ is a global minimizer for (32). □

In what follows we assume that we have at hand an algorithm for solving (32), which either computes a global solution $\bar{x}$ or a descent direction w fulfilling (33). Such an algorithm is e.g. the active set method as described in [8], where we have to rewrite the constraints equivalently in the form $\langle A^T a_i, x\rangle \leq b_i$, $i = 1, \ldots, p$ using the representation of C as the intersection of finitely many half-spaces, $C = \{c \,|\, \langle a_i, c\rangle \leq b_i, \ i = 1, \ldots, p\}$.

Consider now the following algorithm.

Algorithm 1 (Basic algorithm for QPDC)**.**
   Input: starting point x^1 feasible for the QPDC (31).
   1.) Set the iteration counter $k := 1$.
   2.) Select $(\nu^{k,1}, \ldots, \nu^{k,K}) \in \mathcal{Q}(x^k)$ and consider for $l = 1, \ldots, K$ the quadratic programs

$$(QP^{k,l}) \qquad \min q(x) \quad \text{subject to} \quad Ax \in D(\nu^{k,l}).$$

   If one of these programs is unbounded below, stop the algorithm and return the current iterate x^k together with $\bar{\nu} := \nu^{k,l}$ and the descent direction w fulfilling (33). Otherwise let $x^{k,l}$, $l = 1, \ldots, K$ denote the global solutions of $(QP^{k,l})$.
   3.) If $q(x^k) = q(x^{k,l})$, $l = 1, \ldots, K$, stop the algorithm and return x^k together with $\bar{\nu} := \nu^{k,1}$.
   4.) Choose $l \in \{1, \ldots, K\}$ with $q(x^{k,l}) < q(x^k)$, set $x^{k+1} = x^{k,l}$, increase the iteration counter $k := k + 1$ and go to step 2.)

Note that the iterate x^k is feasible for every quadratic subproblem $(QP^{k,l})$. Further note that the number K will also depend on x^k.

Theorem 5. *Algorithm 1 terminates after a finite number of iterations either with some feasible point and some descent direction w indicating that QPDC is unbounded below or with some $\mathcal{Q}$-stationary solution.*

Proof. If Algorithm 1 terminates in step 2.) the output is a feasible point together with some descent direction showing that QPDC is unbounded below. If the algorithm does not terminate in step 2.) the computed sequence of function values $q(x^k)$ is strictly decreasing. Moreover, denoting $\nu^k := \nu^{k-1,l}$ where l is the index chosen in step 4., we see that for each $k \geq 2$ the point x^k is global minimizer of the problem

$$\min q(x) \quad \text{subject to} \quad Ax \in D(\nu^k).$$

This shows that all the vectors ν^k must be pairwise different and since there is only a finite number of possible choices for ν^k, the algorithm must stop in step 3.). We will now show that

the final iterate x^k is $\mathcal{Q}$-stationary with respect to $(\nu^{k,1}, \ldots, \nu^{k,K})$. Since for each $l = 1, \ldots, K$ the point x^k is a global minimizer of the subproblem $(Q^{k,l})$, it also satisfies the first order optimality condition

$$\langle \nabla q(x^k), u \rangle \geq 0 \quad \text{for every } u \in \mathbb{R}^n \text{ satisfying } Au \in T_{D(\nu^{k,l})}(Ax^k).$$

This shows $\mathcal{Q}$-stationarity of x^k and the theorem is proved. $\qquad\square$

6 On verifying $\mathcal{Q}_M$-stationarity for MPDC

The following theorem is crucial for the verification of M-stationarity.

Theorem 6. *(i) Let $\bar{x}$ be feasible for the general program (7). If there exists a B-stationary solution of the program*

$$\min_{(u,v) \in \mathbb{R}^n \times \mathbb{R}^m} \langle \nabla f(\bar{x}), u \rangle + \frac{1}{2} \|v\|^2 \text{ subject to } \nabla F(\bar{x})u + v \in T_D(F(\bar{x})), \qquad (35)$$

then $\bar{x}$ is M-stationary.

(ii) Let $\bar{x}$ be B-stationary for the MPDC (1) and assume that GGCQ holds at $\bar{x}$. Then the program (35) has a global solution.

Proof. (i) Let $(\bar{u}, \bar{v})$ denote a B-stationary solution, i.e. $-(\nabla f(\bar{x}), \bar{v}) \in \widehat{N}_\Gamma(\bar{u}, \bar{v})$, where $\Gamma = (\nabla F(\bar{x}) \mathrel{\vdots} I)^{-1} T_D(F(\bar{x}))$. Since the matrix $(\nabla F(\bar{x}) \mathrel{\vdots} I)$ obviously has full rank, we have $\widehat{N}_\Gamma(\bar{u}, \bar{v}) = (\nabla F(\bar{x}) \mathrel{\vdots} I)^T \widehat{N}_{T_D(F(\bar{x}))}(\nabla F(\bar{x})\bar{u} + \bar{v})$ by [23, Exercise 6.7]. Thus there exists a multiplier $\lambda \in \widehat{N}_{T_D(F(\bar{x}))}(\nabla F(\bar{x})\bar{u} + \bar{v})$ such that $-\nabla f(\bar{x}) = \nabla F(\bar{x})^T \lambda$ and $-\bar{v} = \lambda$. Using [23, Proposition 6.27] we have $\widehat{N}_{T_D(F(\bar{x}))}(\nabla F(\bar{x})\bar{u} + \bar{v}) \subset N_{T_D(F(\bar{x}))}(\nabla F(\bar{x})\bar{u} + \bar{v}) \subset N_{T_D(F(\bar{x}))}(0) \subset N_D(F(\bar{x}))$ establishing M-stationarity of $\bar{x}$.

(ii) Consider for arbitrarily fixed $\nu \in \mathcal{J}(F(\bar{x}))$ the convex quadratic program

$$\min_{(u,v) \in \mathbb{R}^n \times \mathbb{R}^m} \langle \nabla f(\bar{x}), u \rangle + \frac{1}{2} \|v\|^2 \text{ subject to } \nabla F(\bar{x})u + v \in T_{D(\nu)}(F(\bar{x})). \qquad (36)$$

Assuming that this quadratic program does not have a solution, by Lemma 4 we could find a direction (w_u, w_v) satisfying

$$\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} w_u \\ w_v \end{pmatrix} = 0, \ \nabla F(\bar{x})w_u + w_v \in 0^+ T_{D(\nu)}(F(\bar{x})), \ \langle \nabla f(\bar{x}), w_u \rangle + \langle 0, w_v \rangle < 0.$$

This implies $w_v = 0, \nabla F(\bar{x})w_u \in 0^+ T_{D(\nu)}(F(\bar{x})) = T_{D(\nu)}(F(\bar{x})) \subset T_D(F(\bar{x}))$ and $\langle \nabla f(\bar{x}), w_u \rangle < 0$ and thus, together with GGCQ, $-\nabla f(\bar{x}) \notin (T_\Omega^{\mathrm{lin}}(\bar{x}))^\circ = \widehat{N}_D(F(\bar{x}))$ contradicting our assumption that $\bar{x}$ is B-stationary for (1). Hence the quadratic program (36) must possess some global solution (u_ν, v_ν). By choosing $\bar{\nu} \in \mathcal{J}(F(\bar{x}))$ such that $\langle \nabla f(\bar{x}), u_{\bar{\nu}} \rangle = \min\{\langle \nabla f(\bar{x}), u_\nu \rangle \mid \nu \in \mathcal{J}(F(\bar{x}))\}$ it follows from (27) that $(u_{\bar{\nu}}, v_{\bar{\nu}})$ is a global solution of (35). $\qquad\square$

We now want to apply Algorithm 1 to the problem (35). Note that the point $(0,0)$ is feasible for (35) and therefore we can start Algorithm 1 with $(u^1, v^1) = (0,0)$.

Corollary 4. *Let $\bar{x}$ be feasible for the MPDC (1) and apply Algorithm 1 to the QPDC (35). If the algorithm returns an iterate together with some descent direction indicating that (35) is unbounded below and if GGCQ is fulfilled at $\bar{x}$, then $\bar{x}$ is not B-stationary. On the other hand, if the algorithm returns a $\mathcal{Q}$-stationary solution, then $\bar{x}$ is M-stationary.*

Proof. Observe that in case when Algorithm 1 returns a $\mathcal{Q}$-stationary solution, by Theorem 4(ii) this solution is B-stationary because the Jocobian of the constraints $(\nabla F(\bar{x}) \vdots I)$ obviously has full rank. Now the statement follows from Theorem 6. $\qquad \square$

We now want to analyze how the output of Algorithm 1 can be further utilized. Recalling that $T_D(F(\bar{x}))$ has the disjunctive structure

$$T_D(F(\bar{x})) = \prod_{i=1}^{m_D} \Big(\bigcup_{j \in \mathcal{J}_i(F(\bar{x}))} T_{D_i^j}(F_i(\bar{x})) \Big),$$

we define for $y = (y_1, \ldots, y_{m_D}) \in T_D(F(\bar{x}))$ the index sets

$$\mathcal{A}_i^{T_D}(y) := \{j \in \mathcal{J}_i(F(\bar{x})) \mid y_i \in T_{D_i^j}(F_i(\bar{x}))\}, \ i = 1, \ldots, m_D.$$

Further we choose for each $i = 1, \ldots, m_D$ some index set $\mathcal{J}_i^{T_D}(y) \subset \mathcal{A}_i^{T_D}(y)$ such that

$$T_{T_{D_i}(F_i(\bar{x}))}(y_i) = \bigcup_{j \in \mathcal{J}_i^{T_D}(y)} T_{T_{D_i^j}(F_i(\bar{x}))}(y_i) \tag{37}$$

and set

$$\mathcal{J}^{T_D}(y) := \prod_{i=1}^{m_D} \mathcal{J}_i^{T_D}(y).$$

Note that we always have

$$\mathcal{J}^{T_D}(y) \subset \mathcal{J}(F(\bar{x})).$$

In order to verify $\mathcal{Q}$-stationarity for the problem (35) at some feasible point (u, v), we have to consider the set $\mathcal{Q}^{T_D}(u, v)$ consisting of all $(\nu^1, \ldots, \nu^K)$ with $\nu^l \in \mathcal{J}^{T_D}(\nabla F(\bar{x})u + v)$, $l = 1, \ldots, K$ such that

$$\{\nu_i^1, \ldots, \nu_i^K\} = \mathcal{J}_i^{T_D}(\nabla F(\bar{x})u + v), \ i = 1, \ldots, m_D.$$

At the k-th iterate (u^k, v^k) we have to choose $(\nu^{k,1}, \ldots, \nu^{k,K}) \in \mathcal{Q}^{T_D}(u^k, v^k)$ and then for each $l = 1, \ldots, K$ we must analyze the convex quadratic program

$$(QP^{k,l}) \qquad \min_{u,v} \langle \nabla f(\bar{x}), u \rangle + \frac{1}{2}\|v\|^2 \quad \text{subject to} \quad \nabla F(\bar{x})u + v \in T_{D(\nu^{k,l})}(F(\bar{x})).$$

If for some $\bar{l} \in \{1, \ldots, K\}$ this quadratic program is unbounded below then Algorithm 1 returns the index $\bar{\nu} := \nu^{k,\bar{l}}$ together with a descent direction (w_u, w_v) fulfilling, as argued in the proof of Theorem 6(ii),

$$w_v = 0, \ \nabla F(\bar{x})w_u \in 0^+ T_{D(\bar{\nu})}(F(\bar{x})) = T_{D(\bar{\nu})}(F(\bar{x})), \langle \nabla f(\bar{x}), w_u \rangle < 0.$$

Therefore w_u constitutes a feasible descent direction, provided GACQ holds at $\bar{x}$, i.e., for every $\alpha > 0$ sufficiently small the projection of $\bar{x} + \alpha w_u$ on the feasible set $F^{-1}(D)$ yields a

point with a smaller objective function value than $\bar{x}$. If GACQ also holds for the constraint $F(x) \in D(\bar{\nu})$ at $\bar{x}$, then we can also project the point $\bar{x} + \alpha w_u$ on $F^{-1}(D(\bar{\nu}))$ in order to reduce the objective function.

Now assume that the final iterate (u^k, v^k) of Algorithm 1 is $\mathcal{Q}$-stationary for (35) and consequently $\bar{x}$ is M-stationary for the MPDC (1). Setting $\lambda := -v^k$, the first order optimality conditions for the quadratic programs $(QP^{k,l})$ result in

$$
\begin{aligned}
-\nabla f(\bar{x}) &= \nabla F(\bar{x})^T \lambda, \\
\lambda &\in \bigcap_{l=1}^K N_{T_{D(\nu^{k,l})}(F(\bar{x}))}(\nabla F(\bar{x})u^k + v^k) = \widehat{N}_{T_D(F(\bar{x}))}(\nabla F(\bar{x})u^k + v^k) \subset N_D(F(\bar{x})).
\end{aligned}
$$

From this we conclude $-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T(Q_1^\circ \cap N_D(F(\bar{x})))$ with $Q_1 := T_{D(\bar{\nu})}(F(\bar{x})) \subset T_{T_{D(\bar{\nu})}(F(\bar{x}))}(\nabla F(\bar{x})u^k + v^k)$ where $\bar{\nu} = \nu^{k,1}$ is the index vector returned from Algorithm 1. Now choosing $\nu^2, \ldots, \nu^K$ such that $(\bar{\nu}, \nu^2, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$ we can simply check by testing $-\nabla f(\bar{x}) \in N_{D(\nu^l)}(F(\bar{x}))$, $l = 2, \ldots, K$, whether $\bar{x}$ is $\mathcal{Q}_M$ stationary or $\bar{x}$ is not B-stationary.

Further we have the following corollary.

Corollary 5. *Let $\bar{x}$ be B-stationary for the MPDC (1) and assume that GGCQ is fulfilled at $\bar{x}$. Let $\bar{\nu}$ be the index vector returned by Algorithm 1 applied to (35). Then $\bar{\nu} \in \mathcal{J}(F(\bar{x}))$ and for every $\nu^2, \ldots, \nu^K$ with $(\bar{\nu}, \nu^2, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$ the point $\bar{x}$ is $\mathcal{Q}_M$-stationary with respect to $(\bar{\nu}, \nu^2, \ldots, \nu^K)$.*

7 Numerical aspects

In practice the point $\bar{x}$ which should be checked for M-stationarity and $\mathcal{Q}_M$-stationarity, respectively, often is not known exactly. E.g., $\bar{x}$ can be the limit point of a sequence generated by some numerical method for solving MPDC. Hence let us assume that we are given some point $\tilde{x}$ close to $\bar{x}$ and we want to state some rules when we can consider $\tilde{x}$ as approximately M-stationary or $\mathcal{Q}_M$-stationary. Let us assume that the convex polyhedral sets D_i^j have the representation

$$
D_i^j = \{y \mid \langle a_l^{i,j}, y \rangle \leq b_l^{i,j}, \ l = 1, \ldots, p^{i,j}\}, \ i = 1, \ldots, m_D, \ j = 1, \ldots, K_i,
$$

where without loss of generality we assume $\|a_l^{i,j}\| = 1$.

We use here the following approach.

Algorithm 2.
    Input: *A point $\tilde{x}$ and small positive parameters ϵ, σ, η.*
    1.) Calculate the index sets

$$
\tilde{\mathcal{A}}_i(\tilde{x}, \epsilon) := \{j \in \{1, \ldots, K_i\} \mid \mathrm{d}(F_i(\tilde{x}), D_i^j) \leq \epsilon\}, \ i = 1, \ldots, m_D
$$

$$
\tilde{\mathcal{I}}_i^j(\tilde{x}, \epsilon) := \{l \in \{1, \ldots, p^{i,j}\} \mid \langle a_l^{i,j}, F_i(\tilde{x}) \rangle \geq b_l^{i,j} - \epsilon\}, \ i = 1, \ldots, m_D, \ j \in \tilde{\mathcal{A}}_i(\tilde{x}, \epsilon)
$$

and the convex polyhedral cones

$$
T_i^j(\tilde{x}, \epsilon) = \{v \mid \langle a_l^{i,j}, v \rangle \leq 0, l \in \tilde{\mathcal{I}}_i^j(\tilde{x}, \epsilon)\}, \ i = 1, \ldots, m_D, \ j \in \tilde{A}_i(\tilde{x}, \epsilon).
$$

Assume that $\tilde{\mathcal{A}}_i(\tilde{x}, \epsilon) \neq \emptyset$, $i = 1, \ldots, m_D$.

2.) Consider

$$QPDC(\tilde{x}, \epsilon, \sigma) \qquad \min_{u,v} \quad \langle \nabla f(\tilde{x}), u \rangle + \frac{\sigma}{2}\|u\|^2 + \frac{1}{2}\|v\|^2$$

$$subject\ to \quad \nabla F(\tilde{x})u + v \in \prod_{i=1}^{m_D} \Big(\bigcup_{j \in \tilde{\mathcal{A}}_i(\tilde{x},\epsilon)} T_i^j(\tilde{x},\epsilon) \Big).$$

Let $(\tilde{u}, \tilde{v})$ and $\bar{\nu}$ denote the output of Algorithm 1 applied to $QPDC(\tilde{x}, \epsilon, \sigma)$.
3. If $\sigma\|\tilde{u}\| > \eta$ consider the nonlinear programming problem

$$\min f(x) \qquad subject\ to \qquad F(x) \in D(\bar{\nu})$$

in order to improve $\tilde{x}$.
4.) Otherwise consider $\tilde{x}$ as approximately M-stationary and compute $\nu^2, \ldots, \nu^K \in \prod_{i=1}^{m_D} \tilde{\mathcal{A}}_i(\tilde{x}, \epsilon)$ such that

$$T_i^{\bar{\nu}_i}(\tilde{x},\epsilon) \cup \bigcup_{l=2}^{K} T_i^{\nu_i^l}(\tilde{x},\epsilon) = \bigcup_{j \in \tilde{\mathcal{A}}_i(\tilde{x},\epsilon)} T_i^j(\tilde{x},\epsilon), \ i = 1, \ldots, m_D.$$

If

$$\min \left\{ \nabla f(\tilde{x})u \mid \begin{array}{l} \nabla F(\tilde{x})u \in \prod_{i=1}^{m_D} T_i^{\nu_i^l}(\tilde{x},\epsilon), \\ -1 \le u_i \le 1, i = 1, \ldots, n \end{array} \right\} \ge -\eta, \ l = 2, \ldots, K, \qquad (38)$$

accept $\tilde{x}$ as approximately $\mathcal{Q}_M$-stationary. Otherwise consider the nonlinear programming problem

$$\min f(x) \qquad subject\ to \qquad F(x) \in D(\nu^{\bar{l}})$$

in order to improve $\tilde{x}$, where $\bar{l} \in \{2, \ldots, K\}$ denotes some index violating (38).

In the first step of Algorithm 2 we want to estimate the tangent cone $T_D(F(\bar{x}))$. In fact, to calculate $T_D(F(\bar{x}))$ we need not to know the point $F(\bar{x})$, we only need the index sets of constraints active at $\bar{x}$ and these index sets are approximated by ϵ-active constraints. Note that whenever $\tilde{\mathcal{A}}_i(\tilde{x}, \epsilon) = \tilde{\mathcal{A}}_i(\bar{x}, 0) = \mathcal{A}_i(F(\bar{x}))$ and $\tilde{\mathcal{I}}_i^j(\tilde{x}, \epsilon) = \tilde{\mathcal{I}}_i^j(\bar{x}, 0)$, $i = 1, \ldots, m_D$, $j \in \mathcal{A}_i(F(\bar{x}))$ this approach yields the exact tangent cones $T_{D_i^j}(F(\bar{x})) = T_i^j(\tilde{x}, \epsilon)$ for all $i = 1, \ldots, m_D$, $j \in \mathcal{A}_i(F(\bar{x}))$. To be consistent with the notation of Section 4 we make the convention that in this case the index vector $\bar{\nu}$ computed in step 2.) belongs to $\mathcal{J}(\bar{x})$ and also, whenever we determine $\nu^2, \ldots \nu^K$ is step 4.), we have $(\bar{\nu}, \nu^2, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$. The regularization term $\frac{\sigma}{2}\|u\|^2$ in $QPDC(\tilde{x}, \epsilon, \sigma)$ forces the objective to be strictly convex and therefore Algorithm 1 will always terminate with a $\mathcal{Q}$-stationary solution. Further note that the verification of (38) requires the solution of $K - 1$ linear optimization problems.

The following theorem justifies Algorithm 2. Im the sequel we denote by $\mathcal{M}(\bar{x})$ ($\mathcal{M}_{sub}(\bar{x})$) the set of all $\nu \in \mathcal{J}(\bar{x})$ such that the mapping $F(\cdot) - D(\nu)$ is metrically regular near $(\bar{x}, 0)$ (metrically subregular at $(\bar{x}, 0)$).

Theorem 7. *Let $\bar{x}$ be feasible for the MPDC (1) and assume that ∇f and ∇F are Lipschitz near $\bar{x}$. Consider sequences $x_t \to \bar{x}$, $\epsilon_t \downarrow 0$, $\sigma_t \downarrow 0$ and $\eta_t \downarrow 0$ with*

$$\lim_{t \to \infty} \frac{\|x_t - \bar{x}\|}{\epsilon_t} = \lim_{t \to \infty} \frac{\frac{\sigma_t}{\eta_t} + \|x_t - \bar{x}\|}{\eta_t} = 0$$

and let $(\tilde{u}_t, \tilde{v}_t)$, $\bar{\nu}^t$ and eventually $\nu^{t,2} \ldots, \nu^{t,K_t}$ and $\bar{l}_t$ denote the output of Algorithm 2 with input data $(x_t, \epsilon_t, \sigma_t, \eta_t)$.

(i) For all t sufficiently large and for all $i \in \{1, \ldots, m_D\}$ we have

$$\tilde{\mathcal{A}}_i(x_t, \epsilon_t) = \mathcal{A}_i(F(\bar{x})), \ \tilde{\mathcal{I}}_i^j(x_t, \epsilon_t) = \tilde{\mathcal{I}}_i^j(\bar{x}, 0), \ j \in \mathcal{A}_i(F(\bar{x})). \tag{39}$$

(ii) Assume that the mapping $x \rightrightarrows F(x) - D$ is metrically regular near $(\bar{x}, 0)$.

(a) If $\bar{x}$ is B-stationary then for all t sufficiently large the point x_t is accepted as approximately M-stationary and approximately $\mathcal{Q}_M$-stationary.

(b) If for infinitely many t the point x_t is accepted as approximately M-stationary then $\bar{x}$ is M-stationary.

(c) If for infinitely many t the point x_t is accepted as approximately $\mathcal{Q}_M$-stationary and $\{\bar{\nu}^t, \nu^{t,2}, \ldots, \nu^{t,K_t}\} \subset \mathcal{M}(\bar{x})$ then $\bar{x}$ is $\mathcal{Q}_M$-stationary.

(d) For every t sufficiently large such that the point x_t is not accepted as approximately M-stationary and $\bar{\nu}^t \in \mathcal{M}_{sub}(\bar{x})$ we have $\min\{f(x) \,|\, F(x) \in D(\bar{\nu}^t)\} < f(\bar{x})$.

(e) For every t sufficiently large such that the point x_t is not accepted as approximately $\mathcal{Q}_M$-stationary and $\nu^{t,\bar{l}_t} \in \mathcal{M}_{sub}(\bar{x})$ we have $\min\{f(x) \,|\, F(x) \in D(\nu^{t,\bar{l}_t})\} < f(\bar{x})$.

Proof. (i) Let $R > 0$ be chosen such that f, F and their derivatives are Lipschitz on $\mathcal{B}(\bar{x}, R)$ with constant L. It is easy to see that we can choose $\epsilon > 0$ such that for all $i \in \{1, \ldots, m_D\}$ we have $\tilde{\mathcal{A}}_i(\bar{x}, \epsilon) = \tilde{\mathcal{A}}_i(\bar{x}, 0) = \mathcal{A}_i(F(\bar{x}))$ and such that for every $j \in \mathcal{A}_i(F(\bar{x}))$ we have $\tilde{\mathcal{I}}_i^j(\bar{x}, \epsilon) = \tilde{\mathcal{I}}_i^j(\bar{x}, 0)$. Consider t with $\|x_t - \bar{x}\| < R$, $L\|x_t - \bar{x}\| < \epsilon_t < \epsilon/2$ and fix $i \in \{1, \ldots, m_D\}$. For every $j \in \mathcal{A}_i(F(\bar{x}))$ we have

$$\mathrm{d}(F_i(x_t), D_i^j) \leq \|F_i(x_t) - F_i(\bar{x})\| \leq L\|x_t - \bar{x}\| < \epsilon_t,$$

whereas for $j \notin \mathcal{A}_i(F(\bar{x}))$ we have

$$\mathrm{d}(F_i(x_t), D_i^j) \geq \mathrm{d}(F_i(\bar{x}), D_i^j) - \|F_i(x_t) - F_i(\bar{x})\| \geq \epsilon - L\|x_t - \bar{x}\| > \epsilon_t$$

showing $\tilde{\mathcal{A}}_i(x_t, \epsilon_t) = \mathcal{A}_i(F(\bar{x}))$. Now fix $j \in \mathcal{A}_i(F(\bar{x}))$ and let $l \in \tilde{\mathcal{I}}_i^j(\bar{x}, 0)$, i.e. $\langle a_l^{i,j}, F_i(\bar{x}) \rangle = b_l^{i,j}$. By taking into account $\|a_l^{i,j}\| = 1$ we obtain

$$\langle a_l^{i,j}, F_i(x_t) \rangle \geq b_l^{i,j} - \|F_i(x_t) - F_i(\bar{x})\| > b_l^{i,j} - \epsilon_t$$

implying $l \in \tilde{\mathcal{I}}_i^j(x_t, \epsilon_t)$, whereas for $l \notin \tilde{\mathcal{I}}_i^j(\bar{x}, 0) = \tilde{\mathcal{I}}_i^j(\bar{x}, \epsilon)$ we have

$$\langle a_l^{i,j}, F_i(x_t) \rangle \leq \langle a_l^{i,j}, F_i(\bar{x}) \rangle + \|F_i(x_t) - F_i(\bar{x})\| < b_l^{i,j} - \epsilon + \epsilon_t < b_l^{i,j} - \epsilon_t$$

showing $l \notin \tilde{\mathcal{I}}_i^j(x_t, \epsilon_t)$. Hence $\tilde{\mathcal{I}}_i^j(x_t, \epsilon_t) = \tilde{I}(\bar{x}, 0)$. Because of our assumptions we have $\|x_t - \bar{x}\| < R$ and $L\|x_t - \bar{x}\| < \epsilon_t < \epsilon/2$ for all t sufficiently large and this proves (39).

(ii) In view of Proposition 2 we can choose κ large enough such that the mappings $F(\cdot) - D$, $u \rightrightarrows \nabla F(\bar{x})u - T_D(F(\bar{x}))$ and $F(\cdot) - D(\nu)$, $u \rightrightarrows \nabla F(\bar{x})u - T_{D(\nu)}(F(\bar{x}))$, $\nu \in \mathcal{M}(\bar{x})$ are metrically regular near $(\bar{x}, 0)$ with modulus κ. By eventually shrinking R we can assume that for every $x \in \mathcal{B}(\bar{x}, R)$ the mappings $u \rightrightarrows \nabla F(x)u - T_D(F(\bar{x}))$, $u \rightrightarrows \nabla F(x)u - T_{D(\nu)}(F(\bar{x}))$, $\nu \in \mathcal{M}(\bar{x})$ are metrically regular near $(0, 0)$ with modulus $\kappa + 1$.

20

Without loss of generality we can assume that $x_t \in \mathcal{B}(\bar{x}, R)$ and (39) holds for all t implying that $T_{D_i^j}(F(\bar{x})) = T_i^j(\tilde{x}, \epsilon_t)$ holds for all $i = 1, \ldots, m_D$, $j \in \mathcal{A}_i(F(\bar{x}))$. In fact then the problem $QP\tilde{D}C(x_t, \epsilon_t, \sigma_t)$ is the same as

$$\min_{u,v} \langle \nabla f(x_t), u \rangle + \frac{\sigma_t}{2}\|u\|^2 + \frac{1}{2}\|v\|^2 \quad \text{subject to} \quad \nabla F(x_t)u + v \in T_D(F(\bar{x})).$$

The point $(\tilde{u}_t, \tilde{v}_t)$ is $\mathcal{Q}$-stationary for this program and thus also S-stationary by Theorem 4(ii) and the full rank property of the matrix $(\nabla F(x_t) \vdots I)$. Hence there is a multiplier $\lambda_t \in \widehat{N}_{T_D(F(\bar{x}))}(\nabla F(x_t)\tilde{u}_t + \tilde{v}_t) \subset N_{T_D(F(\bar{x}))}(0)$ fulfilling $\tilde{v}_t + \lambda_t = 0$, $\nabla f(x_t) + \sigma_t \tilde{u}_t + \nabla F(x_t)^T \lambda_t = 0$ and we conclude

$$\|\tilde{v}_t\| = \|\lambda_t\| \le (\kappa + 1)\|\nabla f(x_t) + \sigma_t \tilde{u}_t\| \tag{40}$$

from (18).

By $\mathcal{Q}$-stationarity of $(\tilde{u}_t, \tilde{v}_t)$ we know that $(\tilde{u}_t, \tilde{v}_t)$ is the unique solution of the strictly convex quadratic program

$$\min \langle \nabla f(x_t), u \rangle + \frac{\sigma_t}{2}\|u\|^2 + \frac{1}{2}\|v\|^2 \text{ subject to } \nabla F(x_t)u + v \in T_{D(\bar{\nu}^t)}(F(\bar{x})). \tag{41}$$

For every $\alpha \ge 0$ the point $\alpha(\tilde{u}_t, \tilde{v}_t)$ is feasible for this quadratic program and thus $\alpha = 1$ is solution of

$$\min_{\alpha \ge 0} \alpha \langle \nabla f(x_t), \tilde{u}_t \rangle + \alpha^2 \left(\frac{\sigma_t}{2}\|\tilde{u}_t\|^2 + \frac{1}{2}\|\tilde{v}_t\|^2 \right)$$

implying

$$-\langle \nabla f(x_t), \tilde{u}_t \rangle = \sigma_t \|\tilde{u}_t\|^2 + \|\tilde{v}_t\|^2.$$

Hence

$$\sigma_t \|\tilde{u}_t\| \le -\langle \nabla f(x_t), \frac{\tilde{u}_t}{\|\tilde{u}_t\|} \rangle \le \|\nabla f(x_t)\| \tag{42}$$

and from (40) we obtain

$$\|\tilde{v}_t\| = \|\lambda_t\| \le 2(\kappa + 1)\|\nabla f(x_t)\|. \tag{43}$$

(a) Assume on the contrary that $\bar{x}$ is B-stationary but for infinitely many t the point x_t is not accepted as approximately M-stationary and hence $\|\tilde{u}_t\| \ge \eta_t/\sigma_t$. This implies

$$
\begin{aligned}
\mathrm{d}(\nabla F(\bar{x})\frac{\tilde{u}_t}{\|\tilde{u}_t\|}, T_D(F(\bar{x}))) &\le \mathrm{d}(\nabla F(x_t)\frac{\tilde{u}_t}{\|\tilde{u}_t\|}, T_D(F(\bar{x}))) + L\|x_t - \bar{x}\| \le \frac{\|\tilde{v}_t\|}{\|\tilde{u}_t\|} + L\|x_t - \bar{x}\| \\
&\le 2(\kappa + 1)\|f(x_t)\|\frac{\sigma_t}{\eta_t} + L\|x_t - \bar{x}\|
\end{aligned}
$$

and by the metric regularity of $u \rightrightarrows \nabla F(\bar{x})u - T_D(F(\bar{x}))$ near $(0,0)$ we can find $\hat{u}_t \in \nabla F(\bar{x})^{-1}T_D(F(\bar{x}))$ with

$$\|\hat{u}_t - \frac{\tilde{u}_t}{\|\tilde{u}_t\|}\| \le \kappa \left(2(\kappa + 1)\|f(x_t)\|\frac{\sigma_t}{\eta_t} + L\|x_t - \bar{x}\| \right).$$

Our choice of the parameters σ_t, η_t together with (42) ensures that for t sufficiently large we have

$$
\begin{aligned}
\langle \nabla f(\bar{x}), \hat{u}_t \rangle &\leq \langle \nabla f(\bar{x}), \frac{\tilde{u}_t}{\|\tilde{u}_t\|} \rangle + \|\nabla f(\bar{x})\| \|\hat{u}_t - \frac{\tilde{u}_t}{\|\tilde{u}_t\|}\| \\
&\leq \langle \nabla f(x_t), \frac{\tilde{u}_t}{\|\tilde{u}_t\|} \rangle + L\|x_t - \bar{x}\| + \|\nabla f(\bar{x})\| \|\hat{u}_t - \frac{\tilde{u}_t}{\|\tilde{u}_t\|}\| \\
&\leq -\sigma_t \|\tilde{u}_t\| + L\|x_t - \bar{x}\| + \|\nabla f(\bar{x})\| \|\hat{u}_t - \frac{\tilde{u}_t}{\|\tilde{u}_t\|}\| \\
&\leq -\eta_t + L\|x_t - \bar{x}\| + \|\nabla f(\bar{x})\| \kappa \left(2(\kappa+1)\|f(x_t)\| \frac{\sigma_t}{\eta_t} + L\|x_t - \bar{x}\| \right) < 0
\end{aligned}
$$

which contradicts B-stationarity of $\bar{x}$. Hence for all t sufficiently large the point x_t must be accepted as approximately M-stationary.

To prove the statement that x_t is also accepted as approximately $\mathcal{Q}_M$-stationary for all t sufficiently large we can proceed in a similar way. Assume on the contrary that $\bar{x}$ is B-stationary but for infinitely many t the point x_t is not accepted as approximately $\mathcal{Q}_M$-stationary. For those t let w_t denote some element fulfilling $\nabla F(x_t) w_t \in T_{D(\nu^{t,\bar{l}_t})} \subset T_D(F(\bar{x}))$, $\|w_t\|_\infty \leq 1$ and $\langle \nabla f(x_t), w_t \rangle \leq -\eta_t$. Then, similar as before we can find $\hat{w}_t \in \nabla F(\bar{x})^{-1} T_D(F(\bar{x}))$ such that

$$
\|\hat{w}_t - w_t\| \leq \kappa \|\nabla F(\bar{x}) - \nabla F(x_t)\| \|w_t\| \leq \kappa L \sqrt{n} \|x_t - \bar{x}\|
$$

and for large t we obtain

$$
\begin{aligned}
\langle \nabla f(\bar{x}), \hat{w}_t \rangle &\leq \langle \nabla f(x_t), w_t \rangle + \|\nabla f(\bar{x}) - \nabla f(x_t)\| \|w_t\| + \|\nabla f(\bar{x})\| \|\hat{w}_t - w_t\| \\
&\leq -\eta_t + L\sqrt{n}(1 + \kappa \|\nabla f(\bar{x})\|) \|x_t - \bar{x}\| < 0
\end{aligned}
$$

contradicting B-stationarity of $\bar{x}$.

(b) By passing to a subsequence we can assume that for all t the point x_t is accepted as approximately M-stationary and hence $\sigma_t \|u_t\| \leq \eta_t \to 0$. By (43) we have that the sequence $\lambda_t \in N_{T_D(F(\bar{x}))}(0)$ is uniformly bounded and by passing to a subsequence once more we can assume that it converges to some $\bar{\lambda} \in N_{T_D(F(\bar{x}))}(0)$. By [23, Proposition 6.27] we have $\bar{\lambda} \in N_D(F(\bar{x}))$ and together with

$$
0 = \lim_{t \to \infty} \left(\nabla f(x_t) + \nabla F(x_t)^T \lambda_t \right) = \nabla f(\bar{x}) + \nabla F(\bar{x})^T \bar{\lambda}
$$

M-stationarity of $\bar{x}$ is established.

(c) By passing to a subsequence we can assume that for all t the point x_t is accepted as approximately $\mathcal{Q}_M$-stationary and $\{\bar{\nu}^t, \nu^{t,2}, \ldots, \nu^{t,K_t}\} \subset \mathcal{M}(\bar{x})$. Hence for all t the point x_t is also accepted as M-stationary and by passing to a subsequence and arguing as in (b) we can assume that λ_t converges to some $\bar{\lambda} \in N_D(F(\bar{x}))$ fulfilling $\nabla f(\bar{x}) + \nabla F(\bar{x})^T \bar{\lambda} = 0$. Since the set $\mathcal{M}(\bar{x})$ is finite, by passing to a subsequence once more we can assume that there is a number K and elements $\bar{\nu}, \nu^2, \ldots, \nu^K$ such that $K_t = K$, $\bar{\nu}^t = \bar{\nu}$ and $\nu^{t,l} = \nu^l$, $l = 2, \ldots, K$ holds for all t. Since we assume that (39) holds we have $(\bar{\nu}, \nu^2, \ldots, \nu^K) \in \mathcal{Q}(\bar{x})$ and we will now show that $\bar{x}$ is $\mathcal{Q}_M$-stationary with respect to $(\bar{\nu}, \nu^2, \ldots, \nu^K)$. Since $(\tilde{u}_t, \tilde{v}_t)$ also solves (41), it follows that $\lambda_t = -v_t \in N_{T_{D(\bar{\nu})}(F(x_t))}(\nabla F(x_t)\tilde{u}_t + \tilde{v}_t) \subset N_{D(\bar{\nu})}(F(\bar{x}))$ and thus $\bar{\lambda} \in N_D(F(\bar{x})) \cap N_{D(\bar{\nu})}(F(\bar{x}))$ implying $-\nabla f(\bar{x}) \in \nabla F(\bar{x})^T \left(N_D(F(\bar{x})) \cap \left(T_{D(\bar{\nu})}(F(\bar{x})) \right)^\circ \right)$.

There remains to show $-\nabla f(\bar{x}) \in \left(T_{D(\nu^l)}(F(\bar{x}))\right)^\circ = N_{D(\nu^l)}(F(\bar{x}))$, $l = 2, \ldots, K$. Assume on the contrary that $-\nabla f(\bar{x}) \notin \left(T_{D(\nu^{\bar{l}})}(F(\bar{x}))\right)^\circ$ for some index $\bar{l} \in \{2, \ldots, K\}$. Then there is some $u \in \nabla F(\bar{x})^{-1} T_{D(\nu^{\bar{l}})}(F(\bar{x}))$, $\|u\|_\infty = \frac{1}{2}$ such that $\langle \nabla f(\bar{x}), u \rangle =: -\gamma < 0$ and since $\nu^{\bar{l}} \in \mathcal{M}(\bar{x})$, for each t there is some $\hat{u}_t \in \nabla F(x_t)^{-1} T_{D(\nu^{\bar{l}})}(F(\bar{x}))$ with

$$\|u - \hat{u}_t\| \le (\kappa + 1)\|\nabla F(\bar{x}) - \nabla F(x_t)\| \|u\| \le \frac{\sqrt{n}}{2}(\kappa + 1)L\|x_t - \bar{x}\|.$$

It follows that for all t sufficiently large we have $\|\hat{u}_t\|_\infty \le 1$ and

$$
\begin{aligned}
\langle \nabla f(x_t), \hat{u}_t \rangle &\le \langle \nabla f(\bar{x}), u \rangle + \|\nabla f(x_t) - \nabla f(\bar{x})\| \|\hat{u}_t\| + \|\nabla f(\bar{x})\| \|u - \hat{u}_t\| \\
&\le -\gamma + L\sqrt{n}(1 + \frac{\kappa + 1}{2})\|x_t - \bar{x}\| < -\eta_t
\end{aligned}
$$

contradicting our assumption that x_t is accepted as approximately $\mathcal{Q}_M$-stationary.

(d), (e) We assume that κ is chosen large enough such that the mappings $F(\cdot) - D(\nu)$, $\nu \in \mathcal{M}_{sub}(\bar{x})$ are metrically subregular at $(\bar{x}, 0)$ with modulus κ. Then by [10, Proposition 2.1] the mappings $u \rightrightarrows \nabla F(\bar{x})u - T_{D(\nu)}(F(\bar{x}))$, $\nu \in \mathcal{M}_{sub}(\bar{x})$ are metrically subregular at $(0,0)$ with modulus κ as well. Taking into account that $(\tilde{u}_t, \tilde{v}_t)$ solves (41), we can copy the arguments from part (a) with $T_D(F(\bar{x}))$ replaced by $T_{D(\bar{\nu}^t)}(F(\bar{x}))$ to show the existence of $\hat{u}_t \in \nabla F(\bar{x})^{-1} T_{D(\bar{\nu}^t)}(F(\bar{x}))$ with $\langle \nabla f(\bar{x}), \hat{u}_t \rangle < 0$ whenever x_t is not accepted as approximately M-stationary and t is sufficiently large. In doing so we also have to recognize that metric regularity of $u \rightrightarrows \nabla F(\bar{x})u - T_{D(\bar{\nu}^t)}(F(\bar{x}))$ can be replaced by the weaker property of metric subregularity. Since $\bar{\nu}^t \in \mathcal{M}_{sub}(\bar{x})$, $\hat{u}_t$ is a feasible descent direction and for sufficiently small $\alpha > 0$ the projection of $\bar{x} + \alpha \hat{u}_t$ on $F^{-1}(D(\bar{\nu}^t))$ yields a point with a smaller objective function value than $\bar{x}$. This proves (d). In order to show (e) we can proceed in a similar way. Using the same arguments as in part (a) we can prove the existence of a feasible direction $\hat{w}_t \in T_{D(\nu^{t,\bar{l}_t})}$ with $\langle \nabla f(\bar{x}), \hat{w}_t \rangle < 0$, whenever t is sufficiently large and x_t is not accepted as approximately $\mathcal{Q}_M$-stationary. Together with $\nu^{t,\bar{l}_t} \in \mathcal{M}_{sub}(\bar{x})$ the assertion follows. $\qquad \square$

Acknowledgements

References

[1] E. Balas, *Disjunctive programming and a hierarchy of relaxations for discrete optimization problems*, SIAM J. Algebraic Discrete Methods, 6 (1985), pp. 466–486.

[2] M. Benko, H. Gfrerer, *On estimating the regular normal cone to constraint systems and stationary conditions*, Optimization, DOI 10.1080/02331934.2016.1252915.

[3] M. Benko, H. Gfrerer, *An SQP method for mathematical programs with vanishing constraints with strong convergence properties*, 2016, submitted, preprint `http://www.numa.uni-linz.ac.at/publications/List/2016/2016-05.pdf`

[4] S. Ceria, J. Soares, *Convex programming for disjunctive convex optimization*, Math. Program. Ser. A, 86 (1999), pp. 595–614.

[5] A. L. DONTCHEV, R. T. ROCKAFELLAR, *Implicit Functions and Solution Mappings*, Springer, Heidelberg, 2014.

[6] M. L. FLEGEL, C. KANZOW, *A Fritz John approach to first order optimality conditions for mathematical programs with equilibrium constraints*, Optimization, 52 (2003), pp. 277–286.

[7] M. L. FLEGEL, C. KANZOW, J. V. OUTRATA, *Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints*, Set-Valued Anal., 15 (2007), pp. 139–162.

[8] R. FLETCHER, *Practical methods of optimization, Vol. 2: Constrained optimization*, Wiley, Chichester, 1981.

[9] M. FUKUSHIMA, J. S. PANG, *Complementarity constraint qualifications and simplified B-stationary conditions for mathematical programs with equilibrium constraints*, Comput. Optim. Appl., 13 (1999), pp. 111–136.

[10] H. GFRERER, *First order and second order characterizations of metric subregularity and calmness of constraint set mappings*, SIAM J. Optim., 21 (2011), pp. 1439–1474.

[11] H. GFRERER, D. KLATTE, *Lipschitz and Hölder stability of optimization problems and generalized equations*, Math. Program. Ser. A, 158 (2016), pp. 35–75.

[12] H. GFRERER, J. V. OUTRATA, *On computation of generalized derivatives of the normal-cone mapping and their applications*, Math. Oper. Res., 41 (2016), pp. 1535–1556.

[13] H. GFRERER, J. J. YE, *New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis*, 2016, submitted, preprint http://www.numa.uni-linz.ac.at/publications/List/2016/2016-04.pdf

[14] R. HENRION, J. V. OUTRATA, *Calmness of constraint systems with applications*, Math. Program., Ser. B, 104 (2005), pp. 437–464.

[15] T. HOHEISEL, *Mathematical programs with vanishing constraints*, PhD-thesis, Julius–Maximilians–Universität Würzburg, 2009.

[16] R. JEROSLOW, *Representability in mixed integer programming, I: Characterization results*, Discrete Appl. Math., 17 (1977), pp. 223–243.

[17] C. KANZOW, A. SCHWARTZ, *Mathematical programs with equilibrium constraints: enhanced Fritz John conditions, new constraint qualifications and improved exact penalty results*, SIAM J. Optim. 20 (2010), pp. 2730–2753.

[18] C. KANZOW, A. SCHWARTZ, *The price of inexactness: convergence properties of relaxation methods for mathematical programs with equilibrium constraints revisited*, Math. Oper. Res., 40 (2015), pp. 253–275.

[19] J. V. OUTRATA, *Optimality conditions for a class of mathematical programs with equilibrium constraints*, Math. Oper. Res., 24 (1999), pp. 627–644.

[20] J. V. OUTRATA, *A generalized mathematical program with equilibrium constraints*, SIAM J. Control Optim., 38 (2000), pp. 1623–1638.

[21] S. M. ROBINSON, *Some continuity properties of polyhedral multifunctions*, Math. Program. Studies, 14 (1981), pp. 206–214.

[22] R. T. ROCKAFELLAR, *Convex analysis*, Princeton University Press, Princeton, New jersey, 1970.

[23] R. T. ROCKAFELLAR, R. J-B. WETS, *Variational analysis*, Springer, Berlin, 1998.

[24] H. SCHEEL, S. SCHOLTES, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.

[25] J. J. YE, *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM J. Optim., 10 (2000), pp. 943–962.

[26] J. J. YE, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, J. Math. Anal. Appl., 307 (2005), pp. 350–369.

[27] J. J. YE, X. Y. YE, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Math. Oper. Res., 22 (1997), pp. 977–997.