On the linear independence constraint qualification in disjunctive programming

Patrick Mehlitz*

February 6, 2019

Mathematical programs with disjunctive constraints (MPDCs for short) cover several different problem classes from nonlinear optimization including complementarity-, vanishing-, and switching-constrained optimization problems. In this paper, we introduce an abstract but reasonable version of the prominent linear independence constraint qualification which applies to MPDCs. Afterwards, we derive first- and second-order optimality conditions for MPDCs under validity of this constraint qualification based on so-called strongly stationary points. Finally, we apply our findings to some popular classes of disjunctive programs and compare the obtained results to those ones available in the literature. Particularly, new second-order optimality conditions for mathematical programs with switching constraints are by-products of our approach.

Keywords: Constraint qualifications, Disjunctive programming, Linear independence constraint qualification, Strong stationarity, Second-order optimality conditions

MSC: 90C30, 90C33

1 Introduction

In this paper, so-called mathematical programs with disjunctive constraints (MPDCs) are studied. An MPDC is an optimization problem of the form

\[ \begin{align*}
  f(x) & \to \min \\
  F(x) & \in D
\end{align*} \]  

(MPDC)

where \( f: \mathbb{R}^n \to \mathbb{R} \) as well as \( F: \mathbb{R}^n \to \mathbb{R}^m \) are twice continuously differentiable and \( D := \bigcup_{i=1}^r D_i \) is the finite union of given polyhedral sets \( D_1, \ldots, D_r \subset \mathbb{R}^m \). We use

*Brandenburgische Technische Universität Cottbus-Senftenberg, Institute of Mathematics, 03046 Cottbus, Germany, mehlitz@b-tu.de, https://www.b-tu.de/fg-optimale-steuerung/team/dr-patrick-mehlitz, ORCID: 0000-0002-9355-850X
\[ X := \{ x \in \mathbb{R}^n \mid F(x) \in D \} \] in order to denote the feasible set of (MPDC). Clearly, choosing \( r := 1 \) and \( D := \mathbb{R}^p \times \{ 0 \} \) with \( p, q \in \mathbb{N} \) and \( m := p + q \), any standard nonlinear program is a disjunctive program, see Example 3.2. However, it is well known that the model (MPDC) covers mathematical programs with complementarity constraints (MPCCs), see Luo et al. (1996), mathematical programs with vanishing constraints (MPVCs), see Achtzer and Kanzow (2008), a certain reformulation of mathematical programs with cardinality constraints (CCMPs), see Burdakov et al. (2016), and mathematical programs with switching constraints (MPSCs), see Mehlitz (2018), as well. Note that all these problem classes which frequently arise from the mathematical modeling of real-world applications suffer from an inherent lack of regularity. That is why huge effort has been put into the derivation of problem-tailored stationarity notions and constraint qualifications. However, it is clear that any theoretical result which can be derived for the generalized model (MPDC) has a corresponding counterpart for MPCCs, MPVCs, CCMPs, and MPSCs. This observation justifies the theoretical investigation of (MPDC). First ideas on how to study disjunctive structures in nonlinear optimization are presented in Scholtes (2004). Stationarity notions and constraint qualifications for (MPDC) can be found in (Benko and Gfrerer, 2018; Flegel et al., 2007; Gfrerer, 2014). Particularly, second-order necessary and sufficient optimality conditions for disjunctive programs are derived in Gfrerer (2014) with the aid of the celebrated directional limiting calculus, see Gfrerer (2013) as well. Checking (Gfrerer, 2014, Theorems 3.3, 3.17), one can observe that in contrast to classical second-order optimality conditions, the appearing set of multipliers depends on the choice of the particular critical direction.

In this paper, we are going to state an MPDC-tailored version of the linear independence constraint qualification (LICQ) and study its inherent properties. Furthermore, we derive second-order necessary optimality conditions for (MPDC) under validity of this constraint qualification in a completely elementary way using second-order tangent sets. Thus, our approach is related to techniques which were used in Bonnans and Shapiro (2000); Christof and Wachsmuth (2018); Penot (1998); Rockafellar and Wets (1998) to derive second-order conditions for mathematical programs. On the other hand, we present a result which shows the isolatedness of strongly stationary points of (MPDC) where the problem-tailored version of LICQ and a suitable second-order sufficient condition hold. This generalizes some corresponding results for MPCCs, see Guo et al. (2013), and CCMPs, see (Bucher and Schwartz, 2018, Theorem 3.3). Afterwards, we apply our findings to several instances of disjunctive programming. In particular, new second-order optimality conditions for MPSCs will be derived.

The remaining parts of this paper are structured as follows: In Section 2, we comment on the notation used in this manuscript and introduce all the necessary tools from variational analysis which are exploited later. Furthermore, some preliminary results are provided. Section 3 is dedicated to the derivation of an MPDC-tailored version of the linear independence constraint qualification. Some consequences of the validity of this regularity condition are presented. Second-order optimality conditions for (MPDC) are the topic of interest in Section 4. First, it will be shown that a second-order necessary optimality condition holds at the local minimizers of (MPDC) where our new constraint qualification is valid. Afterwards, a second-order sufficient optimality condi-
tion for (MPDC) will be derived. Subsequently, we show that this condition together with the problem-tailored version of LICQ implies that the underlying stationary point is in a certain sense locally isolated. In Section 5, we apply the derived theory to MPCCs, MPVCs, as well as CCMPs and compare our findings to available results from the literature, see Scheel and Scholtes (2000); Hoheisel and Kanzow (2007); Bucher and Schwartz (2018). Furthermore, we obtain new second-order optimality conditions for MPSCs. Some final remarks are presented in Section 6.

2 Preliminaries

2.1 Basic notation

Throughout this paper, \( x \cdot y \) is used to denote the common Euclidean inner product of two vectors \( x, y \in \mathbb{R}^n \). We equip \( \mathbb{R}^n \) with the Euclidean norm \( \| \cdot \|_2 \). The zero vector in \( \mathbb{R}^n \) will be denoted by \( 0^n \) while 0 is used to represent the scalar zero. For \( \varepsilon > 0 \) and some \( \bar{x} \in \mathbb{R}^n \), \( B^\varepsilon(\bar{x}) := \{ x \in \mathbb{R}^n \mid \|x - \bar{x}\|_2 \leq \varepsilon \} \) denotes the closed \( \varepsilon \)-ball around \( \bar{x} \). Similarly, \( U^\varepsilon(\bar{x}) := \{ x \in \mathbb{R}^n \mid \|x - \bar{x}\|_2 < \varepsilon \} \) represents the open \( \varepsilon \)-ball around \( \bar{x} \). Frequently, we will make use of the sets \( \mathbb{R}_+ := \{ t \in \mathbb{R} \mid t \geq 0 \} \) and \( \mathbb{R}_- := \{ t \in \mathbb{R} \mid t \leq 0 \} \). For a given nonempty set \( A \subset \mathbb{R}^n \), we exploit \( \text{cl} A \), \( \text{int} A \), \( \text{conv} A \), \( \text{conv} A \), and \( \text{span} A \) in order to represent the closure of \( A \), the interior of \( A \), the closed convex hull of \( A \), the closed convex hull of \( A \), and the span of \( A \) (i.e. the smallest subspace of \( \mathbb{R}^n \) comprising \( A \)), respectively. We use \( \text{dist}(x, A) := \inf \{ \|z - x\|_2 \mid z \in A \} \) to represent the distance of \( x \in \mathbb{R}^n \) to \( A \). Finally, the Cartesian product \( A \times B \) of two sets \( A \subset \mathbb{R}^n \) and \( B \subset \mathbb{R}^m \) will be interpreted as a subset of \( \mathbb{R}^{n+m} \).

Recall that a set-valued mapping \( \Psi: \mathbb{R}^n \rightrightarrows \mathbb{R}^m \), i.e. a mapping which assigns to any \( x \in \mathbb{R}^n \) a (possibly empty) set \( \Psi(x) \subset \mathbb{R}^m \), is called metrically subregular at some point \( (\bar{x}, \bar{y}) \in \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in \Psi(x)\} \) if there are constants \( \kappa > 0 \) and \( \varepsilon > 0 \) such that

\[
\forall x \in U^\varepsilon(\bar{x}) : \quad \text{dist}(x, \Psi^{-1}(\bar{y})) \leq \kappa \text{dist}(\bar{y}, \Psi(x))
\]

holds true. Here, \( \Psi^{-1}: \mathbb{R}^m \rightrightarrows \mathbb{R}^n \) denotes the inverse set-valued mapping associated with \( \Psi \) which is defined by \( \Psi^{-1}(y) := \{ x \in \mathbb{R}^n \mid y \in \Psi(x) \} \) for all \( y \in \mathbb{R}^m \). It is easily seen that \( \Psi \) is metrically subregular at \( (\bar{x}, \bar{y}) \) if and only if \( \Psi^{-1} \) possesses the so-called calmness property at \( (\bar{y}, \bar{x}) \), see e.g. Henrion and Outrata (2005).

For a twice continuously differentiable mapping \( P: \mathbb{R}^n \to \mathbb{R}^m \), \( \nabla P(x) \in \mathbb{R}^m \times \mathbb{R}^n \) denotes its Jacobian at \( x \in \mathbb{R}^n \). In the particular case \( m = 1 \), the gradient \( \nabla P(x) \) will be interpreted as a column vector. Furthermore, we set

\[
\forall x \in \mathbb{R}^n \forall d, h \in \mathbb{R}^n : \quad \nabla^2 P(x)[d, h] := \begin{pmatrix}
d^\top \nabla^2 P_1(x)h \\
\vdots \\
d^\top \nabla^2 P_m(x)h
\end{pmatrix}
\]

where \( P_1, \ldots, P_m: \mathbb{R}^n \to \mathbb{R} \) are the component mappings associated with \( P \) while the matrices \( \nabla^2 P_1(x), \ldots, \nabla^2 P_m(x) \) are their respective Hessians at \( x \in \mathbb{R}^n \).
2.2 Variational analysis

Here, we introduce the notions of variational analysis which are necessary in order to carry out our later considerations. For terminology and notation, we mainly follow Bonnans and Shapiro (2000) and Rockafellar and Wets (1998).

2.2.1 Polars and annihilators

For a nonempty set $A \subset \mathbb{R}^n$, the polar cone and the annihilator of $A$ are given as stated below:

$$A^\circ := \{ y \in \mathbb{R}^n | \forall x \in A: x \cdot y \leq 0 \}, \quad A^\perp := \{ y \in \mathbb{R}^n | \forall x \in A: x \cdot y = 0 \}. $$

Obviously, $A^\circ$ is a closed, convex cone and one has $A^\perp = A^\circ \cap (-A)^\circ$, i.e. $A^\perp$ is a subspace of $\mathbb{R}^n$. For any two sets $A, B \subset \mathbb{R}^n$, one easily obtains $(A \cup B)^\circ = A^\circ \cap B^\circ$ as well as $(A \cup B)^\perp = A^\perp \cap B^\perp$. For a cone $C \subset \mathbb{R}^n$, the so-called bipolar theorem, see (Rockafellar and Wets, 1998, Corollary 6.21), shows $C^{\circ \circ} = \text{conv } C$. Furthermore, the polarization rule $(C_1 + C_2)^\circ = C_1^\circ \cap C_2^\circ$ follows for any two cones $C_1, C_2 \subset \mathbb{R}^n$, see (Bonnans and Shapiro, 2000, Section 2.1.4). If $C_1, C_2$ are additionally, closed and convex, we have $(C_1 \cap C_2)^\circ = \text{cl}(C_1^\circ + C_2^\circ)$. Particularly, for subspaces $L_1, L_2 \subset \mathbb{R}^n$, $(L_1 \cap L_2)^\perp = L_1^\perp + L_2^\perp$ is valid since each subspace of $\mathbb{R}^n$ is closed. Supposing that $K \subset \mathbb{R}^n$ is a closed, convex cone, one obtains $K^{\circ \perp} = K \cap (-K)$, i.e. $K^{\circ \perp}$ coincides with the so-called lineality space of $K$ which is the largest subspace contained in $K$. Additionally,

$$K^{\circ \perp} = (K^\circ \cap (-K)^\circ)^\perp = (K^\circ \cap (-K)^\circ)^0 = \text{cl}(K^{\circ \circ} + (-K)^{\circ \circ}) = \text{cl}(K - K) = K - K = \text{span } K,$$

follows from the calculation rules provided above.

2.2.2 Tangent and normal cones

Let $A \subset \mathbb{R}^n$ be closed and fix an arbitrary point $\bar{x} \in A$. The closed cone

$$\mathcal{T}_A(\bar{x}) := \left\{ d \in \mathbb{R}^n \mid \exists \{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+: \exists \{d_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n: t_k \downarrow 0, d_k \to d, \bar{x} + t_k d_k \in A \forall k \in \mathbb{N} \right\}$$

is called the tangent (or Bouligand) cone to $A$ at $\bar{x}$. The cone $\hat{\mathcal{N}}_A(\bar{x}) := \mathcal{T}_A(\bar{x})^\circ$ is referred to as Fréchet (or regular) normal cone. By definition, it is closed and convex. Furthermore, we exploit the limiting (or Mordukhovich) normal cone to $A$ at $\bar{x}$ which is given by

$$\mathcal{N}_A(\bar{x}) = \left\{ \eta \in \mathbb{R}^n \mid \exists \{x_k\}_{k \in \mathbb{N}} \subset A \exists \{\eta_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n: x_k \to \bar{x}, \eta_k \to \eta, \eta_k \in \hat{\mathcal{N}}_A(x_k) \forall k \in \mathbb{N} \right\}.$$
Using the notion of the Painlevé-Kuratowski-limit, see e.g. (Rockafellar and Wets, 1998, Section 4.B), we have

\[ \mathcal{N}_A(x) = \limsup_{x \to \bar{x}, x \in A} \hat{N}_A(x). \]

Clearly, \( \hat{N}_A(x) \subset \mathcal{N}_A(x) \) holds and both cones coincide whenever \( A \) is convex. In case where \( A \) is a closed, convex cone, we obtain

\[ \hat{N}_A(x) = (A - \{ \bar{x} \})^\circ = A^\circ \cap \{ \bar{x} \}^\perp. \]

For formal completeness, we set \( T_A(x) := \emptyset, \hat{N}_A(x) := \emptyset, \) and \( \mathcal{N}_A(x) := \emptyset \) for each \( x \notin A \).

**Lemma 2.1.** Let \( Q \subset \mathbb{R}^n \) be a polyhedral set and fix \( \bar{x} \in Q \). Then, for all sufficiently small \( \varepsilon > 0 \), we have

\[ \hat{N}_Q(\bar{x}) = \bigcup_{x \in Q \cap B^\varepsilon(\bar{x})} \hat{N}_Q(x). \]

**Proof.** If \( \bar{x} \in Q \) belongs to the interior of \( Q \) then the statement of the lemma is obvious since both sides of the equality sign reduce to \( \{0^n\} \) for sufficiently small \( \varepsilon > 0 \).

Thus, we may assume that \( \bar{x} \) is a boundary point of \( Q \). Let \( F(\bar{x}) \) denote the set of faces \( S \subset Q \) of \( Q \) which satisfy \( \bar{x} \in S \). Since \( \bar{x} \notin \text{int } Q \) holds, \( F(\bar{x}) \) is nonempty. Noting that \( Q \) possesses only a finite number of faces and due to the fact that all these faces are closed, there exists \( \varepsilon > 0 \) such that \( F(x) \subset F(\bar{x}) \) holds for all \( x \in Q \cap B^\varepsilon(\bar{x}) \). Invoking the outer semicontinuity of the normal cone map \( \mathbb{R}^n \ni x \mapsto \hat{N}_Q(x) \subset \mathbb{R}^n \) associated with the convex set \( Q \), see (Rockafellar and Wets, 1998, Propositions 6.5 and 6.6, Theorem 6.9), and the fact that there only exist finitely many different Fréchet normal cones to a polyhedral set, we now obtain

\[ \hat{N}_Q(\bar{x}) \subset \bigcup_{x \in Q \cap B^\varepsilon(\bar{x})} \hat{N}_Q(x) \subset \limsup_{x \to \bar{x}, x \in Q} \hat{N}_Q(x) \subset \hat{N}_Q(\bar{x}) \]

for any \( \varepsilon \in (0, \varepsilon] \) which completes the proof. \( \square \)

** Lemma 2.2.** Let \( A := \bigcup_{i=1}^r A_i \) be the finite union of closed sets \( A_1, \ldots, A_r \subset \mathbb{R}^n \), choose \( \bar{x} \in A \), and set \( I(\bar{x}) := \{ i \in \{1, \ldots, r\} \mid \bar{x} \in A_i \} \). Then, one has

\[ T_A(\bar{x}) = \bigcup_{i \in I(\bar{x})} T_{A_i}(\bar{x}), \quad \hat{N}_A(\bar{x}) = \bigcap_{i \in I(\bar{x})} \hat{N}_{A_i}(\bar{x}), \quad \mathcal{N}_A(\bar{x}) \subset \bigcup_{i \in I(\bar{x})} \mathcal{N}_{A_i}(\bar{x}). \]

If, additionally, the sets \( A_1, \ldots, A_r \) are convex, then we particularly have

\[ \mathcal{N}_A(\bar{x}) \subset \bigcup_{i \in I(\bar{x})} \hat{N}_{A_i}(\bar{x}). \]

**Proof.** The first property can be found in (Aubin and Frankowska, 2009, Table 4.1) while the second one follows by polarizing the first one. For the proof of the last one, observe
that due to the closedness of all the sets $A_1, \ldots, A_r$, there is some ball $B^\varepsilon(\bar x)$ such that $I(x) \subset I(\bar x)$ holds for all $x \in A \cap B^\varepsilon(\bar x)$. This yields

\[
\mathcal{N}_A(\bar x) = \limsup_{x \to \bar x, x \in A} \bigcap_{i \in I(x)} \hat{\mathcal{N}}_{A_i}(x) \subset \limsup_{x \to \bar x, x \in A} \bigcup_{i \in I(x)} \hat{\mathcal{N}}_{A_i}(x) = \bigcup_{i \in I(\bar x)} \limsup_{x \to \bar x, x \in A_i} \hat{\mathcal{N}}_{A_i}(x) = \bigcup_{i \in I(\bar x)} \mathcal{N}_{A_i}(\bar x)
\]

which completes the proof.

**Lemma 2.3.** Let $K := \bigcup_{i=1}^r K_i$ be the finite union of polyhedral cones $K_1, \ldots, K_r \subset \mathbb{R}^n$. Fix a sequence $\{y_k\}_{k \in \mathbb{N}} \subset K$ converging to $\bar y \in K$. For each $k \in \mathbb{N}$, let $\lambda_k \in \hat{\mathcal{N}}_K(y_k)$ be chosen such that $\lambda_k \to \bar \lambda$ holds true where $\bar \lambda \in \hat{\mathcal{N}}_K(\bar y)$ is valid.

Then, we have

1. $\bar y \cdot \bar \lambda = 0$,
2. $y_k \cdot \lambda_k = 0$ for all $k \in \mathbb{N}$, and
3. $\bar y \cdot \lambda_k = y_k \cdot \bar \lambda = 0$ for sufficiently large $k \in \mathbb{N}$.

**Proof.** Noting that the sets $K_1, \ldots, K_r$ are convex cones, we obtain

\[
\hat{\mathcal{N}}_K(y) = \left( \bigcap_{i \in I(y)} K_i^\circ \right) \cap \{y\}^\perp
\]

for each $y \in K$ where we used $I(y) := \{i \in \{1, \ldots, r\} \mid y \in K_i\}$, see Lemma 2.2. Thereby, the lemma’s first and second assertion are obvious.

Exploiting the closedness of $K_1, \ldots, K_r$ as well as the convergence $y_k \to \bar y$, the inclusion $I(y_k) \subset I(\bar y)$ needs to be valid for all sufficiently large $k \in \mathbb{N}$. Particularly, for each large enough $k_0 \in \mathbb{N}$, there is $i(k_0) \in I(\bar y)$ such that $y_{k_0} \in K_{i(k_0)}$ holds. Now, we can exploit Lemma 2.1 in order to see $\lambda_{k_0} \in \hat{\mathcal{N}}_{K_{i(k_0)}}(y_{k_0}) \subset \hat{\mathcal{N}}_{K_{i(k_0)}}(\bar y) = K_{i(k_0)}^\circ \cap \{\bar y\}^\perp$ for large enough $k_0 \in \mathbb{N}$, i.e. $\bar y \cdot \lambda_{k_0} = 0$ follows.

Noting that $K$ is a finite union, for sufficiently large $k_0 \in \mathbb{N}$, there is a subsequence $\{y_{k_l}\}_{l \in \mathbb{N}}$ of $\{y_k\}_{k \in \mathbb{N}}$ with $k_l \geq k_0$ for all $l \in \mathbb{N}$ and some index $i(k_0) \in I(y_{k_0})$ such that $i(k_l) \in I(y_{k_l}) \subset I(\bar y)$ is valid for all $l \in \mathbb{N}$. Particularly, $\lambda_{k_l} \in \hat{\mathcal{N}}_{K_{i(k_l)}}(y_{k_l}) \subset \hat{\mathcal{N}}_{K_{i(k_l)}}(\bar y)$ holds for all $l \in \mathbb{N}$. Keeping Lemma 2.1 in mind and noting that there exist only finitely many different Fréchet normal cones to a polyhedral set (since it possesses only finitely many faces), we find a subsequence $\{y_{k_{l_\nu}}\}_{\nu \in \mathbb{N}}$ of $\{y_{k_l}\}_{l \in \mathbb{N}}$ and a polyhedral cone $C \subset \mathbb{R}^n$ such that $C = \hat{\mathcal{N}}_{K_{i(k_{l_\nu})}}(y_{k_{l_\nu}})$ holds for all $\nu \in \mathbb{N}$, and for large enough $k_0 \in \mathbb{N}$, we can even guarantee $C = \hat{\mathcal{N}}_{K_{i(k_0)}}(y_{k_0})$. Particularly, $\lambda_{k_{l_\nu}} \in C$ for all $\nu \in \mathbb{N}$ follows. Noting that $\lambda_{k_{l_\nu}} \to \bar \lambda$ holds as $\nu \to \infty$, we have $\bar \lambda \in C$ by closedness of $C$. Finally, we observe $C = \hat{\mathcal{N}}_{K_{i(k_0)}}(y_{k_0}) = K_{i(k_0)}^\circ \cap \{y_{k_0}\}^\perp$, i.e. $y_{k_0} \cdot \bar \lambda = 0$ follows. This completes the proof. \qed
Next, we present some preliminary results on the variational geometry of the feasible set $X$ associated with (MPDC). Fix $\bar{x} \in X$. We call
\[ L_X(\bar{x}) := \{ d \in \mathbb{R}^n | \nabla F(\bar{x})d \in T_D(F(\bar{x})) \} \]
the linearization cone to $X$ at $\bar{x}$. One always has the inclusion $T_X(\bar{x}) \subset L_X(\bar{x})$, see (Rockafellar and Wets, 1998, Theorem 6.31), while equality only holds under validity of a constraint qualification. Similarly, we generally have
\[ \hat{N}_X(\bar{x}) \supset \nabla F(\bar{x})^\top N_D(F(\bar{x})) \]
while the converse inclusion can only be guaranteed postulating additional assumptions. The following result is associated with this issue and a straightforward consequence of (Gfrerer and Outrata, 2016, Theorem 4).

**Proposition 2.4.** Fix a feasible point $\bar{x} \in X$ of (MPDC) and suppose that there exists a subspace $L \subset \mathbb{R}^m$ satisfying $T_D(F(\bar{x}))+L \subset T_D(F(\bar{x}))$, $L^\perp \supset N_D(F(\bar{x}))$, and
\[ \nabla F(\bar{x})\mathbb{R}^n+L = \mathbb{R}^m. \]
Then, we have $T_X(\bar{x}) = L_X(\bar{x})$ and equality holds in (2).

**Proof.** Polarizing equation (3) gives us
\[ 0^n = \nabla F(\bar{x})^\top \lambda, \lambda \in L^\perp \implies \lambda = 0^m. \]
Thus, the assumption $L^\perp \supset N_D(F(\bar{x}))$ guarantees validity of
\[ 0^n = \nabla F(\bar{x})^\top \lambda, \lambda \in N_D(F(\bar{x})) \implies \lambda = 0^m. \]
Invoking e.g. (Gfrerer and Ye, 2017, Section 2), this condition implies that the feasibility mapping $\mathbb{R}^n \ni x \Rightarrow \{ F(x) \} - D \subset \mathbb{R}^m$ of (MPDC) is metrically subregular at $(\bar{x}, 0^m)$. Now, the relation $T_X(\bar{x}) = L_X(\bar{x})$ can be deduced from (Henrion and Outrata, 2005, Proposition 1) while equality in (2) follows from (Gfrerer and Outrata, 2016, Theorem 4).

In the context of certain instances of disjunctive programming, there exist weaker conditions than those ones postulated in Proposition 2.4 which ensure equality in (2), see Benko and Gfrerer (2017).

### 2.2.3 Second-order tangent sets

For the consideration of second-order optimality conditions, we exploit so-called second-order tangent sets. Therefore, let $A \subset \mathbb{R}^n$ be a closed set and fix $\bar{x} \in A$ as well as $d \in T_A(\bar{x})$. The closed sets
\[ \mathcal{T}_A^2(\bar{x}; d) := \left\{ h \in \mathbb{R}^n \left| \exists \{ h_k \}_{k \in \mathbb{N}} \subset \mathbb{R}^n \exists \{ t_k \}_{k \in \mathbb{N}} \subset \mathbb{R}_+ : h_k \to h, t_k \downarrow 0, \bar{x} + t_k d + \frac{1}{2} t_k^2 h_k \in A \forall k \in \mathbb{N} \right\} , \]
\[
\mathcal{T}_{S_i}^{k,2}(x;d) := \left\{ h \in \mathbb{R}^n \mid \forall \{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+ \text{ such that } t_k \downarrow 0 \\exists \{h_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n : h_k \to h, \, x + t_k d + \frac{1}{2} t_k^2 h_k \in A \forall k \in \mathbb{N} \right\}.
\]

are called outer (Bouligand) and inner (adjacent) second-order tangent set to \(A\) at \(\bar{x}\) in direction \(d\), see e.g. (Bonnans and Shapiro, 2000, Definition 3.28). Note that these sets are not conic in general. For \(\bar{d} \notin \mathcal{T}_A(\bar{x})\), we set \(\mathcal{T}_A^2(\bar{x};\bar{d}) = \mathcal{T}_A^{k,2}(\bar{x};\bar{d}) = \emptyset\) for formal completeness. Clearly, we always have \(\mathcal{T}_A^k(\bar{x};d) \subset \mathcal{T}_A^2(\bar{x};d)\). If equality holds, then \(A\) is called \textit{parabolically derivable} at \(\bar{x}\) in direction \(d\). We say that \(A\) is parabolically derivable if it is parabolically derivable at each point \(x \in A\) in each direction \(d \in \mathcal{T}_A(x)\). Note that even convex sets are not parabolically derivable in general. However, it follows from (Bonnans and Shapiro, 2000, Proposition 3.34) that any polyhedral set \(S \subset \mathbb{R}^n\) is parabolically derivable and it holds

\[
\forall x \in S \forall d \in \mathcal{T}_S(x) : \quad \mathcal{T}_S^2(x;d) = \mathcal{T}_S^{k,2}(x;d) = \mathcal{T}_{T_S(x)}(d),
\]

see (Rockafellar and Wets, 1998, Proposition 13.12) as well. In the subsequent lemma, we show that the finite union of polyhedral sets is parabolically derivable.

**Lemma 2.5.** Let \(S_1, \ldots, S_r \subset \mathbb{R}^n\) be polyhedral sets and define \(S := \bigcup_{i=1}^r S_i\). Then, for each \(x \in S\) and \(d \in \mathcal{T}_S(x)\), we have

\[
\mathcal{T}_S^2(x;d) = \mathcal{T}_S^{k,2}(x;d) = \mathcal{T}_{T_S(x)}(d)
\]

and

\[
\mathcal{T}_S^2(x;d) + \bigcap_{i \in I(x)} \mathcal{T}_{S_i}(x)^\perp \subset \mathcal{T}_S^2(x;d)
\]

where we used \(I(x) := \{i \in \{1, \ldots, r\} \mid x \in S_i\}\). In particular, \(S\) is parabolically derivable.

**Proof.** Fix \(x \in S\) and \(d \in \mathcal{T}_S(x)\). We exploit the calculus rules from (Bonnans and Shapiro, 2000, Proposition 3.37) and the fact that polyhedral sets are parabolically derivable in order to obtain

\[
\mathcal{T}_S^2(x;d) = \bigcup_{j \in I(x)} \mathcal{T}_{S_j}^2(x;d) = \bigcup_{j \in I(x)} \mathcal{T}_{S_j}^{k,2}(x;d) \subset \mathcal{T}_S^{k,2}(x;d).
\]

This already shows the parabolic derivability of \(S\) at \(d\) since \(\mathcal{T}_S^{k,2}(x;d) \subset \mathcal{T}_S^2(x;d)\) follows by definition of these second-order tangent sets.

Next, we use formula (4) as well as Lemma 2.2 in order to see

\[
\mathcal{T}_S^2(x;d) = \bigcup_{j \in I(x)} \mathcal{T}_{S_j}^2(x;d) = \bigcup_{j \in I(x)} \mathcal{T}_{S_j}(x)(d) = \bigcup_{j \in I(x)} \mathcal{T}_{S_j}(x)(d) = \mathcal{T}_{I(x)}(d).
\]

In order to prove correctness of the last formula, we first invoke (Rockafellar and Wets, 1998, Proposition 13.12) in order to see that

\[
\mathcal{T}_S^2(x;d) + \mathcal{T}_{S_i}(x) \subset \mathcal{T}_{S_i}^2(x;d)
\]
holds true for all \( i \in I(x) \) since \( S_i \) is a polyhedron. This leads to

\[
T^2_{S_i}(x; d) + \bigcap_{i \in I(x)} T_{S_i}(x)^\perp = \left( \bigcup_{j \in I(x)} T^2_{S_j}(x; d) \right) + \bigcap_{i \in I(x)} T_{S_i}(x)^\perp
\]

\[
= \bigcup_{j \in I(x)} \left( T^2_{S_j}(x; d) + \bigcap_{i \in I(x)} T_{S_i}(x)^\perp \right) \subset \bigcup_{j \in I(x)} \left( T^2_{S_j}(x; d) + T_{S_j}(x)^\perp \right)
\]

\[
\subset \bigcup_{j \in I(x)} \left( T^2_{S_j}(x; d) + T_{S_j}(x) \right) \subset \bigcup_{j \in I(x)} T^2_{S_j}(x; d) = T^2_{S}(x; d)
\]

and completes the proof. \( \square \)

### 3 An MPDC-tailored version of the linear independence constraint qualification

We start this section by defining the constraint qualification of our interest. Recall that \( X \subset \mathbb{R}^n \) denotes the feasible set of \((MPDC)\).

**Definition 3.1.** Let \( \bar{x} \in X \) be an arbitrary feasible point of \((MPDC)\). Then, the linear independence constraint qualification \((MPDC-LICQ)\) is said to hold at \( \bar{x} \) if the following condition is valid:

\[
0^n = \nabla F(\bar{x})^\top \lambda, \lambda \in \bigcap_{i \in I(\bar{x})} \text{span}\, \hat{N}_{D_i}(F(\bar{x})) \implies \lambda = 0^m.
\]

Here, we used \( I(\bar{x}) := \{ i \in \{1, \ldots, r\} \mid F(\bar{x}) \in D_i \} \).

We first note that MPDC-LICQ holds at \( \bar{x} \in X \) whenever the matrix \( \nabla F(\bar{x}) \) possesses full row rank \( m \), i.e. if the gradients \( \nabla F_1(\bar{x}), \ldots, \nabla F_m(\bar{x}) \) of the component mappings \( F_1, \ldots, F_m: \mathbb{R}^n \to \mathbb{R} \) associated with \( F \) are linearly independent. In the example below, it will be demonstrated that MPDC-LICQ reduces to the well-known LICQ whenever standard nonlinear programs are under consideration.

**Example 3.2.** For continuously differentiable functions \( g_1, \ldots, g_p, h_1, \ldots, h_q: \mathbb{R}^n \to \mathbb{R} \), we consider the standard nonlinear program

\[
f(x) \to \min \quad g_j(x) \leq 0 \quad j = 1, \ldots, p \tag{NLP}
\]

\[
h_j(x) \leq 0 \quad j = 1, \ldots, q.
\]

In order to transfer it to a program of type \((MPDC)\), we choose \( r := 1 \), set \( D := \mathbb{R}^p \times \{0^q\} \), and define \( F: \mathbb{R}^n \to \mathbb{R}^{p+q} \) by means of

\[
\forall x \in \mathbb{R}^n: \quad F(x) := \begin{bmatrix} g(x)^\top & h(x)^\top \end{bmatrix}^\top.
\]
Here, the mappings $g : \mathbb{R}^n \to \mathbb{R}^p$ and $h : \mathbb{R}^n \to \mathbb{R}^q$ possess the component mappings $g_1, \ldots, g_p$ and $h_1, \ldots, h_q$, respectively.

Fix a feasible point $\bar{x} \in X$ of (NLP) and define $I^p(\bar{x}) := \{ j \in \{1, \ldots, p\} \mid g_j(\bar{x}) = 0 \}$. Using the calculus rules for the tangent and Fréchet normal cone to Cartesian products of (convex) sets, see \((\text{Rockafellar and Wets}, 1998, \text{Proposition 6.41})\), we have

$$
\mathcal{T}_D(F(\bar{x})) = \mathcal{T}_{\mathbb{R}^n}(g(\bar{x})) \times \mathcal{T}_{\{0\}}(h(\bar{x})), \quad \mathcal{\hat{N}}_D(F(\bar{x})) = \mathcal{\hat{N}}_{\mathbb{R}^n}(g(\bar{x})) \times \mathcal{\hat{N}}_{\{0\}}(h(\bar{x})).
$$

Straightforward calculations lead to the formulas

$$
\begin{align*}
\mathcal{T}_{\mathbb{R}^n}(g(\bar{x})) &= \{ d \in \mathbb{R}^p \mid \forall j \in I^p(\bar{x}) : d_j \leq 0 \}, \\
\mathcal{\hat{N}}_{\mathbb{R}^n}(g(\bar{x})) &= \{ \lambda \in \mathbb{R}^p_+ \mid \forall j \notin I^p(\bar{x}) : \lambda_j = 0 \}, \\
\mathcal{T}_{\{0\}}(h(\bar{x})) &= \{0^q\}, \\
\mathcal{\hat{N}}_{\{0\}}(h(\bar{x})) &= \mathbb{R}^q,
\end{align*}
$$

i.e. we have

$$
\text{span} \mathcal{\hat{N}}_D(F(\bar{x})) = \{ \lambda \in \mathbb{R}^p \mid \forall j \notin I^p(\bar{x}) : \lambda_j = 0 \} \times \mathbb{R}^q.
$$

Thus, MPDC-LICQ from Definition 3.1 takes the following form for (NLP):

$$
0^n = \nabla g(\bar{x})^\top \lambda + \nabla h(\bar{x})^\top \rho, \\
\forall j \notin I^p(\bar{x}) : \lambda_j = 0 \\
\implies \lambda = 0^p, \rho = 0^q.
$$

This is equivalent to the linear independence of the vectors from

$$
\{ \nabla g_j(\bar{x}) \mid j \in I^p(\bar{x}) \} \cup \{ \nabla h_j(\bar{x}) \mid j \in \{1, \ldots, q\} \}
$$

which is precisely the definition of the standard linear independence constraint qualification from nonlinear programming.

In Section 5, we will show that in the particular instances of MPCCs, MPVCs, and MPSCs, MPDC-LICQ coincides with the well-known respective problem-tailored version of LICQ.

We provide an equivalent primal characterization of MPDC-LICQ in the subsequent lemma.

**Lemma 3.3.** Fix $\bar{x} \in X$ arbitrarily. Then, MPDC-LICQ is valid at $\bar{x}$ if and only if the subsequent condition is satisfied:

$$
\nabla F(\bar{x}) \mathbb{R}^n + \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^\perp = \mathbb{R}^m.
$$

**Proof.** First, we note that for any matrix $A \in \mathbb{R}^{m \times n}$ and any subspace $L \subset \mathbb{R}^m$, the equivalence

$$
AR^n + L = \mathbb{R}^m \iff \{ \lambda \in L^\perp \mid A^\top \lambda = 0^n \} = \{0^n\}
$$

follows from the polarization rules provided in Section 2.2. Thus, the statement of the lemma follows setting $A := \nabla F(\bar{x})$ and $L := \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^\perp$ observing that

$$
\left( \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^\perp \right)^\perp = \sum_{i \in I(\bar{x})} \mathcal{\hat{N}}_{D_i}(F(\bar{x}))^\perp = \sum_{i \in I(\bar{x})} \text{span} \mathcal{\hat{N}}_{D_i}(F(\bar{x}))
$$

holds true. \(\square\)
The following lemma will be important for our remaining considerations.

**Lemma 3.4.** For each feasible point \( \bar{x} \in X \) of (MPDC), we have

\[
\mathcal{T}_D(F(\bar{x})) + \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^{\perp} \subset \mathcal{T}_D(F(\bar{x}))
\]

and

\[
\mathcal{N}_D(F(\bar{x})) \subset \left( \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^{\perp} \right)^{\perp}.
\]

**Proof.** Using **Lemma 2.2** and the convexity of \( D_1, \ldots, D_r \), we find

\[
\mathcal{T}_D(F(\bar{x})) + \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^{\perp} = \left( \bigcup_{j \in I(\bar{x})} \mathcal{T}_{D_j}(F(\bar{x})) \right) + \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^{\perp}
\]

\[
= \bigcup_{j \in I(\bar{x})} \left( \mathcal{T}_{D_j}(F(\bar{x})) + \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^{\perp} \right)
\]

\[
\subset \bigcup_{j \in I(\bar{x})} \left( \mathcal{T}_{D_j}(F(\bar{x})) + \mathcal{T}_{D_j}(F(\bar{x}))^{\perp} \right)
\]

\[
= \bigcup_{j \in I(\bar{x})} \left( \mathcal{T}_{D_j}(F(\bar{x})) + \mathcal{T}_{D_j}(F(\bar{x})) \cap (-\mathcal{T}_{D_j}(F(\bar{x}))) \right)
\]

\[
= \bigcup_{j \in I(\bar{x})} \mathcal{T}_{D_j}(F(\bar{x})) = \mathcal{T}_D(F(\bar{x}))
\]

since we have \( C + C \cap (-C) \subset C + C = C \subset C + (-C) \) for any closed, convex cone \( C \subset \mathbb{R}^m \). This shows the validity of the first inclusion. For the proof of the second one, we invoke **Lemma 2.2** and (6) in order to see

\[
\mathcal{N}_D(F(\bar{x})) \subset \bigcup_{i \in I(\bar{x})} \mathcal{\hat{N}}_{D_i}(F(\bar{x})) \subset \sum_{i \in I(\bar{x})} \mathcal{\hat{N}}_{D_i}(F(\bar{x}))
\]

\[
\subset \sum_{i \in I(\bar{x})} \text{span} \mathcal{\hat{N}}_{D_i}(F(\bar{x})) = \left( \bigcap_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))^{\perp} \right)^{\perp}.
\]

This already completes the proof. \( \square \)

We combine the above lemma with **Proposition 2.4** and **Lemma 3.3** in order to obtain the following result.

**Corollary 3.5.** Let \( \bar{x} \in X \) be a feasible point of (MPDC) where MPDC-LICQ is valid. Then, we have

\[
\mathcal{\hat{N}}_X(\bar{x}) = \nabla F(\bar{x})^\top \mathcal{\hat{N}}_D(F(\bar{x})).
\]
Let us briefly interrelate the constraint qualification MPDC-LICQ with other prominent constraint qualifications from disjunctive programming.

**Remark 3.6.** Let \( \bar{x} \in X \) be a feasible point of (MPDC) where MPDC-LICQ is valid. Then, due to Proposition 2.4 as well as Lemmas 3.3 and 3.4, we obtain the relation \( T_X(\bar{x}) = \mathcal{L}_X(\bar{x}) \). Recall that the set \( \mathcal{L}_X(\bar{x}) \) denotes the linearization cone to \( X \) at \( \bar{x} \), see (1). In the literature of disjunctive programming, this condition is called generalized Abadie constraint qualification (GACQ), see (Flegel et al., 2007, Definition 6). Furthermore, we obtain \( \tilde{\mathcal{N}}_X(\bar{x}) = \mathcal{L}_X(\bar{x})^\circ \) by polarization, and the latter condition is called generalized Guignard constraint qualification (GGCQ), see (Flegel et al., 2007, Definition 6).

Now, it is possible to exploit Proposition 2.4 in order to derive necessary optimality conditions of strong stationarity-type for (MPDC).

**Theorem 3.7.** Let \( \bar{x} \in \mathbb{R}^n \) be a locally optimal solution of (MPDC) where MPDC-LICQ is valid. Then, there exists a uniquely determined multiplier \( \lambda \in \mathbb{R}^m \) such that

\[
0^n = \nabla f(\bar{x}) + \nabla F(\bar{x})^\top \lambda, \quad \lambda \in \bigcap_{i \in I(\bar{x})} \tilde{\mathcal{N}}_{D_i}(F(\bar{x}))
\]

holds.

**Proof.** Due to (Rockafellar and Wets, 1998, Theorem 6.12), we have \( -\nabla f(\bar{x}) \in \tilde{\mathcal{N}}_X(\bar{x}) \). Invoking Lemma 2.2 and Corollary 3.5, we obtain

\[
\tilde{\mathcal{N}}_X(\bar{x}) = \nabla F(\bar{x})^\top \tilde{\mathcal{N}}_{D_i}(F(\bar{x})) = \nabla F(\bar{x})^\top \left[ \bigcap_{i \in I(\bar{x})} \tilde{\mathcal{N}}_{D_i}(F(\bar{x})) \right],
\]

i.e. the postulated stationarity system possesses a solution.

It remains to show that the associated multiplier is uniquely determined. Therefore, assume that there are \( \lambda^1, \lambda^2 \in \bigcap_{i \in I(\bar{x})} \tilde{\mathcal{N}}_{D_i}(F(\bar{x})) \) satisfying \( 0^n = \nabla f(\bar{x}) + \nabla F(\bar{x})^\top \lambda^s \), \( s = 1, 2 \). This yields \( 0^n = \nabla F(\bar{x})^\top (\lambda^1 - \lambda^2) \). Moreover, for each \( i \in I(\bar{x}) \), we have

\[
\lambda^1 - \lambda^2 \in \tilde{\mathcal{N}}_{D_i}(F(\bar{x})) - \tilde{\mathcal{N}}_{D_i}(F(\bar{x})) = \text{span}\tilde{\mathcal{N}}_{D_i}(F(\bar{x})).
\]

This yields \( \lambda^1 - \lambda^2 \in \sum_{i \in I(\bar{x})} \text{span}\tilde{\mathcal{N}}_{D_i}(F(\bar{x})) \), and by validity of MPDC-LICQ, \( \lambda^1 = \lambda^2 \) follows. This completes the proof. \( \square \)

Note that the multiplier \( \lambda \) from Theorem 3.7 is chosen from the Fréchet normal cone \( \tilde{\mathcal{N}}_D(F(\bar{x})) \). Keeping (Flegel et al., 2007, Definition 1) in mind, this observation justifies to call the above necessary optimality condition a strong stationarity-type condition.

**Definition 3.8.** A feasible point \( \bar{x} \in X \) of (MPDC) is called strongly stationary (S-stationary for short) if any only if there exists a multiplier \( \lambda \in \bigcap_{i \in I(\bar{x})} \tilde{\mathcal{N}}_{D_i}(F(\bar{x})) \) which satisfies \( 0^n = \nabla f(\bar{x}) + \nabla F(\bar{x})^\top \lambda \).
Some general considerations on S-stationary points of disjunctive programs can be found in Flegel et al. (2007); Benko and Gfrerer (2017, 2018). We note that for prominent classes of disjunctive programs like MPCCs, MPVCs, CCMPs, and MPSCs, there exist respective strong stationarity notions which can be obtained by applying Definition 3.8 to the specific problem setting, see Section 5. With the aid of Example 3.2, it is easily seen that for (NLP), the S-stationarity system equals the classical Karush-Kuhn-Tucker conditions.

Remark 3.9. Assume that the sets $D_1, \ldots, D_r$ are polyhedral cones and let $\bar{x} \in \mathbb{R}^n$ be feasible to (MPDC). Then, we have

$$
\sum_{i \in I(\bar{x})} \text{span} \hat{N}_{D_i}(F(\bar{x})) = \sum_{i \in I(\bar{x})} \text{span}(D_i^0 \cap \{F(\bar{x})\}^\perp) \subset \sum_{i \in I(\bar{x})} \text{span}(D_i^0 \cap \{F(\bar{x})\}^\perp)
$$

$$
\subset \left( \sum_{i \in I(\bar{x})} \text{span} D_i^0 \right) \cap \{F(\bar{x})\}^\perp = \text{span} \left( \sum_{i \in I(\bar{x})} D_i^0 \right) \cap \{F(\bar{x})\}^\perp.
$$

Thus, MPDC-LICQ is implied by validity of

$$
0^n = \nabla F(\bar{x})^\top \lambda, \lambda \in \text{span} \left( \sum_{i \in I(\bar{x})} D_i^0 \right), \quad F(\bar{x}) \cdot \lambda = 0 \implies \lambda = 0^n. \quad (7)
$$

Furthermore, we have

$$
\bigcap_{i \in I(\bar{x})} \hat{N}_{D_i}(F(\bar{x})) = \bigcap_{i \in I(\bar{x})} D_i^0 \cap \{F(\bar{x})\}^\perp = \left( \bigcap_{i \in I(\bar{x})} D_i^0 \right) \cap \{F(\bar{x})\}^\perp,
$$

i.e. $\bar{x}$ is an S-stationary point of (MPDC) if and only if there is some $\lambda \in \mathbb{R}^m$ which satisfies

$$
0^n = \nabla f(\bar{x}) + \nabla F(\bar{x})^\top \lambda, \lambda \in \bigcap_{i \in I(\bar{x})} D_i^0, \quad F(\bar{x}) \cdot \lambda = 0.
$$

For (NLP) from Example 3.2, the constraint qualification (7) takes the form

$$
0^n = \nabla g(\bar{x})^\top \lambda + \nabla h(\bar{x})^\top \rho, \quad g(\bar{x}) \cdot \lambda = 0, \quad \rho = 0^q.
$$

Obviously, this condition is more restrictive than LICQ. Thus, (7) is more restrictive than MPDC-LICQ in general.

Due to Remark 3.6, the validity of MPDC-LICQ at $\bar{x}$ implies that the tangent cone to $X$ at $\bar{x}$ equals the associated linearization cone. As we will see in the lemma below, we also obtain a nice representation of the second-order tangent sets to $X$ at $\bar{x}$ in each direction $d \in T_X(\bar{x})$. For the proof, we follow ideas from (Rockafellar and Wets, 1998, Proposition 13.13).
Lemma 3.10. Let $\bar{x} \in X$ be a feasible point of (MPDC) where MPDC-LICQ is valid. Then, for each $d \in T_X(\bar{x})$, we have

$$T^2_{\bar{x}}(d) = \{ h \in \mathbb{R}^n \mid \nabla F(\bar{x})h + \nabla^2 F(\bar{x})[d, d] \in T_{TD(F(\bar{x}))}(\nabla F(\bar{x})d) \} \neq \emptyset$$

and $X$ is parabolically derivable at $x$ in direction $d$.

Proof. Let us start to prove the formula for the outer second-order tangent set. We will show both inclusions separately.

⊂: Fix $d \in T_X(\bar{x})$ and $h \in T^2_{\bar{x}}(d)$ arbitrarily. Then, we find sequences $\{h_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ and $\{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+$ such that $h_k \to h$, $t_k \downarrow 0$, as well as $F(\bar{x} + t_k d + \frac{1}{2}t_k^2 h_k) \in D$ for all $k \in \mathbb{N}$ hold true. For each $k \in \mathbb{N}$, we now define

$$r_k := \frac{F(\bar{x} + t_k d + \frac{1}{2}t_k^2 h_k) - F(\bar{x}) - t_k \nabla F(\bar{x})d}{\frac{1}{2}t_k^2}.$$

Then, we have $F(\bar{x} + t_k \nabla F(\bar{x})d + \frac{1}{2}t_k^2 r_k) = F(\bar{x} + t_k d + \frac{1}{2}t_k^2 h_k) \in D$ for all $k \in \mathbb{N}$, i.e. supposing that $\{r_k\}_{k \in \mathbb{N}}$ converges, its limit belongs to $T^2_{\bar{x}}(d)$; $\nabla F(\bar{x})d$. On the other hand, we have

$$r_k = \frac{F(\bar{x} + t_k (d + \frac{1}{2}t_k h_k)) - F(\bar{x}) - t_k \nabla F(\bar{x})(d + \frac{1}{2}t_k h_k)}{\frac{1}{2}t_k^2} + \nabla F(\bar{x})h_k$$

and this sum converges to $\nabla^2 F(\bar{x})[d, d] + \nabla F(\bar{x})h$, see e.g. (Rockafellar and Wets, 1998, Example 13.8). Thus, we have $\nabla^2 F(\bar{x})[d, d] + \nabla F(\bar{x})h \in T^2_{\bar{x}}(d)$; $\nabla F(\bar{x})d$. The desired inclusion follows since we have $T^2_{\bar{x}}(d) = T_{TD(F(\bar{x}))}(\nabla F(\bar{x})d)$ by Lemma 2.5.

⊃: Fix $d \in T_X(\bar{x})$ and $h \in \mathbb{R}^n$ satisfying $\nabla F(\bar{x})h + \nabla^2 F(\bar{x})d, d] \in T^2_{\bar{x}}(d)$; $\nabla F(\bar{x})d)$. Since we have $T_{TD(F(\bar{x}))}(\nabla F(\bar{x})d) = T^2_{\bar{x}}(d)$; $\nabla F(\bar{x})d)$ from Lemma 2.5, we find sequences $\{r_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ and $\{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+$ such that $r_k \to \nabla F(\bar{x})h + \nabla^2 F(\bar{x})d$, $t_k \downarrow 0$, and $F(\bar{x} + t_k \nabla F(\bar{x})d + \frac{1}{2}t_k^2 r_k) \in D$ for all $k \in \mathbb{N}$. Noting that $X$ is closed, let us fix $x_k \in \text{argmin}\{||\bar{x} + t_k d + \frac{1}{2}t_k^2 h - x||_2 \mid x \in X\}$ for all $k \in \mathbb{N}$. Combining Lemmas 3.3 and 3.4 with Proposition 2.4, we obtain that the validity of MPDC-LICQ implies that the feasibility mapping $\mathbb{R}^n \ni x \Rightarrow \{F(x)\} - D \subset \mathbb{R}^m$ of (MPDC) is metrically subregular at $(\bar{x}, 0^m)$. Thus, we find constants $\kappa > 0$ and $\varepsilon > 0$ such that

$$\forall x \in U^0(\bar{x}): \quad \text{dist}(x, X) \leq \kappa \text{dist}(F(x), D).$$

We obtain

$$\left\| \frac{x_k - \bar{x} - t_k d}{t_k^2} - h \right\|_2 = \frac{2}{t_k^2} \text{dist}(\bar{x} + t_k d + \frac{1}{2}t_k^2 h, X) \leq \frac{2\kappa}{t_k^2} \text{dist}(F(\bar{x} + t_k d + \frac{1}{2}t_k^2 h), D)$$

$$\leq \kappa \left\| \frac{F(\bar{x} + t_k d + \frac{1}{2}t_k^2 h) - F(\bar{x}) - t_k \nabla F(\bar{x})d}{\frac{1}{2}t_k^2} - r_k \right\|_2$$

$$= \kappa \left\| \frac{F(\bar{x} + t_k (d + \frac{1}{2} t_k h)) - F(\bar{x}) - t_k \nabla F(\bar{x})(d + \frac{1}{2} t_k h)}{\frac{1}{2}t_k^2} + \nabla F(\bar{x})h - r_k \right\|_2$$

14
for sufficiently large \( k \in \mathbb{N} \) and the last term converges to 0 as \( k \to \infty \), see (Rockafellar and Wets, 1998, Example 13.8). Thus, we have shown \( h \in T_X^2(\bar{x}; d) \).

Next, let us show that \( T_X^2(\bar{x}; d) \) is nonempty for each \( d \in T_X(\bar{x}) \). Therefore, fix an arbitrary vector \( d \in T_X(\bar{x}) \). Then, \( d \in \mathcal{L}_X(\bar{x}) \) holds true and, consequently, the set \( T_{T_D(F(\bar{x}))}(\nabla F(\bar{x})d) = T_D^2(\bar{x}; \nabla F(\bar{x})d) \) cannot be empty. We fix an arbitrary vector \( r \in T_D^2(\bar{x}; \nabla F(\bar{x})d) \) and observe by means of Lemma 2.5 that

\[
\{r\} + \bigcap_{i \in I(\bar{x})} T_{D_i}(F(\bar{x}))^{\circ \perp} \subseteq T_D^2(\bar{x}; \nabla F(\bar{x})d) = T_{T_D(F(\bar{x}))}(\nabla F(\bar{x})d)
\]

holds true. By validity of MPDC-LICQ and Lemma 3.3, we now obtain

\[
\nabla F(\bar{x}) \mathbb{R}^n + T_{T_D(F(\bar{x}))}(\nabla F(\bar{x})d) = \mathbb{R}^m.
\]

Particularly, we find some vector \( w \in \mathbb{R}^n \) and some \( h \in T_{T_D(F(\bar{x}))}(\nabla F(\bar{x})d) \) such that we have \( \nabla F(\bar{x})w + h = \nabla^2 F(\bar{x})[d, d] \), i.e. \( -w \in T_X^2(\bar{x}; d) \) is valid.

Similarly as above, we can show

\[
T_X^{0,2}(\bar{x}; d) = \{h \in \mathbb{R}^n \mid \nabla F(\bar{x})h + \nabla^2 F(\bar{x})[d, d] \in T_{T_D(F(\bar{x}))}(\nabla F(\bar{x})d)\},
\]

i.e. \( X \) is parabolically derivable at \( \bar{x} \) in direction \( d \in T_X(\bar{x}) \). \( \Box \)

### 4 Second-order optimality conditions and MPDC-LICQ

Recall that for each feasible point \( \bar{x} \in X \) of (MPDC), \( \mathcal{L}_X(\bar{x}) \) denotes the linearization cone to \( X \) at \( \bar{x} \) and has been defined in (1). For later use, we introduce the so-called critical cone to \( X \) at \( \bar{x} \) by means of

\[
\mathcal{C}_X(\bar{x}) := \{d \in \mathcal{L}_X(\bar{x}) \mid \nabla f(\bar{x}) \cdot d \leq 0\}.
\]  

(8)

Furthermore, we will exploit the so-called Lagrangian function \( L \): \( \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \) of (MPDC) which is given as stated below:

\[
\forall (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m: \quad L(x, \lambda) := f(x) + F(x) \cdot \lambda.
\]

Finally, let us introduce

\[
S(\bar{x}) := \left\{ \lambda \in \bigcap_{i \in I(\bar{x})} \mathcal{N}_{D_i}(F(\bar{x})) \mid \nabla x L(\bar{x}, \lambda) = 0^n \right\},
\]

the set of all multipliers which solve the S-stationarity system associated with (MPDC) at \( \bar{x} \). Clearly, \( \bar{x} \) is an S-stationary point of (MPDC) if and only if \( S(\bar{x}) \) is nonempty.

**Lemma 4.1.** Let \( \bar{x} \in X \) be an S-stationary point of (MPDC). Then, we have

\[
\forall \lambda \in S(\bar{x}): \quad \mathcal{C}_X(\bar{x}) = \left\{d \in \mathbb{R}^n \mid \nabla F(\bar{x})d \in T_D(F(\bar{x})) \cap \{\lambda\}^\perp\right\}.
\]
Proof. For any \( d \in \mathcal{C}(\bar{x}) \) and \( \lambda \in S(\bar{x}) \), we obtain
\[
0 \geq \nabla f(\bar{x}) \cdot d = (-\nabla F(\bar{x})^\top \lambda) \cdot d = -\underbrace{(\nabla F(\bar{x})_d) \cdot \lambda \geq 0}_{\in \mathcal{T}_D(F(\bar{x}))}
\]
from \( \lambda \in \bigcap_{i \in I(\bar{x})} \hat{N}_{D_i}(F(\bar{x})) = \hat{N}_D(F(\bar{x})) = \mathcal{T}_D(F(\bar{x}))^o \), see Lemma 2.2. This yields \( \nabla F(\bar{x}) d \in \{\lambda\}^+ \) and shows the inclusion \( \subset \).

If, on the other hand, \( d \in \mathbb{R}^n \) satisfies \( \nabla F(\bar{x}) d \in \mathcal{T}_D(F(\bar{x})) \cap \{\lambda\}^+ \) for some \( \lambda \in S(\bar{x}) \), then we have \( d \in \mathcal{L}_X(\bar{x}) \) by definition of the linearization cone and
\[
0 = (\nabla F(\bar{x}) d) \cdot \lambda = (\nabla F(\bar{x})^\top \lambda) \cdot d = -\nabla f(\bar{x}) \cdot d
\]
which yields \( d \in \mathcal{C}(\bar{x}) \).

Using the theory on second-order tangent sets provided earlier, we are now in position to state a second-order necessary optimality condition for (MPDC) under validity of MPDC-LICQ. Thus, our approach is closely related to the approaches used in Bonnans and Shapiro (2000); Christof and Wachsmuth (2018); Penot (1998); Rockafellar and Wets (1998) for the derivation of second-order necessary optimality conditions for different classes of mathematical programs in finite- and infinite-dimensional Banach spaces. It seems to be worth mentioning that, in contrast to (Hoheisel and Kanzow, 2007, Theorem 4.3) where a second-order necessary optimality conditions for MPVCs is shown, we do not use an implicit function argument for our proof. Some parts of the upcoming theorem’s proof are inspired by (Christof and Wachsmuth, 2018, Lemma 5.8).

**Theorem 4.2.** Let \( \bar{x} \in X \) be a locally optimal solution of (MPDC) where MPDC-LICQ is valid. Then, we have
\[
\forall d \in \mathcal{C}(\bar{x}): \quad d^\top \nabla^2 x L(\bar{x}, \bar{\lambda}) d \geq 0
\]
where \( \bar{\lambda} \in S(\bar{x}) \) is the uniquely determined multiplier which solves the \( S \)-stationarity system associated with \( \bar{x} \), see Theorem 3.7.

**Proof.** First, we will prove the correctness of
\[
\forall d \in \mathcal{C}(\bar{x}) \forall h \in \mathcal{T}_X^2(\bar{x}; d): \quad \nabla f(\bar{x}) \cdot h + d^\top \nabla^2 f(\bar{x}) d \geq 0. \tag{9}
\]
Therefore, fix \( d \in \mathcal{C}(\bar{x}) \) and \( h \in \mathcal{T}_X^2(\bar{x}; d) \). Then, we find sequences \( \{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+ \) and \( \{h_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n \) such that \( t_k \downarrow 0 \), \( h_k \to h \), and \( \bar{x} + t_k d + \frac{1}{2} t_k^2 h_k \in X \) for all \( k \in \mathbb{N} \). Performing a second-order Taylor expansion of \( f \) at \( \bar{x} \) yields
\[
f(\bar{x} + t_k d + \frac{1}{2} t_k^2 h_k) = f(\bar{x}) + t_k \nabla f(\bar{x}) \cdot d + \frac{1}{2} t_k^2 (\nabla^2 f(\bar{x}) \cdot h_k + d^\top \nabla^2 f(\bar{x}) d) + o(t_k^2)
\]
for all \( k \in \mathbb{N} \). Noting that we have \( f(\bar{x} + t_k d + \frac{1}{2} t_k^2 h_k) \geq f(\bar{x}) \) for sufficiently large \( k \in \mathbb{N} \) from the local optimality of \( \bar{x} \) for (MPDC) while \( \nabla f(\bar{x}) \cdot d \leq 0 \) holds by definition of the critical cone, we obtain
\[
0 \leq \frac{2}{t_k^2} \left( f(\bar{x} + t_k d + \frac{1}{2} t_k^2 h_k) - f(\bar{x}) - t_k \nabla f(\bar{x}) \cdot d \right) = \nabla f(\bar{x}) \cdot h_k + d^\top \nabla^2 f(\bar{x}) d + 2 \frac{o(t_k^2)}{t_k^2}
\]

16
Lemma 2.5. Noting that we have \( \sum \), summarizing these considerations, we have shown \( \lambda \cdot \ell \) for sufficiently large \( k \in \mathbb{N} \). Thus, taking the limit \( k \to \infty \) yields (9).

Due to validity of MPDC-LICQ, (9) implies that

\[
\inf \left\{ \nabla f(\bar{x}) \cdot h \left| \nabla F(\bar{x}) h \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} \right. \right\} + d^T \nabla^2 f(\bar{x}) d \geq 0
\]

holds true for all \( d \in C_X(\bar{x}) \), see Lemma 3.10. By definition of S-stationarity, we have \( \nabla f(\bar{x}) = -\nabla F(\bar{x})^T \lambda \) which yields

\[
\inf \left\{ -\lambda \cdot w \left| w \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} \right. \right\} + d^T \nabla^2 f(\bar{x}) d \geq 0
\]

for each \( d \in C_X(\bar{x}) \).

Due to validity of MPDC-LICQ, for each \( w \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} \), we find \( v \in \mathbb{R}^n \) and \( \ell \in \bigcap_{i \in I(\bar{x})} \mathcal{T}_D_i(F(\bar{x})^o \perp \) such that \( w = \nabla F(\bar{x}) v - \ell \) holds true, see Lemma 3.3. Noting that we have

\[
\bigcap_{i \in I(\bar{x})} \mathcal{T}_D_i(F(\bar{x})^o \perp = \bigcap_{i \in I(\bar{x})} \tilde{\mathcal{N}}_D_i(F(\bar{x}))^\perp \subset \bigcup_{i \in I(\bar{x})} \tilde{\mathcal{N}}_D_i(F(\bar{x}))^\perp = \tilde{\mathcal{N}}_D(F(\bar{x}))^\perp,
\]

see Lemma 2.2, the relation \( \lambda \cdot \ell = 0 \) follows from \( \lambda \in \tilde{\mathcal{N}}_D(F(\bar{x})) \). Furthermore, we infer

\[
\nabla F(\bar{x}) v = w + \ell \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} + \bigcap_{i \in I(\bar{x})} \mathcal{T}_D_i(F(\bar{x})^o \perp
\]

\[
= T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) + \bigcap_{i \in I(\bar{x})} \mathcal{T}_D_i(F(\bar{x})^o \perp - \{\nabla^2 F(\bar{x})[d, d]\}
\]

\[
\subset T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\}
\]

\[
= T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\}
\]

from Lemma 2.5. Summarizing these considerations, we have shown \( \nabla F(\bar{x}) v \in \nabla F(\bar{x}) \mathbb{R}^n \) and \( \nabla F(\bar{x}) v \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} \). Furthermore, we obtain the relation

\[
-\lambda \cdot w = -\lambda(\nabla F(\bar{x}) v - \ell) = -\lambda \cdot (\nabla F(\bar{x}) v). \quad \text{This leads to}
\]

\[
\inf \left\{ -\lambda \cdot w \left| w \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} \right. \right\} \leq \inf \left\{ -\lambda \cdot w \left| w \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} \right. \right\}.
\]

The converse inequality, however, is trivial. Thus, equality holds for the optimal values of the above programs and we obtain

\[
\inf \left\{ -\lambda \cdot w \left| w \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} \right. \right\} + d^T \nabla^2 f(\bar{x}) d \geq 0
\]

for each \( d \in C_X(\bar{x}) \) from (10). Clearly, we have

\[
\inf \left\{ -\lambda \cdot w \left| w \in T_{\mathcal{T}}(\nabla F(\bar{x}))(\nabla F(\bar{x})d) - \{\nabla^2 F(\bar{x})[d, d]\} \right. \right\}
\]

17
Finally, we note that
\[
\mathcal{T}_D(F(\bar{x}))(\nabla F(\bar{x})d) = \bigcup_{i \in I(\bar{x})} \mathcal{T}_{D_i}(F(\bar{x}))(\nabla F(\bar{x})d)
\]
holds true invoking Lemma 2.2 while noticing that the sets \( \mathcal{T}_{D_i}(F(\bar{x})), i \in I(\bar{x}), \) are closed, convex, polyhedral cones. Thus, for any \( w \in \mathcal{T}_D(F(\bar{x}))(\nabla F(\bar{x})d) \), we find a vector \( r \in \mathcal{T}_D(F(\bar{x})) \) and \( \alpha \geq 0 \) such that \( w = r - \alpha \nabla F(\bar{x})d \) holds. Recalling \( \bar{\lambda} \in \hat{N}_D(F(\bar{x})) \) and \( d \in \mathcal{C}_X(\bar{x}) \), we have
\[
-\bar{\lambda} \cdot w = -\bar{\lambda} \cdot (r - \alpha \nabla F(\bar{x})d) \geq \alpha (\nabla F(\bar{x})^\top \bar{\lambda}) \cdot d = \alpha (\nabla f(\bar{x})) \cdot d \geq 0
\]
by definition of S-stationarity, i.e.
\[
\inf \{-\bar{\lambda} \cdot w \mid w \in \mathcal{T}_D(F(\bar{x}))(\nabla F(\bar{x})d)\} + \bar{\lambda} \cdot \nabla^2 F(\bar{x})[d, d] = \bar{\lambda} \cdot \nabla^2 F(\bar{x})[d, d]
\]
follows for each \( d \in \mathcal{C}_X(\bar{x}) \). Combining this with the above arguments, the desired result follows from (11) by definition of the Lagrangian function. This completes the proof. \( \square \)

The above result can be seen as a special instance of (Gfrerer, 2014, Theorem 3.3) where a second-order necessary optimality condition for (MPDC) has been derived using a completely different approach via the variational concepts of the directional limiting normal cone and directional metric subregularity. One can easily check that by demanding validity of MPDC-LICQ at a given local minimizer of (MPDC), the assumptions of (Gfrerer, 2014, Theorem 3.3) hold as well, i.e. the assumptions of Theorem 4.2 are more restrictive. On the other hand, one has to mention that checking validity of MPDC-LICQ and noting that this implies that there is only one S-stationary multiplier, the second-order necessary optimality condition from Theorem 4.2 seems to be much easier to verify than the one from Gfrerer (2014).

Next, we state a second-order sufficient optimality condition for (MPDC). Although this result follows from (Gfrerer, 2014, Theorem 3.21), we provide a completely elementary and simple proof here which generalizes a well-known strategy which has been used to verify second-order sufficient optimality conditions for NLPs, MPCCs, MPVCs, and CCMPs in the past.

**Theorem 4.3.** Let \( \bar{x} \in X \) be an S-stationary point of (MPDC) where the condition
\[
\forall d \in \mathcal{C}_X(\bar{x}) \setminus \{0^n\} \exists \lambda \in S(\bar{x}) : \quad d^\top \nabla^2_{xx} L(\bar{x}, \lambda) d > 0 \tag{12}
\]
holds. Then, there are constants \( \varepsilon > 0 \) and \( C > 0 \) such that the following quadratic-growth-condition is valid:
\[
\forall x \in X \cap B^\varepsilon(\bar{x}) : \quad f(x) \geq f(\bar{x}) + C \| x - \bar{x} \|_2^2.
\]
Particularly, \( \bar{x} \) is a strict local minimizer of (MPDC).
Proof. Assume on the contrary that there is a sequence \( \{x_k\}_{k \in \mathbb{N}} \subset X \) converging to \( \bar{x} \) such that
\[
\forall k \in \mathbb{N}: \quad f(x_k) < f(\bar{x}) + \frac{1}{2k} \|x_k - \bar{x}\|_2^2
\]
holds true. Set \( t_k := \|x_k - \bar{x}\|_2 > 0 \) and observe that \( \{(x_k - \bar{x})/t_k\}_{k \in \mathbb{N}} \) is bounded. We assume w.l.o.g. that \( (x_k - \bar{x})/t_k \to d \) holds for some \( d \in \mathbb{R}^n \setminus \{0^n\} \). By construction, \( d \in T_X(\bar{x}) \subset L_X(\bar{x}) \) is guaranteed. For each \( k \in \mathbb{N} \), we find \( \xi_k \in \text{conv}\{\bar{x}; x_k\} \) which satisfies \( f(x_k) - f(\bar{x}) = \nabla f(\xi_k) \cdot (x_k - \bar{x}) \) by means of the mean value theorem. Dividing by \( t_k \) and taking the limit \( k \to \infty \) while observing that \( \nabla f: \mathbb{R}^n \to \mathbb{R}^n \) is continuous, we have
\[
\nabla f(\bar{x}) \cdot d = \lim_{k \to \infty} \nabla f(\xi_k) \cdot \frac{x_k - \bar{x}}{t_k} = \lim_{k \to \infty} \frac{f(x_k) - f(\bar{x})}{t_k} \leq \lim_{k \to \infty} \frac{1}{2k} \|x_k - \bar{x}\|_2^2 = 0.
\]
This yields \( d \in C_X(\bar{x}) \setminus \{0^n\} \).

Choose \( \lambda \in S(\bar{x}) \) arbitrarily. Then, we have \( \lambda \in \bigcap_{i \in I(\bar{x})} \widehat{N}_D_i(F(\bar{x})) \). For sufficiently large \( k \in \mathbb{N} \), \( I(x_k) \subset I(\bar{x}) \) holds true. Thus, for sufficiently large \( k \in \mathbb{N} \) and \( i \in I(x_k) \), we have \( \lambda \in \bigcap_{i \in I(\bar{x})} (\widehat{N}_D_i(F(\bar{x})))^\circ \) which shows \( (F(x_k) - F(\bar{x})) \cdot \lambda \leq 0 \). This yields
\[
f(\bar{x}) > f(x_k) - \frac{1}{k} \|x_k - \bar{x}\|_2^2 \geq f(x_k) + (F(x_k) - F(\bar{x})) \cdot \lambda - \frac{1}{k} \|x_k - \bar{x}\|_2^2
\]
for sufficiently large \( k \in \mathbb{N} \). Rearranging some terms and applying Taylor’s theorem, we derive
\[
L(\bar{x}, \lambda) > L(x_k, \lambda) - \frac{1}{k} \|x_k - \bar{x}\|_2^2
\]
\[
= L(\bar{x}, \lambda) + \nabla_x L(\bar{x}, \lambda)(x_k - \bar{x}) + \frac{1}{2}(x_k - \bar{x})^\top \nabla^2_{xx} L(\bar{x}, \lambda)(x_k - \bar{x}) + o(\|x_k - \bar{x}\|_2^2).
\]
Now, we exploit the choice \( \lambda \in S(\bar{x}) \) in order to infer
\[
0 > \frac{1}{2}(x_k - \bar{x})^\top \nabla^2_{xx} L(\bar{x}, \lambda)(x_k - \bar{x}) + o(\|x_k - \bar{x}\|_2^2)
\]
for sufficiently large \( k \in \mathbb{N} \). Division by \( t_k^2 \) and taking the limit \( k \to \infty \) yield
\[
0 \geq \frac{1}{2} d^\top \nabla^2_{xx} L(\bar{x}, \lambda)d
\]
which contradicts the theorem’s assumptions since we have shown \( d \in C_X(\bar{x}) \setminus \{0^n\} \) while \( \lambda \in S(\bar{x}) \) was arbitrarily chosen. This completes the proof. \( \square \)

The above result justifies the following definition.

**Definition 4.4.** Let \( \bar{x} \in X \) be an S-stationary point of (MPDC). Then, the MPDC-tailored second-order sufficient condition (MPDC-SOSC for short) holds at \( \bar{x} \) if and only if (12) is valid.

The upcoming theorem shows that S-stationary points of (MPDC), where both MPDC-LICQ and MPDC-SOSC are valid, are locally isolated w.r.t. primal and dual variables. This property does not generally follow from the second-order growth condition as (Guo et al., 2013, Example 4.1), which has been stated in the context of MPCCs, indicates. For the validation of the upcoming result, we generalize the proof of (Guo et al., 2013, Theorem 4.1).
Theorem 4.5. Let $\bar{x} \in X$ be a feasible point of (MPDC) where MPDC-LICQ and MPDC-SOSC are valid. Furthermore, assume that $D_1, \ldots, D_r$ are polyhedral cones. Then, there is some $\varepsilon > 0$ such that

$$\forall x \in \mathbb{B}^\varepsilon(\bar{x}): \lambda \in S(x) \implies x = \bar{x} \land \lambda = \bar{\lambda}$$

where $\bar{\lambda}$ is the uniquely determined vector from $S(\bar{x})$.

Proof. First, we invoke Theorem 3.7 which shows that $\bar{x}$ is an S-stationary point of (MPDC) with uniquely determined S-stationary multiplier $\bar{\lambda}$.

Assume on the contrary, that we can find a sequence $\{x_k\}_{k \in \mathbb{N}} \subset X$ of feasible and S-stationary points of (MPDC) converging to $\bar{x}$. Then, we find $\lambda_k \in S(x_k)$ for each $k \in \mathbb{N}$.

Suppose that $\{\lambda_k\}_{k \in \mathbb{N}}$ is not bounded, i.e. we can assume w.l.o.g. that $\|\lambda_k\|_2 \to \infty$ holds as $k \to \infty$. Thus, we can define $\bar{\lambda}_k := \lambda_k/\|\lambda_k\|_2$ for sufficiently large $k \in \mathbb{N}$ and due to the boundedness of $\{\bar{\lambda}_k\}_{k \in \mathbb{N}}$, we may assume w.l.o.g. that this sequence converges to some nonvanishing vector $\bar{\lambda} \in \mathbb{R}^m$. Furthermore, we have

$$\nabla F(\bar{x})^\top \bar{\lambda} = \lim_{k \to \infty} \nabla F(x_k)^\top \bar{\lambda}_k = \lim_{k \to \infty} \frac{1}{\|\lambda_k\|_2} \left( \nabla f(x_k) + \nabla F(x_k)^\top \lambda_k \right) = 0^n$$

by continuity of $\nabla f: \mathbb{R}^n \to \mathbb{R}^n$ and $\nabla F: \mathbb{R}^n \to \mathbb{R}^{m \times n}$ as well as $\lambda_k \in S(x_k)$. On the other hand, the inclusion $I(x_k) \subset I(\bar{x})$ is valid for all sufficiently large $k \in \mathbb{N}$ and, clearly, $I(x_k) \neq \emptyset$ is true as well since $x_k$ is feasible to (MPDC) for each $k \in \mathbb{N}$. Noting that there are only finitely many indices in $I(\bar{x})$, there must exist some $i_0 \in I(\bar{x})$ such that $i_0 \in I(x_k)$ holds along a subsequence $\{x_{k_l}\}_{l \in \mathbb{N}}$ of $\{x_k\}_{k \in \mathbb{N}}$. The definition of S-stationarity and the fact that the Fréchet normal cone is a cone yield $\bar{\lambda}_{k_l} \in \hat{N}_{D_{i_0}}(F(x_{k_l}))$. Now, the continuity of $F$ can be used to infer

$$\bar{\lambda} \in \mathcal{N}_{D_{i_0}}(F(\bar{x})) = \hat{N}_{D_{i_0}}(F(\bar{x})) \subset \bigcup_{i \in I(\bar{x})} \text{span} \hat{N}_{D_i}(F(\bar{x})).$$

Keeping $\nabla F(\bar{x})^\top \bar{\lambda} = 0^n$ and $\bar{\lambda} \neq 0^m$ in mind, this contradicts MPDC-LICQ.

Due to the above arguments, we may assume w.l.o.g. that $\{\lambda_k\}_{k \in \mathbb{N}}$ converges to some $\lambda \in \mathbb{R}^m$. Similar arguments as above show the existence of $i_0 \in I(\bar{x})$ such that $\lambda \in \hat{N}_{D_{i_0}}(F(\bar{x}))$ holds true. Moreover, from $\nabla f(x_k) + \nabla F(x_k)^\top \lambda_k = 0^n$ we obtain $\nabla f(\bar{x}) + \nabla F(\bar{x})^\top \lambda = 0^n$ since $f$ and $F$ possess continuous derivatives. Keeping $\nabla f(\bar{x}) + \nabla F(\bar{x})^\top \lambda = 0^n$ in mind, we derive $\nabla F(\bar{x})^\top (\lambda - \lambda) = 0^n$. Moreover,

$$\bar{\lambda} - \lambda \in \left( \bigcap_{i \in I(\bar{x})} \hat{N}_{D_i}(F(\bar{x})) \right) - \hat{N}_{D_{i_0}}(F(\bar{x}))$$

$$\subset \hat{N}_{D_{i_0}}(F(\bar{x})) - \hat{N}_{D_{i_0}}(F(\bar{x})) = \text{span} \hat{N}_{D_{i_0}}(F(\bar{x})) \subset \bigcup_{i \in I(\bar{x})} \text{span} \hat{N}_{D_i}(F(\bar{x}))$$

20
follows, and by validity of MPDC-LICQ, \( \lambda = \bar{\lambda} \) is obtained.

We set \( t_k := \|x_k - \bar{x}\| > 0 \) and observe that \( \{(x_k - \bar{x})/t_k\}_{k \in \mathbb{N}} \) is a bounded sequence that converges w.l.o.g. to some nonvanishing direction \( d \in \mathbb{R}^n \). Since \( \{x_k\}_{k \in \mathbb{N}} \subset X \) holds, we infer \( d \in T_X(\bar{x}) \setminus \{0^n\} \subset T_X(\bar{x}) \setminus \{0^n\} \). From \( \lambda_k \in \bar{N}_D(F(x_k)) \) for all \( k \in \mathbb{N} \), \( \bar{\lambda} \in \bar{N}_D(F(\bar{x})) \), \( x_k \to \bar{x} \), and \( \lambda_k \to \bar{\lambda} \), we obtain \( F(\bar{x}) \cdot \bar{\lambda} = 0, F(x_k) \cdot \lambda_k = 0 \) for all \( k \in \mathbb{N} \), as well as \( F(x_k) \cdot \bar{\lambda} = F(\bar{x}) \cdot \lambda_k = 0 \) for all sufficiently large \( k \in \mathbb{N} \), see Lemma 2.3. Here, we used the assumption that \( D_1, \ldots, D_r \) are polyhedral cones. This yields

\[
(\nabla F(\bar{x})d) \cdot \bar{\lambda} = \left( \lim_{k \to \infty} \frac{\nabla F(x_k)(x_k - \bar{x})}{t_k} \right) \cdot \bar{\lambda} = \left( \lim_{k \to \infty} \frac{F(x_k) - F(\bar{x})}{t_k} \right) \cdot \bar{\lambda} = \lim_{k \to \infty} \frac{F(x_k) \cdot \bar{\lambda} - F(\bar{x}) \cdot \bar{\lambda}}{t_k} = 0,
\]

i.e. \( F(\bar{x})d \in \{\bar{\lambda}\}^\perp \) holds true. By means of Lemma 4.1, we deduce \( d \in C_X(\bar{x}) \setminus \{0^n\} \).

For each \( k \in \mathbb{N} \), let us define a continuously differentiable function \( \varphi_k : [0,1] \to \mathbb{R} \) by means of

\[
\forall s \in [0,1]: \quad \varphi_k(s) := \nabla_x L((1 - s)(\bar{x}, \bar{\lambda}) + s(x_k, \lambda_k)) \cdot (x_k - \bar{x}) - L((1 - s)x + sx_k, \lambda_k) + L((1 - s)\bar{x} + s\bar{x}, \bar{\lambda}).
\]

Due to the above remarks, we have

\[
\varphi_k(0) = \nabla_x L(\bar{x}, \bar{\lambda}) \cdot (x_k - \bar{x}) - L(\bar{x}, \lambda_k) + L(\bar{x}, \bar{\lambda}) = -f(\bar{x}) + f(\bar{x}) = 0,
\]

\[
\varphi_k(1) = \nabla_x L(x_k, \lambda_k) \cdot (x_k - \bar{x}) - L(x_k, \lambda_k) + L(x_k, \bar{\lambda}) = -f(x_k) + f(x_k) = 0
\]

for sufficiently large \( k \in \mathbb{N} \). Applying Rolle’s theorem, there must exist \( s_k \in (0,1) \) such that

\[
0 = \varphi_k'(s_k)
\]

\[
= (x_k - \bar{x})^\top \nabla_{xx}^2 L((1 - s_k)(\bar{x}, \bar{\lambda}) + s_k(x_k, \lambda_k))(x_k - \bar{x})
\]

\[
+ (x_k - \bar{x})^\top \nabla_{xx,\lambda}^2 L((1 - s_k)(\bar{x}, \bar{\lambda}) + s_k(x_k, \lambda_k))((x_k - \bar{x}) - \bar{\lambda})
\]

\[
- \nabla_x L((1 - s_k)\bar{x} + s_kx_k, \lambda_k) \cdot (x_k - \bar{x}) + \nabla_x L((1 - s_k)\bar{x} + s_kx_k, \bar{\lambda}) \cdot (x_k - \bar{x})
\]

\[
= (x_k - \bar{x})^\top \nabla_{xx}^2 L((1 - s_k)(\bar{x}, \bar{\lambda}) + s_k(x_k, \lambda_k))(x_k - \bar{x})
\]

\[
+ (x_k - \bar{x})^\top [\nabla F((1 - s_k)x + s_kx_k)^\top (\lambda_k - \bar{\lambda})]
\]

\[
- [\nabla F((1 - s_k)x + s_kx_k)^\top (\lambda_k - \bar{\lambda})] \cdot (x_k - \bar{x})
\]

\[
= (x_k - \bar{x})^\top \nabla_{xx}^2 L((1 - s_k)(\bar{x}, \bar{\lambda}) + s_k(x_k, \lambda_k))(x_k - \bar{x})
\]

holds for all \( k \in \mathbb{N} \) which are sufficiently large. Next, we observe that the relation \((1 - s_k)(\bar{x}, \bar{\lambda}) + s_k(x_k, \lambda_k) \to (\bar{x}, \bar{\lambda})\) holds true as \( k \to \infty \). From above, it follows

\[
0 = \left(\frac{x_k - \bar{x}}{t_k}\right)^\top \nabla_{xx}^2 L((1 - s_k)(\bar{x}, \bar{\lambda}) + s_k(x_k, \lambda_k)) \left(\frac{x_k - \bar{x}}{t_k}\right)
\]

21
for sufficiently large $k \in \mathbb{N}$, i.e. taking the limit $k \to \infty$ yields $0 = d^T \nabla_{xx}^2 L(\bar{x}, \bar{\lambda}) d$. This, however, contradicts the validity of MPDC-SOSC since we already verified that $d \in \mathcal{L}(\bar{x}) \setminus \{0^n\}$ holds true. Thus, the proof is completed.

It remains an open question whether the assertion of Theorem 4.5 even holds in the situation where the disjunctive set $D$ is not conic. In the proof provided above, this assumption is indispensable.

5 Consequences for certain classes of disjunctive programs

In this section, we are going to apply the obtained results to some prominent classes of disjunctive programs, namely MPCCs, MPVCs, CCMPs, and MP SCs in order to check how the above theory relates to existing results in the available literature on these problem classes. Throughout the section, we consider twice continuously differentiable functions $f: \mathbb{R}^n \to \mathbb{R}$, $g: \mathbb{R}^n \to \mathbb{R}^p$, $h: \mathbb{R}^n \to \mathbb{R}^q$, and $G, H: \mathbb{R}^n \to \mathbb{R}^l$. The component mappings of $g, h, G,$ and $H$ will be denoted by $g_j: \mathbb{R}^n \to \mathbb{R}$, $j = 1, \ldots, p$, $h_j: \mathbb{R}^n \to \mathbb{R}$, $j = 1, \ldots, q$, and $G_j, H_j: \mathbb{R}^n \to \mathbb{R}$, $j = 1, \ldots, l$, respectively.

5.1 Application to MPCCs

A mathematical program with complementarity constraints is an optimization problem of the form

$$\begin{align*}
  f(x) & \to \min \\
  g_j(x) & \leq 0 \quad j = 1, \ldots, p \\
  h_j(x) & = 0 \quad j = 1, \ldots, q \\
  0 & \leq G_j(x) \perp H_j(x) \geq 0 \quad j = 1, \ldots, l.
\end{align*}$$

(MPCC)

Due to the frequent appearance of (MPCC) as an abstract model of real-world applications, this problem class has been studied intensively from the theoretical and numerical point of view during the last two decades, see e.g. Gfrerer (2014); Hoheisel et al. (2013); Luo et al. (1996); Outrata et al. (1998); Scheel and Scholtes (2000); Ye (2005) and the references therein.

In order to transfer (MPCC) into a program of type (MPDC), we introduce the sets $S_1^{CC} := \mathbb{R}_+ \times \{0\}$ and $S_2^{CC} := \{0\} \times \mathbb{R}_+$ as well as $\mathcal{J} := \{1, 2\}^l$. Next, we set

$$
\forall \alpha \in \mathcal{J}: \quad D_\alpha^{CC} := \mathbb{R}^p \times \{0^q\} \times \prod_{j=1}^l S_{\alpha_j}^{CC},
$$

and $D^{CC} := \bigcup_{\alpha \in \mathcal{J}} D_\alpha^{CC}$. Furthermore, we introduce $F: \mathbb{R}^n \to \mathbb{R}^{p+q+2l}$ by means of

$$
\forall x \in \mathbb{R}^n: \quad F(x) := \begin{bmatrix} g(x)^T & h(x)^T & G_1(x) & H_1(x) & \ldots & G_l(x) & H_l(x) \end{bmatrix}^T.
$$

Then, the feasible set of (MPCC) is given by $X^{CC} := \{ x \in \mathbb{R}^n \mid F(x) \in D^{CC} \}$, see Figure 1 for an illustration. For a feasible point $\bar{x} \in X^{CC}$ of (MPCC), let us introduce
the following well-known index sets:

\[
\begin{align*}
I^{+0}(\bar{x}) & := \{ j \in \{1, \ldots, l\} \mid G_j(\bar{x}) > 0 \land H_j(\bar{x}) = 0 \}, \\
I^{0+}(\bar{x}) & := \{ j \in \{1, \ldots, l\} \mid G_j(\bar{x}) = 0 \land H_j(\bar{x}) > 0 \}, \\
I^{00}(\bar{x}) & := \{ j \in \{1, \ldots, l\} \mid G_j(\bar{x}) = 0 \land H_j(\bar{x}) = 0 \}.
\end{align*}
\]

We exploit the calculus rules for the tangent and Fréchet normal cone to Cartesian products of (convex) sets, see (Rockafellar and Wets, 1998, Proposition 6.41), in order to obtain

\[
\begin{align*}
\mathcal{T}_D^{\alpha CC}(F(\bar{x})) &= \mathcal{T}_{\mathbb{R}^p}(g(\bar{x})) \times \mathcal{T}_{\{0\}}(h(\bar{x})) \times \prod_{j=1}^l \mathcal{T}_{S^{CC}_{\alpha j}}((G_j(\bar{x}), H_j(\bar{x}))^\top), \\
\hat{\mathcal{N}}_D^{\alpha CC}(F(\bar{x})) &= \hat{\mathcal{N}}_{\mathbb{R}^p}(g(\bar{x})) \times \hat{\mathcal{N}}_{\{0\}}(h(\bar{x})) \times \prod_{j=1}^l \hat{\mathcal{N}}_{S^{CC}_{\alpha j}}((G_j(\bar{x}), H_j(\bar{x}))^\top)
\end{align*}
\]

for each \( \alpha \in \mathcal{J} \). The tangent and Fréchet normal cones to the sets \( \mathbb{R}_-^p \) and \( \{0\} \) have been characterized in Example 3.2 already. A straightforward calculation shows

\[
\begin{align*}
\mathcal{T}_{S^1_{\alpha}}((G_j(\bar{x}), H_j(\bar{x}))^\top) &= \begin{cases} \mathbb{R} \times \{0\} & j \in I^{+0}(\bar{x}), \\
\emptyset & j \in I^{0+}(\bar{x}), \\
\mathbb{R}_+ \times \{0\} & j \in I^{00}(\bar{x}), \end{cases} \\
\mathcal{T}_{S^2_{\alpha}}((G_j(\bar{x}), H_j(\bar{x}))^\top) &= \begin{cases} \emptyset & j \in I^{+0}(\bar{x}), \\
\{0\} \times \mathbb{R} & j \in I^{0+}(\bar{x}), \\
\{0\} \times \mathbb{R}_+ & j \in I^{00}(\bar{x}), \end{cases} \\
\hat{\mathcal{N}}_{S^1_{\alpha}}((G_j(\bar{x}), H_j(\bar{x}))^\top) &= \begin{cases} \{0\} \times \mathbb{R} & j \in I^{+0}(\bar{x}), \\
\emptyset & j \in I^{0+}(\bar{x}), \\
\mathbb{R}_- \times \mathbb{R} & j \in I^{00}(\bar{x}), \end{cases} \\
\hat{\mathcal{N}}_{S^2_{\alpha}}((G_j(\bar{x}), H_j(\bar{x}))^\top) &= \begin{cases} \emptyset & j \in I^{+0}(\bar{x}), \\
\mathbb{R} \times \{0\} & j \in I^{0+}(\bar{x}), \\
\mathbb{R} \times \mathbb{R}_- & j \in I^{00}(\bar{x}). \end{cases}
\end{align*}
\]
Using \( I(\bar{x}) := \{ \alpha \in \mathcal{J} \mid F(\bar{x}) \in D^C_\alpha \} \), we obtain the characterization

\[ \alpha \in I(\bar{x}) \iff \forall j \in I^{+0}(\bar{x}) : \alpha_j = 1 \land \forall j \in I^{0+}(\bar{x}) : \alpha_j = 2 \]

for arbitrary \( \alpha \in \mathcal{J} \). Thus, MPDC-LICQ from Definition 3.1 takes the following form for problem (MPCC) at the reference point \( \bar{x} \):

\[
\begin{align*}
0^0 &= \nabla g(\bar{x})^\top \lambda + \nabla h(\bar{x})^\top \rho + \nabla G(\bar{x})^\top \mu + \nabla H(\bar{x})^\top \nu, \\
\forall j \notin I^0(\bar{x}) : \lambda_j &= 0, \\
\forall j \in I^{+0}(\bar{x}) : \mu_j &= 0, \\
\forall j \in I^{0+}(\bar{x}) : \nu_j &= 0
\end{align*}
\]

Here, the appearing index set \( I^0(\bar{x}) \) has been defined in Example 3.2. The above condition is equivalent to the linear independence of the vectors from

\[
\{ \nabla g_j(\bar{x}) \mid j \in I^0(\bar{x}) \} \cup \{ \nabla h_j(\bar{x}) \mid j \in \{1, \ldots, q\} \}
\]

\[
\cup \{ \nabla G_j(\bar{x}) \mid j \in I^{0+}(\bar{x}) \cup I^{00}(\bar{x}) \} \cup \{ \nabla H_j(\bar{x}) \mid j \in I^{+0}(\bar{x}) \cup I^{00}(\bar{x}) \}.
\]

This, however, is precisely the definition of the prominent constraint qualification MPCC-LICQ, see e.g. (Ye, 2005, Definition 2.8). The associated S-stationarity system from Definition 3.8 reads as

\[
0^0 = \nabla f(\bar{x}) + \nabla g(\bar{x})^\top \lambda + \nabla h(\bar{x})^\top \rho + \nabla G(\bar{x})^\top \mu + \nabla H(\bar{x})^\top \nu,
\]

\[
\lambda \geq 0, \forall j \notin I^0(\bar{x}) : \lambda_j = 0, \\
\forall j \in I^{+0}(\bar{x}) : \mu_j = 0, \\
\forall j \in I^{0+}(\bar{x}) : \nu_j = 0, \\
\forall j \in I^{00}(\bar{x}) : \mu_j, \nu_j \leq 0
\]

and equals the MPCC-tailored system of strong stationarity, see (Ye, 2005, Definition 2.7). Using the above formulas for the appearing tangent cones, the linearization cone from (1) is given by

\[
\mathcal{L}_{Xcc}(\bar{x}) = \begin{Bmatrix}
\{ d \in \mathbb{R}^n \mid \\
\nabla g_j(\bar{x}) \cdot d \leq 0 & j \in I^0(\bar{x}) \\
\nabla h_j(\bar{x}) \cdot d = 0 & j \in \{1, \ldots, q\} \\
\nabla G_j(\bar{x}) \cdot d = 0 & j \in I^{0+}(\bar{x}) \\
\nabla H_j(\bar{x}) \cdot d = 0 & j \in I^{+0}(\bar{x}) \\
0 \leq \nabla G_j(\bar{x}) \cdot d \perp \nabla H(\bar{x}) \cdot d \geq 0 & j \in I^{00}(\bar{x})
\end{Bmatrix}
\]
Theorems 3.7\(\text{Guo et al.}, 2013\) MPVC can be found in e.g. Hoheisel and Kanzow 4.2 Palagachev and Gerdts Izmailov and Solodov Theorem 4.5 Kirches Achtziger and Kanzow arise when searching for the optimal design of a truss structure or in. Thus, Achtziger and Kanzow MPDC recover results from the classical paper Achtziger et al. 2009 provided results of Palagachev and Gerdts 2012 while the stability result from Theorem 4.5 can be found in slightly stronger form in (Guo et al., 2013, Theorem 4.1).

5.2 Application to MPVCs

An optimization problem of type

\[ f(x) \to \min \]
\[ g_j(x) \leq 0 \quad j = 1, \ldots, p \]
\[ h_j(x) = 0 \quad j = 1, \ldots, q \]
\[ H_j(x) \geq 0 \quad j = 1, \ldots, l \]
\[ G_j(x)H_j(x) \leq 0 \quad j = 1, \ldots, l \]

is called a mathematical program with vanishing constraints. The term vanishing reflects the observation that whenever a point \(x \in \mathbb{R}^n\) satisfies \(H_j(x) = 0\) for some \(j \in \{1, \ldots, l\}\), then the constraint \(G_j(x)H_j(x) \leq 0\) is trivially satisfied. Problems of type (MPVC) arise when searching for the optimal design of a truss structure or in the context of mixed-integer optimal control, see Achtziger and Kanzow (2008); Kirches (2011); Palagachev and Gerdts (2015). Theoretical and numerical results on problems of type (MPVC) can be found in e.g. Achtziger and Kanzow (2008); Achtziger et al. (2012); Hoheisel (2009); Hoheisel and Kanzow (2007); Hoheisel et al. (2012); Izmailov and Solodov (2009).

Again, we want to transfer (MPVC) into a problem of type (MPDC). Therefore, we define \(S_1^{VC}, S_2^{VC} \subset \mathbb{R}^2\) by means of \(S_1^{VC} := \mathbb{R}_+ \times \{0\}\) and \(S_2^{VC} := \mathbb{R}_- \times \mathbb{R}_+.\) Furthermore, we set \(\mathcal{J} := \{1, 2\}^l\),

\[ \forall \alpha \in \mathcal{J} : \quad D^{VC}_\alpha := \mathbb{R}_+^p \times \{0^q\} \times \prod_{j=1}^l S_{\alpha_j}^{VC}, \]
as well as $D_{\alpha}^{VC} := \bigcup_{\alpha \in \mathcal{J}} D_{\alpha}^{VC}$. Using the function $F$ defined in (13), the feasible set of (MPVC) can be expressed in the compact form $X^{VC} := \{ x \in \mathbb{R}^n \mid F(x) \in D_{\alpha}^{VC} \}$, see Figure 2. Let us fix a feasible point $\bar{x} \in X^{VC}$ of (MPVC). We will exploit the index sets defined below:

$$
I_{+0}(\bar{x}) := \{ j \in \{1, \ldots, l\} \mid H_j(\bar{x}) > 0 \land G_j(\bar{x}) = 0 \},
$$

$$
I_{+}(\bar{x}) := \{ j \in \{1, \ldots, l\} \mid H_j(\bar{x}) > 0 \land G_j(\bar{x}) < 0 \},
$$

$$
I_{0+}(\bar{x}) := \{ j \in \{1, \ldots, l\} \mid H_j(\bar{x}) = 0 \land G_j(\bar{x}) > 0 \},
$$

$$
I_{0-}(\bar{x}) := \{ j \in \{1, \ldots, l\} \mid H_j(\bar{x}) = 0 \land G_j(\bar{x}) < 0 \},
$$

$$
I_{00}(\bar{x}) := \{ j \in \{1, \ldots, l\} \mid H_j(\bar{x}) = 0 \land G_j(\bar{x}) = 0 \}.
$$

Furthermore, we set $I(\bar{x}) := \{ \alpha \in \mathcal{J} \mid F(\bar{x}) \in D_{\alpha}^{VC} \}$. Then, we obtain the following characterization for any $\alpha \in \mathcal{J}$:

$$
\alpha \in I(\bar{x}) \iff \forall j \in I_{0+}(\bar{x}) : \alpha_j = 1 \land \forall j \in I_{+0}(\bar{x}) \cup I_{+}(\bar{x}) \cup I_{0-}(\bar{x}) : \alpha_j = 2.
$$

Exploiting the formulas

$$
T_{S_1^{VC}}((G_j(\bar{x}), H_j(\bar{x}))^\top) = \begin{cases} 
\emptyset & j \in I_{+0}(\bar{x}) \cup I_{+}(\bar{x}) \cup I_{0-}(\bar{x}), \\
\mathbb{R} \times \{0\} & j \in I_{0+}(\bar{x}), \\
\mathbb{R}_+ \times \{0\} & j \in I_{00}(\bar{x}), \\
\mathbb{R}_- \times \mathbb{R} & j \in I_{+0}(\bar{x}), \\
\mathbb{R}^2 & j \in I_{+}(\bar{x}), \\
\emptyset & j \in I_{0+}(\bar{x}), \\
\mathbb{R} \times \mathbb{R}_+ & j \in I_{0-}(\bar{x}), \\
\mathbb{R}_- \times \mathbb{R}_+ & j \in I_{00}(\bar{x}).
\end{cases}
$$

$$
T_{S_2^{VC}}((G_j(\bar{x}), H_j(\bar{x}))^\top) = \emptyset
$$

$$
\tilde{T}_{S_2^{VC}}((G_j(\bar{x}), H_j(\bar{x}))^\top) = \begin{cases} 
\emptyset & j \in I_{+0}(\bar{x}) \cup I_{+}(\bar{x}) \cup I_{0-}(\bar{x}), \\
\{0\} \times \mathbb{R} & j \in I_{0+}(\bar{x}), \\
\mathbb{R}_- \times \mathbb{R} & j \in I_{00}(\bar{x}).
\end{cases}
$$

Figure 2: Geometric illustrations of $X^{VC}$ (left), $S_1^{VC}$ (middle), and $S_2^{VC}$ (right), respectively.
The constraint qualification MPDC-LICQ takes the following form for (MPVC):

\[
\begin{aligned}
0^n &= \nabla g(\bar{x})^\top \lambda + \nabla h(\bar{x})^\top \rho + \nabla G(\bar{x})^\top \mu + \nabla H(\bar{x})^\top \nu, \\
\forall j \notin I^g(\bar{x}) &: \lambda_j = 0, \\
\forall j \in I_{+-}(\bar{x}) \cup I_{0+}(\bar{x}) \cup I_{0-}(\bar{x}) &: \mu_j = 0, \\
\forall j \in I_{00}(\bar{x}) \cup I_{+-}(\bar{x}) &: \nu_j = 0
\end{aligned}
\]

This condition is equivalent to the linear independence of the vectors from

\[
\{\nabla g_j(\bar{x}) | j \in I^g(\bar{x})\} \cup \{\nabla h_j(\bar{x}) | j \in \{1, \ldots, q\}\} \\
\cup \{\nabla G_j(\bar{x}) | j \in I_{00}(\bar{x}) \cup I_{00}(\bar{x})\} \cup \{\nabla H_j(\bar{x}) | j \in I_{00}(\bar{x}) \cup I_{00}(\bar{x})\}
\]

which is referred to as MPVC-LICQ in the literature, see (Hoheisel and Kanzow, 2007, Definition 4.1). The associated S-stationarity system from Definition 3.8 reads as follows:

\[
\begin{aligned}
0^n &= \nabla f(\bar{x}) + \nabla g(\bar{x})^\top \lambda + \nabla h(\bar{x})^\top \rho + \nabla G(\bar{x})^\top \mu + \nabla H(\bar{x})^\top \nu, \\
\lambda &\geq 0, \forall j \notin I^g(\bar{x}) &: \lambda_j = 0, \\
\forall j \in I_{+-}(\bar{x}) &: \mu_j \geq 0, \\
\forall j \in I_{-+}(\bar{x}) \cup I_{0+}(\bar{x}) \cup I_{0-}(\bar{x}) \cup I_{00}(\bar{x}) &: \mu_j = 0, \\
\forall j \in I_{0+}(\bar{x}) \cup I_{+-}(\bar{x}) &: \nu_j = 0, \\
\forall j \in I_{0-}(\bar{x}) \cup I_{00}(\bar{x}) &: \nu_j \leq 0.
\end{aligned}
\]

We note that this system precisely coincides with the system of strong stationarity for (MPVC) which has been stated in (Hoheisel and Kanzow, 2007, Definition 2.1). One can easily check that the linearization cone from (1) and the critical cone from (8) equal the respective cones from (Hoheisel and Kanzow, 2007, Section 4). Thus, our Theorems 3.7, 4.2 and 4.3 precisely recover (Hoheisel and Kanzow, 2009, Corollary 4.5) and (Hoheisel and Kanzow, 2007, Theorems 4.3, 4.4). Noting that the sets \( S_1^{VC} \) and \( S_2^{VC} \) are polyhedral cones, the stability result from Theorem 4.5 is valid for (MPVC) as well. To the best of our knowledge, this result cannot be found in the available literature on (MPVC).
5.3 Application to CCMPs

Let \( \| \cdot \|_0 : \mathbb{R}^n \to \mathbb{R} \) be the map which assigns to each vector from \( \mathbb{R}^n \) the number of its nonzero components. For some constant \( \kappa \in \{1, \ldots, n-1\} \),

\[
\begin{align*}
   f(x) & \to \min \\
   g_j(x) & \leq 0 \quad j = 1, \ldots, p \\
   h_j(x) & = 0 \quad j = 1, \ldots, q \\
   \|x\|_0 & \leq \kappa
\end{align*}
\]

is a nonlinear optimization problem comprising a so-called cardinality constraint. Problems of the form (14) appear frequently in the context of e.g. portfolio optimization. Following the results from Burdakov et al. (2016); Červinka et al. (2016), \( \bar{x} \in \mathbb{R}^n \) is a global minimizer of (14) if and only if there exists a binary vector \( \bar{y} \in \mathbb{R}^n \) such that \((\bar{x}, \bar{y})\) solves the nonlinear mixed-integer program

\[
\begin{align*}
   f(x) & \to \min \\
   g_j(x) & \leq 0 \quad j = 1, \ldots, p \\
   h_j(x) & = 0 \quad j = 1, \ldots, q \\
   e \cdot y - (n - \kappa) & \geq 0 \\
   x_i y_i & = 0 \quad i = 1, \ldots, n \\
   y_i & \in \{0, 1\} \quad i = 1, \ldots, n
\end{align*}
\]  

(CCMP)
globally. Here, \( e \in \mathbb{R}^n \) denotes the all-ones-vector. Its standard relaxation is given by

\[
\begin{align*}
   f(x) & \to \min \\
   g_j(x) & \leq 0 \quad j = 1, \ldots, p \\
   h_j(x) & = 0 \quad j = 1, \ldots, q \\
   e \cdot y - (n - \kappa) & \geq 0 \\
   x_i y_i & = 0 \quad i = 1, \ldots, n \\
   0 & \leq y_i \leq 1 \quad i = 1, \ldots, n
\end{align*}
\]

and we will refer to (CCMP) as cardinality-constrained mathematical program. Let us briefly note that the global minimizers of (14) and (CCMP) correspond to each other while the local minimizers of (14) can be found (in a certain sense) among the local minimizers of (CCMP). Furthermore, if \((\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^n \) is a local minimizer of (CCMP) where \( \|\bar{x}\|_0 = \kappa \) holds, then \( \bar{x} \) is a local minimizer of (14), see (Burdakov et al., 2016, Section 3) for details.

In order to transfer (CCMP) to a program of type (MPDC), let us introduce the sets \( S_1^C, S_2^C \subset \mathbb{R}^2 \) by means of \( S_1^C := \mathbb{R} \times \{0\} \) and \( S_2^C := \{0\} \times [0, 1] \). Next, we set \( \mathcal{J} := \{1, 2\}^n \),

\[
\forall \alpha \in \mathcal{J} : \quad D_\alpha^C := \mathbb{R}_+^p \times \{0^q\} \times \mathbb{R}_+ \times \prod_{i=1}^n S_{\alpha_i}^C,
\]

28
as well as $D^C := \bigcup_{\alpha \in \mathcal{J}} D^C_\alpha$. Furthermore, let us define $F: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{p+q+1+2n}$ by means of

$$F(x, y) := \begin{bmatrix} g(x)^\top & h(x)^\top & \mathbf{e} \cdot y - (n - \kappa) & x_1 & y_1 & \ldots & x_n & y_n \end{bmatrix}^\top$$

for each $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. Using this, the feasible set of (CCMP) can be represented by $X^C := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n | F(x, y) \in D^C\}$, see Figure 3. Let us fix a feasible pair $(\bar{x}, \bar{y}) \in X^C$ of (CCMP). We will exploit the index sets defined below:

$I_{\pm 0}(\bar{x}, \bar{y}) := \{i \in \{1, \ldots, n\} | x_i \neq 0 \land y_i = 0\}$,

$I_{00}(\bar{x}, \bar{y}) := \{i \in \{1, \ldots, n\} | x_i = 0 \land y_i = 0\}$,

$I_{0+}(\bar{x}, \bar{y}) := \{i \in \{1, \ldots, n\} | x_i = 0 \land 0 < y_i < 1\}$,

$I_{01}(\bar{x}, \bar{y}) := \{i \in \{1, \ldots, n\} | x_i = 0 \land y_i = 1\}$.

Furthermore, we will make use of $I(\bar{x}, \bar{y}) := \{\alpha \in \mathcal{J} | F(\bar{x}, \bar{y}) \in D^C_\alpha\}$. For arbitrary $\alpha \in \mathcal{J}$, we obtain

$$\alpha \in I(\bar{x}, \bar{y}) \iff \forall i \in I_{\pm 0}(\bar{x}, \bar{y}) : \alpha_i = 1 \land \forall i \in I_{0+}(\bar{x}, \bar{y}) \cup I_{01}(\bar{x}, \bar{y}) : \alpha_i = 2.$$ 

We exploit the obvious formulas

\begin{align*}
T_{S^C_1}(\bar{x}, \bar{y}) &= \begin{cases} \mathbb{R} \times \{0\} & i \in I_{\pm 0}(\bar{x}, \bar{y}) \cup I_{00}(\bar{x}, \bar{y}), \\ \emptyset & i \in I_{0+}(\bar{x}, \bar{y}) \cup I_{01}(\bar{x}, \bar{y}) \end{cases}, \\
T_{S^C_2}(\bar{x}, \bar{y}) &= \begin{cases} \{0\} \times \mathbb{R}^+ & i \in I_{00}(\bar{x}, \bar{y}), \\ \{0\} \times \mathbb{R} & i \in I_{0+}(\bar{x}, \bar{y}), \\ \{0\} \times \mathbb{R}^- & i \in I_{01}(\bar{x}, \bar{y}) \end{cases}, \\
\tilde{N}_{S^C_1}(\bar{x}, \bar{y}) &= \begin{cases} \{0\} \times \mathbb{R} & i \in I_{\pm 0}(\bar{x}, \bar{y}) \cup I_{00}(\bar{x}, \bar{y}), \\ \emptyset & i \in I_{0+}(\bar{x}, \bar{y}) \cup I_{01}(\bar{x}, \bar{y}) \end{cases}, \\
\tilde{N}_{S^C_2}(\bar{x}, \bar{y}) &= \begin{cases} \mathbb{R} \times \mathbb{R}^+ & i \in I_{00}(\bar{x}, \bar{y}), \\ \mathbb{R} \times \{0\} & i \in I_{0+}(\bar{x}, \bar{y}), \\ \mathbb{R} \times \mathbb{R}^- & i \in I_{01}(\bar{x}, \bar{y}) \end{cases}.
\end{align*}
in order to see that the constraint qualification MPDC-LICQ takes the subsequently stated form for (CCMP):

\[
0^n = \nabla g(\bar{x})^T \lambda + \nabla h(\bar{x})^T \rho + \mu, \\
0^n = \sigma \mathbf{e} + \nu, \\
\forall j \notin I^q(\bar{x}): \lambda_j = 0, \\
\sigma(\mathbf{e} \cdot \bar{y} - (n - \kappa)) = 0, \\
\forall i \in I_{\pm 0}(\bar{x}, \bar{y}): \mu_i = 0, \\
\forall i \in I_{0+}(\bar{x}, \bar{y}): \nu_i = 0
\]

\begin{equation}
\begin{aligned}
\Rightarrow \left\{ \begin{array}{l}
\lambda = 0^p, \rho = 0^q, \\
\sigma = 0, \\
\mu = \nu = 0^n.
\end{array} \right.
\end{aligned}
\end{equation}

(15)

The following lemma shows that (15) is a stronger constraint qualification than the well-known regularity condition CC-LICQ, see (Červinka et al., 2016, Definition 3.11).

Lemma 5.1. Let \((\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^n\) be a feasible point of (CCMP). Then, the constraint qualification (15) is valid at \((\bar{x}, \bar{y})\) if and only if the vectors from

\[
\begin{array}{c}
\{ \nabla g_j(\bar{x}) \mid j \in I^q(\bar{x}) \} \cup \{ \nabla h_j(\bar{x}) \mid j \in \{1, \ldots, q\} \} \\
\cup \{ \mathbf{e}_i \mid i \in I_{00}(\bar{x}, \bar{y}) \cup I_{0+}(\bar{x}, \bar{y}) \cup I_{01}(\bar{x}, \bar{y}) \}
\end{array}
\]

are linearly independent (i.e. CC-LICQ is valid) and

\[
\begin{array}{c}
0^n = \sigma \mathbf{e} + \nu, \\
\sigma(\mathbf{e} \cdot \bar{y} - (n - \kappa)) = 0, \\
\forall i \in I_{0+}(\bar{x}, \bar{y}): \nu_i = 0
\end{array}
\]

\begin{equation}
\begin{aligned}
\Rightarrow \left\{ \begin{array}{l}
\sigma = 0, \\
\nu = 0^n
\end{array} \right.
\end{aligned}
\end{equation}

(16)

holds. Above, \(\mathbf{e}_i \in \mathbb{R}^n\) denotes the \(i\)-th unit vector from \(\mathbb{R}^n\).

Corollary 5.2. Let \((\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^n\) be a feasible point of (CCMP).

(a) If \(\|\bar{x}\|_0 = \kappa\) holds, then (15) is violated.

(b) If the vectors from (16) are linearly independent and if \(I_{0+}(\bar{x}, \bar{y})\) is nonempty, then (15) is valid.

Proof. For the proof of (a), suppose that \(\|\bar{x}\|_0 = \kappa\) holds. This yields that at least \(\kappa\) entries of \(\bar{y}\) are zero, i.e. \(\mathbf{e} \cdot \bar{y} \leq n - \kappa\) holds. Feasibility of \((\bar{x}, \bar{y})\) for (CCMP) yields \(\mathbf{e} \cdot \bar{y} = n - \kappa\). Noting \(\bar{y}_i \leq 1\) for all \(i = 1, \ldots, n\), we obtain that precisely \(\kappa\) entries of \(\bar{y}\) are zero while the other ones equal 1. This shows that \(I_{0+}(\bar{x}, \bar{y})\) is empty. Choosing \(\sigma := 1\) as well as \(\nu := -\mathbf{e}\) visualizes that (17) is violated. By means of Lemma 5.1, (15) is violated.

In order to show (b), we only need to prove that \(I_{0+}(\bar{x}, \bar{y}) \neq \emptyset\) implies the validity of (17). Then, the assertion follows from Lemma 5.1. Thus, let \(\sigma \in \mathbb{R}\) and \(\nu \in \mathbb{R}^n\) satisfy

\[
0^n = \sigma \mathbf{e} + \nu, \ \sigma(\mathbf{e} \cdot \bar{e} - (n - \kappa)) = 0, \ \nu_i = 0 \text{ for all } i \in I_{0+}(\bar{x}, \bar{y}).
\]

Since the latter set is nonempty, we find some \(\hat{i} \in \{1, \ldots, n\}\) such that \(\nu_{\hat{i}} = 0\) which yields \(\sigma = -\nu_{\hat{i}} = 0\), i.e. \(\nu = 0^n\) follows. This shows that (17) holds. \(\square\)
It is not difficult to check that the strong stationarity system in the sense of Theorem 3.7 associated with (CCMP) takes the following form:

\[
\begin{align*}
0^n &= \nabla f(\bar{x}) + \nabla g(\bar{x})^T \lambda + \nabla h(\bar{x})^T \rho + \mu, \\
0^n &= \sigma e + \nu, \\
\lambda &\geq 0, \forall j \notin I^0(\bar{x}): \lambda_j = 0, \\
\sigma(e \cdot \bar{y} - (n - \kappa)) &= 0, \\
\forall i \in I_{\pm 0}(\bar{x}, \bar{y}) \cup I_{00}(\bar{x}, \bar{y}): \mu_i = 0, \\
\forall i \in I_{0+}(\bar{x}, \bar{y}): \nu_i = 0, \\
\forall i \in I_{00}(\bar{x}, \bar{y}): \nu_i \leq 0, \\
\forall i \in I_{01}(\bar{x}, \bar{y}): \nu_i \geq 0.
\end{align*}
\]

(18)

Noting that the choice \(\sigma = 0\) and \(\nu = 0^n\) is always possible, the above system possesses a solution if and only if the system

\[
\begin{align*}
0^n &= \nabla f(\bar{x}) + \nabla g(\bar{x})^T \lambda + \nabla h(\bar{x})^T \rho + \mu, \\
\lambda &\geq 0, \forall j \notin I^0(\bar{x}): \lambda_j = 0, \\
\forall i \in I_{\pm 0}(\bar{x}, \bar{y}) \cup I_{00}(\bar{x}, \bar{y}): \mu_i = 0
\end{align*}
\]

(19)

possesses a solution. The latter equals the classical strong stationarity system from the literature on (CCMP), see (Červinka et al., 2016, Definition 4.1). It follows from (Červinka et al., 2016, Theorems 3.13, 4.2) that for each local minimizer of (CCMP) where CC-LICQ holds, there exist multipliers which solve the system (19) and it is obvious that these multipliers are uniquely determined, see (Bucher and Schwartz, 2018, Proposition 2.4) as well. However, the associated multipliers which solve the system (18) do not need to be uniquely determined in this situation. This uniqueness (i.e. \(\sigma = 0\) and \(\nu = 0^n\)), which does not seem to be relevant in practical situations, can only be guaranteed under validity of (15) in general, see Theorem 3.7. However, invoking statement (a) of Corollary 5.2, the postulation of (15) rules out many potential candidates for global minimizers of (CCMP). Thus, using the approach of disjunctive programming discussed in Sections 3 and 4 to tackle (CCMP) yields weaker results than a direct consideration of this problem class. This way, our results from Theorems 4.2 and 4.3 are weaker than the second-order conditions for (CCMP) which can be found in (Bucher and Schwartz, 2018, Theorems 3.1, 3.2) although they are based on the same critical cone. Note that our stability result Theorem 4.5 is not applicable here since \(S_C^2\) is no cone. However, in (Bucher and Schwartz, 2018, Theorem 3.3), the local uniqueness of so-called Mordukhovich-stationary points of (CCMP) has been shown under suitable assumptions. Summarizing the above results, the problem class (CCMP) visualizes the limits of this paper’s approach. We would like to mention that the outlined difficulties are caused by the appearance of the artificial variable \(y\) and its associated multipliers.
5.4 Application to MPSCs

Let us consider so-called mathematical programs with switching constraints which are optimization problems of the form

\[
\begin{align*}
    f(x) &\to \min \\
    g_j(x) &\leq 0 \quad j = 1, \ldots, p \\
    h_j(x) &= 0 \quad j = 1, \ldots, q \\
    G_j(x) \cdot H_j(x) &= 0 \quad j = 1, \ldots, l.
\end{align*}
\]

(MPSC)

Models of type (MPSC) arise from the discretization of so-called switching-constrained optimal control problems, see e.g. Clason et al. (2017) and the references therein, as well as the reformulation of logical or-constraints, see (Mehlitz, 2018, Section 7), or semi-continuity conditions on variables, see (Kanzow et al., 2018, Section 5.2.3). First-order necessary optimality conditions as well as numerical relaxation methods for problems of type (MPSC) can be found in Kanzow et al. (2018); Mehlitz (2018).

Let us transfer (MPSC) into a program of type (MPDC). Therefore, we introduce

\[
S^{SC}_1 := \mathbb{R} \times \{0\} \quad \text{and} \quad S^{SC}_2 := \{0\} \times \mathbb{R}
\]

as well as

\[
J := \{1,2\}^l.
\]

We set

\[
\forall \alpha \in J: \quad D^{SC}_\alpha := \mathbb{R}^p \times \{0^q\} \times \prod_{j=1}^l S^{SC}_\alpha
\]

as well as

\[
D^{SC} := \bigcup_{\alpha \in J} D^{SC}_\alpha.
\]

Using the mapping \(F\) defined in (13), the feasible set of (MPSC) can be represented by \(X^{SC} := \{x \in \mathbb{R}^n \mid F(x) \in D^{SC}\}\). The variational geometry of \(X^{SC}\) is visualized in Figure 4. For a feasible point \(\bar{x} \in X^{SC}\) of (MPSC), we introduce the following index sets:

\[
\begin{align*}
    I^G(\bar{x}) &:= \{j \in \{1, \ldots, l\} \mid G_j(\bar{x}) = 0 \land H_j(\bar{x}) \neq 0\}, \\
    I^H(\bar{x}) &:= \{j \in \{1, \ldots, l\} \mid G_j(\bar{x}) \neq 0 \land H_j(\bar{x}) = 0\}, \\
    I^{GH}(\bar{x}) &:= \{j \in \{1, \ldots, l\} \mid G_j(\bar{x}) = 0 \land H_j(\bar{x}) = 0\}.
\end{align*}
\]

Furthermore, we set \(I(\bar{x}) := \{\alpha \in J \mid F(\bar{x}) \in D^{SC}_\alpha\}\). Thus, we have

\[
\alpha \in I(\bar{x}) \iff \forall j \in I^G(\bar{x}): \alpha_j = 2 \land \forall j \in I^H(\bar{x}): \alpha_j = 1
\]
for each $\alpha \in \mathcal{J}$. Clearly, we obtain

$$
\begin{align*}
T_{SC}(G_j(\bar{x}), H_j(\bar{x})) &= \begin{cases} 
\emptyset & j \in \mathcal{I}^G(\bar{x}), \\
\mathbb{R} \times \{0\} & j \in \mathcal{I}^H(\bar{x}) \cup \mathcal{I}^{GH}(\bar{x}),
\end{cases} \\
T_{SC^2}(G_j(\bar{x}), H_j(\bar{x})) &= \begin{cases} 
\{0\} \times \mathbb{R} & j \in \mathcal{I}^G(\bar{x}) \cup \mathcal{I}^{GH}(\bar{x}), \\
\emptyset & j \in \mathcal{I}^H(\bar{x}),
\end{cases}
\end{align*}
$$

Thus, the constraint qualification MPDC-LICQ takes the following form for (MPSC):

$$
\begin{align*}
0^n = \nabla g(\bar{x})^\top \lambda + \nabla h(\bar{x})^\top \rho + \nabla G(\bar{x})^\top \mu + \nabla H(\bar{x})^\top \nu, \\
\forall j \notin \mathcal{I}^g(\bar{x}): \lambda_j = 0, \\
\forall j \in \mathcal{I}^H(\bar{x}): \mu_j = 0, \\
\forall j \in \mathcal{I}^G(\bar{x}): \nu_j = 0
\end{align*}
$$

We note that this is equivalent to the linear independence of all the vectors from

$$
\{\nabla g_j(\bar{x}) \mid j \in \mathcal{I}^g(\bar{x})\} \cup \{\nabla h_j(\bar{x}) \mid j \in \{1, \ldots, q\}\}
\cup \{\nabla G_j(\bar{x}) \mid j \in \mathcal{I}^G(\bar{x}) \cup \mathcal{I}^{GH}(\bar{x})\} \cup \{\nabla H_j(\bar{x}) \mid j \in \mathcal{I}^H(\bar{x}) \cup \mathcal{I}^{GH}(\bar{x})\}
$$

and this condition is called MPSC-LICQ in the literature, see (Mehlitz, 2018, Definition 4.4). The associated S-stationarity system from Theorem 3.7 is given by

$$
\begin{align*}
0^n = \nabla f(\bar{x}) + \nabla g(\bar{x})^\top \lambda + \nabla h(\bar{x})^\top \rho + \nabla G(\bar{x})^\top \mu + \nabla H(\bar{x})^\top \nu, \\
\lambda \geq 0, \forall j \notin \mathcal{I}^g(\bar{x}): \lambda_j = 0, \\
\forall j \in \mathcal{I}^H(\bar{x}) \cup \mathcal{I}^{GH}(\bar{x}): \mu_j = 0, \\
\forall j \in \mathcal{I}^G(\bar{x}) \cup \mathcal{I}^{GH}(\bar{x}): \nu_j = 0
\end{align*}
$$

and equals the problem-tailored system of strong stationarity as it has been stated in (Mehlitz, 2018, Definition 4.3). The results from Theorem 3.7 can be found in (Mehlitz, 2018, Theorem 4.5).

One can easily check that the linearization cone from (1) possesses the form

$$
\mathcal{L}_{\mathcal{X}_{SC}}(\bar{x}) = \left\{ d \in \mathbb{R}^n \right\}
\begin{align*}
\nabla g(\bar{x}) \cdot d &\leq 0 & j \in \mathcal{I}^g(\bar{x}) \\
\nabla h(\bar{x}) \cdot d &\leq 0 & j \in \{1, \ldots, q\} \\
\nabla G(\bar{x}) \cdot d &\leq 0 & j \in \mathcal{I}^G(\bar{x}) \\
\nabla H(\bar{x}) \cdot d &\leq 0 & j \in \mathcal{I}^H(\bar{x}) \\
(\nabla G(\bar{x}) \cdot d)(\nabla H(\bar{x}) \cdot d) &\leq 0 & j \in \mathcal{I}^{GH}(\bar{x})
\end{align*}
$$
while we obtain

$$\mathcal{C}_{X^{SC}}(\bar{x}) = \left\{ d \in \mathbb{R}^n \right\}$$

\[
\begin{align*}
\nabla g(\bar{x}) \cdot d &\leq 0 & j \in I^g(\bar{x}), \lambda_j = 0 \\
\nabla g(\bar{x}) \cdot d & = 0 & j \in I^g(\bar{x}), \lambda_j > 0 \\
\nabla h(\bar{x}) \cdot d & = 0 & j \in \{1, \ldots, q\} \\
\nabla G(\bar{x}) \cdot d & = 0 & j \in I^G(\bar{x}) \\
\nabla H(\bar{x}) \cdot d & = 0 & j \in I^H(\bar{x}) \\
(\nabla G(\bar{x}) \cdot d)(\nabla H(\bar{x}) \cdot d) & = 0 & j \in I^{GH}(\bar{x})
\end{align*}
\]

for the associated critical cone provided \(\bar{x}\) is a strongly stationary point of (MPSC) with associated multipliers \((\lambda, \rho, \mu, \nu) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^l \times \mathbb{R}^l\), see Lemma 4.1. Based on this critical cone, Theorems 4.2 and 4.3 provide a necessary and sufficient second-order optimality condition for (MPSC). Noting that the sets \(S_1^{SC}\) and \(S_2^{SC}\) are polyhedral cones, Theorem 4.5 provides a criterion which ensures local uniqueness of strongly stationary points associated with (MPSC). To the best of our knowledge, these are new results on the problem class (MPSC).

### 6 Final remarks

In this paper, we introduced an abstract version of the prominent linear independence constraint qualification which applies to mathematical programs with disjunctive constraints. We were able to derive first- and second-order optimality conditions based on strongly stationary points under validity of this constraint qualification. Finally, we applied our findings to several different instances of disjunctive programs in order to underline that this new constraint qualification is reasonable. By means of switching-constrained mathematical problems, it has been demonstrated that our theory does not only recover well-known results from the literature but can be used to infer new results on specific instances of disjunctive programming as well. On the other hand, the consideration of a certain reformulation of cardinality-constrained optimization problems has visualized the limits of our approach.

### References

W. Achtziger and C. Kanzow. Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications. *Mathematical Programming, Series A*, 114(1):69–99, 2008. doi: 10.1007/s10107-006-0083-3.

W. Achtziger, C. Kanzow, and T. Hoheisel. On a relaxation method for mathematical programs with vanishing constraints. *GAMM-Mitteilungen*, 35(2):110–130, 2012. doi: 10.1002/gamm.201210009.

J.-P. Aubin and H. Frankowska. *Set-valued Analysis*. Modern Birkhäuser Classics. Birkhäuser, Boston, 2009. Reprint of the 1990 edition.
M. Benko and H. Gfrerer. On estimating the regular normal cone to constraint systems and stationarity conditions. *Optimization*, 66(1):61–92, 2017. doi: 10.1080/02331934.2016.1252915.

M. Benko and H. Gfrerer. New verifiable stationarity concepts for a class of mathematical programs with disjunctive constraints. *Optimization*, 67(1):1–23, 2018. doi: 10.1080/02331934.2017.1387547.

J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, New York, 2000.

M. Bucher and A. Schwartz. Second-Order Optimality Conditions and Improved Convergence Results for Regularization Methods for Cardinality-Constrained Optimization Problems. *Journal of Optimization Theory and Applications*, 178(2):383–410, 2018. doi: 10.1007/s10957-018-1320-7.

O. P. Burdakov, C. Kanzow, and A. Schwartz. Mathematical programs with cardinality constraints: reformulation by complementarity-type conditions and a regularization method. *SIAM Journal on Optimization*, 26(1):397–425, 2016. doi: 10.1137/140978077.

M. Červinka, C. Kanzow, and A. Schwartz. Constraint qualifications and optimality conditions for optimization problems with cardinality constraints. *Mathematical Programming, Series A*, 160(1):353–377, 2016. doi: 10.1007/s10107-016-0986-6.

C. Christof and G. Wachsmuth. No-Gap Second-Order Conditions via a Directional Curvature Functional. *SIAM Journal on Optimization*, 28(3):2097–2130, 2018. doi: 10.1137/17M1140418.

C. Clason, A. Rund, and K. Kunisch. Nonconvex penalization of switching control of partial differential equations. *Systems & Control Letters*, 106:1–8, 2017. doi: 10.1016/j.sysconle.2017.05.006.

M. L. Flegel, C. Kanzow, and J. V. Outrata. Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints. *Set-Valued Analysis*, 15(2):139–162, 2007. doi: 10.1007/s11228-006-0033-5.

H. Gfrerer. On Directional Metric Subregularity and Second-Order Optimality Conditions for a Class of Nonsmooth Mathematical Programs. *SIAM Journal on Optimization*, 23(1):632–665, 2013. doi: 10.1137/120891216.

H. Gfrerer. Optimality Conditions for Disjunctive Programs Based on Generalized Differentiation with Application to Mathematical Programs with Equilibrium Constraints. *SIAM Journal on Optimization*, 24(2):898–931, 2014. doi: 10.1137/130914449.

H. Gfrerer and J. V. Outrata. On Computation of Generalized Derivatives of the Normal-Cone Mapping and Their Applications. *Mathematics of Operations Research*, 41(4):1535–1556, 2016. doi: 10.1287/moor.2016.0789.
H. Gfrerer and J. J. Ye. New Constraint Qualifications for Mathematical Programs with Equilibrium Constraints via Variational Analysis. *SIAM Journal on Optimization*, 27(2):842–865, 2017. doi: 10.1137/16M1088752.

L. Guo, G.-H. Lin, and J. J. Ye. Second-Order Optimality Conditions for Mathematical Programs with Equilibrium Constraints. *Journal of Optimization Theory and Applications*, 158(1):33–64, 2013. doi: 10.1007/s10957-012-0228-x.

R. Henrion and J. V. Outrata. Calmness of constraint systems with applications. *Mathematical Programming*, 104(2):437–464, 2005. doi: 10.1007/s10107-005-0623-2.

T. Hoheisel. *Mathematical Programs with Vanishing Constraints*. PhD thesis, University of Würzburg, 2009.

T. Hoheisel and C. Kanzow. First- and second-order optimality conditions for mathematical programs with vanishing constraints. *Applications of Mathematics*, 52(6):495–514, 2007. doi: 10.1007/s10492-007-0029-y.

T. Hoheisel and C. Kanzow. On the Abadie and Guignard constraint qualifications for mathematical programmes with vanishing constraints. *Optimization*, 58(4):431–448, 2009. doi: 10.1080/02331930701763405.

T. Hoheisel, C. Kanzow, and A. Schwartz. Convergence of a local regularization approach for mathematical programmes with complementarity or vanishing constraints. *Optimization Methods and Software*, 27(3):483–512, 2012. doi: 10.1080/10556788.2010.535170.

T. Hoheisel, C. Kanzow, and A. Schwartz. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. *Mathematical Programming*, 137(1):257–288, 2013. doi: 10.1007/s10107-011-0488-5.

A. F. Izmailov and M. V. Solodov. Mathematical Programs with Vanishing Constraints: Optimality Conditions, Sensitivity, and a Relaxation Method. *Journal of Optimization Theory and Applications*, 142(3):501–532, 2009. doi: 10.1007/s10957-009-9517-4.

C. Kanzow, P. Mehlitz, and D. Steck. Relaxation schemes for mathematical programs with switching constraints. *Preprint arXiv*, 2018. URL https://arxiv.org/abs/1809.02388.

C. Kirches. *Fast Numerical Methods for Mixed-Integer Nonlinear Model-Predictive Control*. PhD thesis, University of Heidelberg, 2011.

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

P. Mehlitz. Stationarity conditions and constraint qualifications for mathematical programs with switching constraints. *Preprint TU Freiberg*, 2018. URL https://tu-freiberg.de/sites/default/files/media/fakultaet-fuer-mathematik-und-informat
J. V. Outrata, M. Kočvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic, Dordrecht, 1998.

K. Palagachev and M. Gerdts. Mathematical Programs with Blocks of Vanishing Constraints Arising in Discretized Mixed-Integer Optimal Control Problems. *Set-Valued and Variational Analysis*, 23(1):149–167, 2015. doi: 10.1007/s11228-014-0297-0.

J. Penot. Second-Order Conditions for Optimization Problems with Constraints. *SIAM Journal on Control and Optimization*, 37(1):303–318, 1998. doi: 10.1137/S0363012996311095.

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

S. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, 2000. doi: 10.1287/moor.25.1.1.15213.

S. Scholtes. Nonconvex Structures in Nonlinear Programming. *Operations Research*, 52(3):368–383, 2004. doi: 10.1287/opre.1030.0102.

J. J. Ye. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *Journal of Mathematical Analysis and Applications*, 307(1):350 – 369, 2005. doi: 10.1016/j.jmaa.2004.10.032.