Optimality Conditions for Abs-Normal NLPs

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des akademischen Grades
Doktorin der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation von

M.Sc. Lisa Christine Hegerhorst-Schultchen

2020

Referent:	Prof. Dr. Marc C. Steinbach, Gottfried Wilhelm Leibniz Universität Hannover
Korreferent:	Prof. Dr. Christian Kirches, Technische Universität Carolo-Wilhelmina zu Braunschweig
Korreferentin:	Prof. Dr. Andrea Walther, Humboldt-Universität zu Berlin
Tag der Promotion:	29. Mai 2020

Abstract

Structured nonsmoothness is widely present in practical optimization problems. A particularly attractive class of nonsmooth problems, both from a theoretical and from an algorithmic perspective, are nonsmooth NLPs with equality and inequality constraints in abs-normal form, so-called abs-normal NLPs.

In this thesis optimality conditions for this particular class are obtained. To this aim, first the theory for the case of unconstrained optimization problems in abs-normal form of Andreas Griewank and Andrea Walther is extended. In particular, similar necessary and sufficient conditions of first and second order are obtained that are directly based on classical Karush-Kuhn-Tucker (KKT) theory for smooth NLPs.

Then, it is shown that the class of abs-normal NLPs is equivalent to the class of Mathematical Programs with Equilibrium Constraints (MPECs). Hence, the regularity assumption LIKQ introduced for the abs-normal NLP turns out to be equivalent to MPEC-LICQ. Moreover, stationarity concepts and optimality conditions under these regularity assumptions of linear independece type are equivalent up to technical assumptions.

Next, well established constraint qualifications of Mangasarian Fromovitz, Abadie and Guignard type for MPECs are used to define corresponding concepts for abs-normal NLPs. Then, it is shown that kink qualifications and MPEC constraint qualifications of Mangasarian Fromovitz resp. Abadie type are equivalent. As it remains open if this holds for Guignard type kink and constraint qualifications, branch formulations for abs-normal NLPs and MPECs are introduced. Then, equivalence of Abadie's and Guignard's constraint qualifications for all branch problems hold.

Throughout a reformulation of inequalities with absolute value slacks is considered. It preserves constraint qualifications of linear independence and Abadie type but *not* of Mangasarian Fromovitz type. For Guignard type it is still an open question but ACQ and GCQ are preserved passing over to branch problems. Further, M-stationarity and B-stationarity concepts for abs-normal NLPs are introduced and corresponding first order optimality conditions are proven using the corresponding concepts for MPECs.

Moreover, a reformulation to extend the optimality conditions for abs-normal NLPs to those with additional nonsmooth objective functions is given and the preservation of regularity assumptions is considered. Using this, it is shown that the unconstrained abs-normal NLP always satisfies constraint qualifications of Abadie and thus Guignard type. Hence, in this special case every local minimizer satisfies the M-stationarity and B-stationarity concepts for abs-normal NLPs.

Keywords: nonsmooth optimization, abs-normal NLPs, MPECs, kink and constraint qualifications, optimality conditions

Kurzzusammenfassung

Nichtglattheiten, die einer gewissen Struktur gehorchen, treten in zahlreichen Anwendungsproblemen auf. Eine besonders interessante Klasse, sowohl von der theoretischen als auch von der algorithmischen Seite her, sind nichtglatte Optimierungsprobleme mit Gleichungs- und Ungleichungsbedingungen in Abs-Normal Form, sogenannte Abs-Normal NLPs.

In der vorliegenden Dissertation werden Optimalitätsbedingungen für eben diese Klasse hergeleitet. Dazu wird erst einmal die Theorie für unbeschränkte Abs-Normal NLPs von Andreas Griewank und Andrea Walther erweitert. Insbesondere erhält man notwendige und hinreichende Optimalitätsbedingungen erster und zweiter Ordnung, welche direkt auf den klassisichen Karush-Kuhn-Tucker (KKT) Bedingungen basieren.

Danach wird gezeigt, dass die Klasse der Abs-Normal NLPs äquivalent zur Klasse der MPECs ist. Daher folgt, dass die Regularitätsbedingung LIKQ, welche für Abs-Normal NLPs eingeführt wurde, äquivalent zur Regularitätsbedingung MPEC-LICQ ist. Außerdem sind die Stationaritäts- und Optimalitätsbedingungen unter diesen Regularitätsbedingungen bis auf technische Voraussetzungen äquivalent.

Im Weiteren werden dann bekannte MPEC-Regularitätsbedingungen im Sinne von Mangasarian Fromovitz, Abadie und Guignard genutzt, um entsprechende Konzepte für Abs-Normal NLPs zu formulieren. Dann wird gezeigt, dass die Regularitätsbedingungen im Sinne von Mangasarian Fromovitz und Abadie für MPECs und Abs-Normal NLPs äquivalent sind. Da dieses jedoch für Bedingungen im Sinne von Guignard bisher weder gezeigt noch widerlegt werden kann, werden Branch-Probleme für Abs-Normal NLPs und MPECs eingeführt. Dann kann Äquivalenz zwischen ACQ and GCQ für alle Branch-Probleme gezeigt werden.

Im Verlauf wird wiederholt eine Reformulierung von Ungleichungen mithilfe von Betragsslacks betrachtet. Diese erhält Regularitätsbedingungen im Sinne linearer Unabhängigkeit und von Abadie, aber nicht im Sinne von Mangasarian Fromovitz. Für Regularität im Sinne von Guignard kann dies bisher weder gezeigt noch widerlegt worden. Wird hier jedoch zu den Branch-Problemen übergegangen, gilt die Äquivalenz für ACQ and GCQ. Im Weiteren werden dann M- und B-Stationaritätskonzepte für Abs-Normal NLPs eingeführt und entsprechende Optimalitätsbedingungen erster Ordnung bewiesen. Hierzu werden die entsprechenden Konzepte für MPECs benutzt.

Außerdem wird eine Reformulierung, um die Optimalitätsbedingungen auf Abs-Normal NLPs mit nichtglatter Zielfunktion zu übertragen, angegeben und die Erhaltung der Regularität untersucht. Damit wird dann für das unbeschränkte Abs-Normal NLP gezeigt, dass Regularitätsbedingungen im Sinne von Abadie und Guignard stets gegeben sind. Somit erfüllt in diesem Spezialfall jedes lokales Minimum die Bedingungen für die Abs-Normal NLP Konzepte von M-Stationarität und B-Stationarität.

Stichworte: Nichtglatte Optimierung, Abs-Normal NLPs, MPECs, Regularitätsbedingungen, Optimalitätsbedingungen

Acknowledgements

I would like to thank my supervisor Marc Steinbach for the opportunity to be part of his work group and for a lot of discussions that guided me during my work on this thesis.

I would like to thank Christian Kirches for the nice collaboration in comparing abs-normal NLPs to MPECs. Moreover, I would like to thank Andrea Walther and Andreas Griewank for prolific discussions concerning abs-normal forms.

I am also very grateful to all my current and former colleagues at the Institute of Applied Mathematics. In particular, I would like to thank Jan Philipp Thiele and Jan Thiedau for proofreading drafts of my thesis.

Additionally, I would like to thank my family and friends for their support without really understanding what I have been working on most of the time.

Above all I would like to thank my husband Marius Schultchen: on one hand for proof-reading drafts of my thesis and on the other hand for his encouragement, support and belief in me that made this thesis possible.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

As a motivating example for nonsmooth constraints in optimization models, gas networks are considered and discussed in more detail here; this is based on [18]. The network topology is represented as a directed graph $G = (V, A)$ with node set V and arc set A. Every node $u \in V$ has a pressure p_u, a temperature T_u and gas quality parameters X_u. Here for example the heating value H_u is contained. Similarly, every arc $a \in A$ has a mass flow q_a, two temperatures $T_{a,in}$ and $T_{a,out}$ (at head and tail) as well as a gas quality parameter X_a. A first example is the mixing and propagation of gas quality parameters. Here, it is assumed that gas flows entering node u mix perfectly and that the mixed gas quality parameter X_u is a convex combination. Moreover, for ease of notation, it is assumed that no external inflow resp. outflow in nodes occur. Then, the mixing equation reads

$$0 = X_u \left(\sum_{a \in \mathcal{I}_u} \hat{q}_a \right) - \sum_{a \in \mathcal{I}_u} \hat{q}_a X_a,$$

where $\mathcal{I}_u$ denotes the set of inflow arcs and $\hat{q}_a = \frac{q_a}{m_a}$ the molar inflows. This can be reformulated with help of the maximum function and gives

$$0 = X_u \left(\sum_{a \in \delta_u} \max(0, \hat{q}_a) \right) - \left(\sum_{a \in \delta_u} \max(0, \hat{q}_a) X_a \right),$$

where δ_u denotes the set of all incident arcs. Then, the mixed gas parameters are propagated to all outflow arcs. This reads

$$0 = X_u - X_a, \ a \in \mathcal{O}_u,$$

where $\mathcal{O}_u$ denotes the set of all outflow arcs. Similarly, this can be restated using the minimum function and gives

$$0 = \min(\hat{q}_a, 0)(X_u - X_a), \ a \in \delta_u.$$

Analogous conditions are obtained for the mixing and propagation of temperature. This gives

$$0 = T_u \left(\sum_{a \in \delta_u} \max(0, \hat{q}_a) \right) - \left(\sum_{a \in \delta_u} \max(0, \hat{q}_a) T_{a,in} \right) \quad \text{and} \quad 0 = \min(\hat{q}_a, 0)(T_u - T_{a,in}).$$

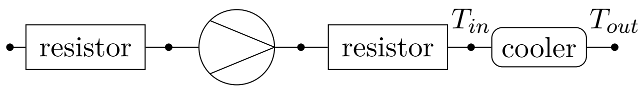

Figure 1.1: Compressor Station (schematic overview); based on [18]

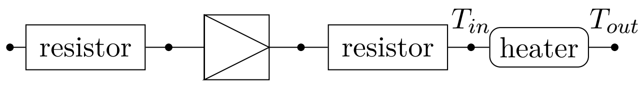

Figure 1.2: Control Valve Station (schematic overview); based on [18]

Here, the additional assumption of identical heat capacities for entering gas are made.
A second example is the sensitivity control of temperature which occurs for example in compressor and control valve stations. The schematic overviews are depicted in figure 1.1 and figure 1.2. If gas is transported through a pipeline network it experiences pressure drop. This is due to friction loss which is caused by roughness of the inner pipe wall. Then, compressor stations increase the pressure which leads to increasing temperature caused by the Joule-Thomson-effect. Thus, an additional cooler (here modeled as cooling of the outlet gas) is needed to keep the temperature below a treshold $\overline{T}_{out}$:

$$T_{out} = \min(\overline{T}_{out}, T_{in}).$$

The similar situation occurs in control valve stations. They reduce the inflow pressure in a controllable way and thus the temperature is decreased. Hence, a gas heater is needed to keep the temperature above a treshold $\underline{T}_{out}$:

$$T_{out} = \max(\underline{T}_{out}, T_{in}).$$

Similar constraints involving the absolute value, maximum and minimum functions arise in many applications from engineering, economics and other areas. Then, combining such constraints and models with a task (for example cost minimization) gives rise to so-called level-1 nonsmooth optimization problems

$$\begin{aligned} \min_{x \in D^x} \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \\ & h(x) \geq 0, \end{aligned}$$

where $f\colon D^x \to \mathbb{R}$ smooth and $g\colon D^x \to \mathbb{R}^{m_1}$, $h\colon D^x \to \mathbb{R}^{m_2}$ for $D^x \subseteq \mathbb{R}^n$ are composed of smooth functions and the absolute value. Here, only the absolute value is considered as the maximum and the minimum function can be recast in terms of it. Then, g and h can be formulated in abs-normal form, which was introduced in [7]. This leads to so-called abs-normal NLPs

$$\begin{aligned}
\min_{(x,z)} \quad & f(x) \\
\text{s.t.} \quad & c_{\mathcal{E}}(x,|z|) = 0, \\
& c_{\mathcal{I}}(x,|z|) \geq 0, \\
& c_{\mathcal{Z}}(x,|z|) - z = 0,
\end{aligned}$$

where $\partial_2 c_{\mathcal{Z}}(x,|z|)$ is strictly lower triangular and $c_{\mathcal{E}}$, $c_{\mathcal{I}}$ and $c_{\mathcal{Z}}$ are smooth with respect to x and $|z|$. In order to solve these problems, one is interested in optimality conditions for this particular class of nonsmooth optimization problems.

1.2 Literature

Nonsmooth optimization deals with problems where the objective function and/or constraints are not differentiable everywhere. It dates back to the 1960s and early 1970s where subdifferentials and optimality conditions were introduced by R.T. Rockafellar in his book [22] which is still a standard reference. In the mid-1970s and later on more subdifferentials and quasidifferentials as well as corresponding optimality conditions were introduced, for example by Clarke or Mordukhovich. These concepts can be found in the books [3] and [20].

More recently, Griewank and Walther considered in [8] unconstrained finite-dimensional nonlinear optimization problems where the nonsmoothness is caused by possibly nested occurrences of the absolute value function. As both the Clarke and the Mordukhovich necessary condition fail at the simple example $\min(x, 0)$, they derived different optimality conditions without involving subdifferentials. The key idea in this approach is to formulate the function in the so-called abs-normal form which was introduced by Griewank in [7]. Then, they extended the LICQ of smooth optimization to the so-called linear independence kink qualification (LIKQ) and also first and second order optimality conditions to their setting. In [10], the regularity concept of LIKQ was relaxed to the Mangasarian-Fromovitz kink qualification which is a regularity assumption for first order convexity. Additionally, Griewank and Walther obtained in [11] optimality conditions without any regularity assumption. Moreover, the formulation in abs-normal form leads to a natural algorithm of successive abs-linear minimization with a proximal term (SALMIN) which achieves a linear rate of convergence under suitable assumptions [10]. It is based on piecewise linearizations and uses algorithmic differentiation techniques. A first version of this algorithm called LiPsMin was implemented and tested in [4]. Then, in [9] the inner solver was improved using an active signature strategy which is similar to active set strategies in smooth optimization. This provides further opportunities for development of the solver.

Another important class in optimization are Mathematical Programs with Equilibrium Constraints (MPECs). Therein, all functions are smooth but complementary constraints $0 \leq u \perp v \geq 0$ occur. This leads to kinked feasible sets and thus standard theory for smooth optimization problems cannot be applied. Thus, new constraint qualifications and corresponding optimality conditions were introduced. By now, there is a huge amount of these concepts. Therefore only those needed later will be mentioned here and it is referred to [25] for the definition of even more MPEC constraint qualifications. One important concept is a constraint qualification of linear independence type, MPEC-LICQ, under which local minimizers are strongly stationary. A weaker constraint qualification of Mangasarian Fromovitz type is MPEC-MFCQ and a weaker stationarity concept is B-stationarity for MPECs. These concepts were introduced by Scheel and Scholtes in [23] and by Luo, Pang and Ralph in [19]. Other weaker constraint qualifications are Abadie's and Guignard's constraint qualifications for MPECs. Under these all local minimizers are M-stationary points as was shown by Flegel in [5] which is another weaker stationarity concept.

There is a substantial body of literature on other areas of nonsmooth optimization in finite and infinite dimensions. None of them are adressed here since the focus of this thesis is entirely on the abs-normal problems and MPECs mentioned above.

1.3 Contributions

This thesis provides optimality conditions for the class of abs-normal NLPs.

First, the theory of Griewank and Walther in [8] is generalized to abs-normal NLPs. In particular, the linear independence kink qualification (LIKQ) is extended and first and second order necessary and sufficient optimality conditions for the abs-normal NLP are obtained. As in the unconstrained case, this is achieved by applying standard KKT theory to suitable trunk and branch problems that satisfy the LICQ whenever the abs-normal NLP satisfies the LIKQ. Here, inequality constraints can be tackled via a reformulation as equality constraints using absolute value slacks. This is due to the fact, that LIKQ is preserved under this reformulation.

Next, it is shown that the class of abs-normal NLPs is equivalent to the class of MPECs. Then, regularity and stationarity concepts of both classes are compared. In particular, equivalence between LIKQ and MPEC-LICQ is proven in both formulations of inequality constraints. Next, equivalence of kink stationarity and strong stationarity and thus of first order necessary conditions is shown. Then, equivalence of positive (semi-)definiteness of the associated reduced Hessians is proven under additional assumptions. This gives correspondences of second order necessary and sufficient conditions.

Then, kink qualifications of Mangasarian Fromovitz (IDKQ), Abadie (AKQ) and Guignard (GKQ) type both for the standard formulation and for the reformulation with absolute value slacks are considered. In particular, it is shown that constraint qualifications of Mangasarian Fromovitz type are equivalent for abs-normal NLPs and MPECs but that they are not preserved under the slack reformulation. Whereas, constraint qualifications of Abadie type are equivalent for abs-normal NLPs and MPECs and preserved under the slack reformulation. For constraint qualifications of Guignard type equivalence cannot be proven so far but only certain implications. However, when considering branch problems of abs-normal NLPs and MPECs, equivalence of constraint qualifications of Abadie and Guignard type is obtained, even under the slack reformulation. Next, Mordukhovich and linearized Bouligand type stationarity concepts for abs-normal NLPs are introduced and first order optimality conditions are proven using the corresponding concepts for MPECs.

Further, the possible extension to level-1 nonsmooth objective functions is considered throughout this thesis. Here, the idea is to replace the objective function by an additional variable. This leads to an additional equality constraint and thus the problem is reformulated as an abs-normal NLP. In particular, it is proven that LIKQ, IDKQ, AKQ and GCQ for all branches are preserved under this reformulation. Thus, optimality conditions for abs-normal NLPs with nonsmooth objective functions can be obtained using the results for MPECs.

Also, so-called unconstrained abs-normal NLPs (the setting of Griewank and Walther in [8]) are considered. In this case it is proven that MFKQ is weaker than MPEC-MFCQ and that AKQ as well as GKQ hold always. The latter implies directly that every local minimizer is stationary in the sense of Mordukhovich and linearized Bouligand.

Publications During the work on this thesis, the publications [17, 14, 13] and the preprints [15, 16] were contributed. Note that [13] as a conference paper summarizes results of [14] and is therefore not referenced in the following. Parts of chapter 2 are based on [17], parts of chapter 3 are based on [17, 15], parts of chapter 4 are based on [14, 15] and parts of chapter 5 are based on [14, 15, 16].

1.4 Outline

This thesis is structured as follows. In chapter 2 basic concepts of nonlinear smooth optimization, unconstrained abs-normal NLPs and MPECs are introduced. Then, the theory for unconstrained abs-normal NLPs is extended to general abs-normal NLPs in chapter 3. First, abs-normal NLPs with a smooth objective function and without inequality constraints are considered. In particular, the linear independence kink qualification (LIKQ) as a fundamental regularity condition is formulated. Next, necessary and sufficient optimality conditions of first and second order for the abs-normal NLP are derived. Then, the results are transfered to abs-normal NLPs with inequality constraints using a reformulation as equality constraints via absolute value slacks. Further, abs-normal NLPs with a nonsmooth objective function are considered and it is shown that and how the optimality conditions can be transfered. In chapter 4 the equivalence of the class of abs-normal NLPs and MPECs is shown. Moreover, relations between regularity assumptions of linear independence type for the direct handling of inequality constraints as well as for the reformulation with absolute value slacks are proven. Then, the stationarity concept for abs-normal NLPs is compared to S-stationarity for MPECs and also optimality conditions for both problem classes. Next, LIKQ and MFKQ are considered for the special case of an unconstrained abs-normal NLP and relations to MPEC-LICQ and MPEC-MFCQ are given. In chapter 5 the well-known theory for MPECs is used to formulate weaker kink qualifications of Mangasarian-Fromovitz, Abadie and Guignard type. Relations between them are proven for both formulations of inequality constraints. Then, M- and abs-normal-linearized B-Stationarity for the abs-normal NLP are formulated and corresponding first order conditions are proven using MPECs. Further, the extension of these weaker kink qualifications to abs-normal NLPs with a nonsmooth objective function is considered. Using these results the special case of unconstrained abs-normal NLPs are examined next. Finally, in chapter 6 the results of this work are summarized and future research directions are sketched.

1.5 Notation

By ∂_i the partial derivative with respect to the i-th argument will be denoted and by ∂_y the derivative with respect to the variable y. Further, $P_{\mathcal{S}} \in \mathbb{R}^{|\mathcal{S}| \times s}$ denotes the projector onto the subspace defined by $\mathcal{S} \subseteq \{1, \ldots, s\}$. The notation $f \in C^d$ is used to denote that a function f is d times continuous differentiable. Thus, it is assumed that $d \in \mathbb{N}_{\geq 0} \cup \{\infty\}$ holds.

Throughout this thesis open domains D of functions are considered. A superscript index denotes the corresponding variable, for example $D^x \subseteq \mathbb{R}^n$ is the open domain of the variable $x \in \mathbb{R}^n$. Further, the term $D^{|z|} \subseteq \mathbb{R}^s$ is used to denote the open and *symmetric* domain of a so-called switching variable $z \in \mathbb{R}^s$. Here, *symmetric* means that $\Sigma z \in D^{|z|}$ holds for all $z \in D^{|z|}$ and $\Sigma = \operatorname{diag}(\sigma)$ with $\sigma \in \{-1, 0, 1\}^s$. Thus, $0 \in D^{|z|}$ and $|z| \in D^{|z|}$ hold in particular. Further, the short-hand $D^{x,|z|}$ for $D^x \times D^{|z|}$ is used.

In the following, the relations $<$, $>$, $\leq$ and $\geq$ as well as the functions $|\cdot|$, $\max(\cdot, \cdot)$ and $\min(\cdot, \cdot)$ are considered componentwise.

Chapter 2

Basic Concepts

This chapter introduces the basic concepts of the fields of optimization which are part of this thesis. In particular, smooth nonlinear programs (smoothNLPs) are studied in section 2.1. Then, unconstrained abs-normal NLPs (unNLPs) are presented in section 2.2 and mathematical programs with equilibrium constraints (MPECs) in section 2.3. In all three sections the problem class is introduced and regularity assumptions as well as optimality conditions are stated.

2.1 Nonlinear Smooth Optimization

This section presents the basic concepts of nonlinear smooth optimization. Herein, the smooth nonlinear program is considered which is presented in subsection 2.1.1. Then, the linear independence constraint qualification is defined in subsection 2.1.2 and optimality conditions are given in subsection 2.1.3. Finally, weaker constraint qualifications are discussed in subsection 2.1.4.

This section is based on [21], [2] and [6].

2.1.1 Formulation

First, the smooth nonlinear program is defined.

Definition 2.1 (Smooth NLP)**.** Let D be an open subset of $\mathbb{R}^n$ and $\mathcal{E}, \mathcal{I}$ disjoint finite index sets. The *smooth NLP* reads

$$\begin{aligned} \min_{x \in D} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x) = 0, \\ & c_{\mathcal{I}}(x) \geq 0, \end{aligned} \tag{smoothNLP}$$

with the *objective function* $f \in C^d(D, \mathbb{R})$, the vector of *equality constraints* $c_{\mathcal{E}} \in C^d(D, \mathbb{R}^{m_1})$ and the vector of *inequality constraints* $c_{\mathcal{I}} \in C^d(D, \mathbb{R}^{m_2})$ for $d \geq 1$. Here, $m_1 = |\mathcal{E}|$ and $m_2 = |\mathcal{I}|$ hold and the relation $\geq$ is defined componentwise. The *feasible set* of (smoothNLP) is defined as

$$\mathcal{F} := \{x \in D \colon c_{\mathcal{E}}(x) = 0,\ c_{\mathcal{I}}(x) \geq 0\}.$$

A point $x \in D$ is called *feasible* if $x \in \mathcal{F}$.

A local minimizer of (smoothNLP) is defined next.

Definition 2.2 (Local Minimizer)**.** A point $x^* \in D$ is a *local minimizer* of (smoothNLP) if $x^* \in \mathcal{F}$ and there exists a neighborhood $\mathcal{N}$ of x^* such that $f(x^*) \leq f(x)$ for all $x \in \mathcal{N} \cap \mathcal{F}$. A point $x^* \in D$ is a *strict local minimizer* of (smoothNLP) if $x^* \in \mathcal{F}$ and there exists a neighborhood $\mathcal{N}$ of x^* such that $f(x^*) < f(x)$ for all $x \in \mathcal{N} \cap \mathcal{F}$ with $x \neq x^*$.

2.1.2 LICQ

To state optimality conditions for (smoothNLP) a regularity assumption is needed. This leads to the definition of the linear independence constraint qualification which is given in this subsection.

For this purpose, the active inequality set is defined.

Definition 2.3 (Active Inequality Set)**.** Consider a feasible point x of (smoothNLP). The inequality constraint $i \in \mathcal{I}$ is called *active* if $c_i(x) = 0$ holds and *inactive* otherwise. The *active inequality set* $\mathcal{A}(x)$ consists of all indices of active inequality constraints,

$$\mathcal{A}(x) := \{i \in \mathcal{I} \colon c_i(x) = 0\}.$$

The previous definition deviates from standard literature in considering only inequality constraints. This is done to simplify notation in the following chapters.

Definition 2.4 (Active Jacobian)**.** Consider a feasible point x of (smoothNLP) and set $\mathcal{A} = \mathcal{A}(x)$ and $c_{\mathcal{A}} = [c_i]_{i \in \mathcal{A}}$. The *active Jacobian* is

$$J(x) := \begin{bmatrix} J_{\mathcal{E}}(x) \\ J_{\mathcal{A}}(x) \end{bmatrix} \in \mathbb{R}^{(m_1 + |\mathcal{A}|) \times n}.$$

It is composed of the *equality-constraints Jacobian*

$$J_{\mathcal{E}}(x) := \partial_x c_{\mathcal{E}}(x) \in \mathbb{R}^{m_1 \times n}$$

and of the *active inequality-constraints Jacobian*

$$J_{\mathcal{A}}(x) := \partial_x c_{\mathcal{A}}(x) \in \mathbb{R}^{|\mathcal{A}| \times n}.$$

Definition 2.5 (LICQ)**.** Consider a feasible point x of (smoothNLP). One says that the *linear independence constraint qualification (LICQ)* holds at x if the active Jacobian

$$J(x) = \begin{bmatrix} J_{\mathcal{E}}(x) \\ J_{\mathcal{A}}(x) \end{bmatrix} \in \mathbb{R}^{(m_1 + |\mathcal{A}|) \times n}$$

has full row rank $m_1 + |\mathcal{A}|$.

If LICQ holds at x, a matrix whose columns are a basis of $\ker(J(x))$ will be denoted by $U(x) \in \mathbb{R}^{n \times [n - (m_1 + |\mathcal{A}|)]}$. Thus, $\ker(J(x)) = \operatorname{im}(U(x))$ holds.

2.1.3 Optimality Conditions

In this subsection optimality conditions under LICQ are stated.

For that, the Lagrangian function of (smoothNLP) is introduced.

Definition 2.6 (Lagrangian function)**.** The *Lagrangian function* of (smoothNLP) is defined as

$$\mathcal{L}(x, \lambda, \mu) := f(x) + \lambda^T c_{\mathcal{E}}(x) - \mu^T c_{\mathcal{I}}(x).$$

The vectors $\lambda \in \mathbb{R}^{m_1}$ and $\mu \in \mathbb{R}^{m_2}$ are called *Lagrange multiplier vectors.*

Note that this deviates from [21] in sign of the Lagrange multiplier vector λ. This has no consequences on the following theory as the vector of equality constraints $c_{\mathcal{E}}$ can be multiplied by -1 without changing (smoothNLP).

Now, optimality conditions for (smoothNLP) can be stated. They form the basis to formulate optimality conditions for particular nonsmooth optimization problems in chapter 3.

Theorem 2.7 (First Order Necessary Conditions)**.** *Assume that x^* is a local minimizer of (smoothNLP) and that LICQ holds at x^*. Then, there exist Lagrange multiplier vectors λ^* and μ^* such that the following conditions are satisfied:*

$$\begin{aligned}
\partial_x \mathcal{L}(x^*, \lambda^*, \mu^*) &= 0 \quad \textit{(stationarity)}, && (2.1a)\\
c_{\mathcal{E}}(x^*) &= 0, && (2.1b)\\
c_{\mathcal{I}}(x^*) &\geq 0 \quad \textit{(primal feasibility)}, && (2.1c)\\
\mu^* &\geq 0 \quad \textit{(nonnegativity)}, && (2.1d)\\
(\mu^*)^T c_{\mathcal{I}}(x^*) &= 0 \quad \textit{(complementarity)}. && (2.1e)
\end{aligned}$$

Proof. A proof may be found in [21, Section 12.4] and more details in subsection 2.1.4. □

The conditions (2.1) are also called Karush-Kuhn-Tucker conditions (in short KKT conditions). If the conditions (2.1) are satisfied, x^* is called stationary point and the triple (x^*, λ^*, μ^*) KKT point.

Lemma 2.8. *Assume that (x^*, λ^*, μ^*) is a KKT point and that LICQ holds at x^*. Then, the Lagrange multiplier vectors λ^* and μ^* are unique.*

Proof. The condition (2.1a) reads

$$f'(x^*) + (\lambda^*)^T \partial_x c_{\mathcal{E}}(x^*) - (\mu^*)^T \partial_x c_{\mathcal{I}}(x^*) = 0.$$

Conditions (2.1d) and (2.1e) imply $\mu_i^* = 0$ for $i \notin \mathcal{A}(x^*)$. With the notation $\mathcal{A} = \mathcal{A}(x^*)$, $\mu_{\mathcal{A}}^* = [\mu_i^*]_{i \in \mathcal{A}}$ and $c_{\mathcal{A}} = [c_i]_{i \in \mathcal{A}}$ this gives

$$f'(x^*) + (\lambda^*)^T \partial_x c_{\mathcal{E}}(x^*) - (\mu_{\mathcal{A}}^*)^T \partial_x c_{\mathcal{A}}(x^*) = 0$$

which can be written as

$$-f'(x^*)^T = \begin{bmatrix} \partial_x c_{\mathcal{E}}(x^*)^T & \partial_x c_{\mathcal{A}}(x^*)^T \end{bmatrix} \begin{bmatrix} \lambda^* \\ -\mu_{\mathcal{A}}^* \end{bmatrix} = J(x^*)^T \begin{bmatrix} \lambda^* \\ -\mu_{\mathcal{A}}^* \end{bmatrix}.$$

Then, the vectors λ^* and $\mu_{\mathcal{A}}^*$ are unique because the matrix $J(x^*)^T$ has full column rank by LICQ. □

Before second order conditions are presented, a special case of complementarity is defined.

Definition 2.9 (Strict Complementarity)**.** Assume that (x^*, λ^*, μ^*) is a KKT point. One says that *strict complementarity* holds if $\mu_i^* > 0$ for all $i \in \mathcal{A}(x^*)$. In other words, either $\mu_i^* = 0$ or $c_i(x^*) = 0$ holds for each index $i \in \mathcal{I}$ but not both.

The necessary and sufficient second order optimality conditions for (smoothNLP) are given in the next two theorems.

Theorem 2.10 (Second Order Necessary Conditions)**.** *Consider* (smoothNLP) *for $d \geq 2$. Assume that x^* is a local minimizer and that LICQ holds at x^*. Denote by λ^* and μ^* the unique Lagrange multiplier vectors and assume further that strict complementarity holds at (x^*, λ^*, μ^*). Then,*

$$U(x^*)^T H(x^*, \lambda^*, \mu^*) U(x^*) \geq 0,$$

where $H(x^*, \lambda^*, \mu^*) = \partial_{xx}^2 \mathcal{L}(x^*, \lambda^*, \mu^*) \in \mathbb{R}^{n \times n}$.

Proof. A proof may be found in [21, Section 12.5]. □

Theorem 2.11 (Second Order Sufficient Conditions)**.** *Consider* (smoothNLP) *for $d \geq 2$. Given a feasible point x^*, assume that LICQ holds at x^*. Assume further that Lagrange multiplier vectors λ^* and μ^* exist such that the first order necessary conditions* (2.1) *are satisfied with strict complementarity at (x^*, λ^*, μ^*) and that*

$$U(x^*)^T H(x^*, \lambda^*, \mu^*) U(x^*) > 0.$$

Then, x^* is a strict local minimizer of (smoothNLP)*.*

Proof. A proof may be found in [21, Section 12.5]. □

Under some additional assumptions the second order conditions are not needed, for example, if all functions are linear. Then, the necessary first order conditions are sufficient for x^* to be a global minimizer due to convexity of (smoothNLP). Besides, LICQ is no longer required as the desired regularity follows directly by linearity of the constraints.

Theorem 2.12 (First Order Necessary and Sufficient Conditions for linear functions)**.** *Given* (smoothNLP)*, assume that f, $c_{\mathcal{E}}$ and $c_{\mathcal{I}}$ are linear. Then, x^* is a global minimizer if and only if there exist λ^* and μ^* such that the conditions* (2.1) *are satisfied.*

Proof. A proof may be found in [21, Section 13.1]. □

Another special case is that the number of equality and active inequality constraints is equal to the dimension. Then, the second order conditions are trivial and thus always satisfied.

Corollary 2.13 (First Order Sufficient Conditions for $m_1 + |\mathcal{A}| = n$)**.** *Given* (smoothNLP) *for $d \geq 2$, assume that $m_1 + |\mathcal{A}| = n$ holds. Consider a feasible point x^* and assume that LICQ holds at x^*. Then, x^* is a strict local minimizer of* (smoothNLP) *if there exist λ^* and μ^* such that the conditions* (2.1) *with strict complementarity at (x^*, λ^*, μ^*) are satisfied.*

Proof. As $m_1 + |\mathcal{A}| = n$ holds, the matrix $U(x)$ is empty. Thus, Theorem 2.10 gives the result. □

2.1.4 Weaker CQs

In this subsection the regularity assumption LICQ is weakened. To this aim a closer look at concepts which are substantial for the proof of Theorem 2.7 is taken.

First, the tangential and the linearized cone are defined.

Definition 2.14 (Tangential and Linearized Cone)**.** Given (smoothNLP), consider $x \in \mathcal{F}$. The *tangential cone* to $\mathcal{F}$ at x is defined as

$$\mathcal{T}(x) := \{\delta \in \mathbb{R}^n \colon \exists\ \tau_k \searrow 0,\ \mathcal{F} \ni x_k \to x \text{ such that } \tau_k^{-1}(x_k - x) \to \delta\}.$$

Here, the notation $\tau_k \searrow 0$ means that the sequence $\{\tau_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}_{>0}$ converges to 0. Further, $\mathcal{F} \ni x_k \to x$ is short-hand for that the sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq \mathcal{F}$ converges to x.
Set $\mathcal{A} = \mathcal{A}(x)$ and $c_{\mathcal{A}} = [c_i]_{i\in\mathcal{A}}$. The *linearized cone* is defined as

$$\mathcal{T}^{\text{lin}}(x) := \{\delta \in \mathbb{R}^n \colon \partial_x c_{\mathcal{E}}(x)^T\delta = 0,\ \partial_x c_{\mathcal{A}}(x)^T\delta \geq 0\}.$$

In the previous definition the term *cones* was used for the sets $\mathcal{T}(x)$ and $\mathcal{T}^{\text{lin}}(x)$. The formal definition of a cone is given next; here 0 is always an element of the cone. But, note that this is handled differently across literature.

Definition 2.15 (Cone)**.** A set $C \subseteq \mathbb{R}^n$ is called *cone* if $\alpha c \in C$ for every element $c \in C$ and all $\alpha \in \mathbb{R}_{\geq 0}$.

By definition of the sets $\mathcal{T}(x)$ and $\mathcal{T}^{\text{lin}}(x)$ it follows directly that they are cones in the sense of the above definition.

Without any regularity assumption the following necessary condition holds.

Theorem 2.16. *Assume that x^* is a local minimizer of* (smoothNLP)*. Then,*

$$f'(x^*)^T\delta \geq 0 \quad \textit{for all } \delta \in \mathcal{T}(x^*). \tag{2.2}$$

Proof. A proof may be found in [21, Theorem 12.3] or in [2, Proposition 3.3.15]. □

The condition (2.2) is called Bouligand stationarity or B-stationarity.

The next theorem converts the KKT conditions into a condition on the linearized cone.

Theorem 2.17. *Assume that x^* is a local minimizer of* (smoothNLP)*. Then, there exist Lagrange multiplier vectors λ^* and μ^* such that the KKT conditions* (2.1) *are satisified if and only if*

$$f'(x^*)^T\delta \geq 0 \quad \textit{for all } \delta \in \mathcal{T}^{lin}(x^*).$$

Proof. A proof may be found in [2, Proposition 3.3.14]. □

To combine the previous two theorems the definition of the dual cone is needed.

Definition 2.18 (Dual Cone)**.** Given an arbitrary cone $C \subseteq \mathbb{R}^n$, the *dual cone* C^* is defined as

$$C^* := \{\omega \in \mathbb{R}^n \colon \omega^T\delta \geq 0 \text{ for all } \delta \in C\}.$$

Thus, under the assumption that $\mathcal{T}(x^*)^* = \mathcal{T}^{\text{lin}}(x^*)^*$, Theorem 2.16 and Theorem 2.17 show that a local minimizer x^* of (smoothNLP) is a stationary point. This motivates the next definition.

Definition 2.19 (ACQ and GCQ)**.** Consider a feasible point x of (smoothNLP). One says that *Abadie's Constraint Qualification (ACQ)* is satisfied at x if

$$\mathcal{T}(x) = \mathcal{T}^{\text{lin}}(x)$$

and *Guignard's Constraint Qualification (GCQ)* if

$$\mathcal{T}(x)^* = \mathcal{T}^{\text{lin}}(x)^*.$$

It follows directly that ACQ implies GCQ. Moreover, LICQ implies ACQ as can be seen in [21, Lemma 12.2]. Thus, the proof of Theorem 2.7 follows directly from Theorem 2.16 and Theorem 2.17.

ACQ and GCQ involve the tangential cone and hence both conditions are difficult to check. A constraint qualification which is easier to check is MFCQ.

Definition 2.20 (MFCQ)**.** Consider a feasible point x of (smoothNLP). One says that the *Mangasarian Fromovitz constraint qualification (MFCQ)* holds at x if the equality constraints Jacobian

$$J_{\mathcal{E}}(x) \in \mathbb{R}^{m_1 \times n}$$

has full row rank m_1 and if there exists a vector $d \in \mathbb{R}^n$ such that

$$J_{\mathcal{E}}(x)d = 0 \quad \text{and} \quad J_{\mathcal{A}}(x)d > 0.$$

It is weaker than LICQ but stronger than ACQ as the next lemma shows. This also implies that under MFCQ every local minimizer of (smoothNLP) is a stationary point.

Lemma 2.21. *Consider a feasible point x of* (smoothNLP)*. Then, the following chain of implications holds at x:*

$$\textit{LICQ} \implies \textit{MFCQ} \implies \textit{ACQ} \implies \textit{GCQ}$$

Proof. Due to LICQ at x the system

$$J(x)w = \begin{bmatrix} J_{\mathcal{E}}(x) \\ J_{\mathcal{A}}(x) \end{bmatrix} w = \begin{bmatrix} 0 \\ e \end{bmatrix}$$

with $0 \in \mathbb{R}^{m_1}$ and $e = (1, \dots, 1)^T \in \mathbb{R}^{|\mathcal{A}|}$ has a solution $w \in \mathbb{R}^n$. Thus, MFCQ holds at x with the choice $d = w$ (see [21, section 12.6]). The second implication can be found in [6, Satz 2.39] and the third implication follows directly by definition of the dual cone. $\square$

2.2 Unconstrained Abs-Normal NLP

In this section the class of unconstrained abs-normal NLPs is presented. Therefore, the abs-normal form is introduced in subsection 2.2.1 and used in subsection 2.2.2 to formulate the unconstrained abs-normal NLP. Then, the linear independence kink qualification (LIKQ) is stated in subsection 2.2.3 and optimality conditions under this regularity assumption are given in subsection 2.2.4. Moreover, the Mangasarian Fromovitz kink qualification is formulated in subsection 2.2.5.

This section is based on the work of Griewank and Walther in [7, 8, 10] but deviates in notation as $c_{\mathcal{Z}}$ instead of F is used in the abs-normal form and no short-hand notations for partial derivatives are introduced. Besides, open domains D^x and $D^{|z|}$ are considered as continuous derivatives and the implicit function theorem are needed.

The extension to an m-dimensional abs-normal form (in subsection 2.2.1) and the added proofs are published in [17].

2.2.1 Abs-normal form

This subsection introduces level-1 nonsmooth functions and their representation in abs-normal form. It is based on the work of Griewank and Walther in [8]. But here, the definitions of the case $m = 1$ are extended to functions $\varphi\colon D^x \to \mathbb{R}^m$ in the spirit of Griewank's m-dimensional abs-normal form [7].

First, a level-1 nonsmooth function is defined.

Definition 2.22 (Level-1 Nonsmooth Function)**.** Let D^x be an open subset of $\mathbb{R}^n$. A function $\varphi\colon D^x \to \mathbb{R}$ is called *level-1 nonsmooth* if it can be formulated as a composition of smooth functions and the absolute value function. A function $\varphi\colon D^x \to \mathbb{R}^m$ is called level-1 nonsmooth if all component functions $\varphi_i\colon D^x \to \mathbb{R}$ for $i = 1, \dots, m$ are level-1 nonsmooth.

Example 2.23. The maximum and the minimum function are level-1 nonsmooth as both can be rewritten with help of the absolute value function:

$$\max(x, y) = \frac{1}{2}(x + y + |x - y|) \quad \text{and} \quad \min(x, y) = \frac{1}{2}(x + y - |x - y|).$$

For an arbitrary level-1 nonsmooth function $\varphi\colon D^x \to \mathbb{R}^m$, one can replace every argument of an absolute value by variables z_i, $i = 1, \dots, s$. This can be done from left to right and from inside to outside if nested absolute value evaluations occur. Moreover, already defined variables z_i can be reused if arguments repeat. Then, φ can be expressed as the system

$$\begin{aligned} \varphi(x) &= f(x, |z|) \quad \text{for all } x \in D^x, \\ z &= c_{\mathcal{Z}}(x, |z|), \end{aligned}$$

where $f\colon D^x \times D^{|z|} \to \mathbb{R}^m$ and $c_{\mathcal{Z}}\colon D^x \times D^{|z|} \to \mathbb{R}^s$ are smooth with $D^{|z|} \subseteq \mathbb{R}^s$.

Note that w.l.o.g. $0 \in D^{|z|}$ can be assumed because otherwise the terms $|\cdot|$ would be irrelevant. Moreover, the partial derivative $\partial_2 c_{\mathcal{Z}}(x, |z|)$ is strictly lower triangular by construction because z_i can influence z_j only for $i < j$, in other words $z_j = c_j(x, |z_1|, \dots, |z_{j-1}|)$ holds.

The next example illustrates this procedure. Note, that it is based on the second Nesterov variant considered in [11] and modified to illustrate handling of repeating arguments.

Example 2.24. Consider the level-1 nonsmooth function

$$\varphi\colon \mathbb{R}^n \to \mathbb{R},\ \varphi(x) = \frac{1}{4}|x_1| + \sum_{k=1}^{n-1} |x_{k+1} - 2|x_k| + 1|.$$

Following the described method, it can be rewritten as

$$\varphi(x) = \frac{1}{4}|z_1| + \sum_{k=1}^{n-1} |z_{2k}| \text{ with } z_{2k} = x_{k+1} - 2|z_{2k-1}| + 1,\ z_{2k-1} = x_k.$$

The partial derivative $\partial_2 c_{\mathcal{Z}}(x, |z|) \in \mathbb{R}^{2(n-1)\times 2(n-1)}$ is strictly lower triangular as

$$[\partial_2 c_{\mathcal{Z}}(x,|z|)]_{i,j} = \begin{cases} -2, & \text{if } (i,j) = (2k, 2k-1) \\ 0, & \text{otherwise} \end{cases} \quad \text{for } k = 1, \dots, n-1.$$

This reformulation motivates the definition of the abs-normal form.

Definition 2.25 (Abs-Normal Form)**.** One says that a function $\varphi\colon D^x \to \mathbb{R}^m$ is given in *abs-normal form* if $f\colon D^{x,|z|} \to \mathbb{R}^m$ and $c_{\mathcal{Z}}\colon D^{x,|z|} \to \mathbb{R}^s$, with $D^{|z|} \subseteq \mathbb{R}^s$ open and symmetric, exist such that

$$\varphi(x) = f(x, |z|), \tag{2.3a}$$
$$z = c_{\mathcal{Z}}(x, |z|) \quad \text{with } \partial_2 c_{\mathcal{Z}}(x, |z|) \text{ strictly lower triangular.} \tag{2.3b}$$

The variables z_i, $i = 1, \dots, s$, are called *switching variables* and (2.3b) is called *switching system*.

Recall section 1.5: $D^{|z|}$ symmetric means that $\Sigma z \in D^{|z|}$ holds for all $z \in D^{|z|}$ and $\Sigma = \operatorname{diag}(\sigma)$ with $\sigma \in \{-1, 0, 1\}^s$.

Further, note that the assumption of $D^{|z|}$ open and symmetric is made here. This is due to the implicit function theorem which is needed in Lemma 2.31.

The switching system (2.3b) provides two ways of computing and characterizing the switching variables. On one hand the switching variables z_j can be computed one by one from x (and z_i with $i < j$) as the partial derivative $\partial_2 c_{\mathcal{Z}}(x, |z|)$ is strictly lower triangular. In the following the notation $z(x)$ is used to denote this dependence on x explicitly. On the other hand, z is implicitly defined by the switching system (2.3b).

Definition 2.26 (Class $C^d_{\text{abs}}(D^x)$)**.** Let D^x be an open subset of $\mathbb{R}^n$. The set of functions $\varphi\colon D^x \to \mathbb{R}^m$ in abs-normal form (2.3) with $f \in C^d(D^{x,|z|}, \mathbb{R}^m)$ and $c_{\mathcal{Z}} \in C^d(D^{x,|z|}, \mathbb{R}^s)$ is denoted by $C^d_{\text{abs}}(D^x, \mathbb{R}^m)$.

Example 2.27. The maximum and the minimum function are level-1 nonsmooth and can be reformulated in abs-normal form:

$$\max(x, y) = \frac{1}{2}(x + y + |z|) \text{ and } \min(x, y) = \frac{1}{2}(x + y - |z|) \text{ both with } z = x - y.$$

Moreover, they are elements of C^∞_{abs}.

Note that the abs-normal form itself is not unique. To show this a closer look at example 2.24 is taken.

Example 2.28. The function $\varphi(x) = \frac{1}{4}|x_1| + \sum_{k=1}^{n-1} |x_{k+1} - 2|x_k| + 1|$ can also be stated as

$$\varphi(x) = \frac{1}{4}|z_1| + \sum_{k=1}^{n-1} |z_{n+k-1}| \text{ with } z_k = x_k,\ z_{n+k-1} = x_{k+1} - 2|z_k| + 1$$

with strictly lower triangular partial derivative $\partial_2 c_{\mathcal{Z}}(x, |z|) \in \mathbb{R}^{2(n-1)\times 2(n-1)}$ given as

$$[\partial_2 c_{\mathcal{Z}}(x, |z|)]_{i,j} = \begin{cases} -2, & \text{if } (i, j) = (n + k - 1, k) \\ 0, & \text{otherwise} \end{cases} \quad \text{for } k = 1, \dots, n - 1.$$

2.2.2 Formulation

Griewank and Walther considered in [8] the unconstrained optimization problem

$$\min_{x \in D^x} \varphi(x)$$

with $D^x \subseteq \mathbb{R}^n$ open and $\varphi \in C^d_{\text{abs}}(D^x, \mathbb{R})$ for $d \geq 1$.

Then, $\varphi(x)$ can be rewritten in abs-normal form:

$$\varphi(x) = f(x, |z|),\ c_{\mathcal{Z}}(x, |z|) = z.$$

This leads to the definition of an unconstrained abs-normal NLP.

Definition 2.29 (Unconstrained Abs-Normal NLP)**.** A nonsmooth unconstrained optimization problem is called an *unconstrained abs-normal NLP* if functions $f \in C^d(D^x, \mathbb{R})$ and $c_{\mathcal{Z}} \in C^d(D^{x,|z|}, \mathbb{R}^s)$ for $d \geq 1$ exist such that the NLP can equivalently be stated as

$$\begin{aligned} \min_{(x,z)\in D^{x,|z|}} \quad & f(x, |z|) \\ \text{s.t.} \quad & c_{\mathcal{Z}}(x, |z|) - z = 0, \end{aligned} \tag{unNLP}$$

where $D^{|z|}$ is symmetric and $\partial_2 c_{\mathcal{Z}}(x, |z|)$ is strictly lower triangular. The feasible set of (unNLP) is denoted by

$$\mathcal{F}_{\text{un}} := \{(x, z) \in D^{x,|z|} \colon c_{\mathcal{Z}}(x, |z|) = z\}.$$

As $\partial_2 c_{\mathcal{Z}}(x, |z|)$ is strictly lower triangular, z can be computed successively from x via $c_{\mathcal{Z}}$. Thus, the feasible set of (unNLP) can be rewritten as

$$\mathcal{F}_{\text{un}} = \{(x, z(x)) \colon x \in D^x\}.$$

Hence, for every $x \in D^x$ the tuple $(x, z(x))$ is feasible for (unNLP) without any additional assumptions.

2.2.3 LIKQ

Griewank and Walther introduced in [8] the linear independence kink qualification (LIKQ) as the fundamental nonsmooth regularity condition.

First, the solvability of the nonlinear switching system $z = c_{\mathcal{Z}}(x, |z|)$ is considered. To this end, the reformulation $|z_i| = \text{sign}(z_i)z_i$ is used. This leads to the definition of a signature.

Definition 2.30 (Signature of z)**.** For $x \in D^x$ the *signature* $\sigma(x)$ and the associated *signature matrix* $\Sigma(x)$ are defined as

$$\sigma(x) := \text{sign}(z(x)) \in \{-1, 0, 1\}^s \quad \text{and} \quad \Sigma(x) := \text{diag}(\sigma(x)).$$

If $\sigma(x) \in \{-1, 1\}^s$ holds, $\sigma(x)$ and $\Sigma(x)$ are called *definite*, otherwise *indefinite*.

For signatures $\sigma, \hat{\sigma} \in \{-1, 0, 1\}^s$, the partial order

$$\hat{\sigma} \succeq \sigma :\Leftrightarrow \hat{\sigma}_i \sigma_i \geq \sigma_i^2 \text{ for } i = 1, \dots, s$$

is used, i.e., $\hat{\sigma}_i$ is arbitrary if $\sigma_i = 0$ and $\hat{\sigma}_i = \sigma_i$ otherwise.

Observe that the absolute value can be expressed as $|z| = \hat{\Sigma} z$ where $\hat{\sigma}_i = \text{sign}(z_i)$ for $z_i \neq 0$ and arbitrary $\hat{\sigma}_i \in \{-1, 1\}$ otherwise. The resulting vectors $\hat{\sigma}$ are exactly the *definite* signatures that satisfy the partial order $\hat{\sigma} \succeq \sigma(x)$. Those will be used to define so-called branch NLPs below. More generally, one can allow arbitrary $\hat{\sigma}_i \in \{-1, 0, 1\}$ for $z_i = 0$. Application of the implicit function theorem to the switching system for z then leads to the next lemma. Note, that the formal assumptions and the proof were omitted in [8].

Lemma 2.31. *Consider* (unNLP) *and fix* $\tilde{x} \in D^x$*. Let* $\tilde{z} = z(\tilde{x})$*,* $\tilde{\sigma} = \sigma(\tilde{x})$ *and* $\Sigma = \text{diag}(\sigma)$ *for arbitrary* $\sigma \in \{-1, 0, 1\}^s$ *with* $\sigma \succeq \tilde{\sigma}$*. Then, the switching system* $z = c_{\mathcal{Z}}(x, \Sigma z)$ *has locally a unique solution* $z^\sigma(x)$ *which is* d *times continuously differentiable with Jacobian*

$$\partial_x z^\sigma(x) = [I - \partial_2 c_{\mathcal{Z}}(x, \Sigma z^\sigma(x))\Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(x, \Sigma z^\sigma(x)) \in \mathbb{R}^{s \times n}.$$

Proof. Set $h(x, z) := c_{\mathcal{Z}}(x, \Sigma z) - z$. Then, $h(x, z) \in C^d(D^{x,|z|}, \mathbb{R}^s)$ and $h(\tilde{x}, \tilde{z}) = 0$ holds. Finally, the partial derivative $\partial_2 h(\tilde{x}, \tilde{z}) = \partial_2 c_{\mathcal{Z}}(\tilde{x}, \Sigma\tilde{z})\Sigma - I = \partial_2 c_{\mathcal{Z}}(\tilde{x}, |\tilde{z}|)\Sigma - I$ is regular as $\partial_2 c_{\mathcal{Z}}(\tilde{x}, |\tilde{z}|)\Sigma$ is strictly lower triangular. Thus, the assumptions of the implicit function theorem (see, e.g, [1, Corollary 8.4 in Chapter VII]) are satisfied and it can be applied. Hence, there exist neighborhoods $\tilde{\mathcal{N}}^x$ and $\tilde{\mathcal{N}}^z$ of $\tilde{x}$ and $\tilde{z}$ resp. such that $c_{\mathcal{Z}}(x, \Sigma z) - z = 0$ has a unique solution $z^\sigma(x) \in \tilde{\mathcal{N}}^z$ for all $x \in \tilde{\mathcal{N}}^x$. Furthermore, z^σ is d times continuously differentiable, i.e. $z^\sigma \in C^d(\tilde{\mathcal{N}}^x, \tilde{\mathcal{N}}^z)$, and the Jacobian is given as

$$\begin{aligned}
\partial_x z^\sigma(x) &= -\partial_2 h(x, z^\sigma(x))^{-1} \partial_1 h(x, z^\sigma(x)) \\
&= -\left(\partial_2 c_{\mathcal{Z}}(x, \Sigma z^\sigma(x))\Sigma - I\right)^{-1} \partial_1 c_{\mathcal{Z}}(x, \Sigma z^\sigma(x)) \\
&= \left(I - \partial_2 c_{\mathcal{Z}}(x, \Sigma z^\sigma(x))\Sigma\right)^{-1} \partial_1 c_{\mathcal{Z}}(x, \Sigma z^\sigma(x)).
\end{aligned}$$

□

By construction, $z^\sigma(\tilde{x}) = z(\tilde{x})$ and $\Sigma z^\sigma(\tilde{x}) = |z(\tilde{x})|$ hold and thus

$$\partial_x z^\sigma(\tilde{x}) = [I - \partial_2 c_{\mathcal{Z}}(\tilde{x}, |z(\tilde{x})|)\Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(\tilde{x}, |z(\tilde{x})|).$$

Since, $z^\sigma(x)$ and its Jacobian $\partial_x z^\sigma(x)$ are only considered at the point $\tilde{x}$ with associated signature vector $\tilde{\sigma}$, the index σ on z^σ and $\partial_x z^\sigma$ is dropped in the following, writing $z(\tilde{x})$ and $\partial_x z(\tilde{x})$.

Definition 2.32 (Active Switching Variables)**.** Given (unNLP), consider $x \in D^x$. The switching variable z_i is called *active* if $z_i(x) = 0$ and *inactive* otherwise. The *active switching set* $\alpha(x)$ collects all indices of active switching variables,

$$\alpha(x) := \{i \in \{1, \dots, s\} \colon z_i(x) = 0\}.$$

Thus, there are $|\alpha(x)|$ active switching variables and $|\sigma(x)| := s - |\alpha(x)|$ inactive ones.

Definition 2.33 (Active Switching Jacobian)**.** Given (unNLP), consider $x \in D^x$ and set $\alpha = \alpha(x)$, $\sigma = \sigma(x)$ and $\Sigma = \operatorname{diag}(\sigma)$. The *active switching Jacobian* is defined as

$$J_\alpha(x) := \left[e_i^T \partial_x z(x)\right]_{i \in \alpha} = \left[e_i^T [I - \partial_2 c_{\mathcal{Z}}(x, |z(x)|)\Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(x, |z(x)|)\right]_{i \in \alpha}.$$

Similar to the smooth case, optimality conditions for (eqNLP) require certain regularity assumptions. The linear independence kink qualification (LIKQ) provides a strong regularity guarantee. It reduces to the classical LICQ in the smooth case.

Definition 2.34 (LIKQ)**.** Given (unNLP), consider $x \in D^x$. One says that the *linear independence kink qualification (LIKQ)* holds at x if the active Jacobian

$$J_{\text{un}}(x) := J_\alpha(x) = \left[e_i^T \partial_x z(x)\right]_{i \in \alpha} \in \mathbb{R}^{|\alpha| \times n}$$

has full row rank $|\alpha|$.

If LIKQ holds at x a matrix which columns are a basis for the nullspace of $J_{\text{un}}(x)$ is needed in the following. It is denoted by $U_{\text{un}}(x) \in \mathbb{R}^{n \times (n - |\alpha|)}$.

2.2.4 Optimality Conditions

Griewank and Walther derived in [8] optimality conditions for the minimization of a level-1 nonsmooth function. Note, that their theory is summarized very shortly in this subsection as the unconstrained abs-normal NLP is a special case of the equality-constrained abs-normal NLP in chapter 3. There, all concepts are studied in more detail and proofs for the optimality conditions are given.

Localized Case To obtain optimality conditions, they considered at first a special case: the localized switching.

Definition 2.35 (Localization)**.** Given (unNLP), consider $x \in D^x$. One says that the switching is *localized* at x if $z(x) = 0$ and *non-localized* otherwise.

Then, the idea is to consider two types of smooth subproblems. The first one is the smooth trunk problem which is obtained by inserting $z = 0$ in (unNLP):

$$\begin{aligned} \min_{x \in D^x} \quad & f(x, 0) \\ \text{s.t.} \quad & c_{\mathcal{Z}}(x, 0) = 0. \end{aligned}$$

By construction, local minimizers of (unNLP) are inherited by the trunk problem. Thus, standard theory can be applied to obtain necessary conditions of first and second order.

The second type are the 2^s branch problems where the feasible set is divided into the areas of different possible signs of z. With $\Sigma = \operatorname{diag}(\sigma)$ for $\sigma \in \{-1, 1\}$ they read as follows:

$$\begin{aligned} \min_{(x,z) \in D^{x,|z|}} \quad & f(x, \Sigma z) \\ \text{s.t.} \quad & c_{\mathcal{Z}}(x, \Sigma z) - z = 0, \\ & \Sigma z \geq 0. \end{aligned}$$

Again by construction, local minimizers of (unNLP) are inherited by every branch problem. Moroever, a local minimizer of all branch problems has to be a local minimizer of (unNLP). Thus, standard theory can be applied to obtain first and second order necessary and second order sufficient conditions.

An appropiate combination of these two sets of conditions leads to the following optimality conditions for (unNLP).

Note that the switching feasibility is not mentioned explicitly in [8] as it follows directly from the context. Moreover, formal proofs were omitted.

Theorem 2.36 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0)$ is a local minimizer of (unNLP) and that LIKQ holds at x^*. Then, there exists λ^* such that the following conditions are satisfied:*

$$\begin{aligned} \partial_1 f(x^*, 0) + (\lambda^*)^T \partial_1 c_{\mathcal{Z}}(x^*, 0) &= 0 && \textit{(tangential stationarity)}, \\ \partial_2 f(x^*, 0) + (\lambda^*)^T \partial_2 c_{\mathcal{Z}}(x^*, 0) &\geq |\lambda^*|^T && \textit{(normal growth)}, \\ c_{\mathcal{Z}}(x^*, 0) &= 0 && \textit{(switching feasibility)}. \end{aligned}$$

Theorem 2.37 (Second Order Necessary Condition)**.** *Consider (unNLP) for $d \geq 2$. Assume that $(x^*, 0)$ is a local minimizer of (unNLP) and that LIKQ holds at x^*. Denote by λ^* the unique Lagrange multiplier vector. Then,*

$$U_{un}(x^*)^T H_{un}(x^*, \lambda^*) U_{un}(x^*) \geq 0,$$

where $H_{un}(x^*, \lambda^*) := \partial^2_{xx} \mathcal{L}_{un}(x^*, \lambda^*) \in \mathbb{R}^{n \times n}$ with $\mathcal{L}_{un}(x, \lambda) := f(x) + \lambda^T c_{\mathcal{Z}}(x, 0)$.

Theorem 2.38 (Second Order Sufficient Conditions)**.** *Consider (unNLP) for $d \geq 2$. Assume that $(x^*, 0)$ is feasible and that LIKQ holds at x^*. Assume further that a Lagrange multiplier vector λ^* exists such that the first order necessary conditions are satisfied with strict normal growth,*

$$\partial_2 f(x^*, 0) + (\lambda^*)^T \partial_2 c_{\mathcal{Z}}(x^*, 0) > |\lambda^*|^T,$$

and that

$$U_{un}(x^*)^T H_{un}(x^*, \lambda^*) U_{un}(x^*) > 0.$$

Then, $(x^*, 0)$ is a strict local minimizer of (unNLP).

Note that the second order conditions are always satisfied if the number of variables is equal to the number of switchings. Moreover, the second order sufficient conditions need not to be checked at all if f and $c_{\mathcal{Z}}$ are linear.

Generalized Case In the non-localized case there are active and inactive switching variables at the point of interest. For $x^* \in D^x$ the shorthand notations $z^* := z(x^*)$, $\sigma^* := \sigma(x^*)$ and $\alpha^* := \alpha(x^*)$ are used. For the inactive components

$$z_+ := (\sigma_i^* z_i)_{i \notin \alpha^*} = (|z_i|)_{i \notin \alpha^*} \in \mathbb{R}^{|\sigma^*|} \quad \text{and} \quad \sigma_+^* := (\sigma_i^*)_{i \notin \alpha^*}$$

are defined. By construction $z_+(x^*) > 0$ holds and z_+ keeps its positive sign in some neighborhood of z^* for x in some neighborhood B of x^*, by continuity of $c_{\mathcal{Z}}$. This leads to the relation

$$\operatorname{diag}(\sigma_+^*) z_+(x) = (z_i(x))_{i \notin \alpha^*} \quad \text{for} \quad x \in B.$$

For the active components the notations

$$z_0 := (z_i)_{i \in \alpha^*} \in \mathbb{R}^{|\alpha^*|} \quad \text{and} \quad \bar{z}_0 := |z_0|$$

are used. Then, $z_0(x^*) = 0$ holds but no additional information on the sign of z_0 in the neighborhood B is given. This leads to the definitions

$$\sigma_0(x) := \operatorname{sign}(z_0(x)) \in \{-1, 0, 1\}^{|\alpha^*|} \quad \text{and} \quad \alpha_0(x) := \{i \in \{1, \dots, |\alpha^*|\} \colon (z_0(x))_i = 0\}.$$

Further, the shorthand notations $z_+^* := z_+(x^*)$, $z_0^* := z_0(x^*) = 0$, $\bar{z}_0^* := \bar{z}_0(x^*) = 0$ and $\sigma_0^* := \sigma_0(x^*) = 0$ are used. To partition the switching constraints, the domains $D^{|z_0|}$ and D^{z_+} are defined via

$$D^{|z_0|} := P_{\alpha^*} D^{|z|} \quad \text{and} \quad D^{z_+} := \{z_+ \in \Sigma_+^* P_{(\alpha^*)^c} D^{|z|} \colon z_+ > 0\}$$

where $P_{\mathcal{S}} \in \mathbb{R}^{|\mathcal{S}| \times s}$ denotes the projector onto the subspace defined by $\mathcal{S} \subseteq \{1, \dots, s\}$ and $\Sigma_+^* = \operatorname{diag}(\sigma_+^*)$. Using the shorthand notation $D^{x,|z_0|,z_+} := B \times D^{|z_0|} \times D^{z_+}$, set

$$\tilde{c}_{\mathcal{Z}}(x, |z_0|, z_+) := c_{\mathcal{Z}}(x, |z|) \quad \text{with} \quad |z| = \Pi \begin{bmatrix} |z_0| \\ z_+ \end{bmatrix}$$

and further

$$c_+ := (\sigma_i^* e_i^T \tilde{c}_{\mathcal{Z}})_{i \notin \alpha^*} \in C^d(D^{x,|z_0|,z_+}, \mathbb{R}^{|\sigma^*|}) \quad \text{and} \quad c_0 := (e_i^T \tilde{c}_{\mathcal{Z}})_{i \in \alpha^*} \in C^d(D^{x,|z_0|,z_+}, \mathbb{R}^{|\alpha^*|}).$$

Here, Π denotes an appropiate permutation matrix.

Then, (unNLP) can be rewritten in a neighborhood of x^* in two different ways. For the first one, the variables and functions are split. This gives

$$\begin{aligned} \min_{(x,z_0,z_+)\in D^{x,|z_0|,z_+}} \quad & f(x,z_0,z_+) \\ \text{s.t.} \quad & c_0(x,|z_0|,z_+) - z_0 = 0, \\ & c_+(x,|z_0|,z_+) - z_+ = 0. \end{aligned} \tag{2.4}$$

To obtain the second one, the implicit function theorem is applied to $c_+(x,\bar{z}_0,z_+) = z_+$. Note, that the proof was omitted in [8].

Lemma 2.39. *For given (unNLP), the switching system $c_+(x,\bar{z}_0,z_+) = z_+$ with $\bar{z}_0 := |z_0|$ has locally around $(x^*,\bar{z}_0^*) = (x^*,0)$ a unique solution $z_+(x,\bar{z}_0)$ which is d times continuously differentiable with Jacobians*

$$\begin{aligned} \partial_x z_+(x,\bar{z}_0) &= [I - \partial_3 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0))]^{-1}\partial_1 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0)) \in \mathbb{R}^{|\sigma^*|\times n}, \\ \partial_{\bar{z}_0} z_+(x,\bar{z}_0) &= [I - \partial_3 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0))]^{-1}\partial_2 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0)) \in \mathbb{R}^{|\sigma^*|\times|\alpha^*|}. \end{aligned}$$

Proof. Set $h(x,\bar{z}_0,z_+) := c_+(x,\bar{z}_0,z_+) - z_+$, then $h(x,\bar{z}_0,z_+) \in C^d(D^{x,|z_0|,z_+},\mathbb{R}^{|\sigma^*|})$ with $D^{x,|z_0|,z_+}$ open. Moreover, $h(x^*,0,z_+^*) = 0$ holds and the partial derivative $\partial_3 h(x^*,0,z_+^*) = \partial_3 c_+(x^*,0,z_+^*) - I$ is regular as $\partial_3 c_+(x^*,0,z_+^*)$ is strictly lower triangular. Thus, the Implicit Function Theorem (see, e.g, [1, Corollary 8.4 in Chapter VII]) gives the existence of neighborhoods $\mathcal{N}^{x,|z_0|}$ and $\mathcal{N}^{z_+}$ of $(x^*,0)$ resp. z_+^* such that $c_+(x,\bar{z}_0,z_+) - z_+$ has a unique solution $z_+(x,\bar{z}_0) \in \mathcal{N}^{z_+}$ for all $(x,\bar{z}_0) \in \mathcal{N}^{x,|z_0|}$. Additionally, $z_+ \in C^d(\mathcal{N}^{x,|z_0|},\mathcal{N}^{z_+})$ holds with first order derivatives

$$\begin{aligned} \partial_x z_+(x,\bar{z}_0) &= -\partial_3 h(x,\bar{z}_0,z_+(x,\bar{z}_0))^{-1}\partial_1 h(x,\bar{z}_0,z_+(x,\bar{z}_0)) \\ &= -\left(\partial_3 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0)) - I\right)^{-1}\partial_1 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0)) \\ &= \left(I - \partial_3 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0))\right)^{-1}\partial_1 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0)) \\ \partial_{\bar{z}_0} z_+(x,\bar{z}_0) &= -\partial_3 h(x,\bar{z}_0,z_+(x,\bar{z}_0))^{-1}\partial_2 h(x,\bar{z}_0,z_+(x,\bar{z}_0)) \\ &= -\left(\partial_3 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0)) - I\right)^{-1}\partial_2 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0)) \\ &= \left(I - \partial_3 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0))\right)^{-1}\partial_2 c_+(x,\bar{z}_0,z_+(x,\bar{z}_0)) \end{aligned}$$

□

Note that $z_+(x^*,0) = z_+(x^*)$ holds. Inserting $z_+(x,\bar{z}_0)$ into (2.4), the problem reduces to

$$\begin{aligned} \min_{(x,z_0)\in\mathcal{N}^{x,|z_0|}} \quad & f(x,z_0,z_+(x,|z_0|)) \\ \text{s.t.} \quad & c_0(x,|z_0|,z_+(x,|z_0|)) - z_0 = 0. \end{aligned} \tag{2.5}$$

As in the localized case, trunk and branch problems can be considered for both formulations and standard theory for the smooth case leads to optimality conditions for (2.4) resp. (2.5). Then, the conditions for both formulations are combined in a suitable way to obtain a complete set of optimality conditions for (unNLP).

Thus, the first order conditions are derived using the split formulation. Once again, the switching feasibility is not mentioned explicitly in [8] as it follows directly from the context.

Theorem 2.40 (First Order Necessary Conditions). *Assume that* $(x^*, 0, z_+^*)$ *is a local minimizer of* (2.4) *and that LIKQ holds at* x^*. *Then, there exists* $\lambda^* = (\lambda_0^*, \lambda_+^*)$ *such that the following conditions are satisfied, where all functions and their partial derivatives are evaluated at* $(x^*, 0, z_+^*)$:

$$\begin{aligned} \partial_1 f + (\lambda_0^*)^T \partial_1 c_0 + (\lambda_+^*)^T \partial_1 c_+ &= 0, \\ \partial_3 f + (\lambda_0^*)^T \partial_3 c_0 + (\lambda_+^*)^T [\partial_3 c_+ - I] &= 0 && \textit{(tangential stationarity)}, \\ \partial_2 f + (\lambda_0^*)^T \partial_2 c_0 + (\lambda_+^*)^T \partial_2 c_+ &\geq |\lambda_0^*|^T && \textit{(normal growth)}, \\ c_+ - z_+^* &= 0 && \textit{(switching feasibility)}. \end{aligned}$$

The Lagrange multiplier vector λ^* *is unique.*

Proof. A proof may be found in [8, Proposition 4]. □

Whereas, the reduced formulation is chosen for the second order conditions. Note that formal proofs were omitted.

Theorem 2.41 (Second Order Necessary Condition). *Consider* (2.5) *for* $d \geq 2$. *Assume that* $(x^*, 0)$ *is a local minimizer of* (2.5) *and that LIKQ holds at* x^*. *Denote by* $\lambda^* = (\lambda_0^*, \lambda_+^*)$ *the unique Lagrange multiplier vector from Theorem 2.40. Then,*

$$U_{un}(x^*)^T H_{un}^r(x^*, \lambda_0^*) U_{un}(x^*) \geq 0$$

where $H_{un}^r(x, \lambda_0) = \partial_{xx}^2 \mathcal{L}_{un}^r(x, \lambda_0)$ *with* $\mathcal{L}_{un}^r(x, \lambda_0) = f(x, 0, z_+(x, 0)) + \lambda_0^T c_0(x, 0, z_+(x, 0))$.

Theorem 2.42 (Second Order Sufficient Conditions). *Consider* (2.5) *for* $d \geq 2$. *Assume that* $(x^*, 0)$ *is feasible and that LIKQ holds at* x^*. *Assume further that a Lagrange multiplier vector* $\lambda^* = (\lambda_0^*, \lambda_+^*) \in \mathbb{R}^{|\alpha^*| + |\sigma^*|}$ *exists such that the first order necessary conditions at* $(x^*, 0, z_+(x^*, 0))$ *are satisfied with strict normal growth,*

$$\partial_2 f(x^*) + (\lambda_0^*)^T \partial_2 c_0(x^*, 0, z_+(x^*, 0)) + (\lambda_+^*)^T \partial_2 c_+(x^*, 0, z_+(x^*, 0)) > |\lambda_0^*|^T$$

and that

$$U_{un}(x^*)^T H_{un}^r(x^*, \lambda_0^*) U_{un}(x^*) > 0.$$

Then, $(x^*, 0)$ *is a strict local minimizer of* (2.5).

2.2.5 MFKQ

Moreover, Griewank and Walther defined in [10] the Mangasarian Fromovitz kink qualification to weaken LIKQ. It was used to obtain equivalence between first order convexity and subdifferential regularity. These two concepts are not introduced here as they are not needed later. The definitions and details can be found in [10].

Definition 2.43 (MFKQ). Given (unNLP), consider $x \in D^x$ and set $\Sigma = \operatorname{diag}(\sigma)$ for $\sigma = \operatorname{sign}(z(x))$. One says that the *Mangasarian Fromovitz kink qualification (MFKQ)* holds at x if for all *definite* $\tilde{\sigma} \succeq \sigma$ the linear inequality system $\left[e_i^T \tilde{\Sigma} \partial_x z(x)\right]_{i \in \alpha} w > 0$ admits a solution $w \in \mathbb{R}^n$, unless $\left[e_i^T \tilde{\Sigma} \partial_x z(x)\right]_{i \in \alpha} w \geq 0$ admits only the solution $w = 0$.

Indefinite signatures $\tilde{\sigma} \succeq \sigma$ must be excluded in the definition of MFKQ: if $\tilde{\sigma}_k = \sigma_k = 0$, then $k \in \alpha(x)$ and w_k can be chosen arbitrarily since row k of the matrix is zero. Then, MFKQ could never be satisfied in the localized case. This is not stated formally in [10] but follows from the context.

That MFKQ is indeed weaker than LIKQ is shown in the next lemma. The proof of the implication was omitted in [10].

Lemma 2.44. *Given (unNLP), consider $x \in D^x$. Then, MFKQ holds at x if LIKQ holds at x. The converse is not true.*

Proof. This follows as in the proof of lemma 2.21. LIKQ at x implies that the system

$$[e_i^T \partial_x z(x)]_{i\in\alpha} w = [\tilde{\sigma}_i]_{i\in\alpha}$$

has a solution $w \in \mathbb{R}^n$ for every definite $\tilde{\sigma} \succeq \sigma(x)$. Thus, $w \in \mathbb{R}^n$ also solves the system

$$[e_i^T \tilde{\Sigma} \partial_x z(x)]_{i\in\alpha} w = [1 \dots 1]^T > 0$$

and MFKQ holds at x. A counterexample is given in [10, Lemma 2.11]. □

2.3 MPEC

This section presents mathematical programs with equilibrium constraints (MPECs) which are introduced in subsection 2.3.1. Then, different constraint qualifications are formulated in subsection 2.3.2 and corresponding first order conditions are given in subsection 2.3.3. Second order conditions are just stated under the strongest constraint qualification (MPEC-LICQ) in subsection 2.3.4.

This section is based on [23], [24] and [5].

2.3.1 Formulation

First, a mathematical program with equilibrium constraints is defined.

Definition 2.45 (MPEC)**.** Consider open sets $D^x \subseteq \mathbb{R}^n$, $D^u \subseteq \mathbb{R}^s$ with $0 \in D^u$ and $D^v \subseteq \mathbb{R}^s$ with $0 \in D^v$. An optimization problem of the form

$$\begin{aligned} \min_{(x,u,v)\in D^{x,u,v}} \quad & f(x,u,v) \\ \text{s.t.} \quad & g(x,u,v) = 0, \\ & h(x,u,v) \geq 0, \\ & 0 \leq u \perp v \geq 0, \end{aligned} \tag{MPEC}$$

with $f \in C^d(D^{x,u,v}, \mathbb{R})$, $g \in C^d(D^{x,u,v}, \mathbb{R}^{m_1})$ and $h \in C^d(D^{x,u,v}, \mathbb{R}^{m_2})$ for $d \geq 1$ is called a *Mathematical Program with Equilibrium Constraints* (MPEC). Here, the short-hand notation $0 \leq u \perp v \geq 0$, which is called complementarity condition, is an abbreviation for

$$u^T v = 0, \;\; u \geq 0, \;\; v \geq 0.$$

The feasible set of (MPEC) is denoted by

$$\mathcal{F}_{\text{mpec}} := \{(x,u,v) \in D^{x,u,v} \colon g(x,u,v) = 0, \;\; h(x,u,v) \geq 0, \;\; 0 \leq u \perp v \geq 0\}.$$

Note that in standard MPEC literature the sets $D^u = D^v = \mathbb{R}^s$ are considered. Here, the setting is modified to simplify notation in the following chapters as abs-normal forms are defined on open sets.

In the following, the short-hand notation $\bar{s} := \{1, \dots, s\}$ is used.

Definition 2.46 (Index Sets)**.** Consider a feasible point (x, u, v) of (MPEC). The set of indices of active inequalities $u_i \geq 0$ is denoted by $\mathcal{U}_0 := \{i \in \bar{s} \colon u_i = 0\}$ and the set of indices of inactive inequalities $u_i \geq 0$ by $\mathcal{U}_+ := \{i \in \bar{s} \colon u_i > 0\}$. Analogous definitions hold of $\mathcal{V}_0$ and $\mathcal{V}_+$. The set of indices of non-strict (degenerate) complementarity pairs is denoted by $\mathcal{D} := \mathcal{U}_0 \cap \mathcal{V}_0$.

Note that this deviates from contemporary MPEC literature, which frequently makes reference to the sets

$$\mathcal{I}_{+0} = \mathcal{U}_+ \cap \mathcal{V}_0, \quad \mathcal{I}_{0+} = \mathcal{U}_0 \cap \mathcal{V}_+, \quad \mathcal{I}_{00} = \mathcal{U}_0 \cap \mathcal{V}_0.$$

Note further that by complementarity the relations $\mathcal{U}_+ \cap \mathcal{V}_0 = \mathcal{U}_+$, $\mathcal{U}_0 \cap \mathcal{V}_+ = \mathcal{V}_+$ and $\mathcal{U}_+ \cap \mathcal{V}_+ = \emptyset$ hold and hence the partitioning $\bar{s} = \mathcal{D} \cup \mathcal{U}_+ \cup \mathcal{V}_+$.

2.3.2 Constraint Qualifications

Every MPEC can equivalently be rewritten as a smooth NLP. But, due to the structure of the complementarity condition, the standard constraint qualifications have some problems. This is illustrated in the following example.

Example 2.47. Consider the simplest possible MPEC

$$\begin{aligned} \min_{u\in\mathbb{R}, v\in\mathbb{R}} \quad & f(u,v) \\ \text{s.t.} \quad & 0 \leq u \perp v \geq 0 \end{aligned}$$

and the corresponding smooth NLP

$$\begin{aligned} \min_{u\in\mathbb{R}, v\in\mathbb{R}} \quad & f(u,v) \\ \text{s.t.} \quad & uv = 0 \\ & u \geq 0, \\ & v \geq 0. \end{aligned}$$

Then, LICQ and MFCQ are violated at every feasible point and ACQ is violated at $(0, 0)$. Only GCQ is satisfied at every feasible point. A proof may be found in [24].

Thus, weaker constraint qualifications especially for MPECs are needed. Versions of LICQ and MFCQ will be defined via a related NLP which is defined next.

Definition 2.48 (Tightened NLP)**.** Given a feasible point $(\hat{x}, \hat{u}, \hat{v})$ of (MPEC) with associated index sets $\mathcal{U}_+$, $\mathcal{V}_+$ and $\mathcal{D}$. The *tightened NLP* is defined as

$$\begin{aligned}
\min_{(x,u,v)\in D^{x,u,v}} \quad & f(x,u,v) \\
\text{s.t.} \quad & g(x,u,v) = 0, \\
& h(x,u,v) \geq 0, \\
& 0 \leq u_i,\ 0 = v_i \quad \text{if } \hat{u}_i > 0,\ \hat{v}_i = 0\ (i \in \mathcal{U}_+), \\
& 0 = u_i,\ 0 \leq v_i \quad \text{if } \hat{u}_i = 0,\ \hat{v}_i > 0\ (i \in \mathcal{V}_+), \\
& 0 = u_i,\ 0 = v_i \quad \text{if } \hat{u}_i = 0,\ \hat{v}_i = 0\ (i \in \mathcal{D}).
\end{aligned} \tag{TNLP}$$

Definition 2.49 (MPEC-LICQ and MPEC-MFCQ)**.** Consider a feasible point (x, u, v) of (MPEC). One says that (x, u, v) satisfies *MPEC-LICQ* resp. *MPEC-MFCQ* if it satisfies LICQ resp. MFCQ of the tightened NLP (TNLP).

In general MPEC-LICQ implies, but is stronger than, MPEC-MFCQ. The purely equality constrained case (without $h(x, u, v) \geq 0$) is an exception: then, MPEC-MFCQ is equivalent to MPEC-LICQ because the tightened NLP has no active inequalities.

To introduce weaker CQs in the spirit of Abadie and Guignard the tangential cone and the MPEC-linearized cone are defined.

Definition 2.50 (Tangential Cone and MPEC-Linearized Cone for (MPEC))**.** Consider a feasible point (x, u, v) of (MPEC) with associated index sets $\mathcal{U}_+$, $\mathcal{V}_+$ and $\mathcal{D}$. The tangential cone to $\mathcal{F}_{\text{mpec}}$ at (x, u, v) is

$$\mathcal{T}_{\text{mpec}}(x,u,v) = \left\{ (\delta x, \delta u, \delta v) \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\text{mpec}} \ni (x_k, u_k, v_k) \to (x,u,v): \\ \tau_k^{-1}(x_k - x, u_k - u, v_k - v) \to (\delta x, \delta u, \delta v) \end{array} \right\}.$$

The *complementarity cone* at (u, v) is

$$\mathcal{T}_\perp(u,v) = \left\{ \begin{pmatrix} \delta u \\ \delta v \end{pmatrix} \;\middle|\; \begin{array}{r} \delta v_i = 0,\ i \in \mathcal{U}_+, \\ \delta u_i = 0,\ i \in \mathcal{V}_+, \\ 0 \leq \delta u_i \perp \delta v_i \geq 0,\ i \in \mathcal{D} \end{array} \right\}.$$

Set $\mathcal{A} = \mathcal{A}(x, u, v)$ and $h_\mathcal{A} = [h_i]_{i\in\mathcal{A}}$. The *MPEC-linearized cone* at (x, u, v) is

$$\mathcal{T}^{\text{lin}}_{\text{mpec}}(x,u,v) := \left\{ \begin{pmatrix} \delta x \\ \delta u \\ \delta v \end{pmatrix} \;\middle|\; \begin{array}{r} \partial_1 g \delta x + \partial_2 g \delta u + \partial_3 g \delta v = 0, \\ \partial_1 h_\mathcal{A} \delta x + \partial_2 h_\mathcal{A} \delta u + \partial_3 h_\mathcal{A} \delta v \geq 0, \\ (\delta u, \delta v) \in \mathcal{T}_\perp(u,v) \end{array} \right\}.$$

Here all partial derivatives are evaluated at (x, u, v).

The linearized cone for the smooth reformulation reads

$$\mathcal{T}^{\text{lin}}(x,u,v) := \left\{ \begin{pmatrix} \delta x \\ \delta u \\ \delta v \end{pmatrix} \;\middle|\; \begin{array}{r} \partial_1 g \delta x + \partial_2 g \delta u + \partial_3 g \delta v = 0, \\ \partial_1 h_\mathcal{A} \delta x + \partial_2 h_\mathcal{A} \delta u + \partial_3 h_\mathcal{A} \delta v \geq 0, \\ \delta v_i = 0,\ i \in \mathcal{U}_+, \\ \delta u_i = 0,\ i \in \mathcal{V}_+, \\ \delta u_i \geq 0,\ \delta v_i \geq 0,\ i \in \mathcal{D} \end{array} \right\}.$$

Thus, the additional condition $\delta u_i \perp \delta v_i$ for $i \in \mathcal{D}$ has to hold in the MPEC-linearized cone. Hence, it is always contained in the linearized cone.

Lemma 2.51. *The complementarity cone $\mathcal{T}_\perp(\hat{u}, \hat{v})$ is both, the tangential cone and the MPEC-linearized cone to the complementarity set $\{(u, v) : 0 \le u \perp v \ge 0\}$ at $(\hat{u}, \hat{v})$.*

Proof. Given a tangent vector $(\delta u, \delta v) = \lim \tau_k^{-1}(u_k - \hat{u}, v_k - \hat{v})$ where $0 \le u_k \perp v_k \ge 0$ and $\tau_k \searrow 0$. Then, the following hold for k large enough:

$$\begin{aligned} u_{ki} > 0,\ v_{ki} = 0, &\qquad i \in \mathcal{U}_+\ (\hat{u}_i > 0,\ \hat{v}_i = 0), \\ u_{ki} = 0,\ v_{ki} > 0, &\qquad i \in \mathcal{V}_+\ (\hat{u}_i = 0,\ \hat{v}_i > 0), \\ 0 \le u_{ki} \perp v_{ki} \ge 0, &\qquad i \in \mathcal{D}\ (\hat{u}_i = 0,\ \hat{v}_i = 0). \end{aligned}$$

This implies $(\delta u, \delta v) \in \mathcal{T}_\perp(\hat{u}, \hat{v})$. Conversely, every $(\delta u, \delta v) \in \mathcal{T}_\perp(\hat{u}, \hat{v})$ is a tangent vector generated by the sequence $(u_k, v_k) = (\hat{u}, \hat{v}) + \tau_k(\delta u, \delta v)$ with $\tau_k = 1/k$, $k \in \mathbb{N}_{>0}$. ☐

To show that the tangential cone is a subset of the MPEC-linearized cone and not only of the linearized cone, the following subproblems of (MPEC) are needed.

Definition 2.52 (Branch NLPs)**.** Consider a feasible point $(\hat{x}, \hat{u}, \hat{v})$ of (MPEC) with associated index sets $\mathcal{U}_+$, $\mathcal{V}_+$ and $\mathcal{D}$. For $\mathcal{P} \subseteq \mathcal{D}$ and $\bar{\mathcal{P}} = \mathcal{D} \setminus \mathcal{P}$, the branch problem NLP($\mathcal{P}$) is defined as

$$\begin{aligned} \min_{(x,u,v) \in D^{x,u,v}} \quad & f(x, u, v) \\ \text{s.t.} \quad & g(x, u, v) = 0, \\ & h(x, u, v) \ge 0, \\ & 0 \le u_i,\ 0 = v_i,\ i \in \mathcal{U}_+ \cup \bar{\mathcal{P}}, \\ & 0 = u_i,\ 0 \le v_i,\ i \in \mathcal{V}_+ \cup \mathcal{P}. \end{aligned} \tag{NLP($\mathcal{P}$)}$$

The feasible set of (NLP($\mathcal{P}$)), which always contains $(\hat{x}, \hat{u}, \hat{v})$, is denoted by

$$\mathcal{F}_\mathcal{P} := \left\{ (x, u, v) \in D^{x,u,v} \;\middle|\; \begin{array}{l} g(x, u, v) = 0,\ h(x, u, v) \ge 0, \\ 0 \le u_i \text{ and } 0 = v_i \text{ for } i \in \mathcal{U}_+ \cup \bar{\mathcal{P}}, \\ 0 = u_i \text{ and } 0 \le v_i \text{ for } i \in \mathcal{V}_+ \cup \mathcal{P} \end{array} \right\}.$$

Lemma 2.53 (Tangential Cone and Linearized Cone for (NLP($\mathcal{P}$)))**.** *Consider (NLP($\mathcal{P}$)) at a feasible point (x, u, v) of (MPEC). The tangential cone to $\mathcal{F}_\mathcal{P}$ at (x, u, v) reads*

$$\mathcal{T}_\mathcal{P}(x, u, v) := \left\{ (\delta x, \delta u, \delta v) \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_\mathcal{P} \ni (x_k, u_k, v_k) \to (x, u, v): \\ \tau_k^{-1}(x_k - x, u_k - u, v_k - v) \to (\delta x, \delta u, \delta v) \end{array} \right\}.$$

Set $\mathcal{A} = \mathcal{A}(x)$ and $h_\mathcal{A} = [h_i]_{i \in \mathcal{A}}$. The linearized cone reads

$$\mathcal{T}_\mathcal{P}^{lin}(x, u, v) := \left\{ \begin{pmatrix} \delta x \\ \delta u \\ \delta v \end{pmatrix} \;\middle|\; \begin{array}{l} \partial_1 g \delta x + \partial_2 g \delta u + \partial_3 g \delta v) = 0, \\ \partial_1 h_\mathcal{A} \delta x + \partial_2 h_\mathcal{A} \delta u + \partial_3 h_\mathcal{A} \delta v \ge 0, \\ 0 = \delta u_i \text{ for } i \in \mathcal{V}_+ \cup \mathcal{P},\ 0 = \delta v_i \text{ for } i \in \mathcal{U}_+ \cup \bar{\mathcal{P}}, \\ 0 \le \delta u_i \text{ for } i \in \bar{\mathcal{P}},\ 0 \le \delta v_i \text{ for } i \in \mathcal{P} \end{array} \right\}.$$

Here, all partial derivatives are evaluated at (x, u, v).

Proof. This follows directly from Definition 2.14. ☐

Lemma 2.54. *Consider a feasible point* (x, u, v) *of* (MPEC) *with associated branch problems* (NLP($\mathcal{P}$))*. Then, the following decompositions of the tangential cone and of the MPEC-linearized cone of* (MPEC) *hold:*

$$\mathcal{T}_{mpec}(x, u, v) = \bigcup_{\mathcal{P}} \mathcal{T}_{\mathcal{P}}(x, u, v) \quad \text{and} \quad \mathcal{T}^{lin}_{mpec}(x, u, v) = \bigcup_{\mathcal{P}} \mathcal{T}^{lin}_{\mathcal{P}}(x, u, v).$$

Proof. A proof may be found in [5, Lemma 2.16 and Lemma 3.6]. ☐

Lemma 2.55. *Let* (x, u, v) *be feasible for* (MPEC)*. Then,*

$$\mathcal{T}_{mpec}(x, u, v) \subseteq \mathcal{T}^{lin}_{mpec}(x, u, v) \quad \text{and} \quad \mathcal{T}_{mpec}(x, u, v)^* \supseteq \mathcal{T}^{lin}_{mpec}(x, u, v)^*.$$

Proof. A proof may be found in [5, Lemma 3.7]. ☐

In general, the converses do not hold. This motivates the definition of MPEC-ACQ and MPEC-GCQ.

Definition 2.56 (MPEC-ACQ)**.** Consider a feasible point (x, u, v) of (MPEC). We say that *Abadie's Constraint Qualification for MPEC (MPEC-ACQ)* holds at (x, u, v) if

$$\mathcal{T}_{\mathrm{mpec}}(x, u, v) = \mathcal{T}^{\mathrm{lin}}_{\mathrm{mpec}}(x, u, v).$$

Definition 2.57 (MPEC-GCQ)**.** Consider a feasible point (x, u, v) of (MPEC). We say that *Guignard's Constraint Qualification for MPEC (MPEC-GCQ)* holds at (x, u, v) if

$$\mathcal{T}_{\mathrm{mpec}}(x, u, v)^* = \mathcal{T}^{\mathrm{lin}}_{\mathrm{mpec}}(x, u, v)^*.$$

Both MPEC-CQs are implied if the corresponding CQ holds for all branch problems.

Theorem 2.58 (ACQ for all (NLP($\mathcal{P}$)) implies MPEC-ACQ)**.** *Consider a feasible point* (x, u, v) *of* (MPEC)*. If ACQ holds at* (x, u, v) *for all* (NLP($\mathcal{P}$))*, then MPEC-ACQ holds at* (x, u, v)*.*

Proof. This follows directly from Lemma 2.54. ☐

Theorem 2.59 (GCQ for all (NLP($\mathcal{P}$)) implies MPEC-GCQ)**.** *Consider a feasible point* (x, u, v) *of* (MPEC)*. If GCQ holds at* (x, u, v) *for all* (NLP($\mathcal{P}$))*, then MPEC-GCQ holds at* (x, u, v)*.*

Proof. This follows directly from Lemma 2.54 as it yields

$$\mathcal{T}_{\mathrm{mpec}}(x, u, v)^* = \bigcap_{\mathcal{P}} \mathcal{T}_{\mathcal{P}}(x, u, v)^* \quad \text{and} \quad \mathcal{T}^{\mathrm{lin}}_{\mathrm{mpec}}(x, u, v)^* = \bigcap_{\mathcal{P}} \mathcal{T}^{\mathrm{lin}}_{\mathcal{P}}(x, u, v)^*$$

by dualization of the decompositions. ☐

2.3.3 Stationarity and First Order Conditions

This subsection presents stationarity concepts and first order conditions corresponding to the MPEC-CQs which were introduced in the previous subsection.

Definition 2.60 (MPEC-Lagrangian function)**.** The MPEC-Lagrangian function is defined as

$$\mathcal{L}_\perp(x,u,v,\lambda,\mu) := f(x,u,v) + \lambda_\mathcal{E}^T g(x,u,v) - \lambda_\mathcal{I}^T h(x,u,v) - \mu_\mathrm{u}^T u - \mu_\mathrm{v}^T v$$

with Lagrange multiplier vectors $\lambda = (\lambda_\mathcal{E}, \lambda_\mathcal{I}) \in \mathbb{R}^{m_1+m_2}$ and $\mu = (\mu_\mathrm{u}, \mu_\mathrm{v}) \in \mathbb{R}^{2s}$.

Note, that this deviates from standard literature once again in sign of the Lagrange multiplier vector $\lambda_\mathcal{E}$.

First, strong stationarity is introduced which is the strongest stationarity concept for MPECs.

Definition 2.61 (Strong Stationarity)**.** Consider a feasible point (x^*, u^*, v^*) of (MPEC) with associated index sets $\mathcal{U}_+$, $\mathcal{V}_+$ and $\mathcal{D}$. It is *strongly stationary or S-stationary* if there exist Lagrange multiplier vectors $\lambda^* = (\lambda_\mathcal{E}^*, \lambda_\mathcal{I}^*)$ and $\mu^* = (\mu_\mathrm{u}^*, \mu_\mathrm{v}^*)$ such that the following conditions are satisfied:

$$\begin{aligned}
\partial_{x,u,v}\mathcal{L}_\perp(x^*,u^*,v^*,\lambda^*,\mu^*) &= 0 & & \text{(stationarity)},\\
(\mu_\mathrm{u}^*)_i &= 0, \quad i \in \mathcal{U}_+, & &\\
(\mu_\mathrm{v}^*)_i &= 0, \quad i \in \mathcal{V}_+, & &\\
(\mu_\mathrm{u}^*)_i \geq 0,\ (\mu_\mathrm{v}^*)_i &\geq 0, \quad i \in \mathcal{D} & & \text{(strong stationarity)},\\
\lambda_\mathcal{I}^* &\geq 0 & & \text{(nonnegativity)},\\
(\lambda_\mathcal{I}^*)^T h(x^*,u^*,v^*) &= 0 & & \text{(complementarity)}.
\end{aligned}$$

Originally the definition of strong stationarity is due to Scheel and Scholtes in [23]. However, the conditions of vanishing Lagrange multipiers $(\mu_\mathrm{u}^*)_i$ and $(\mu_\mathrm{v}^*)_i$ are missing there and thus it is referred to [5] instead.

In [5, Proposition 4.5] it is shown that these conditions are equivalent to the KKT conditions applied to the smooth NLP which corresponds to (MPEC). Moreover, a strong stationary point can equivalently be defined as a KKT point of the relaxed NLP, which is defined as

$$\begin{aligned}
\min_{(x,u,v)\in D^{x,u,v}} \quad & f(x,u,v)\\
\text{s.t.} \quad & g(x,u,v) = 0,\\
& h(x,u,v) \geq 0,\\
& 0 \leq u_i,\ 0 = v_i \quad \text{if } \hat{u}_i > 0,\ \hat{v}_i = 0\ (i \in \mathcal{U}_+),\\
& 0 = u_i,\ 0 \leq v_i \quad \text{if } \hat{u}_i = 0,\ \hat{v}_i > 0\ (i \in \mathcal{V}_+),\\
& 0 \leq u_i,\ 0 \leq v_i \quad \text{if } \hat{u}_i = 0,\ \hat{v}_i = 0\ (i \in \mathcal{D}).
\end{aligned} \tag{RNLP}$$

This is due to the fact that for indices $i \in \mathcal{U}_+$ (and analogously for indices $i \in \mathcal{V}_+$) the complementarity $(\mu_\mathrm{u}^*)_i u_i = 0$ together with the nonnegativity $(\mu_\mathrm{u}^*)_i \geq 0$ is exactly the condition $(\mu_\mathrm{u}^*)_i = 0$.

Theorem 2.62. *Assume that* (x^*, u^*, v^*) *is a local minimizer of* (MPEC) *and that MPEC-LICQ holds. Then,* (x^*, u^*, v^*) *is a strongly stationary point with unique Lagrange multiplier vectors* λ^* *and* μ^**.*

Proof. A proof may be found in [5, Theorem 4.9]. □

Weaker stationarity concepts are defined next. Here, Mordukhovich and MPEC-linearized Bouligand stationarity are considered.

Definition 2.63 (Mordukhovich Stationarity)**.** Consider a feasible point (x^*, u^*, v^*) of (MPEC) with associated index sets $\mathcal{U}_+$, $\mathcal{V}_+$ and $\mathcal{D}$. It is *Mordukhovich stationary* or *M-stationary* if there exist Lagrange multiplier vectors $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{I}})$ and $\mu^* = (\mu^*_{\mathrm{u}}, \mu^*_{\mathrm{v}})$ such that the following conditions are satisfied:

$$
\begin{aligned}
\partial_{x,u,v}\mathcal{L}_\perp(x^*, u^*, v^*, \lambda^*, \mu^*) &= 0 && \text{(stationarity)},\\
(\mu^*_{\mathrm{u}})_i &= 0,\ \ i \in \mathcal{U}_+, &&\\
(\mu^*_{\mathrm{v}})_i &= 0,\ \ i \in \mathcal{V}_+, &&\\
((\mu^*_{\mathrm{u}})_i > 0,\ (\mu^*_{\mathrm{v}})_i > 0)\ \vee\ (\mu^*_{\mathrm{u}})_i(\mu^*_{\mathrm{v}})_i &= 0,\ \ i \in \mathcal{D} && \text{(M-stationarity)},\\
\lambda^*_{\mathcal{I}} &\geq 0 && \text{(nonnegativity)},\\
(\lambda^*_{\mathcal{I}})^T h(x^*, u^*, v^*) &= 0 && \text{(complementarity)}.
\end{aligned}
$$

Theorem 2.64. *Assume that* (x^*, u^*, v^*) *is a local minimizer of* (MPEC) *and that MPEC-GCQ holds. Then,* (x^*, u^*, v^*) *is a Mordukhovich stationary point.*

Proof. A proof may be found in [5, Theorem 5.28]. □

Definition 2.65 (MPEC-linearized Bouligand stationary)**.** A feasible point (x^*, u^*, v^*) of (MPEC) is an *MPEC-linearized Bouligand stationary* or *MPEC-linearized B-stationary* point if

$$
\partial_{x,u,v} f(x^*, u^*, v^*)^T \delta \geq 0 \quad \text{for all } \delta \in \mathcal{T}^{\mathrm{lin}}_{\mathrm{mpec}}(x^*, u^*, v^*).
$$

Theorem 2.66. *Assume that* (x^*, u^*, v^*) *is a local minimizer of* (MPEC) *and that MPEC-GCQ holds. Then,* (x^*, u^*, v^*) *is an MPEC-linearized Bouligand stationary point.*

Proof. The result follows directly from Theorem 2.16. □

MPEC-linearized Bouligand stationarity can equivalently be defined using the branch problems.

Lemma 2.67 (MPEC-linearized Bouligand Stationarity)**.** *A feasible point* (x^*, u^*, v^*) *of* (MPEC) *with associated index sets* $\mathcal{U}_+$*,* $\mathcal{V}_+$ *and* $\mathcal{D}$ *is an* MPEC-linearized Bouligand stationary *point if and only if it is a stationary point of all branch problems* (NLP($\mathcal{P}$))*. In other words, if for all subsets* $\mathcal{P} \subseteq \mathcal{D}$*, there exist Lagrange multiplier vectors* $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{I}})$

and $\mu^* = (\mu_u^*, \mu_v^*)$ *such that the following conditions are satisfied:*

$$\begin{aligned} \partial_{x,u,v}\mathcal{L}_\perp(x^*, u^*, v^*, \lambda^*, \mu^*) &= 0, \\ (\mu_u^*)_i &= 0, \ i \in \mathcal{U}+, \\ (\mu_v^*)_i &= 0, \ i \in \mathcal{V}+, \\ (\mu_u^*)_i &\geq 0, \ i \in \mathcal{P}, \\ (\mu_v^*)_i &\geq 0, \ i \in \bar{\mathcal{P}}, \\ \lambda_{\mathcal{I}}^* &\geq 0, \\ (\lambda_{\mathcal{I}}^*)^T h(x^*, u^*, v^*) &= 0. \end{aligned}$$

Proof. By Lemma 2.54 the MPEC-linearized cone can be written as the decomposition of the linearized cones of the branch problems. Thus, the condition for MPEC-linearized Bouligand stationarity reads:

$$\partial_{x,u,v} f(x^*, u^*, v^*)^T \delta \geq 0 \quad \text{for all } \delta \in \bigcup_{\mathcal{P}} \mathcal{T}_{\mathcal{P}}^{\text{lin}}(x^*, u^*, v^*).$$

In other words for every $\mathcal{P} \subseteq \mathcal{D}$ the following condition holds:

$$\partial_{x,u,v} f(x^*, u^*, v^*)^T \delta \geq 0 \quad \text{for all } \delta \in \mathcal{T}_{\mathcal{P}}^{\text{lin}}(x^*, u^*, v^*).$$

This is equivalent to the assertion that x^* is a stationary point for every branch problem by Theorem 2.17.
The formulation of the KKT conditions makes use of the fact that for indices $i \in \mathcal{U}_+$ (and analogously for indices $i \in \mathcal{V}_+$) the complementarity $(\mu_{\mathrm{u}}^*)_i u_i = 0$ together with the nonnegativity $(\mu_{\mathrm{u}}^*)_i \geq 0$ is identical to the condition $(\mu_{\mathrm{u}}^*)_i = 0$. □

Based on this equivalent formulation of MPEC-linearized Bouligand stationarity it seems natural to require GCQ for all branch problems as a suitable CQ.

Corollary 2.68. *Assume that* (x^*, u^*, v^*) *is a local minimizer of* (MPEC) *and that GCQ holds for all* (NLP($\mathcal{P}$))*. Then,* (x^*, u^*, v^*) *is an MPEC-linearized Bouligand stationary point.*

Proof. The result follows directly from Theorem 2.66 as GCQ for all (NLP($\mathcal{P}$)) implies MPEC-GCQ by Theorem 2.59. □

2.3.4 Second Order Conditions

In this section second order necessary and sufficient conditions under MPEC-LICQ are stated.

It is based on the paper [23] by Scheel and Scholtes. Note that in their definition of strong stationarity the conditions of vanishing Lagrange multipiers $(\mu_{\mathrm{u}})_i$ and $(\mu_{\mathrm{v}})_i$ are missing. Nevertheless, the results cited here still hold as strong stationarity always occurs as an assumption.

Note further that a critical direction is defined here at a strong stationarity point instead of a weak stationary point as the latter is not used in this thesis.

Definition 2.69 (Critical direction)**.** Consider a strongly stationary point $y^* = (x^*, u^*, v^*)$ of (MPEC) with associated index sets $\mathcal{U}_+$, $\mathcal{V}_+$ and $\mathcal{D}$. A vector $d = (dx, du, dv) \in \mathbb{R}^{n+2s}$ is called a *critical direction* at y^* if

$$\begin{aligned} \partial_1 f(y^*)dx + \partial_2 f(y^*)du + \partial_3 f(y^*)dv &= 0, \\ \partial_1 g(y^*)dx + \partial_2 g(y^*)du + \partial_3 g(y^*)dv &= 0, \\ \partial_1 h_i(y^*)dx + \partial_2 h_i(y^*)du + \partial_3 h_i(y^*)dv &\geq 0, \quad i \in \mathcal{A}, \\ dv_i &= 0, \quad i \in \mathcal{U}_+, \\ du_i &= 0, \quad i \in \mathcal{V}_+, \\ \min(du_i, dv_i) &= 0, \quad i \in \mathcal{D}. \end{aligned}$$

Theorem 2.70 (Second Order Necessary Conditions)**.** *Consider* (MPEC) *for $d \geq 2$. Assume that $y^* = (x^*, u^*, v^*)$ is a local minimizer and that MPEC-LICQ holds. Denote by λ^* and μ^* the unique Lagrange multiplier vectors and assume further that MPEC-strict complementarity holds. Then, every critical direction d satisfies*

$$d^T H_{mpec}(y^*, \lambda^*)d \geq 0$$

where $H_{mpec}(y^*, \lambda^*) := \partial^2_{yy}\mathcal{L}_\perp(y^*, \lambda^*, \mu^*)$. (Note that $\partial^2_{yy}\mathcal{L}_\perp$ does not depend on μ^*.)

Proof. A proof may be found in [23, Theorem 7]. □

In [23] MPEC-SMFCQ instead of MPEC-LICQ was originally assumed. As MPEC-LICQ is stronger it can be used here instead.

Theorem 2.71 (Second Order Sufficient Conditions)**.** *Consider* (MPEC) *for $d \geq 2$. Assume that $y^* = (x^*, u^*, v^*)$ is strongly stationary for* (MPEC)*. Assume further that for every critical direction $d \neq 0$ there exist Lagrange multiplier vectors λ^* and μ^* such that*

$$d^T H_{mpec}(y^*, \lambda^*)d > 0.$$

Then, y^* is a strict local minimizer of (MPEC)*.*

Proof. A proof may be found in [23, Theorem 7]. □

Chapter 3

Optimality Conditions for Abs-Normal NLPs under LIKQ

This chapter extends the optimality conditions of Griewank and Walther in [8] to an optimization problem with level-1 nonsmooth constraints. To begin with, a smooth objective function and level-1 nonsmooth equality constraints are considered in section 3.1. Then, section 3.2 deals with additional level-1 nonsmooth inequality constraints as well as a level-1 nonsmooth objective function.

Parts of section 3.1 are published in [17] and parts of section 3.2 can be found in [15].

3.1 Equality-Constrained Abs-Normal NLP

In this section optimality conditions for the equality-constrained abs-normal NLP are obtained. This particular class is defined in section 3.1.1. Then, the LIKQ of Griewank and Walther is extended to this problem class in section 3.1.2. Optimality conditions are obtained next and following the approach in [8] it is distinguished into the localized case in section 3.1.3 and the generalized case in section 3.1.4.

3.1.1 Formulation

First, consider the level-1 nonsmooth equality-constrained NLP

$$\begin{aligned} \min_{x \in D^x} \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \end{aligned} \tag{3.1}$$

where $D^x \subseteq \mathbb{R}^n$ is open, $f \in C^d(D^x, \mathbb{R})$ and $g \in C^d_{\text{abs}}(D^x, \mathbb{R}^m)$ with $d \geq 1$.

To obtain the desired equality-constrained abs-normal NLP associated with NLP (3.1), the nonsmooth constraint $g(x) = 0$ is stated in abs-normal form:

$$g(x) = c_{\mathcal{E}}(x, |z|) = 0, \quad c_{\mathcal{Z}}(x, |z|) = z.$$

Note, that the term $c_{\mathcal{E}}(x, |z|)$ is used to fit common notation of NLPs.

Definition 3.1 (Equality-Constrained Abs-Normal NLP)**.** A nonsmooth optimization problem is called an *equality-constrained abs-normal NLP* if functions $f \in C^d(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in$

$C^d(D^{x,|z|}, \mathbb{R}^m)$ and $c_{\mathcal{Z}} \in C^d(D^{x,|z|}, \mathbb{R}^s)$ for $d \geq 1$ exist such that the NLP (3.1) can equivalently be stated as

$$\begin{aligned} \min_{(x,z)\in D^{x,|z|}} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x,|z|) = 0, \\ & c_{\mathcal{Z}}(x,|z|) - z = 0, \end{aligned} \tag{eqNLP}$$

where $D^{|z|}$ is symmetric and $\partial_2 c_{\mathcal{Z}}(x,|z|)$ is strictly lower triangular. The feasible set of (eqNLP) is denoted by

$$\mathcal{F}_{\text{eq}} := \{(x,z) \in D^{x,|z|} \colon c_{\mathcal{E}}(x,|z|) = 0, c_{\mathcal{Z}}(x,|z|) = z\}.$$

As in the unconstrained case, the feasible set of (eqNLP) in Definition 3.1 can be rewritten using $z(x)$:

$$\mathcal{F}_{\text{eq}} = \{(x, z(x)) \colon x \in D^x,\ c_{\mathcal{E}}(x, |z(x)|) = 0\}.$$

By construction, x^* is a local minimizer of the nonsmooth NLP (3.1) if and only if $(x^*, z^*) = (x^*, z(x^*))$ is a local minimizer of (eqNLP). Therefore, only (eqNLP) is considered hereafter.

The following small example will be used to illustrate the basic setting and the subsequent theory.

Example 3.2. Consider the level-1 nonsmooth NLP

$$\begin{aligned} \min_{x\in\mathbb{R}^3} \quad & x_1 + x_2^2 + x_3^2 \\ \text{s.t.} \quad & x_1 - |x_2(1+x_3)| = 0. \end{aligned}$$

This NLP has only one local minimizer: the strict global solution $x^* = (0,0,0)$. The constraint reformulation requires a single switching variable z to obtain the associated abs-normal NLP:

$$\begin{aligned} \min_{x\in\mathbb{R}^3, z\in\mathbb{R}} \quad & x_1 + x_2^2 + x_3^2 \\ \text{s.t.} \quad & x_1 - |z| = 0, \\ & x_2(1+x_3) - z = 0. \end{aligned}$$

Here the switching variable vanishes at the solution: $z^* := z(x^*) = 0$.

3.1.2 LIKQ

In this section the definitions of the active Jacobian and the LIKQ of Griewank and Walther in subsection 2.2 will be extended to (eqNLP). Therefore, the equality-constraints Jacobian $J_{\mathcal{E}}(x)$ is needed.

Definition 3.3 (Active Jacobian)**.** Given (eqNLP), consider $(x, z(x)) \in \mathcal{F}_{\text{eq}}$ and set $\alpha = \alpha(x)$, $\sigma = \sigma(x)$ and $\Sigma = \operatorname{diag}(\sigma)$. The *active Jacobian* is

$$J_{\text{eq}}(x) := \begin{bmatrix} J_{\mathcal{E}}(x) \\ J_{\alpha}(x) \end{bmatrix} \in \mathbb{R}^{(m+|\alpha|)\times n}.$$

It consists of the *equality-constraints Jacobian*

$$\begin{aligned} J_{\mathcal{E}}(x) := \partial_x c_{\mathcal{E}}(x, \Sigma z(x)) &= \partial_1 c_{\mathcal{E}}(x, \Sigma z(x)) + \partial_2 c_{\mathcal{E}}(x, \Sigma z(x))\Sigma \partial_x z(x) \\ &= \partial_1 c_{\mathcal{E}}(x, |z(x)|) + \partial_2 c_{\mathcal{E}}(x, |z(x)|)\Sigma \partial_x z(x) \end{aligned}$$

and of the active switching Jacobian (see Definition 2.33)

$$J_\alpha(x) = \left[e_i^T \partial_x z(x)\right]_{i\in\alpha} = \left[e_i^T [I - \partial_2 c_{\mathcal{Z}}(x, |z(x)|)\Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(x, |z(x)|)\right]_{i\in\alpha}.$$

As will be shown later, the following extension of Definiton 3.67 provides enough regularity to prove optimality conditions for (eqNLP). It reduces also to the classical LICQ in the smooth case.

Definition 3.4 (LIKQ)**.** Given (eqNLP), consider $(x, z(x)) \in \mathcal{F}_{\text{eq}}$. One says that the *linear independence kink qualification (LIKQ)* holds at x if the active Jacobian

$$J_{\text{eq}}(x) = \begin{bmatrix} J_{\mathcal{E}}(x) \\ J_\alpha(x) \end{bmatrix} \in \mathbb{R}^{(m+|\alpha|)\times n}$$

has full row rank $m + |\alpha|$.

If LIKQ holds at x, a matrix whose columns are a basis of $\ker(J_{\text{eq}}(x))$ will be denoted by $U_{\text{eq}}(x) \in \mathbb{R}^{n\times[n-(m+|\alpha|)]}$, i.e., $\ker(J_{\text{eq}}(x)) = \text{im}(U_{\text{eq}}(x))$ holds.

Next, example 3.2 is reviewed and checked whether LIKQ holds at the solution.

Example 3.5 (LIKQ for example 3.2)**.** Recall the abs-normal NLP

$$\begin{aligned} \min_{x\in\mathbb{R}^3, z\in\mathbb{R}} \quad & x_1 + x_2^2 + x_3^2 \\ \text{s.t.} \quad & x_1 - |z| = 0, \\ & x_2(1 + x_3) - z = 0, \end{aligned}$$

with solution $x^* = (0, 0, 0)$ and $z^* = 0$. Then, $\sigma^* = \sigma(x^*) = 0$ holds and the active Jacobian at x^* is

$$J_{\text{eq}}(x^*) = \begin{bmatrix} J_{\mathcal{E}}(x^*) \\ J_\alpha(x^*) \end{bmatrix} = \begin{bmatrix} \partial_1 c_{\mathcal{E}}(x^*, 0) \\ \partial_1 c_{\mathcal{Z}}(x^*, 0) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 + x_3^* & x_2^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

This matrix has full row rank. Hence, LIKQ is satisfied at $x^* = (0, 0, 0)$ with $z^* = 0$. A basis of the nullspace of $J_{\text{eq}}(x^*)$ is given by the matrix $U_{\text{eq}}(x^*) = [0,\ 0,\ 1]^T$.

3.1.3 Localized Case

As in the unconstrained case in subsection 2.2, it makes sense to distinguish the situation where all switching variables are active from the more intricate situation where some of them are nonzero. Given a point of interest x^*, in this section it is assumed that $z(x^*) = 0$ holds. Thus, the active switching set is $\alpha(x^*) = \{1, \ldots, s\}$ and of cardinality $|\alpha(x^*)| = s$.

Trunk Problem Substituting $z = 0$ into (eqNLP) yields the so-called localized trunk problem.

Definition 3.6 (Localized Trunk Problem)**.** The *localized trunk problem* reads

$$\begin{aligned} \min_{x \in D^x} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, 0) = 0, \\ & c_{\mathcal{Z}}(x, 0) = 0. \end{aligned} \tag{3.2}$$

Its feasible set is denoted by $\mathcal{F}_t := \{x \in D^x : c_{\mathcal{E}}(x, 0) = 0, c_{\mathcal{Z}}(x, 0) = 0\}$.

The following lemma is immediately clear by construction as $\mathcal{F}_t \times \{0\} \subseteq \mathcal{F}_{\text{eq}}$.

Lemma 3.7. *If a point $(x^*, 0)$ is a local minimizer of (eqNLP), then x^* is a local minimizer of the localized trunk problem (3.2).*

Moreover, the trunk problem is obviously smooth and standard theory can be applied to derive necessary optimality conditions. The LICQ for the localized trunk problem requires full row rank of the matrix

$$J_t(x^*) = \begin{bmatrix} \partial_1 c_{\mathcal{E}}(x^*, 0) \\ \partial_1 c_{\mathcal{Z}}(x^*, 0) \end{bmatrix} \in \mathbb{R}^{(m+s)\times n}.$$

The following lemma shows that LIKQ reduces to this condition.

Lemma 3.8. *Assume that $(x^*, 0)$ is feasible for (eqNLP). Then, LIKQ at x^* is LICQ at x^* for the localized trunk problem.*

Proof. With $\sigma = \operatorname{sign}(z(x^*)) = 0$ and $\Sigma = \operatorname{diag}(\sigma) = 0$, Definition 3.3 yields

$$\begin{aligned} J_{\mathcal{E}}(x^*) &= \partial_1 c_{\mathcal{E}}(x^*, \Sigma z(x^*)) + \partial_2 c_{\mathcal{E}}(x^*, \Sigma z(x^*))\Sigma \partial_x z(x^*) = \partial_1 c_{\mathcal{E}}(x^*, 0), \\ J_{\alpha}(x^*) &= \left[e_i^T \partial_x z(x^*) \right]_{i \in \alpha} = \partial_x z(x^*) = [I - \partial_2 c_{\mathcal{Z}}(x^*, 0)\Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(x^*, 0) = \partial_1 c_{\mathcal{Z}}(x^*, 0). \end{aligned}$$

This proves the claim. □

If LIKQ holds at x^*, a basis of the nullspace of $J_t(x^*)$ is given by the columns of the matrix $U_t(x^*) = U_{\text{eq}}(x^*) \in \mathbb{R}^{n \times [n-(m+s)]}$. First and second order necessary conditions of the localized trunk problem are now readily stated in terms of the Lagrangian for the localized trunk problem which is denoted by

$$\mathcal{L}_t(x, \lambda) := f(x) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x, 0) + \lambda_{\mathcal{Z}}^T c_{\mathcal{Z}}(x, 0)$$

with Lagrange multiplier vector $\lambda = (\lambda_{\mathcal{E}}, \lambda_{\mathcal{Z}}) \in \mathbb{R}^{m+s}$.

Theorem 3.9 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0)$ is a local minimizer of (eqNLP) and that LIKQ holds at x^*. Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_{\mathcal{Z}}^*)$ such that the following conditions are satisfied:*

$$\begin{aligned} f'(x^*) + (\lambda_{\mathcal{E}}^*)^T \partial_1 c_{\mathcal{E}}(x^*, 0) + (\lambda_{\mathcal{Z}}^*)^T \partial_1 c_{\mathcal{Z}}(x^*, 0) &= 0 \quad \textit{(tangential stationarity)}, \\ c_{\mathcal{E}}(x^*, 0) &= 0 \quad \textit{(primal feasibility)}, \\ c_{\mathcal{Z}}(x^*, 0) &= 0 \quad \textit{(switching feasibility)}. \end{aligned} \tag{3.3}$$

Proof. Because of Lemma 3.7, the smooth localized trunk problem (3.2) is considered. As LICQ holds at x^* by Lemma 3.8 the first order necessary conditions for smooth NLPs in Theorem 2.7 can be applied. These are directly the conditions (3.3). Moreover, λ^* is unique by Lemma 2.8 since LICQ holds. $\square$

Remark 3.10. In contrast to standard theory for smooth NLPs, these conditions are *not* sufficient in the linear case since the converse of Lemma 3.7 does not hold: $(x^*, 0)$ is not necessarily a local minimizer of the abs-normal NLP if x^* is a local minimizer of the localized trunk problem.

Theorem 3.11 (Second Order Necessary Condition)**.** *Consider* (eqNLP) *for* $d \geq 2$. *Assume that* $(x^*, 0)$ *is a local minimizer and that LIKQ holds at* x^*. *Denote by* λ^* *the unique Lagrange multiplier vector. Then,*

$$U_t(x^*)^T H_t(x^*, \lambda^*) U_t(x^*) \geq 0$$

where $H_t(x^*, \lambda^*) := \partial^2_{xx} \mathcal{L}_t(x^*, \lambda^*) \in \mathbb{R}^{n \times n}$.

Proof. As before, the smooth localized trunk problem is used and LICQ holds at x^* by Lemma 3.8. Here, strict complementarity does not have to be checked as $\mathcal{I} = \emptyset$. Thus, the second order necessary conditions for smooth NLPs in Theorem 2.10 can be applied which leads to the result. $\square$

Again, Example 3.2 is used to illustrate the trunk problem and its necessary conditions.

Example 3.12 (Trunk problem of Example 3.2)**.** With $z = 0$ in Example 3.5 one obtains the localized trunk problem

$$\begin{aligned} \min_{x \in \mathbb{R}^3} \quad & x_1 + x_2^2 + x_3^2 \\ \text{s.t.} \quad & x_1 = 0, \\ & x_2(1 + x_3) = 0. \end{aligned}$$

LICQ holds by Lemma 3.8 since LIKQ is satisfied at $x^* = (0, 0, 0)$ as was shown in Example 3.5. The Lagrangian reads

$$\mathcal{L}_t(x, \lambda) = x_1 + x_2^2 + x_3^2 + \lambda_{\mathcal{E}} x_1 + \lambda_{\mathcal{Z}} x_2(1 + x_3)$$

where $\lambda = (\lambda_{\mathcal{E}}, \lambda_{\mathcal{Z}})$. Then, the first order necessary conditions at $(x, 0)$ are

$$\begin{aligned} \partial_{x_1} \mathcal{L}_t(x, \lambda) = 1 + \lambda_{\mathcal{E}} &= 0, \\ \partial_{x_2} \mathcal{L}_t(x, \lambda) = 2x_2 + \lambda_{\mathcal{Z}}(1 + x_3) &= 0, \\ \partial_{x_3} \mathcal{L}_t(x, \lambda) = 2x_3 + \lambda_{\mathcal{Z}} x_2 &= 0 \quad \text{(tangential stationarity)}, \\ x_1 &= 0 \quad \text{(primal feasibility)}, \\ x_2(1 + x_3) &= 0 \quad \text{(switching feasibility)}. \end{aligned}$$

They are satisfied at $x^* = (0, 0, 0)$ with $\lambda^*_{\mathcal{E}} = -1$ and $\lambda^*_{\mathcal{Z}} = 0$.

The second order necessary conditions involve the Hessian

$$H_t(x^*, \lambda^*) = \partial^2_{xx}\mathcal{L}_t(x^*, \lambda^*) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & \lambda^*_{\mathcal{Z}} \\ 0 & \lambda^*_{\mathcal{Z}} & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

and the matrix $U_{\text{eq}}(x^*) = [0,\ 0,\ 1]^T$ from Example 3.5:

$$U_{\text{eq}}(x^*)^T H_t(x^*, \lambda^*) U_{\text{eq}}(x^*) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 2 \geq 0.$$

Thus, second order necessary conditions of Theorem 3.11 hold at $x^* = (0, 0, 0)$ with $z^* = 0$.

Branch Problems One cannot derive sufficient conditions for the abs-normal NLP using the localized trunk problem alone since the converse of Lemma 3.7 does not hold. Rather, the local behavior of the abs-normal NLP in a neighborhood of x^* with $z(x^*) = 0$ has to be considered for all possible combinations of signs of the switching variables. This leads to the definition of 2^s branch problems.

Definition 3.13 (Localized Branch Problems)**.** Choose $\sigma \in \{-1, 1\}^s$ and set $\Sigma = \operatorname{diag}(\sigma)$. The *localized branch problem* associated with σ reads

$$\begin{aligned} \min_{(x,z) \in D^{x,|z|}} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, \Sigma z) = 0, \\ & c_{\mathcal{Z}}(x, \Sigma z) - z = 0, \\ & \Sigma z \geq 0. \end{aligned}$$

The feasible set is denoted by

$$\mathcal{F}_\Sigma := \{(x, z) \in D^{x,|z|} \colon c_{\mathcal{E}}(x, \Sigma z) = 0, c_{\mathcal{Z}}(x, \Sigma z) = z, \Sigma z \geq 0\}.$$

Using the notation $\bar{z} = \Sigma z$, the branch problem takes the equivalent form

$$\begin{aligned} \min_{x,\bar{z}} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, \bar{z}) = 0, \\ & c_{\mathcal{Z}}(x, \bar{z}) - \Sigma \bar{z} = 0, \\ & \bar{z} \geq 0. \end{aligned} \tag{3.4}$$

By construction, the following inclusions of feasible sets hold:

$$\mathcal{F}_t \times \{0\} \subseteq \mathcal{F}_\Sigma \subseteq \mathcal{F}_{\text{eq}} \quad \text{for all } \sigma \in \{-1, 1\}^s. \tag{3.5}$$

Moreover, the branch problems coincide in the localized trunk problem and provide a decomposition of the feasible set of (eqNLP).

Lemma 3.14. *The feasible sets satisfy the following relations:*

$$\bigcap_{\sigma\in\{-1,1\}^s} \mathcal{F}_\Sigma = \mathcal{F}_t \times \{0\} \quad \textit{and} \quad \bigcup_{\sigma\in\{-1,1\}^s} \mathcal{F}_\Sigma = \mathcal{F}_{eq}.$$

Proof. With respect to (3.5) one has to show that

$$\bigcap_{\sigma\in\{-1,1\}^s} \mathcal{F}_\Sigma \subseteq \mathcal{F}_t \times \{0\} \quad \text{and} \quad \bigcup_{\sigma\in\{-1,1\}^s} \mathcal{F}_\Sigma \supseteq \mathcal{F}_{\mathrm{eq}}.$$

To prove the first claim, consider $(x, z) \in \bigcap_{\sigma\in\{-1,1\}^s} \mathcal{F}_\Sigma$. Then, $\Sigma z \geq 0$ has to hold for every definite Σ which implies $z = 0$ and thus $(x, 0) \in \mathcal{F}_t \times \{0\}$.
For the second claim, consider $(x, z) \in \mathcal{F}_{\mathrm{eq}}$. Set $\tilde{\sigma}_i = 1$ for $i \in \alpha(x)$ and $\tilde{\sigma}_i = (\sigma(x))_i$ otherwise as well as $\tilde{\Sigma} = \mathrm{diag}(\tilde{\sigma})$. Then, $(x, z) \in \mathcal{F}_{\tilde{\Sigma}} \subseteq \bigcup_{\sigma\in\{-1,1\}^s} \mathcal{F}_\Sigma$. ☐

This decomposition implies the following equivalence, which provides an approach to formulate sufficient optimality conditions.

Lemma 3.15. *A pair $(x^*, 0)$ is a local minimizer of (eqNLP) if and only if it is a local minimizer of the localized branch problem (3.4) for every definite $\Sigma = \mathrm{diag}(\sigma)$.*

Proof. Assume that $(x^*, 0)$ is a local minimizer of (eqNLP). Then, it is feasible for every branch problem as Lemma 3.7 and (3.5) imply:

$$(x^*, 0) \in \mathcal{F}_t \times \{0\} \subseteq \mathcal{F}_\Sigma.$$

Thus, the pair $(x^*, 0)$ has to be a local minimizer of every branch problem (3.4) as $\mathcal{F}_\Sigma \subseteq \mathcal{F}_{\mathrm{eq}}$ holds for all definite σ by (3.5). Otherwise, it could not be a local minimizer of (eqNLP) and this would contradict the assumption.
Conversely, assume that $(x^*, 0)$ is a local minimizer of every branch problem (3.4). Then, it is feasible for (eqNLP) by (3.5):

$$(x^*, 0) \in \mathcal{F}_\Sigma \subseteq \mathcal{F}_{\mathrm{eq}} \quad \text{for all } \sigma \in \{-1, 1\}^s.$$

Finally, $(x^*, 0)$ has to be a local minimizer of (eqNLP) as $\bigcup_{\sigma\in\{-1,1\}^s} \mathcal{F}_\Sigma = \mathcal{F}_{\mathrm{eq}}$ holds by Lemma 3.14. Otherwise, there exists at least one branch problem where $(x^*, 0)$ is not a local minimizer and this would be a contradiction to the assumption. ☐

As every branch problem (3.4) is smooth, standard NLP theory can be applied to formulate optimality conditions. Here, LICQ at $(x^*, 0)$ means full row rank of the matrix

$$J_\Sigma(x^*) = \begin{bmatrix} \partial_1 c_\mathcal{E}(x^*, 0) & \partial_2 c_\mathcal{E}(x^*, 0) \\ \partial_1 c_\mathcal{Z}(x^*, 0) & \partial_2 c_\mathcal{Z}(x^*, 0) - \Sigma \\ 0 & I \end{bmatrix} \in \mathbb{R}^{(m+2s)\times(n+s)}.$$

Block elimination yields the following implication.

Lemma 3.16. *Assume that $(x^*, 0)$ is feasible for (eqNLP). Then, LICQ holds at $(x^*, 0)$ for all branch problems if LIKQ holds at x^*.*

Proof. Block elimination immediately yields that LICQ at $(x^*, 0)$ holds for all branch problems if LICQ at x^* holds for the trunk problem. By Lemma 3.8, LICQ for the trunk problem is implied by LIKQ at x^*. □

If LIKQ holds at x^*, a basis of the nullspace of $J_\Sigma(x^*)$ is given by the columns of the matrix $U_\Sigma(x^*) = [U_{\text{eq}}(x^*)^T,\ 0]^T \in \mathbb{R}^{(n+s)\times[n-(m+s)]}$. The Lagrangians of the branch problems are denoted by

$$\mathcal{L}_\Sigma(x, \bar{z}, \lambda, \mu) := f(x) + \lambda_\mathcal{E}^T c_\mathcal{E}(x, \bar{z}) + \lambda_\mathcal{Z}^T[c_\mathcal{Z}(x, \bar{z}) - \Sigma\bar{z}] - \mu^T \bar{z}$$

with Lagrange multiplier vectors $\lambda = (\lambda_\mathcal{E}, \lambda_\mathcal{Z}) \in \mathbb{R}^{m+s}$ and $\mu \in \mathbb{R}^s$.

Theorem 3.17 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0)$ is a local minimizer of (eqNLP) and that LIKQ holds at x^*. Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda_\mathcal{E}^*, \lambda_\mathcal{Z}^*)$ such that the following conditions are satisfied for every definite $\Sigma = \operatorname{diag}(\sigma)$:*

$$\begin{aligned}
f'(x^*) + (\lambda_\mathcal{E}^*)^T\partial_1 c_\mathcal{E}(x^*, 0) + (\lambda_\mathcal{Z}^*)^T\partial_1 c_\mathcal{Z}(x^*, 0) &= 0 \quad \textit{(tangential stationarity)},\\
(\lambda_\mathcal{E}^*)^T\partial_2 c_\mathcal{E}(x^*, 0) + (\lambda_\mathcal{Z}^*)^T[\partial_2 c_\mathcal{Z}(x^*, 0) - \Sigma] &\geq 0 \quad \textit{(normal growth)},\\
c_\mathcal{E}(x^*, 0) &= 0 \quad \textit{(primal feasibility)},\\
c_\mathcal{Z}(x^*, 0) &= 0 \quad \textit{(switching feasibility)}.
\end{aligned}$$

Proof. Given $\sigma \in \{-1, 1\}^s$, applying the first order necessary conditions of smooth NLPs in Theorem 2.7 yields the existence of $\lambda^* = (\lambda_\mathcal{E}^*, \lambda_\mathcal{Z}^*) \in \mathbb{R}^{m+s}$ and $\mu^* \in \mathbb{R}^s$ such that

$$\begin{aligned}
\partial_x\mathcal{L}_\Sigma(x^*, 0, \lambda^*, \mu^*) = f'(x^*) + (\lambda_\mathcal{E}^*)^T\partial_1 c_\mathcal{E}(x^*, 0) + (\lambda_\mathcal{Z}^*)^T\partial_1 c_\mathcal{Z}(x^*, 0) &= 0,\\
\partial_{\bar{z}}\mathcal{L}_\Sigma(x^*, 0, \lambda^*, \mu^*) = (\lambda_\mathcal{E}^*)^T\partial_2 c_\mathcal{E}(x^*, 0) + (\lambda_\mathcal{Z}^*)^T[\partial_2 c_\mathcal{Z}(x^*, 0) - \Sigma] - (\mu^*)^T &= 0,\\
c_\mathcal{E}(x^*, 0) &= 0,\\
c_\mathcal{Z}(x^*, 0) &= 0,\\
\mu^* &\geq 0.
\end{aligned}$$

The first, third and fourth condition are precisely the necessary conditions of the trunk problem from Theorem 3.9. Thus, the Lagrange multiplier vector λ^* is unique and identical for all branch problems and for the trunk problem. Only the second and fifth condition (and μ^* itself) depend on Σ. Combining these two conditions eliminates $\mu^* \geq 0$ and yields the normal growth condition

$$(\lambda_\mathcal{E}^*)^T\partial_2 c_\mathcal{E}(x^*, 0) + (\lambda_\mathcal{Z}^*)^T[\partial_2 c_\mathcal{Z}(x^*, 0) - \Sigma] \geq 0,$$

which holds for every definite $\Sigma = \operatorname{diag}(\sigma)$. □

Verification of the necessary conditions just stated appears to require that 2^s cases must be checked, one for each branch problem, which differ only in the normal growth condition. Fortunately, since the Lagrange multiplier vectors λ^* coincide for all cases and μ^* does not appear explicitly, it turns out that only one of the strongest of the 2^s conditions needs to be checked: the one associated with one of the branch problems that satisfy $\Sigma\lambda^*_{\mathcal{Z}} = |\lambda^*_{\mathcal{Z}}|$. This is because of the obvious equivalence

$$\begin{aligned} & (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, 0) \geq |\lambda^*_{\mathcal{Z}}|^T \\ \iff & (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, 0) \geq (\lambda^*_{\mathcal{Z}})^T (\pm\Sigma) \text{ for every definite } \Sigma. \end{aligned}$$

Lemma 3.18. *The equivalence just stated immediately implies:*

$$\begin{aligned} & (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, 0) \geq |\lambda^*_{\mathcal{Z}}|^T \\ \iff & (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T [\partial_2 c_{\mathcal{Z}}(x^*, 0) - \Sigma] \geq 0 \text{ for every definite } \Sigma. \end{aligned}$$

As in the unconstrained case, checking of the first order necessary conditions has no combinatorial complexity contrary to what Theorem 3.17 might suggest. Instead a single set of conditions for *all* 2^s branch problems is obtained.

Corollary 3.19 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0)$ is a local minimizer of (eqNLP) and that LIKQ holds at x^*. Then, there exists a unique Lagrange multipier vector $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{Z}})$ such that the following conditions are satisfied:*

$$\begin{aligned} f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_1 c_{\mathcal{Z}}(x^*, 0) &= 0 && \textit{(tangential stationarity)}, \\ (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, 0) &\geq |\lambda^*_{\mathcal{Z}}|^T && \textit{(normal growth)}, \\ c_{\mathcal{E}}(x^*, 0) &= 0 && \textit{(primal feasibility)}, \\ c_{\mathcal{Z}}(x^*, 0) &= 0 && \textit{(switching feasibility)}. \end{aligned} \tag{3.6}$$

Observe that the normal growth condition provides a considerable strengthening of the otherwise identical first order conditions of the trunk problem derived in Theorem 3.9. Later it will be shown that, as in the smooth case, (3.6) provides sufficient first order conditions under special assumptions, possibly requiring normal growth in strict form. Strict normal growth in (3.6) is also needed for the following second order sufficient conditions.

Theorem 3.20 (Second Order Sufficient Conditions)**.** *Consider (eqNLP) for $d \geq 2$. Assume that $(x^*, 0)$ is feasible and that LIKQ holds at x^*. Assume further that a Lagrange multiplier vector λ^* exists such that the first order necessary conditions (3.6) are satisfied with strict normal growth,*

$$(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, 0) > |\lambda^*_{\mathcal{Z}}|^T,$$

and that

$$U_{eq}(x^*)^T H_t(x^*, \lambda^*) U_{eq}(x^*) > 0.$$

Then, $(x^*, 0)$ is a strict local minimizer of (eqNLP).

Proof. The strict normal growth condition implies $\mu^* > 0$ for all branch problems. Thus, strict complementarity holds in every branch problem and by LIKQ the Lagrange multiplier vectors are unique. To apply standard second order conditions for smooth NLPs it has to be shown that

$$U_\Sigma(x^*)^T H_\Sigma(x^*, \lambda^*) U_\Sigma(x^*) > 0$$

holds for every branch problem. The Hessians of all branch problems coincide since H_Σ is independent of μ^*: they read

$$H_\Sigma(x^*, \lambda^*) = \begin{bmatrix} \partial_{xx}\mathcal{L}_\Sigma(x^*, 0, \lambda^*, \mu^*) & \partial_{x\bar{z}}\mathcal{L}_\Sigma(x^*, 0, \lambda^*, \mu^*) \\ \partial_{\bar{z}x}\mathcal{L}_\Sigma(x^*, 0, \lambda^*, \mu^*) & \partial_{\bar{z}\bar{z}}\mathcal{L}_\Sigma(x^*, 0, \lambda^*, \mu^*) \end{bmatrix} \in \mathbb{R}^{(n+s)\times(n+s)}$$

with $\partial_{xx}\mathcal{L}_\Sigma(x^*, 0, \lambda^*, \mu^*) = \partial_{xx}\mathcal{L}_t(x^*, \lambda^*) = H_t(x^*, \lambda^*)$. Substituting $U_\Sigma(x^*)$ yields

$$U_\Sigma(x^*)^T H_\Sigma(x^*, \lambda^*) U_\Sigma(x^*) = U_{\mathrm{eq}}(x^*)^T H_t(x^*, \lambda^*) U_{\mathrm{eq}}(x^*) > 0,$$

and the assertion follows by Theorem 2.11. □

Again a closer look at Example 3.2 is taken.

Example 3.21. With $\bar{z} = \Sigma z$ for definite $\Sigma = \operatorname{diag}(\sigma) = \sigma$ the following branch problems are obtained:

$$\begin{aligned} \min_{x\in\mathbb{R}^2, \bar{z}\in\mathbb{R}} \quad & x_1 + x_2^2 + x_3^2 \\ \text{s.t.} \quad & x_1 - \bar{z} = 0, \\ & x_2(1 + x_3) - \Sigma\bar{z} = 0, \\ & \bar{z} \geq 0. \end{aligned}$$

Here, LICQ holds at $x^* = (0, 0, 0)$ by Lemma 3.16 since LIKQ holds at x^*, see Example 3.5. Applying Theorem 3.17 with the Lagrangian

$$\mathcal{L}_\Sigma(x, \bar{z}, \lambda, \mu) = x_1 + x_2^2 + x_3^2 + \lambda_{\mathcal{E}}(x_1 - \bar{z}) + \lambda_{\mathcal{Z}}(x_2(1 + x_3) - \Sigma\bar{z}) - \mu\bar{z}$$

leads to the following conditions:

$$\begin{aligned} 1 + \lambda_{\mathcal{E}} &= 0, & & \\ 2x_2 + \lambda_{\mathcal{Z}}(1 + x_3) &= 0, & & \\ 2x_3 + \lambda_{\mathcal{Z}} x_2 &= 0 & & \text{(tangential stationarity)}, \\ -\lambda_{\mathcal{E}} &\geq |\lambda_{\mathcal{Z}}| & & \text{(normal growth)}, \\ x_1 &= 0 & & \text{(primal feasibility)}, \\ x_2(1 + x_3) &= 0 & & \text{(switching feasibility)}. \end{aligned}$$

They hold at $x^* = (0, 0, 0)$ with $\lambda_{\mathcal{E}}^* = -1$ and $\lambda_{\mathcal{Z}}^* = 0$. Moreover, strict normal growth is satisfied,

$$-\lambda_{\mathcal{E}}^* = 1 > 0 = |\lambda_{\mathcal{Z}}^*|,$$

and the reduced Hessian is positive definite,

$$U_{\mathrm{eq}}(x^*)^T H_t(x^*, \lambda^*) U_{\mathrm{eq}}(x^*) = 2 > 0.$$

Thus, Theorem 3.20 can be applied, saying that $x^* = (0, 0, 0)$ with $z^* = 0$ is a strict local minimizer of the abs-normal NLP.

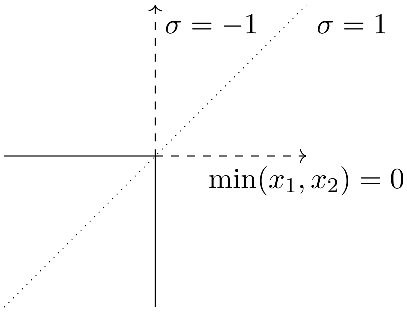

Figure 3.1: Illustration of Example 3.23. Dashed line: feasible set; dotted line: $z = 0$.

There are two special cases where first order necessary conditions are already sufficient and second order conditions are not needed. The first case pertains to linear functions where the result is due to the convexitiy of all branch problems.

Theorem 3.22 (First Order Necessary and Sufficient Conditions for Linear Functions)**.** *Given* (eqNLP)*, assume that f, $c_{\mathcal{E}}$ and $c_{\mathcal{Z}}$ are linear. Assume further that $(x^*, 0)$ is feasible and that LIKQ holds at x^*. Then, $(x^*, 0)$ is a local minimizer of* (eqNLP) *if and only if there exists a Lagrange multiplier vector $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{Z}})$ such that the conditions* (3.6) *are satisfied.*

Proof. By the assumptions every branch problem is linear and $(x^*, 0)$ satisfies the first order necessary conditions with the same set of Lagrange multiplier vectors. The latter is implied by the LIKQ. Thus, Corollary 2.12 can be applied to every smooth branch problem and yields that $(x^*, 0)$ is a global minimizer in there. Then, $(x^*, 0)$ is a local minimizer of (eqNLP) by Lemma 3.15. □

In contrast to the smooth case some regularity is needed above. It is provided by the LIKQ which ensures that the Lagrange multiplier vectors of all branch problems coincide. Moreover, it cannot be concluded that the minimizer is global as it is in the smooth case.

Next, a closer look at a very simple example is taken.

Example 3.23. Consider the problem depicted in Figure 3.1:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & x_1 + x_2 \\ \text{s.t.} \quad & \min(x_1, x_2) = 0. \end{aligned}$$

The L-shaped feasible set consists of the two nonnegative axes. There is only one local minimizer: the strict global solution $x^* = (0, 0)$. The associated abs-normal NLP reads

$$\begin{aligned} \min_{x \in \mathbb{R}^2, z \in \mathbb{R}} \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1 + x_2 - |z| = 0, \\ & x_1 - x_2 - z = 0. \end{aligned}$$

Here, the reformulation $\min(x_1, x_2) = \frac{1}{2}(x_1 + x_2 - |x_1 - x_2|)$ was used. Clearly, the switching at x^* is localized as $z^* = z(x^*) = 0$ and LIKQ holds at x^* by full row rank of

$$J_{\mathrm{eq}}(x^*) = \begin{bmatrix} J_{\mathcal{E}}(x^*) \\ J_{\alpha}(x^*) \end{bmatrix} = \begin{bmatrix} \partial_1 c_{\mathcal{E}}(x^*, 0) \\ \partial_1 c_{\mathcal{Z}}(x^*, 0) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Since all functions are linear, the assumptions of Theorem 3.22 are satified. Therefore, the following first order conditions are both necessary and sufficient:

$$\begin{aligned} 1 + \lambda_{\mathcal{E}} + \lambda_{\mathcal{Z}} &= 0, && \\ 1 + \lambda_{\mathcal{E}} - \lambda_{\mathcal{Z}} &= 0 && \text{(tangential stationarity)}, \\ -\lambda_{\mathcal{E}} &\geq |\lambda_{\mathcal{Z}}| && \text{(normal growth)}, \\ x_1 + x_2 &= 0 && \text{(primal feasibility)}, \\ x_1 - x_2 &= 0 && \text{(switching feasibility)}. \end{aligned}$$

They are satisfied at $x^* = (0, 0)$ with $\lambda_{\mathcal{E}}^* = -1$ and $\lambda_{\mathcal{Z}}^* = 0$.

In the second case the number of constraints plus the number of switching variables is equal to the dimension. Then, the matrix $U_{\mathrm{eq}}(x^*)$ is empty and the second order conditions become trivial.

Theorem 3.24 (First Order Sufficient Conditions for $s+m = n$)**.** *Given* (eqNLP) *for $d \geq 2$, assume that $s + m = n$ holds. Assume further that $(x^*, 0)$ is feasible and that LIKQ holds at x^*. Then, $(x^*, 0)$ is a strict local minimizer of* (eqNLP) *if there exists a Lagrange multiplier vector $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_{\mathcal{Z}}^*)$ such that the conditions* (3.6) *with strict normal growth,*

$$(\lambda_{\mathcal{E}}^*)^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda_{\mathcal{Z}}^*)^T \partial_2 c_{\mathcal{Z}}(x^*, 0) > |\lambda_{\mathcal{Z}}^*|^T,$$

are satisfied.

Proof. As $s + m = n$ holds, the matrix $U_{\mathrm{eq}}(x)$ is empty. Thus, the result follows directly from Theorem 3.20. □

Next a slightly more complicated example is looked at.

Example 3.25. Consider the abs-normal NLP depicted in Figure 3.2:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & x_1^2 + x_2 \\ \text{s.t.} \quad & |x_1^2 + x_1 - x_2| - x_2 = 0. \end{aligned}$$

The feasible set consists of two unconnected components and there are two local minimizers, $(0, 0)$ and $(-1, 0)$. Here, the global solution $x^* = (0, 0)$ is considered. The abs-normal NLP reads

$$\begin{aligned} \min_{x \in \mathbb{R}^2, z \in \mathbb{R}} \quad & x_1^2 + x_2 \\ \text{s.t.} \quad & |z| - x_2 = 0, \\ & x_1^2 + x_1 - x_2 - z = 0. \end{aligned}$$

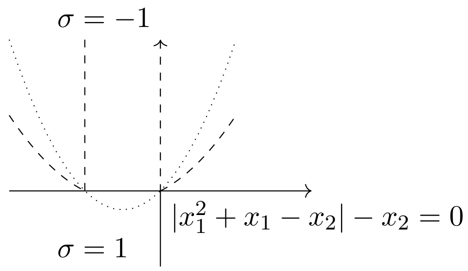

Figure 3.2: Illustration of Example 3.25. Dashed lines: feasible set; dotted line: $z = 0$.

Once again the switching is localized at x^* as $z^* = 0$ and the LIKQ holds at x^* by full rank of

$$J_{\text{eq}}(x^*) = \begin{bmatrix} J_{\mathcal{E}}(x^*) \\ J_{\alpha}(x^*) \end{bmatrix} = \begin{bmatrix} \partial_1 c_{\mathcal{E}}(x^*, 0) \\ \partial_1 c_{\mathcal{Z}}(x^*, 0) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 2x_1^* + 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}.$$

Here, the number of active switchings plus the number of constraints is equal to the dimension. Thus, the nullspace of the active Jacobian is trivial and second order conditions are always satisfied. Applying Theorem 3.24 yields the first order sufficient conditions:

$$\begin{aligned} 2x_1 + \lambda_{\mathcal{Z}}(2x_1 + 1) &= 0, \\ 1 - \lambda_{\mathcal{E}} - \lambda_{\mathcal{Z}} &= 0 && \text{(tangential stationarity)}, \\ \lambda_{\mathcal{E}} &> |\lambda_{\mathcal{Z}}| && \text{(strict normal growth)}, \\ -x_2 &= 0 && \text{(primal feasibility)}, \\ x_1^2 + x_1 - x_2 &= 0 && \text{(switching feasibility)}. \end{aligned}$$

They are satisfied at $x^* = (0, 0)$ with $\lambda_{\mathcal{E}}^* = 1$ and $\lambda_{\mathcal{Z}}^* = 0$. Note that the second local minimizer $x^* = (-1, 0)$ is also strict: the switching is localized, LIKQ holds and the first order sufficient conditions are satisfied with $\lambda_{\mathcal{E}}^* = 3$ and $\lambda_{\mathcal{Z}}^* = -2$.

3.1.4 General Non-Localized Case

As in the unconstrained case, the switching variables and constraints are divided into active and inactive ones if the switching is not localized. This is briefly recalled here; details can be found in subsection 2.2.4. The variables are partitioned:

$$z_+ = (\sigma_i^* z_i)_{i \notin \alpha^*} = (|z_i|)_{i \notin \alpha^*} \quad \text{and} \quad z_0 = (z_i)_{i \in \alpha^*}$$

where $z^* = z(x^*)$, $\sigma^* = \sigma(x^*)$ and $\alpha^* = \alpha(x^*)$. Moreover, the switching constraints are divided:

$$c_+ = (\sigma_i^* e_i^T \tilde{c}_{\mathcal{Z}})_{i \notin \alpha^*} \in C^d(D^{x,|z_0|,z_+}, \mathbb{R}^{|\sigma^*|}) \quad \text{and} \quad c_0 = (e_i^T \tilde{c}_{\mathcal{Z}})_{i \in \alpha^*} \in C^d(D^{x,|z_0|,z_+}, \mathbb{R}^{|\alpha^*|})$$

where

$$\tilde{c}_{\mathcal{Z}}(x, |z_0|, z_+) = c_{\mathcal{Z}}(x, |z|) \quad \text{with} \quad |z| = \Pi \begin{bmatrix} |z_0| \\ z_+ \end{bmatrix}$$

and an appropiate permutation matrix Π. The equality constraint in (eqNLP) is analogously partitioned:

$$\tilde{c}_{\mathcal{E}}(x, |z_0|, z_+) := c_{\mathcal{E}}(x, |z|) \quad \text{with} \quad |z| = \Pi \begin{bmatrix} |z_0| \\ z_+ \end{bmatrix}.$$

Thus, $c_{\mathcal{E}} \in C^d(D^{x,|z|}, \mathbb{R}^m)$ is interpreted as an element $\tilde{c}_{\mathcal{E}} \in C^d(D^{x,|z_0|,z_+}, \mathbb{R}^m)$ and in this sense usually just $c_{\mathcal{E}}(x, |z_0|, z_+)$ will be written in the following to simplify notation.

Using this partition, (smoothNLP) can be rewritten in a neighborhood of x^* in two different ways. For the first one, the split variables and functions are inserted in (eqNLP). This leads to the split abs-normal NLP which is defined next.

Definition 3.26 (Split Abs-Normal NLP)**.** The *split abs-normal NLP* reads

$$\begin{aligned} \min_{(x,z_0,z_+)\in D^{x,|z_0|,z_+}} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, |z_0|, z_+) = 0, \\ & c_0(x, |z_0|, z_+) - z_0 = 0, \\ & c_+(x, |z_0|, z_+) - z_+ = 0. \end{aligned} \tag{sNLP}$$

Its feasible set is denoted by

$$\mathcal{F}^s := \{(x, z_0, z_+) \in D^{x,|z_0|,z_+} \colon c_{\mathcal{E}}(x, |z_0|, z_+) = 0, c_0(x, |z_0|, z_+) = z_0, c_+(x, |z_0|, z_+) = z_+\}.$$

By construction, (x^*, z^*) with $z^* = \Pi \begin{bmatrix} 0 \\ \Sigma_+^* z_+^* \end{bmatrix}$ is a local minimizer of (eqNLP) if and only if $(x^*, 0, z_+^*)$ is a local minimizer of (sNLP).

To obtain the second one, the locally unique solution $z_+(x, \bar{z}_0)$ of $c_+(x, \bar{z}_0, z_+) = z_+$ from Lemma 2.39 is inserted. The resulting NLP is called reduced abs-normal NLP and formulated in the next definition.

Definition 3.27 (Reduced Abs-Normal NLP)**.** The *reduced abs-normal NLP* reads

$$\begin{aligned} \min_{(x,z_0)\in \mathcal{N}^{x,|z_0|}} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, |z_0|, z_+(x, |z_0|)) = 0, \\ & c_0(x, |z_0|, z_+(x, |z_0|)) - z_0 = 0. \end{aligned} \tag{rNLP}$$

Its feasible set is denoted by

$$\mathcal{F}^r := \{(x, z_0) \in \mathcal{N}^{x,|z_0|} \colon c_{\mathcal{E}}(x, |z_0|, z_+(x, |z_0|)) = 0, c_0(x, |z_0|, z_+(x, |z_0|) = z_0\}.$$

Clearly, $(x^*, 0, z_+^*)$ with $z_+^* = z_+(x^*, 0)$ is a local minimizer of (sNLP) if and only if $(x^*, 0)$ is a local minimizer of (rNLP).

The partial derivatives of c_0 and $c_{\mathcal{E}}$ at $(x, \bar{z}_0, z_+(x, \bar{z}_0))$ are obtained by the chain rule:

$$\begin{aligned} \partial_x c_0 &= \partial_1 c_0 + \partial_3 c_0 \partial_x z_+(x, \bar{z}_0), & \partial_x c_{\mathcal{E}} &= \partial_1 c_{\mathcal{E}} + \partial_3 c_{\mathcal{E}} \partial_x z_+(x, \bar{z}_0), \\ \partial_{\bar{z}_0} c_0 &= \partial_2 c_0 + \partial_3 c_0 \partial_{\bar{z}_0} z_+(x, \bar{z}_0), & \partial_{\bar{z}_0} c_{\mathcal{E}} &= \partial_2 c_{\mathcal{E}} + \partial_3 c_{\mathcal{E}} \partial_{\bar{z}_0} z_+(x, \bar{z}_0). \end{aligned}$$

In the following two subsections the trunk and branch problems will be studied again to derive optimality conditions for the split abs-normal NLP and for the reduced abs-normal NLP, yielding equivalent but different formulations.

3.1.4.1 Optimality Conditions of Reduced Abs-Normal NLP

The reduced abs-normal NLP looks like the abs-normal NLP in the localized case, except that it involves the implicit function $z_+(x, |z_0|)$. This leads to more complicated tangential stationarity and normal growth conditions that involve the above total derivatives.

Reduced Trunk Problem Setting $z_0 = 0$ in (rNLP) yields the reduced trunk problem.

Definition 3.28 (Reduced Trunk Problem)**.** The *reduced trunk problem* reads

$$\begin{aligned} \min_{x \in \mathcal{N}^x} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, 0, z_+(x, 0)) = 0, \\ & c_0(x, 0, z_+(x, 0)) = 0. \end{aligned} \tag{3.7}$$

Its feasible set is denoted by

$$\mathcal{F}_t^r := \{x \in \mathcal{N}^x \colon c_{\mathcal{E}}(x, 0, z_+(x, 0)) = 0, c_0(x, 0, z_+(x, 0)) = 0\}.$$

As in the localized case, the inclusion $\mathcal{F}_t^r \times \{0\} \subseteq \mathcal{F}^r$ directly implies the next lemma.

Lemma 3.29. *If $(x^*, 0)$ is a local minimizer of (rNLP), then x^* is a local minimizer of the reduced trunk problem (3.7).*

Again, the trunk problem is smooth and it is shown that LIKQ at x^* reduces to LICQ at x^*, which is full row rank of

$$J_t^r(x^*) = \begin{bmatrix} \partial_x c_{\mathcal{E}}(x^*, 0, z_+(x^*, 0)) \\ \partial_x c_0(x^*, 0, z_+(x^*, 0)) \end{bmatrix} \in \mathbb{R}^{(m+|\alpha^*|)\times n}.$$

Lemma 3.30. *Assume that $(x^*, z(x^*))$ is feasible for (eqNLP). Then, LIKQ at x^* is LICQ at x^* for the reduced trunk problem.*

Proof. Let $\Sigma^* = \operatorname{diag}(\sigma^*)$ and split $z^* = z(x^*)$. This gives

$$z^* = \Pi \begin{bmatrix} z_0^* \\ \Sigma_+^* z_+^* \end{bmatrix} = \Pi \begin{bmatrix} 0 \\ \Sigma_+^* z_+^* \end{bmatrix} \quad \text{and} \quad \Sigma^* = \Pi \begin{bmatrix} \Sigma_0^* & 0 \\ 0 & \Sigma_+^* \end{bmatrix} \Pi^T = \Pi \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_+^* \end{bmatrix} \Pi^T$$

with an appropiate permutation matrix Π as well as

$$c_{\mathcal{Z}}(x^*, \Sigma z^*) = \Pi \begin{bmatrix} c_0(x^*, 0, z_+^*) \\ \Sigma_+^* c_+(x^*, 0, z_+^*) \end{bmatrix} \quad \text{and} \quad c_{\mathcal{E}}(x^*, \Sigma z^*) = \tilde{c}_{\mathcal{E}}(x^*, 0, z_+^*)$$

using the explicit notation $\tilde{c}_{\mathcal{E}}$. This implies the following equalities for partial derivatives omitting the argument $(x^*, 0, z_+^*)$ on all partial derivatives of c_0, c_+ and $\tilde{c}_{\mathcal{E}}$:

$$\begin{aligned} \partial_1 c_{\mathcal{Z}}(x^*, \Sigma^* z^*) &= \Pi \begin{bmatrix} \partial_1 c_0 \\ \Sigma_+^* \partial_1 c_+ \end{bmatrix}, & \partial_2 c_{\mathcal{Z}}(x^*, \Sigma^* z^*) &= \Pi \begin{bmatrix} \partial_2 c_0 & \partial_3 c_0 \\ \Sigma_+^* \partial_2 c_+ & \Sigma_+^* \partial_3 c_+ \end{bmatrix} \Pi^T, \\ \partial_1 c_{\mathcal{E}}(x^*, \Sigma^* z^*) &= \partial_1 \tilde{c}_{\mathcal{E}}, & \partial_2 c_{\mathcal{E}}(x^*, \Sigma^* z^*) &= \begin{bmatrix} \partial_2 \tilde{c}_{\mathcal{E}} & \partial_3 \tilde{c}_{\mathcal{E}} \end{bmatrix} \Pi^T. \end{aligned}$$

Then, the Jacobian of the implicit function reads

$$\begin{aligned}
\partial_x z(x^*) =&(I - \partial_2 c_{\mathcal{Z}}(x^*, \Sigma^* z(x^*))\Sigma^*)^{-1}\partial_1 c_{\mathcal{Z}}(x^*, \Sigma^* z(x^*)) \\
=&\left(\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \Pi \begin{bmatrix} \partial_2 c_0 & \partial_3 c_0 \\ \Sigma_+^* \partial_2 c_+ & \Sigma_+^* \partial_3 c_+ \end{bmatrix} \Pi^T \Pi \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_+^* \end{bmatrix} \Pi^T \right)^{-1} \Pi \begin{bmatrix} \partial_1 c_0 \\ \Sigma_+^* \partial_1 c_+ \end{bmatrix} \\
=&\Pi \left(\begin{bmatrix} I & -\partial_3 c_0 \Sigma_+^* \\ 0 & I - \Sigma_+^* \partial_3 c_+ \Sigma_+^* \end{bmatrix}\right)^{-1} \begin{bmatrix} \partial_1 c_0 \\ \Sigma_+^* \partial_1 c_+ \end{bmatrix} \\
=&\Pi \begin{bmatrix} I & \partial_3 c_0 \Sigma_+^* \left[I - \Sigma_+^* \partial_3 c_+ \Sigma_+^*\right]^{-1} \\ 0 & \left[I - \Sigma_+^* \partial_3 c_+ \Sigma_+^*\right]^{-1} \end{bmatrix} \begin{bmatrix} \partial_1 c_0 \\ \Sigma_+^* \partial_1 c_+ \end{bmatrix} \\
=&\Pi \begin{bmatrix} I & \partial_3 c_0 \left[I - \partial_3 c_+\right]^{-1} \\ 0 & \Sigma_+^* \left[I - \partial_3 c_+\right]^{-1} \end{bmatrix} \begin{bmatrix} \partial_1 c_0 \\ \partial_1 c_+ \end{bmatrix}.
\end{aligned}$$

Here, the explicit inverses $\Pi^{-1} = \Pi^T$ and $(\Sigma_+^*)^{-1} = \Sigma_+^*$ and thus the relations $\Pi^T \Pi = I$ and $\Sigma_+^* \Sigma_+^* = I$ were used. Inserting this in Definition 3.3 and further omitting the argument $(x^*, 0, z_+(x^*))$ on all partial derivatives of $c_{\mathcal{E}}, c_0$ and c_+, one obtains

$$\begin{aligned}
J_{\mathcal{E}}(x^*) =&\partial_1 c_{\mathcal{E}}(x^*, \Sigma^* z(x^*)) + \partial_2 c_{\mathcal{E}}(x^*, \Sigma^* z(x^*))\Sigma^* \partial_x z(x^*) \\
=&\partial_1 \tilde{c}_{\mathcal{E}} + \begin{bmatrix} \partial_2 \tilde{c}_{\mathcal{E}} & \partial_3 \tilde{c}_{\mathcal{E}} \end{bmatrix} \Pi^T \Pi \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_+^* \end{bmatrix} \Pi^T \partial_x z(x^*) \\
=&\partial_1 \tilde{c}_{\mathcal{E}} + \partial_3 \tilde{c}_{\mathcal{E}} \left[I - \partial_3 c_+\right]^{-1} \partial_1 c_+, \\
J_\alpha(x^*) =&[e_i^T \partial_x z(x^*)]_{i \in \alpha} \\
=&\partial_1 c_0 + \partial_3 c_0 [I - \partial_3 c_+]^{-1} \partial_1 c_+
\end{aligned}$$

and with $z_+(x^*) = z_+(x^*, 0)$ finally

$$\begin{aligned}
J_{\mathcal{E}}(x^*) &= \partial_x \tilde{c}_{\mathcal{E}}(x^*, 0, z_+(x^*, 0)), \\
J_\alpha(x^*) &= \partial_x c_0(x^*, 0, z_+(x^*, 0)).
\end{aligned}$$

Thus, $J_t^r(x^*)$ has the form above using the notation $c_{\mathcal{E}}$ instead of $\tilde{c}_{\mathcal{E}}$. □

Note that the matrix Π was omitted in [17] to simplify notation.

Thus, the columns of the matrix $U_t^r(x^*) = U_{\mathrm{eq}}(x^*) \in \mathbb{R}^{n \times [n-(m+|\alpha^*|)]}$ are a basis of the nullspace of $J_t^r(x^*)$.

Again first and second order necessary conditions are obtained by standard NLP theory for the smooth case. In contrast to the localized case, the implicit function $z_+(x, 0)$ and therefore total derivatives occur.

The Lagrangian is denoted by

$$\mathcal{L}_t^r(x, \lambda) := f(x) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x, 0, z_+(x, 0)) + \lambda_0^T c_0(x, 0, z_+(x, 0))$$

with Lagrange multiplier vector $\lambda = (\lambda_{\mathcal{E}}, \lambda_0) \in \mathbb{R}^{m+|\alpha^*|}$.

Theorem 3.31 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0)$ is a local minimizer of (rNLP) and that LIKQ holds at x^*. Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_0)$ such that the following conditions are satisfied, where $c_{\mathcal{E}}, c_0, \partial_x c_{\mathcal{E}}$ and $\partial_x c_0$ are evaluated at $(x^*, 0, z_+(x^*, 0))$:*

$$
\begin{aligned}
f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_x c_{\mathcal{E}} + (\lambda^*_0)^T \partial_x c_0 &= 0 \quad \textit{(tangential stationarity)},\\
c_{\mathcal{E}} &= 0 \quad \textit{(primal feasibility)},\\
c_0 &= 0 \quad \textit{(switching feasibility)}.
\end{aligned}
\tag{3.8}
$$

Proof. This follows as in the proof of Theorem 3.9: Lemma 3.29 tells that local minimizers are inherited by the trunk problem. Then, Theorem 2.7 can be applied since the trunk problem is smooth and LICQ holds by Lemma 3.30. Further, the latter implies that the Lagrange multiplier vector λ^* is unique. □

Note that these conditions are *not* sufficient for linear functions since the converse of Lemma 3.29 does not hold.

Theorem 3.32 (Second Order Necessary Condition)**.** *Consider (rNLP) for $d \geq 2$. Assume that $(x^*, 0)$ is a local minimizer and that LIKQ holds at x^*. Denote by λ^* the unique Lagrange multiplier vector. Then,*

$$
U_{eq}(x^*)^T H^r_t(x^*, \lambda^*) U_{eq}(x^*) \geq 0
$$

with $H^r_t(x, \lambda) = \partial^2_{xx} \mathcal{L}^r_t(x, \lambda)$.

Proof. Again, the proof of the localized case, i.e. the proof of Theorem 3.9, can be adapted: Thus, the smooth trunk problem is considered. Then, Theorem 2.10 can be applied as LICQ is satisfied and as $I = \emptyset$ holds. □

Here, the Hessian is more difficult to compute than in the localized case because an implicit function occurs in the Lagrangian.

Reduced branch problems To obtain sufficient conditions, the $2^{|\alpha^*|}$ branch problems are considered.

Definition 3.33 (Reduced Branch Problems)**.** Choose $\sigma_0 \in \{-1, 1\}^{|\alpha^*|}$ and set $\Sigma_0 = \operatorname{diag}(\sigma_0)$. The *reduced branch problem* associated with σ_0 reads

$$
\begin{aligned}
\min_{(x, z_0) \in \mathcal{N}^{x, |z_0|}} \quad & f(x)\\
\text{s.t.} \quad & c_{\mathcal{E}}(x, \Sigma_0 z_0, z_+(x, \Sigma_0 z_0)) = 0,\\
& c_0(x, \Sigma_0 z_0, z_+(x, \Sigma_0 z_0)) - z_0 = 0,\\
& \Sigma_0 z_0 \geq 0.
\end{aligned}
$$

Its feasible set is denoted by

$$
\mathcal{F}^r := \left\{ (x, z_0) \in \mathcal{N}^{x, |z_0|} \;\middle|\; \begin{array}{l} c_{\mathcal{E}}(x, \Sigma_0 z_0, z_+(x, \Sigma_0 z_0)) = 0, \\ c_0(x, \Sigma_0 z_0, z_+(x, \Sigma_0 z_0)) = z_0, \;\; \Sigma_0 z_0 \geq 0 \end{array} \right\}.
$$

With $\bar{z}_0 = \Sigma_0 z_0$ it takes the equivalent form

$$\begin{aligned} \min_{x,\bar{z}_0} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, \bar{z}_0, z_+(x, \bar{z}_0)) = 0, \\ & c_0(x, \bar{z}_0, z_+(x, \bar{z}_0)) - \Sigma_0 \bar{z}_0 = 0, \\ & \bar{z}_0 \geq 0. \end{aligned} \tag{3.9}$$

Again, the feasible sets have some closer relations. By construction:

$$\mathcal{F}_t^r \times \{0\} \subseteq \mathcal{F}_{\Sigma_0}^r \subseteq \mathcal{F}^r \quad \text{for all } \sigma_0 \in \{-1, 1\}^{|\alpha^*|} \tag{3.10}$$

Lemma 3.34. *The feasible sets satisfy the following relations:*

$$\bigcap_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}_{\Sigma_0}^r = \mathcal{F}_t^r \times \{0\} \quad \textit{and} \quad \bigcup_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}_{\Sigma_0}^r = \mathcal{F}^r$$

Proof. The proof of Lemma 3.14 can be modified to the reduced case by using z_0 and Σ_0 instead of z and Σ: Due to the implications (3.10) one just needs to show

$$\bigcap_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}_{\Sigma_0}^r \subseteq \mathcal{F}_t^r \times \{0\} \quad \text{and} \quad \bigcup_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}_{\Sigma_0}^r \supseteq \mathcal{F}^r.$$

Thus, consider $(x, z_0) \in \bigcap_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}_{\Sigma_0}^r$ to prove the first inclusion. Then, $z_0 = 0$ has to hold as $\Sigma_0 z_0 \geq 0$ is satisfied for every definite σ_0. This implies directly $(x, 0) \in \mathcal{F}_t^r \times \{0\}$. For the second inclusion, consider $(x, z_0) \in \mathcal{F}^r$. Then, $(x, z_0) \in \mathcal{F}_{\tilde{\Sigma}}^r \subseteq \bigcup_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}_{\Sigma_0}^r$ for the choice $(\tilde{\sigma}_0)_i = 1$ for $i \in \alpha_0(x)$ and $(\tilde{\sigma}_0)_i = (\sigma_0(x))_i$ otherwise. □

As in the localized case, the relation between feasible sets implies a relation between minimizers.

Lemma 3.35. *The pair $(x^*, 0)$ is a local minimizer of (rNLP) if and only if it is a local minimizer of the reduced branch problem (3.9) for every definite $\Sigma_0 = \operatorname{diag}(\sigma_0)$.*

Proof. The proof is analogous to the proof of Lemma 3.15 except that the reduced feasible sets have to be considered: Assume that $(x^*, 0)$ is a local minimizer of (rNLP), then $(x^*, 0) \in \mathcal{F}_t^r \times \{0\} \subseteq \mathcal{F}_{\Sigma_0}^r$ by Lemma 3.29 and (3.10). As $\mathcal{F}_{\Sigma_0}^r \subseteq \mathcal{F}^r$ holds for all σ_0 by (3.10), the pair $(x^*, 0)$ has to be a local minimizer of every branch problem (3.9).
Conversely, assume that $(x^*, 0)$ is a local minimizer of every branch problem (3.9). Then, $(x^*, 0) \in \mathcal{F}_{\Sigma_0}^r \subseteq \mathcal{F}^r$ for all σ_0 by (3.10) and $(x^*, 0)$ has to be a local minimizer of (rNLP) since $\bigcup_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}_{\Sigma_0}^r = \mathcal{F}^r$ by Lemma 3.34. □

By construction, every branch problem is smooth and once again standard NLP theory can be applied. As before, LIKQ at x^* implies LICQ at $(x^*, 0)$, which is full row rank of the matrix

$$J_{\Sigma_0}^r(x^*) = \begin{bmatrix} \partial_x c_{\mathcal{E}}(x^*, 0, z_+(x^*, 0)) & \partial_{\bar{z}_0} c_{\mathcal{E}}(x^*, 0, z_+(x^*, 0)) \\ \partial_x c_0(x, 0, z_+(x^*, 0)) & \partial_{\bar{z}_0} c_0(x^*, 0, z_+(x^*, 0)) - \Sigma_0 \\ 0 & I \end{bmatrix} \in \mathbb{R}^{(m+2|\alpha^*|) \times (n+|\alpha^*|)}.$$

Lemma 3.36. *Assume that $(x^*, 0)$ is feasible for (rNLP). Then, LICQ holds at $(x^*, 0)$ for all branch problems if LIKQ holds at x^*.*

Proof. LIKQ at x^* implies LICQ at $(x^*, 0)$ for the trunk problem by Lemma 3.30. Then, block elimination in $J^r_{\Sigma_0}$ gives the result. □

A basis of the nullspace of $J^r_{\Sigma_0}(x^*)$ is given by the columns of the matrix $U^r_{\Sigma_0}(x^*) = \left[U_{\mathrm{eq}}(x^*)^T \; 0\right]^T \in \mathbb{R}^{(n+|\alpha^*|)\times[n-(m+|\alpha^*|)]}$. The Lagrangian of the branch problem (3.9) is denoted by

$$\mathcal{L}^r_{\Sigma_0}(x, \bar{z}_0, \lambda, \mu_0) := f(x) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x, \bar{z}_0, z_+(x, \bar{z}_0)) + \lambda_0^T[c_0(x, \bar{z}_0, z_+(x, \bar{z}_0)) - \Sigma_0 \bar{z}_0] - \mu_0^T \bar{z}_0$$

with Lagrange multiplier vectors $\lambda = (\lambda_{\mathcal{E}}, \lambda_0) \in \mathbb{R}^{m+|\alpha^*|}$ and $\mu_0^* \in \mathbb{R}^{|\alpha^*|}$. Using this Lagrangian, first order conditions are obtained in the next theorem. Here, total derivatives are needed as in the trunk problem.

Theorem 3.37 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0)$ is a local minimizer of (rNLP) and that LIKQ holds at x^* Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_0)$ such that the following conditions are satisfied for every definite $\Sigma_0 = \mathrm{diag}(\sigma_0)$:*

$$\begin{aligned}
f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_x c_{\mathcal{E}} + (\lambda^*_0)^T \partial_x c_0 &= 0 && \textit{(tangential stationarity)},\\
(\lambda^*_{\mathcal{E}})^T \partial_{\bar{z}_0} c_{\mathcal{E}} + (\lambda^*_0)^T [\partial_{\bar{z}_0} c_0 - \Sigma_0] &\ge 0 && \textit{(normal growth)},\\
c_{\mathcal{E}} &= 0 && \textit{(primal feasibility)},\\
c_0 &= 0 && \textit{(switching feasibility)}.
\end{aligned}$$

Here, $c_{\mathcal{E}}, \partial_x c_{\mathcal{E}}, \partial_{\bar{z}_0} c_{\mathcal{E}}, c_0, \partial_x c_0$ and $\partial_{\bar{z}_0} c_0$ are evaluated at $(x^*, 0, z_+(x^*, 0))$.

Proof. This follows as in the proof of Theorem 3.17: For fixed $\sigma_0 \in \{-1, 1\}^{|\alpha^*|}$ the branch problem is smooth and Theorem 2.7 can be applied. Thus, $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_0) \in \mathbb{R}^{m+|\alpha^*|}$ and $\mu_0^* \in \mathbb{R}^{|\alpha^*|}$ exist such that

$$\begin{aligned}
\partial_x \mathcal{L}^r_{\Sigma_0}(x^*, 0, \lambda^*, \mu_0^*) = f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_x c_{\mathcal{E}} + (\lambda^*_0)^T \partial_x c_0 &= 0,\\
\partial_{\bar{z}} \mathcal{L}^r_{\Sigma_0}(x^*, 0, \lambda^*, \mu^*) = (\lambda^*_{\mathcal{E}})^T \partial_{\bar{z}_0} c_{\mathcal{E}} + (\lambda^*_0)^T [\partial_{\bar{z}_0} c_0 - \Sigma_0] - (\mu_0^*)^T &= 0,\\
c_{\mathcal{E}} &= 0,\\
c_0 &= 0,\\
\mu_0^* &\ge 0,
\end{aligned}$$

where $c_{\mathcal{E}}, \partial_x c_{\mathcal{E}}, \partial_{\bar{z}_0} c_{\mathcal{E}}, c_0, \partial_x c_0, \partial_{\bar{z}_0} c_0$ are evaluated at $(x^*, 0, z_+(x^*, 0))$. The first, third and fourth condition are the first order necessary conditions for the trunk problem (see Theorem 3.31). This implies that the Lagrange multiplier vector λ^* is unique and coincides for all branch problems and for the trunk problem. Again, only the second and fifth condition (and μ_0^* itself) depend on Σ_0. Combination of these gives the normal growth condition

$$(\lambda^*_{\mathcal{E}})^T \partial_{\bar{z}_0} c_{\mathcal{E}} + (\lambda^*_0)^T [\partial_{\bar{z}_0} c_0(x^*, 0) - \Sigma_0] \ge 0,$$

for every definite $\Sigma_0 = \mathrm{diag}(\sigma_0)$. □

As in the localized case it suffices to consider one branch with $\Sigma_0\lambda_0^* = |\lambda_0^*|$. This is due to the equivalence

$$\begin{aligned}
&(\lambda_{\mathcal{E}}^*)^T \left[\partial_{\bar{z}_0} c_{\mathcal{E}}\right] + (\lambda_0^*)^T \left[\partial_{\bar{z}_0} c_0\right] \geq |\lambda_0^*|^T \\
\iff\ &(\lambda_{\mathcal{E}}^*)^T \partial_{\bar{z}_0} c_{\mathcal{E}} + (\lambda_0^*)^T \partial_{\bar{z}_0} c_0 \geq (\lambda_0^*)^T(\pm\Sigma_0) \text{ for every definite } \Sigma_0.
\end{aligned}$$

and the next lemma.

Lemma 3.38. *The equivalence stated before directly implies:*

$$\begin{aligned}
&(\lambda_{\mathcal{E}}^*)^T \left[\partial_{\bar{z}_0} c_{\mathcal{E}}\right] + (\lambda_0^*)^T \left[\partial_{\bar{z}_0} c_0\right] \geq |\lambda_0^*|^T \\
\iff\ &(\lambda_{\mathcal{E}}^*)^T \partial_{\bar{z}_0} c_{\mathcal{E}} + (\lambda_0^*)^T \left[\partial_{\bar{z}_0} c_0 - \Sigma_0\right] \geq 0 \textit{ for every definite } \Sigma_0,
\end{aligned}$$

where $\partial_x c_{\mathcal{E}}, \partial_{\bar{z}_0} c_{\mathcal{E}}, \partial_x c_0, \partial_{\bar{z}_0} c_0$ are evaluated at $(x^*, 0, z_+(x^*, 0))$.

Thus, the first order necessary conditions can be stated with a single normal growth condition rather than $2^{|\alpha^*|}$ of them.

Corollary 3.39 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0)$ is a local minimizer of* (rNLP) *and that LIKQ holds at x^*. Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_0^*)$ such that the following conditions are satisfied, where $c_{\mathcal{E}}, c_0, \partial_x c_{\mathcal{E}}, \partial_x c_0, \partial_{\bar{z}_0} c_{\mathcal{E}}$ and $\partial_{\bar{z}_0} c_0$ are evaluated at $(x^*, 0, z_+(x^*, 0))$:*

$$\begin{aligned}
f'(x^*) + (\lambda_{\mathcal{E}}^*)^T \partial_x c_{\mathcal{E}} + (\lambda_0^*)^T \partial_x c_0 &= 0 && \textit{(tangential stationarity)}, \\
(\lambda_{\mathcal{E}}^*)^T \partial_{\bar{z}_0} c_{\mathcal{E}} + (\lambda_0^*)^T \partial_{\bar{z}_0} c_0 &\geq |\lambda_0^*|^T && \textit{(normal growth)}, \\
c_{\mathcal{E}} &= 0 && \textit{(primal feasibility)}, \\
c_0 &= 0 && \textit{(switching feasibility)}.
\end{aligned} \tag{3.11}$$

Moreover, second order sufficient conditions are obtained using the branch problems.

Theorem 3.40 (Second Order Sufficient Conditions)**.** *Consider* (rNLP) *for $d \geq 2$. Assume that $(x^*, 0)$ is feasible and that LIKQ holds at x^*. Assume further that a Lagrange multiplier vector $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_0^*)$ exists such that the first order necessary conditions* (3.11) *are satisfied with strict normal growth,*

$$(\lambda_{\mathcal{E}}^*)^T \partial_{\bar{z}_0} c_{\mathcal{E}}(x^*, 0, z_+(x^*, 0)) + (\lambda_0^*)^T \partial_{\bar{z}_0} c_0(x^*, 0, z_+(x^*, 0)) > |\lambda_0^*|^T,$$

and that

$$U_{eq}(x^*)^T H_t^r(x^*, \lambda^*) U_{eq}(x^*) > 0.$$

Then, $(x^*, 0)$ is a strict local minimizer of (rNLP)*.*

Proof. Once again, the proof of the localized case (see Theorem 3.20) can be adapted: In every branch problem strict complementarity holds as the strict normal growth condition is satisfied. Moreover, LIKQ implies that the Lagrange multiplier vectors are unique. Thus, all technical assumptions in Theorem 2.11 are fullfilled and it is left to show that

$$U_{\Sigma_0}^r(x^*)^T H_{\Sigma_0}^r(x^*, \lambda_0^*) U_{\Sigma_0}^r(x^*) > 0$$

holds for every branch problem. As μ_0 appears only in linear terms in the Lagrangian, it is independent of μ_0 and thus all Hessians coincide:

$$H^r_{\Sigma_0}(x^*, \lambda^*) = \begin{bmatrix} \partial_{xx}\mathcal{L}^r_{\Sigma_0}(x^*, 0, \lambda^*, \mu_0^*) & \partial_{x\bar{z}}\mathcal{L}^r_{\Sigma_0}(x^*, 0, \lambda^*, \mu_0^*) \\ \partial_{\bar{z}x}\mathcal{L}^r_{\Sigma_0}(x^*, 0, \lambda^*, \mu_0^*) & \partial_{\bar{z}\bar{z}}\mathcal{L}^r_{\Sigma_0}(x^*, 0, \lambda^*, \mu_0^*) \end{bmatrix} \in \mathbb{R}^{(n+|\alpha^*|)\times(n+|\alpha^*|)}$$

with $\partial_{xx}\mathcal{L}^r_{\Sigma_0}(x^*, 0, \lambda^*, \mu_0^*) = \partial_{xx}\mathcal{L}^r_t(x^*, \lambda^*) = H^r_t(x^*, \lambda^*)$. Substituting $U^r_{\Sigma_0}(x^*)$ yields

$$U^r_{\Sigma_0}(x^*)^T H^r_{\Sigma_0}(x^*, \lambda^*) U^r_{\Sigma_0}(x^*) = U_{\text{eq}}(x^*)^T H^r_t(x^*, \lambda^*) U_{\text{eq}}(x^*) > 0$$

and the assertion follows by Theorem 2.11. □

As in the localized case, the first order conditions in Theorem 3.31 are not sufficient under special circumstances. For this, the (strict) normal growth is additionally needed which is also a necessary condition as was shown in Corollary 3.39. Again, even in the linear case some regularity is required to ensure that the Lagrange multiplier vectors of all branch problems coincide.

Theorem 3.41 (First Order Necessary and Sufficient Conditions for Linear Functions)**.** *Given (rNLP), assume that f, $c_{\mathcal{E}}$ and c_0 are linear. Assume further that $(x^*, 0)$ is feasible and that LIKQ holds at x^*. Then, $(x^*, 0)$ is a local minimizer of (rNLP) if and only if there exist $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_0)$ such that the conditions (3.11) are satisfied.*

Proof. This is done analogously to the proof of Theorem 3.22: The assumptions provide that every branch problem is linear and that $(x^*, 0)$ satisfies the first order necessary conditions with the same set of Lagrange multiplier vectors. Thus, by Corollary 2.12 $(x^*, 0)$ is a global minimizer in every smooth branch problem. Consequently, $(x^*, 0)$ is a local minimizer of (rNLP) by Lemma 3.35. □

Under some assumptions on the problem dimension, the matrix $U_{\text{eq}}(x^*)$ is empty and thus the second order conditions become trivial.

Corollary 3.42 (First Order Sufficient Conditions for $|\alpha^*| + m = n$)**.** *Given (rNLP) for $d \geq 2$, assume that $|\alpha^*| + m = n$ holds. Assume further that $(x^*, 0)$ is feasible and that LIKQ holds at x^*. Then, $(x^*, 0)$ is a strict local minimizer of (rNLP) if there exists $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_0)$ such that the conditions (3.11) are satisfied with strict normal growth*

$$(\lambda^*_{\mathcal{E}})^T \partial_{\bar{z}_0} c_{\mathcal{E}} + (\lambda^*_0)^T \partial_{\bar{z}_0} c_0 > |\lambda^*_0|^T.$$

Proof. The matrix $U_{\text{eq}}(x)$ is empty as $|\alpha^*| + m = n$ and the result follows directly from Theorem 3.40. □

3.1.4.2 Optimality Conditions of Split Abs-Normal NLP

Here the implicit function of the reduced form is avoided with an additional switching constraint $c_+ - z_+ = 0$ and associated switching variables z_+. This causes an additional term in the Lagrangian, an additional tangential stationarity condition and a more complicated nullspace in the second order conditions.

Split trunk problem The split trunk problem is obtained from (sNLP) by setting $z_0 = 0$.

Definition 3.43 (Split Trunk Problem)**.** The *split trunk problem* reads

$$\begin{aligned} \min_{(x,z_+)\in D^{x,z_+}} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x,0,z_+) = 0, \\ & c_0(x,0,z_+) = 0, \\ & c_+(x,0,z_+) - z_+ = 0. \end{aligned} \tag{3.12}$$

Its feasible set is denoted by

$$\mathcal{F}_t^s := \{(x,z_+) \in D^{x,z_+} \colon c_{\mathcal{E}}(x,0,z_+) = 0, c_0(x,0,z_+) = 0, c_+(x,0,z_+) = z_+\}.$$

Note that $(x,z_+) \in \mathcal{F}_t^s$ implies that $(x,0,z_+) \in \mathcal{F}^s$. In the following the notation $\mathcal{F}_t^s \times \{0\} \subseteq \mathcal{F}^s$ is used to denote this relation. Once again, the next lemma follows directly by this inclusion.

Lemma 3.44. *If $(x^*,0,z_+^*)$ is a local minimizer of (sNLP), then (x^*,z_+^*) is a local minimizer of the split trunk problem (3.12).*

As before, the trunk problem is used to obtain necessary conditions.
Here, LICQ at (x^*,z_+^*) means full row rank of

$$J_t^s(x^*,z_+^*) = \begin{bmatrix} \partial_1 c_{\mathcal{E}}(x^*,0,z_+^*) & \partial_3 c_{\mathcal{E}}(x^*,0,z_+^*) \\ \partial_1 c_0(x^*,0,z_+^*) & \partial_3 c_0(x^*,0,z_+^*) \\ \partial_1 c_+(x^*,0,z_+^*) & \partial_3 c_+(x^*,0,z_+^*) - I \end{bmatrix} \in \mathbb{R}^{(m+s)\times(n+|\sigma^*|)}.$$

This is equivalent to LIKQ at x^*.

Lemma 3.45. *Assume that $(x^*,z(x^*))$ is feasible for (eqNLP). Then, LIKQ holds at x^* if and only if LICQ holds at (x^*,z_+^*) for the split trunk problem.*

Proof. Block elimination and using the representations of the Jacobians in the proof of Lemma 3.30 yields

$$\begin{aligned} \operatorname{rank}(J_t^s(x^*,z_+^*)) &= \operatorname{rank} \begin{bmatrix} \partial_1 c_{\mathcal{E}} - \partial_3 c_{\mathcal{E}}(\partial_3 c_+ - I)^{-1}\partial_1 c_+ & 0 \\ \partial_1 c_0 - \partial_3 c_0(\partial_3 c_+ - I)^{-1}\partial_1 c_+ & 0 \\ \partial_1 c_+ & \partial_3 c_+ - I \end{bmatrix} \\ &= \operatorname{rank} \begin{bmatrix} J_{\mathcal{E}}(x^*) & 0 \\ J_\alpha(x^*) & 0 \\ \partial_1 c_+ & \partial_3 c_+ - I \end{bmatrix} \\ &= \operatorname{rank}(J_{\mathrm{eq}}(x^*)) + |\sigma^*|, \end{aligned}$$

where all partial derivatives of $c_{\mathcal{E}}, c_0$ and c_+ are evaluated at $(x^*,0,z_+^*)$. Note, that the last equality holds since $\partial_3 c_+(x^*,0,z_+^*)$ is strictly lower triangular. □

If LIKQ holds at x^*, a matrix whose columns are a basis for $J_t^s(x^*,z_+^*)$ is denoted by $U_t^s(x^*,z_+^*) \in \mathbb{R}^{(n+|\sigma^*|)\times[n-(m+|\alpha^*|)]}$. The particular form is given in the next lemma.

Lemma 3.46. *The matrix $U_t^s(x^*, z_+^*)$ has the following form:*

$$U_t^s(x^*, z_+^*) = \begin{bmatrix} U_{eq}(x^*) \\ (I - \partial_3 c_+(x^*, 0, z_+^*))^{-1}\partial_1 c_+(x^*, 0, z_+^*) U_{eq}(x^*) \end{bmatrix}.$$

Proof. Consider the equation

$$0 = J_t^s(x^*, z_+^*) \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} \partial_1 c_{\mathcal{E}}(x^*, 0, z_+^*) & \partial_3 c_{\mathcal{E}}(x^*, 0, z_+^*) \\ \partial_1 c_0(x^*, 0, z_+^*) & \partial_3 c_0(x^*, 0, z_+^*) \\ \partial_1 c_+(x^*, 0, z_+^*) & \partial_3 c_+(x^*, 0, z_+^*) - I \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}.$$

Then, the third line gives the correspondence $U_2 = (I - \partial_3 c_+(x^*, 0, z_+^*))^{-1}\partial_1 c_+(x^*, 0, z_+^*) U_1$ as $\partial_3 c_+(x^*, 0, z_+^*)$ is strictly lower triangular. Inserting this representation of U_2 into the first and the second equations leads to

$$0 = \begin{bmatrix} J_{\mathcal{E}}(x^*) \\ J_\alpha(x^*) \end{bmatrix} U_1.$$

Then, the result follows directly as $U_{\mathrm{eq}}(x^*)$ is defined as the proper matrix. □

Note that the nullspace matrix is more complicated than in the reduced abs-normal NLP. This is due to the fact that the variable z_+ is explicitly handled.

Using once again the theory for the smooth case, necessary optimality conditions can be obtained. Here, the Lagrangian is denoted by

$$\mathcal{L}_t^s(x, z_+, \lambda) = f(x) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x, 0, z_+) + \lambda_0^T c_0(x, 0, z_+) + \lambda_+^T [c_+(x, 0, z_+) - z_+]$$

with Lagrange multiplier vector $\lambda = (\lambda_{\mathcal{E}}, \lambda_0, \lambda_+) \in \mathbb{R}^{m+|\alpha^*|+|\sigma^*|}$.

The additional term leads to a further tangential stationarity condition.

Theorem 3.47 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0, z_+^*)$ is a local minimizer of (sNLP) and that LIKQ holds at x^*. Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_0^*, \lambda_+^*)$ such that the following conditions are satisfied, where the constraints $c_{\mathcal{E}}, c_0, c_+$ and all their partial derivatives are evaluated at $(x^*, 0, z_+^*)$:*

$$\begin{aligned} f'(x^*) + (\lambda_{\mathcal{E}}^*)^T \partial_1 c_{\mathcal{E}} + (\lambda_0^*)^T \partial_1 c_0 + (\lambda_+^*)^T \partial_1 c_+ &= 0, \\ (\lambda_{\mathcal{E}}^*)^T \partial_3 c_{\mathcal{E}} + (\lambda_0^*)^T \partial_3 c_0 + (\lambda_+^*)^T [\partial_3 c_+ - I] &= 0 \quad \textit{(tangential stationarity)}, \\ c_{\mathcal{E}} &= 0 \quad \textit{(primal feasibility)}, \\ c_0 &= 0, \\ c_+ - z_+^* &= 0 \quad \textit{(switching feasibility)}. \end{aligned} \tag{3.13}$$

Proof. Once again, this follows as in the proof of Theorem 3.9. Due to Lemma 3.44, the necessary conditions are obtained by applying Theorem 2.7 to the trunk problem. By Lemma 3.45 the assumptions of the latter are satisfied and the Lagrange multiplier vector λ^* is unique. □

Due to the explicit handling of z_+ the Hessian contains four partial derivatives of the Lagrangian $\mathcal{L}_t^s$.

Theorem 3.48 (Second Order Necessary Condition)**.** *Consider* (sNLP) *for $d \geq 2$. Assume that $(x^*, 0, z_+^*)$ is a local minimizer and that LIKQ holds at x^*. Denote by λ^* the unique Lagrange multiplier vector. Then,*

$$U_t^s(x^*, z_+^*)^T H_t^s(x^*, z_+^*, \lambda^*) U_t^s(x^*, z_+^*) \geq 0$$

with $H_t^s(x^*, z_+^*, \lambda^*) = \partial^2_{(x,z_+),(x,z_+)} \mathcal{L}_t^s(x^*, z_+^*, \lambda^*)$.

Proof. Again, the proof of Theorem 3.11 is adapted: Using Lemma 3.44, the smooth trunk problem is considered. Then, Theorem 2.10 gives the result as LICQ holds by Lemma 3.45 and $\mathcal{I} = \emptyset$. □

As one cannot obtain sufficient conditions by using the split trunk problem, once again branch problems in the neighborhood of the local minimizer are studied.

Split branch problems The splitting yields $2^{|\alpha^*|}$ different branch problems.

Definition 3.49 (Split Branch Problem)**.** Choose $\sigma_0 \in \{-1, 1\}^{|\alpha^*|}$ and set $\Sigma_0 = \operatorname{diag}(\sigma_0)$. The *split branch problem* associated with σ_0 reads

$$\begin{aligned} \min_{(x,z_0,z_+) \in D^{x,|z_0|,z_+}} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, \Sigma_0 z_0, z_+) = 0, \\ & c_0(x, \Sigma_0 z_0, z_+) - z_0 = 0, \\ & c_+(x, \Sigma_0 z_0, z_+) - z_+ = 0, \\ & \Sigma_0 z_0 \geq 0. \end{aligned}$$

Its feasible set is denoted by

$$\mathcal{F}_{\Sigma_0}^s := \left\{ (x, z_0, z_+) \in D^{x,|z_0|,z_+} \;\middle|\; \begin{array}{l} c_{\mathcal{E}}(x, \Sigma_0 z_0, z_+) = 0, \;\; c_0(x, \Sigma_0 z_0, z_+) = z_0, \\ c_+(x, \Sigma_0 z_0, z_+) = z_+, \;\; \Sigma_0 z_0 \geq 0 \end{array} \right\}.$$

With $\bar{z}_0 = \Sigma_0 z_0$ it takes the equivalent form

$$\begin{aligned} \min_{x,\bar{z}_0,z_+} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, \bar{z}_0, z_+) = 0, \\ & c_0(x, \bar{z}_0, z_+) - \Sigma_0 \bar{z}_0 = 0, \\ & c_+(x, \bar{z}_0, z_+) - z_+ = 0, \\ & \bar{z}_0 \geq 0. \end{aligned} \tag{3.14}$$

By construction, the following implications are given:

$$\mathcal{F}_t^s \times \{0\} \subseteq \mathcal{F}_{\Sigma_0}^s \subseteq \mathcal{F}^s \quad \text{for all } \sigma_0 \in \{-1, 1\}^{|\alpha^*|} \tag{3.15}$$

Moreover, the feasible set of the trunk problem is the intersection of the feasible sets of the branch problems and the set $\mathcal{F}^s$ is the union.

Lemma 3.50. *The feasible sets satisfy the following relation:*

$$\bigcap_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}^s_{\Sigma_0} = \mathcal{F}^s_t \times \{0\} \quad \textit{and} \quad \bigcup_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}^s_{\Sigma_0} = \mathcal{F}^s$$

Proof. Once again, the proof of Lemma 3.14 can be modified. Here, z_0 and Σ_0 are considered instead of z and Σ: It is sufficient to show

$$\bigcap_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}^s_{\Sigma_0} \subseteq \mathcal{F}^s_t \times \{0\} \quad \text{and} \quad \bigcup_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}^s_{\Sigma_0} \supseteq \mathcal{F}^s$$

due to the implications (3.15). For the first inclusion, consider $(x, z_0) \in \bigcap_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}^s_{\Sigma_0}$. Then, $z_0 = 0$ has to hold as $\Sigma_0 z_0 \geq 0$ is satisfied for every definite σ_0. This implies directly $(x, 0) \in \mathcal{F}^s_t \times \{0\}$.
To prove the second inclusion, consider $(x, z_0) \in \mathcal{F}^s$. Then, $(x, z_0) \in \mathcal{F}^s_{\tilde{\Sigma}} \subseteq \bigcup_{\sigma \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}^s_{\Sigma_0}$ for the choice $(\tilde{\sigma}_0)_i = 1$ for $i \in \alpha_0(x)$ and $(\tilde{\sigma}_0)_i = (\sigma_0(x))_i$ otherwise. ☐

Once again, the smooth branch problems can be used to derive sufficient conditions.

Lemma 3.51. *The point $(x^*, 0, z_+^*)$ is a local minimizer of (sNLP) if and only if it is a local minimizer of the split branch problem (3.14) for every definite $\Sigma_0 = \operatorname{diag}(\sigma_0)$.*

Proof. This is proven like Lemma 3.15 except for the split feasible sets that occur: Assume that $(x^*, 0, z_+^*)$ is a local minimizer of (sNLP). Then, $(x^*, z_+^*) \in \mathcal{F}^s_t \times \{0\} \subseteq \mathcal{F}^s_{\Sigma_0}$ as well as $\mathcal{F}^s_{\Sigma_0} \subseteq \mathcal{F}^s$ hold for all σ_0 by Lemma 3.44 and (3.15). Thus, the pair $(x^*, 0, z_+^*)$ has to be a local minimizer of every branch problem (3.14).
Conversely, assume that $(x^*, 0, z_+^*)$ is a local minimizer of every branch problem (3.14). Then, $(x^*, 0, z_+^*) \in \mathcal{F}^s_{\Sigma_0} \subseteq \mathcal{F}^s$ for all σ_0 by (3.15) and $\bigcup_{\sigma_0 \in \{-1,1\}^{|\alpha^*|}} \mathcal{F}^s_{\Sigma_0} = \mathcal{F}^s$ by Lemma 3.50. Consequently, $(x^*, 0, z_+^*)$ has to be a local minimizer of (sNLP). ☐

LICQ at $(x^*, 0, z_+^*)$ for the split branch problems is full row rank of the matrix

$$J^s_{\Sigma_0}(x^*, z_+^*) = \begin{bmatrix} \partial_1 c_{\mathcal{E}}(x^*, 0, z_+^*) & \partial_2 c_{\mathcal{E}}(x^*, 0, z_+^*) & \partial_3 c_{\mathcal{E}}(x^*, 0, z_+^*) \\ \partial_1 c_0(x^*, 0, z_+^*) & \partial_2 c_0(x^*, 0, z_+^*) - \Sigma_0 & \partial_3 c_0(x^*, 0, z_+^*) \\ \partial_1 c_+(x^*, 0, z_+^*) & \partial_2 c_+(x^*, 0, z_+^*) & \partial_3 c_+(x^*, 0, z_+^*) - I \\ 0 & I & 0 \end{bmatrix}$$
$$\in \mathbb{R}^{(m+s+|\alpha^*|)\times(n+s)}.$$

As before, LIKQ at x^* ensures this property.

Lemma 3.52. *Assume that $(x^*, 0, z_+^*)$ is a feasible point for (sNLP). Then, LICQ holds at $(x^*, 0, z_+^*)$ for all branch problems if LIKQ holds at x^*.*

Proof. Block elimination directly gives that LICQ at $(x^*, 0, z_+^*)$ for the branch problems is implied by LICQ at (x^*, z_+^*) for the trunk problem. The latter is in turn implied by LIKQ at x^* as was shown in Lemma 3.45. ☐

Then, a basis for $J^s_{\Sigma_0}(x^*, z^*_+)$ is given by the columns of the matrix

$$U^s_{\Sigma_0}(x^*, z^*_+) = \Big[U_{\mathrm{eq}}(x^*)^T \quad 0 \quad (I - \partial_3 c_+)^{-1} \partial_1 c_+ U_{\mathrm{eq}}(x^*)^T \Big]^T \in \mathbb{R}^{(n+|\sigma^*|) \times [n-(m+|\alpha^*|)]},$$

where the partial derivatives of c_+ are evaluated at $(x^*, 0, z^*_+)$. This follows directly from the form of U^s_t exploring the identity block in $J^s_{\Sigma_0}$.

To state optimality conditions, the Lagrangian is needed. It is denoted by

$$\begin{aligned} \mathcal{L}^s_{\Sigma_0}(x, \bar z_0, z_+, \lambda, \mu_0) := f(x) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x, \bar z_0, z_+) + \lambda_0^T [c_0(x, \bar z_0, z_+) - \Sigma_0 \bar z_0] \\ + \lambda_+^T [c_+(x, \bar z_0, z_+) - z_+] - \mu_0^T \bar z_0 \end{aligned}$$

with Lagrange multiplier vectors $\lambda = (\lambda_{\mathcal{E}}, \lambda_0, \lambda_+) \in \mathbb{R}^{m+|\alpha^*|+|\sigma^*|}$ and $\mu_0 \in \mathbb{R}^{|\alpha^*|}$.

Theorem 3.53 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0, z^*_+)$ is a local minimizer of (sNLP) and that LIKQ holds at x^*. Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_0, \lambda^*_+)$ such that the following conditions are satisfied for every definite $\Sigma_0 = \operatorname{diag}(\sigma_0)$:*

$$\begin{aligned} f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 c_{\mathcal{E}} + (\lambda^*_0)^T \partial_1 c_0 + (\lambda^*_+)^T \partial_1 c_+ &= 0, \\ (\lambda^*_{\mathcal{E}})^T \partial_3 c_{\mathcal{E}} + (\lambda^*_0)^T \partial_3 c_0 + (\lambda^*_+)^T [\partial_3 c_+ - I] &= 0 \quad \textit{(tangential stationarity)}, \\ (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}} + (\lambda^*_0)^T [\partial_2 c_0 - \Sigma_0] + (\lambda^*_+)^T \partial_2 c_+ &\geq 0 \quad \textit{(normal growth)}, \\ c_{\mathcal{E}} &= 0 \quad \textit{(primal feasibility)}, \\ c_0 &= 0, \\ c_+ - z^*_+ &= 0 \quad \textit{(switching feasibility)}. \end{aligned}$$

Here, the constraints $c_{\mathcal{E}}, c_0, c_+$ and all their partial derivatives are evaluated at $(x^*, 0, z^*_+)$.

Proof. This follows as in the proof of Theorem 3.17: Every branch problem is smooth and thus Theorem 2.7 can be applied. This gives the existence of $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_0, \lambda^*_+)$ and μ^*_0 such that:

$$\begin{aligned} \partial_x \mathcal{L}^s_{\Sigma_0} = f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 c_{\mathcal{E}} + (\lambda^*_0)^T \partial_1 c_0 + (\lambda^*_+)^T \partial_1 c_+(x^*, 0, z_+) &= 0, \\ \partial_{z_+} \mathcal{L}^s_{\Sigma_0} = f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_3 c_{\mathcal{E}} + (\lambda^*_0)^T \partial_3 c_0 + (\lambda^*_+)^T \left[\partial_3 c_+ - I\right] + (\lambda^*_+)^T \partial_2 c_+ &= 0, \\ \partial_{\bar z_0} \mathcal{L}^s_{\Sigma_0} = (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}} + (\lambda^*_0)^T \left[\partial_2 c_0 - \Sigma_0\right] - \mu^*_0 &= 0, \\ c_{\mathcal{E}} &= 0, \\ c_0 &= 0, \\ c_+ - z^*_+ &= 0 \\ \mu^*_0 &\geq 0, \end{aligned}$$

where the Lagrangian is evaluated at $(x^*, 0, z^*_+, \lambda^*, \mu^*_0)$ and the constraints $c_{\mathcal{E}}, c_0, c_+$ and all their partial derivatives are evaluated at $(x^*, 0, z^*_+)$. The first, second, fourth, fifth and sixth condition are the first order necessary conditions for the trunk problem which were obtained in Theorem 3.47. Thus, the Lagrange multiplier vectors λ^* coincide for all branch problems and for the trunk problem as they are unique. Only the third and seventh condition (and μ_0^* itself) depend on Σ_0. They can be combined which gives the normal growth condition

$$(\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} + (\lambda_0^*)^T [\partial_2 c_0(x^*, 0) - \Sigma_0] + (\lambda_+^*)^T \partial_2 c_+ \geq 0,$$

for every definite $\Sigma_0 = \operatorname{diag}(\sigma_0)$. □

Even with the additional tangential stationarity condition it is sufficient to consider one normal growth condition instead of $2^{|\alpha^*|}$. Once again, it is one branch which satisfies $\Sigma_0 \lambda_0^* = |\lambda_0^*|$ as the following equivalence holds and gives the next lemma.

$$\begin{aligned} & (\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} + (\lambda_0^*)^T \partial_2 c_0 + (\lambda_+^*)^T \partial_2 c_+ \geq |\lambda_0^*|^T \\ \iff & (\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} + (\lambda_0^*)^T \partial_2 c_0 + (\lambda_+^*)^T \partial_2 c_+ \geq (\lambda_0^*)^T (\pm \Sigma_0) \text{ for every definite } \Sigma_0. \end{aligned}$$

Lemma 3.54. *The equivalence stated before directly implies:*

$$\begin{aligned} & (\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} + (\lambda_0^*)^T \partial_2 c_0 + (\lambda_+^*)^T \partial_2 c_+ \geq |\lambda_0^*|^T \\ \iff & (\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} + (\lambda_0^*)^T \left[\partial_2 c_0 - \Sigma_0\right] + (\lambda_+^*)^T \partial_2 c_+ \geq 0 \textit{ for every definite } \Sigma_0, \end{aligned}$$

where the constraints $c_\mathcal{E}, c_0, c_+$ and all their partial derivatives are evaluated at $(x^*, 0, z_+^*)$

Corollary 3.55 (First Order Necessary Conditions)**.** *Assume that $(x^*, 0, z_+^*)$ is a local minimizer of (sNLP) and that LIKQ holds at x^*. Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda_\mathcal{E}^*, \lambda_0^*, \lambda_+^*)$ such that the following conditions are satisfied, where the constraints $c_\mathcal{E}, c_0, c_+$ and all their partial derivatives are evaluated at $(x^*, 0, z_+^*)$:*

$$\begin{aligned} f'(x^*) + (\lambda_\mathcal{E}^*)^T \partial_1 c_\mathcal{E} + (\lambda_0^*)^T \partial_1 c_0 + (\lambda_+^*)^T \partial_1 c_+ &= 0, \\ (\lambda_\mathcal{E}^*)^T \partial_3 c_\mathcal{E} + (\lambda_0^*)^T \partial_3 c_0 + (\lambda_+^*)^T [\partial_3 c_+ - I] &= 0 && \textit{(tangential stationarity)}, \\ (\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} + (\lambda_0^*)^T \partial_2 c_0 + (\lambda_+^*)^T \partial_2 c_+ &\geq |\lambda_0^*|^T && \textit{(normal growth)}, \\ c_\mathcal{E} &= 0 && \textit{(primal feasibility)}, \\ c_0 &= 0, \\ c_+ - z_+^* &= 0 && \textit{(switching feasibility)}. \end{aligned} \tag{3.16}$$

Under the additional assumption of strict normal growth, second order sufficient conditions are obtained. Here, the more complicated nullspace matrix U_t^s is needed again.

Theorem 3.56 (Second Order Sufficient Conditions)**.** *Consider (sNLP) for $d \geq 2$. Assume that $(x^*, 0, z_+^*)$ is feasible and that LIKQ holds at x^*. Assume further that a Lagrange multiplier vector λ^* exists such that the first order necessary conditions (3.16) are satisfied with strict normal growth,*

$$(\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E}(x^*, 0, z_+^*) + (\lambda_0^*)^T \partial_2 c_0(x^*, 0, z_+^*) + (\lambda_+^*)^T \partial_2 c_+(x^*, 0, z_+^*) > |\lambda_0^*|^T,$$

and that

$$U_t^s(x^*, z_+^*)^T H_t^s(x^*, z_+^*, \lambda^*) U_t^s(x^*, z_+^*) > 0.$$

Then, $(x^*, 0, z_+^*)$ is a strict local minimizer of (sNLP).

Proof. This follows again the proof of the localized case (see Theorem 3.20): The strict normal growth condition implies that strict complementarity holds in every branch problem and LIKQ implies that the Lagrange multiplier vectors are unique. Thus, Theorem 2.11 can be applied if the reduced Hessian is positive definite for every branch problem:

$$U^s_{\Sigma_0}(x^*)^T H^s_{\Sigma_0}(x^*, z^*_+, \lambda^*) U^s_{\Sigma_0}(x^*) > 0.$$

The multiplier μ_0 appears only in linear terms in the Lagrangian. Thus, all Hessians coincide as they are independent of μ_0 and read

$$H^s_{\Sigma_0}(x^*, z^*_+, \lambda^*) = \begin{bmatrix} \partial_{xx}\mathcal{L}^s_{\Sigma_0} & \partial_{x\bar{z}}\mathcal{L}^s_{\Sigma_0} & \partial_{xz_+}\mathcal{L}^s_{\Sigma_0} \\ \partial_{\bar{z}x}\mathcal{L}^s_{\Sigma_0} & \partial_{\bar{z}\bar{z}}\mathcal{L}^s_{\Sigma_0} & \partial_{\bar{z}z_+}\mathcal{L}^s_{\Sigma_0} \\ \partial_{z_+x}\mathcal{L}^s_{\Sigma_0} & \partial_{z_+\bar{z}}\mathcal{L}^s_{\Sigma_0} & \partial_{z_+z_+}\mathcal{L}^s_{\Sigma_0} \end{bmatrix} \in \mathbb{R}^{(n+s)\times(n+s)}$$

where all derivatives of $\mathcal{L}^s_{\Sigma_0}$ are evaluated at $(x^*, 0, z^*_+, \lambda^*, \mu^*_0)$. Here, H^s_t is contained in $H^s_{\Sigma_0}$, in particular:

$$\begin{bmatrix} \partial_{xx}\mathcal{L}^s_{\Sigma_0} & \partial_{xz_+}\mathcal{L}^s_{\Sigma_0} \\ \partial_{z_+x}\mathcal{L}^s_{\Sigma_0} & \partial_{z_+z_+}\mathcal{L}^s_{\Sigma_0} \end{bmatrix} = \begin{bmatrix} \partial_{xx}\mathcal{L}^s_t & \partial_{xz_+}\mathcal{L}^s_t \\ \partial_{z_+x}\mathcal{L}^s_t & \partial_{z_+z_+}\mathcal{L}^s_t \end{bmatrix} = H^s_t(x^*, z^*_+, \lambda^*).$$

Inserting the particular form of $U^s_{\Sigma_0}(x^*, z^*_+)$ leads to

$$U^s_{\Sigma_0}(x^*, z^*_+)^T H^s_{\Sigma_0}(x^*, z^*_+, \lambda^*) U^s_{\Sigma_0}(x^*, z^*_+) = U^s_t(x^*, z^*_+)^T H^s_t(x^*, z^*_+, \lambda^*) U^s_t(x^*, z^*_+) > 0.$$

Thus, Theorem 2.11 gives the claim. □

Once again, the first order conditions in Theorem 3.53 are sufficient under some additional assumptions. The key is still the (strict) normal growth condition.

Theorem 3.57 (First Order Necessary and Sufficient Conditions for Linear Functions)**.** *Given* (sNLP)*, assume that f, $c_\mathcal{E}$, c_+ and c_0 are linear. Assume further that $(x^*, 0, z^*_+)$ is feasible and that LIKQ holds at x^*. Then, $(x^*, 0, z^*_+)$ is a local minimizer of* (sNLP) *if and only if there exists $\lambda^* = (\lambda^*_\mathcal{E}, \lambda^*_0, \lambda^*_+)$ such that the conditions* (3.16) *are satisfied.*

Proof. This is done as in the proof of Theorem 3.22: By the assumptions every branch problem is linear and $(x^*, 0, z^*_+)$ satisfies the first order necessary conditions with the identical set of Lagrange multiplier vectors. Then, $(x^*, 0)$ is a global minimizer in every smooth branch problem by Corollary 2.12 and $(x^*, 0)$ is a local minimizer of (sNLP) by Lemma 3.51. □

Again, even in the linear case LIKQ is needed to ensure that the Lagrange multiplier vectors of all branch problems coincide.

Corollary 3.58 (First Order Sufficient Conditions for $|\alpha^*| + m = n$)**.** *Given* (sNLP) *for $d \geq 2$, assume that $|\alpha^*| + m = n$ holds. Assume further that $(x^*, 0, z^*_+)$ is feasible and that LIKQ holds at x^*. Then, $(x^*, 0, z^*_+)$ is a strict local minimizer of* (sNLP) *if there exists $\lambda^* = (\lambda^*_\mathcal{E}, \lambda^*_0, \lambda^*_+)$ such that the conditions* (3.16) *are satisfied with strict normal growth*

$$(\lambda^*_\mathcal{E})^T \partial_2 c_\mathcal{E} + (\lambda^*_0)^T \partial_2 c_0 + (\lambda^*_+)^T \partial_2 c_+ > |\lambda^*_0|^T.$$

Proof. The matrix $U^s_t(x, z_+)$ is empty as $|\alpha^*| + m = n$ and the result follows directly from Theorem 3.56. □

3.1.4.3 Combined Optimality Conditions

In the two preceding sections optimality conditions for the non-localized abs-normal NLP in split form and in reduced form were derived. Now a suitable combination of these conditions is selected and more compactly stated. Before, the connection between the Lagrange multiplier vectors in both formulations is formulated and proven.

Lemma 3.59. *A point $(x^*, 0, \lambda^*_{\mathcal{E}}, \lambda^*_0)$ satisfies the first order necessary conditions of Corollary 3.39 for (rNLP) if and only if the point $(x^*, 0, z^*_+, \lambda^*_{\mathcal{E}}, \lambda^*_0, \lambda^*_+)$ with $z^*_+ = z_+(x^*, 0)$ and $\lambda^*_+ = [(\lambda^*_{\mathcal{E}})^T \partial_3 c_{\mathcal{E}}(x^*, 0, z_+) + (\lambda^*_0)^T \partial_3 c_+(x^*, 0, z_+)][I - \partial_3 c_+(x^*, 0, z_+)]^{-T}$ satisfies the first order necessary conditions of Corollary 3.55 for (sNLP).*

Proof. Due to construction $z_+(x^*) = z^*_+ = z_+(x^*, 0)$ holds and the second tangential stationarity condition in Corollary 3.55 gives the relation

$$(\lambda^*_+)^T = [(\lambda^*_{\mathcal{E}})^T \partial_3 c_{\mathcal{E}}(x^*, 0, z_+) + (\lambda^*_0)^T \partial_3 c_+(x^*, 0, z_+)][I - \partial_3 c_+(x^*, 0, z_+)]^{-1}.$$

Using $z_+(x^*) = z^*_+ = z_+(x^*, 0)$ and the definition of $z^*_+(x^*, 0)$ the primal feasibilities and the switching feasibilities can be converted from one formulation into the other. Thus, it is left to show equivalence between the tangential stationarity and normal growth conditions. For this, note that by the definitions of z^*_+ and λ^*_+ the following relations are satisfied where all partial derivatives are evaluated at $(x^*, 0, z_+(x^*, 0))$:

$$\begin{aligned}
(\lambda^*_+)^T \partial_1 c_+ =& (\lambda^*_{\mathcal{E}})^T \partial_3 c_{\mathcal{E}} \partial_x z_+(x^*, 0) + (\lambda^*_0)^T \partial_3 c_0 \partial_x z_+(x^*, 0), \\
(\lambda^*_+)^T \partial_2 c_+ =& (\lambda^*_{\mathcal{E}})^T \partial_3 c_{\mathcal{E}} \partial_{\bar{z}_0} z_+(x^*, 0) + (\lambda^*_0)^T \partial_3 c_0 \partial_{\bar{z}_0} z_+(x^*, 0).
\end{aligned}$$

Then, it is immediately clear that the conditions are equivalent. □

Lemma 3.60. *A point $(x^*, 0, \lambda^*_{\mathcal{E}}, \lambda^*_0)$ satisfies the strict normal growth condition for (rNLP) if and only if the point $(x^*, 0, z^*_+, \lambda^*_{\mathcal{E}}, \lambda^*_0, \lambda^*_+)$ with $z^*_+ = z_+(x^*, 0)$ and $\lambda^*_+ = [(\lambda^*_{\mathcal{E}})^T \partial_3 c_{\mathcal{E}}(x^*, 0, z_+) + (\lambda^*_0)^T \partial_3 c_+(x^*, 0, z_+)][I - \partial_3 c_+(x^*, 0, z_+)]^{-T}$ satisfies the strict normal growth condition for (sNLP).*

Proof. The equivalence in the previous proof holds also with $>$ instead of $\geq$ in the normal growth condition. □

Thus, $\lambda^*_{\mathcal{E}}$ and λ^*_0 coincide in both formulations and λ^*_+ is a linear combination of them.

In [17] the optimality conditions were combined following Griewank and Walther in [8] (which is stated in paragraph 2.2.4). Thus, the first order conditions of the split abs-normal NLP were chosen as the more convenient ones. The reason was that the first order necessary conditions of the reduced form involve complicated total derivatives due to the implicit function whereas the corresponding conditions of the split form involve only partial derivatives. In contrast, for second order necessary and sufficient conditions the ones of the reduced abs-normal NLP were selected. Here, it was argued that the second order conditions in split form involve the complicated nullspace matrix $U^s_t(x^*, z^*_+)$ whereas the corresponding conditions in reduced form involve the basic nullspace matrix $U_{\text{eq}}(x^*)$ of $J_{\text{eq}}(x^*)$.

In contrast to [17], in this thesis all conditions of the split abs-normal NLP are chosen as the convenient ones. Due to the explicit handling of z_+ they can be compared directly to the optimality conditions for MPECs in the following two chapters. For that, they are stated again in compact form which means without split variables.

Thus, Corollary 3.55 is rewritten in the following.

Corollary 3.61 (First Order Necessary Conditions)**.** *Assume that (x^*, z^*) is a local minimizer of (eqNLP) and that LIKQ holds at x^*. Then, there exists a unique Lagrange multiplier vector $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{Z}})$ such that the following conditions are satisfied:*

$$
\begin{aligned}
f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 c_{\mathcal{E}}(x^*, |z^*|) + (\lambda^*_{\mathcal{Z}})^T \partial_1 c_{\mathcal{Z}}(x^*, |z^*|) &= 0, \\
[(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, |z^*|) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, |z^*|)]_i &\geq |(\lambda^*_{\mathcal{Z}})_i|, \quad i \in \alpha(x^*), \\
[(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, |z^*|) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, |z^*|)]_i &= (\lambda^*_{\mathcal{Z}})_i \sigma^*_i, \quad i \notin \alpha(x^*), \\
c_{\mathcal{E}}(x^*, |z^*|) &= 0, \\
c_{\mathcal{Z}}(x^*, |z^*|) - z^* &= 0.
\end{aligned} \tag{3.17}
$$

Proof. Recall that

$$
z^* = \Pi \begin{bmatrix} z^*_0 \\ \Sigma^*_+ z^*_+ \end{bmatrix} = \Pi \begin{bmatrix} 0 \\ \Sigma^*_+ z^*_+ \end{bmatrix} \quad \text{and} \quad \Sigma^* = \Pi \begin{bmatrix} \Sigma^*_0 & 0 \\ 0 & \Sigma^*_+ \end{bmatrix} \Pi^T = \Pi \begin{bmatrix} 0 & 0 \\ 0 & \Sigma^*_+ \end{bmatrix} \Pi^T
$$

with an appropiate permutation matrix Π as well as

$$
c_{\mathcal{Z}}(x^*, \Sigma z^*) = \Pi \begin{bmatrix} c_0(x, 0, z^*_+) \\ \Sigma^*_+ c_+(x, 0, z^*_+) \end{bmatrix} \quad \text{and} \quad c_{\mathcal{E}}(x^*, \Sigma z^*) = \tilde{c}_{\mathcal{E}}(x, 0, z^*_+)
$$

with the explicit notation $\tilde{c}_{\mathcal{E}}$. Further, the following equalities for partial derivatives hold:

$$
\begin{aligned}
\partial_1 c_{\mathcal{Z}}(x^*, \Sigma^* z^*) &= \Pi \begin{bmatrix} \partial_1 c_0 \\ \Sigma^*_+ \partial_1 c_+ \end{bmatrix}, &\quad \partial_2 c_{\mathcal{Z}}(x^*, \Sigma^* z^*) &= \Pi \begin{bmatrix} \partial_2 c_0 & \partial_3 c_0 \\ \Sigma^*_+ \partial_2 c_+ & \Sigma^*_+ \partial_3 c_+ \end{bmatrix} \Pi^T, \\
\partial_1 c_{\mathcal{E}}(x^*, \Sigma^* z^*) &= \partial_1 \tilde{c}_{\mathcal{E}}, &\quad \partial_2 c_{\mathcal{E}}(x^*, \Sigma^* z^*) &= \begin{bmatrix} \partial_2 \tilde{c}_{\mathcal{E}} & \partial_3 \tilde{c}_{\mathcal{E}} \end{bmatrix} \Pi^T.
\end{aligned}
$$

Note that the argument $(x^*, 0, z^*_+)$ is omitted on all partial derivatives of $c_0, c_+, \tilde{c}_{\mathcal{E}}$. Then, the first condition of (3.16) with $\lambda^*_{\mathcal{Z}} = \Pi \ [(\lambda^*_0)^T \ (\Sigma^*_+ \lambda^*_+)^T]^T$ reads

$$
\begin{aligned}
0 &= f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 \tilde{c}_{\mathcal{E}} + \begin{bmatrix} (\lambda^*_0)^T & (\lambda^*_+)^T \end{bmatrix} \begin{bmatrix} \partial_1 c_0 \\ \partial_1 c_+ \end{bmatrix} \\
&= f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 \tilde{c}_{\mathcal{E}} + \begin{bmatrix} (\lambda^*_0)^T & (\lambda^*_+)^T \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \Sigma^*_+ \end{bmatrix} \Pi^T \Pi \begin{bmatrix} I & 0 \\ 0 & \Sigma^*_+ \end{bmatrix} \begin{bmatrix} \partial_1 c_0 \\ \partial_1 c_+ \end{bmatrix} \\
&= f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 c_{\mathcal{E}}(x^*, |z^*|) + \lambda^T_{\mathcal{Z}} \partial_1 c_{\mathcal{Z}}(x^*, |z^*|).
\end{aligned}
$$

Here, the special identities $\Pi^T \Pi = I$ and $\Sigma^*_+ \Sigma^*_+ = I$ were used. Analogously, the relation

$$
\begin{aligned}
&(\lambda^*_{\mathcal{E}})^T \begin{bmatrix} \partial_2 \tilde{c}_{\mathcal{E}} & \partial_3 \tilde{c}_{\mathcal{E}} \end{bmatrix} + \begin{bmatrix} (\lambda^*_0)^T & (\lambda^*_+)^T \end{bmatrix} \begin{bmatrix} \partial_2 c_0 & \partial_3 c_0 \\ \partial_2 c_+ & \partial_3 c_+ \end{bmatrix} \\
&= (\lambda^*_{\mathcal{E}})^T \begin{bmatrix} \partial_2 \tilde{c}_{\mathcal{E}} & \partial_3 \tilde{c}_{\mathcal{E}} \end{bmatrix} + \begin{bmatrix} (\lambda^*_0)^T & (\lambda^*_+)^T \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \Sigma^*_+ \end{bmatrix} \Pi^T \Pi \begin{bmatrix} I & 0 \\ 0 & \Sigma^*_+ \end{bmatrix} \begin{bmatrix} \partial_2 c_0 & \partial_3 c_0 \\ \Sigma^*_+ \partial_2 c_+ & \Sigma^*_+ \partial_3 c_+ \end{bmatrix} \\
&= (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, |z^*|) \Pi + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, |z^*|) \Pi
\end{aligned}
$$

follows. Inserting this into the conditions (3.16) and using that λ_0^* belongs to the active components and λ_+^* to the inactive ones, gives

$$\begin{aligned}
&[(\lambda_{\mathcal{E}}^*)^T \partial_2 c_{\mathcal{E}}(x^*, |z^*|) + (\lambda_{\mathcal{Z}}^*)^T \partial_2 c_{\mathcal{Z}}(x^*, |z^*|)]_i \geq |(\lambda_{\mathcal{Z}}^*)_i|, \quad i \in \alpha(x^*),\\
&[(\lambda_{\mathcal{E}}^*)^T \partial_2 c_{\mathcal{E}}(x^*, |z^*|) + (\lambda_{\mathcal{Z}}^*)^T \partial_2 c_{\mathcal{Z}}(x^*, |z^*|)]_i = (\lambda_{\mathcal{Z}}^*)_i \sigma_i^*, \quad i \notin \alpha(x^*).
\end{aligned}$$

□

To state second order conditions the Lagrangian $\mathcal{L}_t^s$ and the matrix U_t^s without split variables are needed. Setting $\lambda_{\mathcal{Z}} := \Pi\ [\lambda_0^T\ (\Sigma_+^* \lambda_+)^T]^T$, the Lagrangian $\mathcal{L}_t^s$ can be rewritten as

$$\tilde{\mathcal{L}}_t^s(x, |z|, \lambda) := f(x) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x, |z|) + \lambda_{\mathcal{Z}}^T \left[c_{\mathcal{Z}}(x, |z|) - \mathcal{P}_{(\alpha^*)^c}^T \mathcal{P}_{(\alpha^*)^c} \Sigma^* |z|\right]$$

with $\lambda = (\lambda_{\mathcal{E}}, \lambda_{\mathcal{Z}}) \in \mathbb{R}^{m+s}$. Here, $\mathcal{P}_{(\alpha^*)^c} \in \mathbb{R}^{|\sigma^*| \times s}$ denotes the projector onto the inactive components. Due to the form of $\partial_x z(x)$ which was derived in the proof of Lemma 3.30, the matrix $U_t^s(x^*)$ reads

$$\tilde{U}_t^s(x^*) := \begin{bmatrix} U_{\text{eq}}(x^*) \\ [e_i^T \Sigma^* \partial_x z(x^*)]_{i \notin \alpha^*} U_{\text{eq}}(x^*) \end{bmatrix}.$$

Thus, the second order conditions in Theorem 3.48 and in Theorem 3.56 can be stated without split variables.

Corollary 3.62 (Second Order Necessary Condition)**.** *Consider (eqNLP) for $d \geq 2$. Assume that (x^*, z^*) is a local minimizer and that LIKQ holds at x^*. Denote by λ^* the unique Lagrange multiplier. Then,*

$$\tilde{U}_t^s(x^*)^T \tilde{H}_t^s(x^*, z^*, \lambda^*) \tilde{U}_t^s(x^*) \geq 0.$$

with

$$\tilde{H}_t^s(x^*, z^*, \lambda^*) := \begin{bmatrix} I & 0 \\ 0 & P_{(\alpha^*)^c} \end{bmatrix} \begin{bmatrix} \partial_{xx} \tilde{\mathcal{L}}_t^s(x^*, |z^*|, \lambda^*) & \partial_{x|z|} \tilde{\mathcal{L}}_t^s(x^*, |z^*|, \lambda^*) \\ \partial_{|z|x} \tilde{\mathcal{L}}_t^s(x^*, |z^*|, \lambda^*) & \partial_{|z||z|} \tilde{\mathcal{L}}_t^s(x^*, |z^*|, \lambda^*) \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & P_{(\alpha^*)^c}^T \end{bmatrix}$$

and where $P_{(\alpha^*)^c} \in \mathbb{R}^{|\sigma^*| \times s}$ denotes the projector onto the inactive components.

Proof. By construction, $\tilde{\mathcal{L}}_t^s(x^*, |z^*|, \lambda_{\mathcal{E}}^*, \lambda_{\mathcal{Z}}^*) = \mathcal{L}_t^s(x^*, z_+^*, \lambda_{\mathcal{E}}^*, \lambda_0^*, \lambda_+^*)$ is satisfied for $|z^*| = \Pi[0^T\ (z_+^*)^T]^T$ and $\lambda_{\mathcal{Z}}^* = \Pi\ [(\lambda_0^*)^T\ (\Sigma_+^* \lambda_+^+)^T]^T$. As z_+^* contains the inactive components the following hold:

$$\tilde{H}_t^s = \begin{bmatrix} I & 0 \\ 0 & P_{(\alpha^*)^c} \end{bmatrix} \begin{bmatrix} \partial_{xx} \tilde{\mathcal{L}}_t^s & \partial_{xz} \tilde{\mathcal{L}}_t^s \\ \partial_{zx} \tilde{\mathcal{L}}_t^s & \partial_{zz} \tilde{\mathcal{L}}_t^s \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & P_{(\alpha^*)^c}^T \end{bmatrix} = \begin{bmatrix} \partial_{xz_+} \mathcal{L}_t^s & \partial_{xz_+} \mathcal{L}_t^s \\ \partial_{z_+x} \mathcal{L}_t^s & \partial_{z_+z_+} \mathcal{L}_t^s \end{bmatrix} = H_t^s$$

Here, $\tilde{H}_t^s$ and $\tilde{\mathcal{L}}_t^s$ are evaluated at $(x^*, |z^*|, \lambda_{\mathcal{E}}^*, \lambda_{\mathcal{Z}}^*)$ and H_t^s and $\mathcal{L}_t^s$ at $(x^*, z_+^*, \lambda_{\mathcal{E}}^*, \lambda_0^*, \lambda_+^*)$. Then, the result follows directly from Theorem 3.48. □

Corollary 3.63 (Second Order Sufficient Condition)**.** *Consider (eqNLP) for $d \geq 2$. Assume that (x^*, z^*) is feasible and that LIKQ holds at x^*. Assume further that a Lagrange multiplier* *vector λ^* exists such that the first order necessary conditions (3.17) are satisfied with strict normal growth,*

$$[(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, |z^*|) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, |z^*|)]_i > |(\lambda^*_{\mathcal{Z}})_i|, \quad i \in \alpha(x^*),$$

and that

$$\tilde{U}^s_t(x^*)^T \tilde{H}^s_t(x^*, z^*, \lambda^*) \tilde{U}^s_t(x^*) > 0.$$

Then, (x^*, z^*) is a strict local minimizer of (eqNLP).

Proof. This follows directly by inserting the results of Corollary 3.61 and Corollary 3.62 in Theorem 3.56. □

The first order sufficient conditions will not be repeated here without split notation as they are not needed in the following. Moreover, it is directly clear from Corollary 3.61 how they read without the split variables.

3.2 Handling of Nonsmooth Objective Function and Inequality Constraints

In this section the general level-1 nonsmooth NLP is considered:

$$\begin{aligned} \min_{x \in D^x} \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \\ & h(x) \geq 0, \end{aligned}$$

where $D^x \subseteq \mathbb{R}^n$ is open, $f \in C^d_{\text{abs}}(D^x, \mathbb{R})$, $g \in C^d_{\text{abs}}(D^x, \mathbb{R}^{m_1})$ and $h \in C^d_{\text{abs}}(D^x, \mathbb{R}^{m_2})$ for $d \geq 1$. There are two possibilities to handle these problems. On one hand the existing theory can be extended. Then, the complete previous theory has to be adapted to a level-1 nonsmooth objective function and additional level-1 nonsmooth inequality constraints. On the other hand it can be reformulated as an equality-constrained abs-normal NLP:

$$\begin{aligned} \min_{x,v,w} \quad & v \\ \text{s.t.} \quad & f(x) - v = 0, \\ & g(x) = 0, \\ & h(x) - |w| = 0, \end{aligned}$$

with $v \in \mathbb{R}$, slack variables $w \in \mathbb{R}^{m_2}$, $f(x) - v \in C^d_{\text{abs}}(D^x \times \mathbb{R}^{m_2}, \mathbb{R})$ and $h(x) - |w| \in C^d_{\text{abs}}(D^x \times \mathbb{R}^{m_2}, \mathbb{R}^{m_2})$. Note that, the conversion of inequalities $h(x) \geq 0$ to equalities $h(x) - |w| = 0$ has been suggested by Griewank in a personal discussion [12]. Then, the previous results can directly be applied to this formulation. Nevertheless, the question occurs if one formulation provides better regularity than the other.

Here, level-1 nonsmooth inequalities and a level-1 nonsmooth objective function are considered separately. Thus, in subsection 3.2.1 level-1 nonsmooth inequalities are considered and the LIKQ is formulated for the direct approach and for the reformulation as an equality-constrained abs-normal NLP. Then, equivalence is proven between LIKQ for both formulations and thus optimality conditions can be obtained using the reformulation and the results of the previous section. Further, a level-1 nonsmooth objective function is studied in subsection 3.2.2 via both approaches.

3.2.1 Handling of Level-1 Nonsmooth Inequality Constraints

To begin with, only additional level-1 nonsmooth inequality constraints are considered in the following. Then, the problem reads:

$$\begin{aligned} \min_{x \in D^x} \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \\ & h(x) \geq 0, \end{aligned} \tag{NLP}$$

with $f \in C^d(D^x, \mathbb{R})$, $g \in C^d_{\text{abs}}(D^x, \mathbb{R}^{m_1})$ and $h \in C^d_{\text{abs}}(D^x, \mathbb{R}^{m_2})$ for $d \geq 1$.

Direct Handling In this paragraph abs-normal NLPs with equality and inequality constraints are considered. They are obtained by substituting the constraints representation in abs-normal form (2.3b) into the general nonsmooth problem (NLP). Note that the variables $t \in \mathbb{R}^{n_t}$ and $z^t \in \mathbb{R}^{s_t}$ instead of $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^s$ are used. Thus, also $\sigma^t(t)$ and $\alpha^t(t)$ is written instead of $\sigma(x)$ and $\alpha(x)$.

Definition 3.64 (Abs-Normal NLP)**.** A nonsmooth optimization problem is called an *abs-normal NLP* if functions $f \in C^d(D^t, \mathbb{R})$, $c_\mathcal{E} \in C^d(D^{t,|z^t|}, \mathbb{R}^{m_1})$, $c_\mathcal{I} \in C^d(D^{t,|z^t|}, \mathbb{R}^{m_2})$ and $c_\mathcal{Z} \in C^d(D^{t,|z^t|}, \mathbb{R}^{s_t})$ for $d \geq 1$ exist such that the NLP can equivalently be stated as

$$\begin{aligned} \min_{(t,z^t) \in D^{t,|z^t|}} \quad & f(t) \\ \text{s.t.} \quad & c_\mathcal{E}(t, |z^t|) = 0, \\ & c_\mathcal{I}(t, |z^t|) \geq 0, \\ & c_\mathcal{Z}(t, |z^t|) - z^t = 0, \end{aligned} \tag{I-NLP}$$

where $D^{|z^t|}$ is symmetric and $\partial_2 c_\mathcal{Z}(t, |z^t|)$ is strictly lower triangular. The feasible set of (I-NLP) is denoted by

$$\mathcal{F}_{\text{i-abs}} := \{(t, z^t) \in D^{t,|z^t|} \colon c_\mathcal{E}(t, |z^t|) = 0,\ c_\mathcal{I}(t, |z^t|) \geq 0,\ c_\mathcal{Z}(t, |z^t|) - z^t = 0\}.$$

The feasible set of (I-NLP) can be rewritten using $z^t(t)$:

$$\mathcal{F}_{\text{i-abs}} = \{(t, z^t(t)) : t \in D^t,\ c_\mathcal{E}(t, |z^t(t)|) = 0,\ c_\mathcal{I}(t, |z^t(t)|) \geq 0\}.$$

As in the smooth case, the set of active inequality constraints is defined. Once again, equalities are not counted as active constraints in contrast to standard NLP theory.

Definition 3.65 (Active Inequality Set)**.** Given (I-NLP), consider $(t, z^t(t)) \in \mathcal{F}_{\text{i-abs}}$. The inequality constraint $i \in \mathcal{I}$ is called *active* if $c_i(t, |z^t(t)|) = 0$ holds and *inactive* otherwise. The *active set* $\mathcal{A}(t)$ consists of all indices of active constraints,

$$\mathcal{A}(t) := \{i \in \mathcal{I} \colon c_i(t, |z^t(t)|) = 0\}.$$

The number of active inequality constraints is denoted by $|\mathcal{A}(t)|$.

To define the linear independence kink qualification as well as the interior direction kink qualification for (I-NLP) its Jacobian is needed.

Definition 3.66 (Active Jacobian)**.** Given (I-NLP), consider $(t, z^t(t)) \in \mathcal{F}_{\text{i-abs}}$ and set $\mathcal{A} = \mathcal{A}(t)$, $\alpha^t = \alpha^t(t)$, $\sigma^t = \sigma^t(t)$, $\Sigma^t = \operatorname{diag}(\sigma^t)$ and $c_{\mathcal{A}} = [c_i]_{i\in\mathcal{A}}$. The active Jacobian is

$$J_{\text{i-abs}}(t) := \begin{bmatrix} J_{\mathcal{E}}(t) \\ J_{\mathcal{A}}(t) \\ J_{\alpha^t}(t) \end{bmatrix} \in \mathbb{R}^{(m_1+|\mathcal{A}|+|\alpha^t|)\times n_t}.$$

It consists of the equality-constraints Jacobian (see Definition 3.3)

$$J_{\mathcal{E}}(t) = \partial_t c_{\mathcal{E}}(t, \Sigma^t z^t(t)) = \partial_1 c_{\mathcal{E}}(t, |z^t(t)|) + \partial_2 c_{\mathcal{E}}(t, |z^t(t)|)\Sigma^t \partial_t z^t(t),$$

the *active inequality Jacobian*

$$\begin{aligned} J_{\mathcal{A}}(t) := \partial_t c_{\mathcal{A}}(t, \Sigma^t z^t(t)) &= \partial_1 c_{\mathcal{A}}(t, \Sigma^t z^t(t)) + \partial_2 c_{\mathcal{A}}(t, \Sigma^t z^t(t))\Sigma^t \partial_t z^t(t) \\ &= \partial_1 c_{\mathcal{A}}(t, |z^t(t)|) + \partial_2 c_{\mathcal{A}}(t, |z^t(t)|)\Sigma^t \partial_t z^t(t) \end{aligned}$$

and the active switching Jacobian (see Definition 2.33)

$$J_{\alpha^t}(t) = \left[e_i^T \partial_t z^t(t)\right]_{i\in\alpha^t} = \left[e_i^T [I - \partial_2 c_{\mathcal{Z}}(t, |z^t(t)|)\Sigma^t]^{-1} \partial_1 c_{\mathcal{Z}}(t, |z^t(t)|)\right]_{i\in\alpha^t}.$$

Definition 3.67 (Linear Independence Kink Qualification (LIKQ))**.** Given (I-NLP), consider $(t, z^t(t)) \in \mathcal{F}_{\text{i-abs}}$. One says that the *linear independence kink qualification (LIKQ)* holds at t if the active Jacobian

$$J_{\text{i-abs}}(t) = \begin{bmatrix} J_{\mathcal{E}}(t) \\ J_{\mathcal{A}}(t) \\ J_{\alpha^t}(t) \end{bmatrix} = \begin{bmatrix} \partial_t c_{\mathcal{E}}(t, |z^t(t)|) \\ \partial_t c_{\mathcal{A}}(t, |z^t(t)|) \\ [e_i^T \partial_t z^t(t)]_{i\in\alpha^t} \end{bmatrix} \in \mathbb{R}^{(m_1+|\mathcal{A}|+|\alpha^t|)\times n_t}$$

has full row rank $m_1 + |\mathcal{A}| + |\alpha^t|$.

Inequality Slacks In this paragraph abs-normal NLPs with slack variables introduced for all inequalities are studied. Use of the absolute value of a slack variable is made, an idea proposed by Griewank [12]. This results in a class of purely equality-constrained abs-normal NLPs, which simplifies the derivation of optimality conditions under the LIKQ.

Using slack variables $w \in \mathbb{R}^{m_2}$, the following reformulation of (NLP) is obtained:

$$\begin{aligned} \min_{(t,w)\in D^{t,w}} \quad & f(t) \\ \text{s.t.} \quad & g(t) = 0, \\ & h(t) - |w| = 0, \end{aligned}$$

where $D^{t,w} = D^t \times \mathbb{R}^{m_2}$. Then, g and h can be expressed in abs-normal form as in (2.3b) and additional switching variables z^w can be introduced to handle $|w|$. This approach leads to the next definition.

Definition 3.68 (Abs-Normal NLP with Inequality Slacks)**.** An abs-normal NLP posed in the following form is called an *abs-normal NLP with inequality slacks*:

$$\begin{aligned} \min_{(t,w,z^t,z^w)\in D^{t,w,|z^t|,|z^w|}} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t,|z^t|) = 0, \\ & c_{\mathcal{I}}(t,|z^t|) - |z^w| = 0, \\ & c_{\mathcal{Z}}(t,|z^t|) - z^t = 0, \\ & w - z^w = 0, \end{aligned} \tag{E-NLP}$$

where $D^{t,w,|z^t|,|z^w|} = D^t \times \mathbb{R}^{m_2} \times D^{|z^t|} \times \mathbb{R}^{m_2}$. The feasible set of (E-NLP) is a lifting of $\mathcal{F}_{\text{i-abs}}$ and is denoted by

$$\begin{aligned} \mathcal{F}_{\text{e-abs}} &:= \left\{ (t,w,z^t,z^w) \in D^{t,w,|z^t|,|z^w|} \;\middle|\; \begin{array}{l} c_{\mathcal{E}}(t,|z^t|) = 0,\; c_{\mathcal{I}}(t,|z^t|) - |z^w| = 0, \\ c_{\mathcal{Z}}(t,|z^t|) - z^t = 0,\; w - z^w = 0 \end{array} \right\} \\ &= \left\{ (t,w,z^t,z^w) \colon (t,z^t) \in \mathcal{F}_{\text{i-abs}},\; w = z^w,\; |z^w| = c_{\mathcal{I}}(t,|z^t|) \right\}. \end{aligned}$$

Using the dependence of z^t and z^w of t and w, the feasible set can be written as

$$\begin{aligned} \mathcal{F}_{\text{e-abs}} &= \left\{ (t,w,z^t(t),z^w(w)) \colon (t,w) \in D^{t,w},\; c_{\mathcal{E}}(t,|z^t(t)|) = 0,\; c_{\mathcal{I}}(t,|z^t(t)|) - |z^w(w)| = 0 \right\} \\ &= \left\{ (t,w,z^t(t),z^w(w)) \colon (t,z^t(t)) \in \mathcal{F}_{\text{i-abs}},\; |z^w(w)| = c_{\mathcal{I}}(t,|z^t(t)|) \right\}. \end{aligned}$$

The signature and the active switching are split into components for variables t and w, i.e. $\sigma = (\sigma^t, \sigma^w)$ and $\alpha = (\alpha^t, \alpha^w)$. Analogously, the associated signature matrix is partitioned:

$$\Sigma = \begin{bmatrix} \Sigma^t & 0 \\ 0 & \Sigma^w \end{bmatrix}.$$

Remark 3.69. Note that introducing $|w|$ converts inequalities to pure equalities without a nonnegativity condition for the slack variables w. However, the slack reformulation has some subtle issues. In section 5.3 it will be shown that, in contrast to LIKQ, a weaker KQ (in particular IDKQ) is *not* preserved. Moreover, one cannot eliminate the equation $w - z^w = 0$ (and hence z^w or w) in (E-NLP) since this would destroy the abs-normal form. Finally, the slack w is not uniquely determined since the signs of nonzero components w_i can be chosen arbitrarily, yielding a set of $2^{m_2-|\alpha^w|}$ choices, $W(t) := \{w \colon |w| = c_{\mathcal{I}}(t,|z^t(t)|)\}$.

Lemma 3.70. *Given* (E-NLP)*, consider* $(t,w,z^t(t),z^w(w)) \in \mathcal{F}_{\textit{e-abs}}$*. Then, LIKQ at* (t,w) *is full row rank of*

$$J_{\textit{e-abs}}(t,w) := \begin{bmatrix} \partial_t c_{\mathcal{E}}(t,|z^t(t)|) & 0 \\ \partial_t c_{\mathcal{I}}(t,|z^t(t)|) & -\Sigma^w \\ [e_i^T \partial_t z^t(t)]_{i\in\alpha^t} & 0 \\ 0 & [e_i^T I]_{i\in\alpha^w} \end{bmatrix} \in \mathbb{R}^{(m_1+m_2+|\alpha^t|+|\alpha^w|)\times(n_t+m_2)}.$$

Proof. Set $x = (t, w)$, $z = (z^t, z^w)$, $\bar{f}(x) = f(t)$, $\bar{c}_{\mathcal{E}}(x, |z|) = (c_{\mathcal{E}}(t, |z^t|), c_{\mathcal{I}}(t, |z^t|) - |z^w|)$ and $\bar{c}_{\mathcal{Z}}(x, |z|) = (c_{\mathcal{Z}}(t, |z^t|), w)$. Then, (E-NLP) can be written compactly as

$$\begin{aligned} \min_{x,z} \quad & \bar{f}(x) \\ \text{s.t.} \quad & \bar{c}_{\mathcal{E}}(x, |z|) = 0, \\ & \bar{c}_{\mathcal{Z}}(x, |z|) - z = 0, \end{aligned} \tag{$\overline{\text{E-NLP}}$}$$

and $\bar{J}_{\mathcal{E}}$ and $\bar{J}_\alpha$ can be computed from Definition 3.66 using the special structure of (E-NLP). This leads to

$$\begin{aligned} \partial_x z(x) &= (I - \partial_2 \bar{c}_{\mathcal{Z}}(x, |z(x)|)\Sigma)^{-1} \partial_1 \bar{c}_{\mathcal{Z}}(x, |z(x)|) \\ &= \left(\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} \partial_2 \bar{c}_{\mathcal{Z}}(t, |z^t(t)|) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Sigma^t & 0 \\ 0 & \Sigma^w \end{bmatrix} \right)^{-1} \begin{bmatrix} \partial_1 \bar{c}_{\mathcal{Z}}(t, |z^t(t)|) & 0 \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} (I - \partial_2 \bar{c}_{\mathcal{Z}}(t, |z^t(t)|)\Sigma^t)^{-1} \partial_1 \bar{c}_{\mathcal{Z}}(t, |z^t(t)|) & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \partial_t z^t(t) & 0 \\ 0 & I \end{bmatrix}. \end{aligned}$$

Then, the active switching Jacobian reads

$$\bar{J}_\alpha(x) = \begin{bmatrix} [e_i^T \partial_t z^t(t)]_{i \in \alpha^t} & \\ & [e_i^T I]_{i \in \alpha^w} \end{bmatrix}$$

and the equality-constraints Jacobian

$$\begin{aligned} \bar{J}_{\mathcal{E}}(x) &= \partial_1 c_{\mathcal{E}}(x, |z(x)|) + \partial_2 c_{\mathcal{E}}(x, |z(x)|)\Sigma \partial_x z(x) \\ &= \begin{bmatrix} \partial_1 c_{\mathcal{E}}(t, |z^t(t)|) & 0 \\ \partial_1 c_{\mathcal{I}}(t, |z^t(t)|) & 0 \end{bmatrix} + \begin{bmatrix} \partial_2 c_{\mathcal{E}}(t, |z^t(t)|) & 0 \\ \partial_2 c_{\mathcal{I}}(t, |z^t(t)|) & -I \end{bmatrix} \begin{bmatrix} \Sigma^t & 0 \\ 0 & \Sigma^w \end{bmatrix} \begin{bmatrix} \partial_t z^t(t) & 0 \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} \partial_1 c_{\mathcal{E}}(t, |z^t(t)|) + \partial_2 c_{\mathcal{E}}(t, |z^t(t)|)\Sigma^t \partial_t z^t(t) & 0 \\ \partial_1 c_{\mathcal{I}}(t, |z^t(t)|) + \partial_2 c_{\mathcal{I}}(t, |z^t(t)|)\Sigma^t \partial_t z^t(t) & -\Sigma^w \end{bmatrix} \\ &= \begin{bmatrix} \partial_t c_{\mathcal{E}}(t, |z^t(t)|) & 0 \\ \partial_t c_{\mathcal{I}}(t, |z^t(t)|) & -\Sigma^w \end{bmatrix}. \end{aligned}$$

The resulting matrix $J_{\text{e-abs}}(x) = \begin{bmatrix} \bar{J}_{\mathcal{E}}(x)^T & \bar{J}_\alpha(x)^T \end{bmatrix}^T$ in Definition 3.67 has the form above. □

Remark 3.71. Clearly, the rank of $J_{\text{e-abs}}$ does not depend on the signs of ± 1 entries in Σ^w but only on their positions. Hence, LIKQ does not depend on the particular choice of w. Otherwise it would not make sense to consider (E-NLP).

Relation of LIKQ In this paragraph the relations of LIKQ for the two different formulations of abs-normal NLPs are discussed. Here, the set $W(t)$ from above is used.

Theorem 3.72. *LIKQ for (I-NLP) holds at $(t, z^t(t)) \in \mathcal{F}_{\textit{i-abs}}$ if and only if LIKQ for (E-NLP) holds at $(t, w, z^t(t), z^w(w)) \in \mathcal{F}_{\textit{e-abs}}$ for any (and hence all) $w \in W(t)$.*

Proof. This follows immediately by comparison of $J_{\text{i-abs}}$ and $J_{\text{e-abs}}$ using the relation

$$\alpha^w(w) = \{i \in \mathcal{I}\colon w_i = 0\} = \{i \in \mathcal{I}\colon c_i(t, z^t(t)) = 0\} = \mathcal{A}(t)$$

and the particular form of

$$\Sigma^w = \operatorname{diag}(\sigma^w) \quad \text{with} \quad \sigma_i^w = \operatorname{sign}(w_i) = \begin{cases} 0, & i \in \mathcal{A}(t), \\ \pm 1, & i \notin \mathcal{A}(t). \end{cases}$$

□

Optimality Conditions In this paragraph the optimality conditions for (I-NLP) are obtained. This is done using the slack formulation (E-NLP) as LIKQ is preserved under this reformulation.

Definition 3.73 (Kink Stationarity)**.** A feasible point $(t^*, (z^t)^*)$ of (I-NLP) is *kink stationary* if there exist a Lagrange multiplier vector $\lambda = (\lambda_\mathcal{E}, \lambda_\mathcal{I}, \lambda_\mathcal{Z}) \in \mathbb{R}^{m_1+m_2+s_t}$ such that the following conditions are satisfied:

$$f'(t^*) + \lambda_\mathcal{E}^T \partial_1 c_\mathcal{E} - \lambda_\mathcal{I}^T \partial_1 c_\mathcal{I} + \lambda_\mathcal{Z}^T \partial_1 c_\mathcal{Z} = 0, \tag{3.18a}$$

$$[\lambda_\mathcal{E}^T \partial_2 c_\mathcal{E} - \lambda_\mathcal{I}^T \partial_2 c_\mathcal{I} + \lambda_\mathcal{Z}^T \partial_2 c_\mathcal{Z}]_i \geq |(\lambda_\mathcal{Z})_i|, \quad i \in \alpha^t(t^*), \tag{3.18b}$$

$$[\lambda_\mathcal{E}^T \partial_2 c_\mathcal{E} - \lambda_\mathcal{I}^T \partial_2 c_\mathcal{I} + \lambda_\mathcal{Z}^T \partial_2 c_\mathcal{Z}]_i = (\lambda_\mathcal{Z})_i (\sigma^t)_i^*, \quad i \notin \alpha^t(t^*), \tag{3.18c}$$

$$\lambda_\mathcal{I} \geq 0, \tag{3.18d}$$

$$\lambda_\mathcal{I}^T c_\mathcal{I} = 0. \tag{3.18e}$$

Here, the constraints and the partial derivatives are evaluated at $(t^*, |(z^t)^*|)$.

Theorem 3.74 (First Order Necessary Conditions for (I-NLP))**.** *Assume that $(t^*, (z^t)^*)$ is a local minimizer of (I-NLP) and that LIKQ holds at t^*. Then, $(t^*, (z^t)^*)$ is a kink stationary point with unique Lagrange multiplier vector.*

Proof. By Lemma 3.72 the slack reformulation (E-NLP) can be considered. Then, Corollary 3.61 gives the following first order conditions for (E-NLP):

$$\begin{aligned} \bar{f}'(x^*) + \bar{\lambda}_\mathcal{E}^T \partial_1 \bar{c}_\mathcal{E}(x^*, |z^*|) + \bar{\lambda}_\mathcal{Z}^T \partial_1 \bar{c}_\mathcal{Z}(x^*, |z^*|) &= 0, \\ [\bar{\lambda}_\mathcal{E}^T \partial_2 \bar{c}_\mathcal{E}(x^*, |z^*|) + \bar{\lambda}_\mathcal{Z}^T \partial_2 \bar{c}_\mathcal{Z}(x^*, |z^*|)]_i &\geq |(\bar{\lambda}_\mathcal{Z})_i|, \quad i \in \alpha(x^*), \\ [\bar{\lambda}_\mathcal{E}^T \partial_2 \bar{c}_\mathcal{E}(x^*, |z^*|) + \bar{\lambda}_\mathcal{Z}^T \partial_2 \bar{c}_\mathcal{Z}(x^*, |z^*|)]_i &= (\bar{\lambda}_\mathcal{Z})_i \sigma_i^*, \quad i \notin \alpha(x^*). \end{aligned}$$

In the original notation of (E-NLP) with $\bar{\lambda}_\mathcal{E} = (\lambda_\mathcal{E}, -\lambda_\mathcal{I}) \in \mathbb{R}^{m_1+m_2}$ and $\bar{\lambda}_\mathcal{Z} = (\lambda_\mathcal{Z}, \lambda_\mathcal{Z}^w) \in \mathbb{R}^{s_t+m_2}$, where all derivatives are evaluated at $(t^*, |(z^t)^*|)$, these conditions read:

$$\begin{aligned} f'(t^*) + \lambda_\mathcal{E}^T \partial_1 c_\mathcal{E} - \lambda_\mathcal{I}^T \partial_1 c_\mathcal{I} + \lambda_\mathcal{Z}^T \partial_1 c_\mathcal{Z} &= 0, \\ (\lambda_\mathcal{Z}^w)^T &= 0, \\ [\lambda_\mathcal{E}^T \partial_2 c_\mathcal{E} - \lambda_\mathcal{I}^T \partial_2 c_\mathcal{I} + \lambda_\mathcal{Z}^T \partial_2 c_\mathcal{Z}]_i &\geq |(\lambda_\mathcal{Z})_i|, \quad i \in \alpha^t(t^*), \\ [\lambda_\mathcal{E}^T \partial_2 c_\mathcal{E} - \lambda_\mathcal{I}^T \partial_2 c_\mathcal{I} + \lambda_\mathcal{Z}^T \partial_2 c_\mathcal{Z}]_i &= (\lambda_\mathcal{Z})_i (\sigma^t)_i^*, \quad i \notin \alpha^t(t^*), \\ \lambda_\mathcal{I} &\geq |(\lambda_\mathcal{Z}^w)_i|, \quad i \in \alpha^w(w^*), \\ \lambda_\mathcal{I} &= (\lambda_\mathcal{Z}^w)_i (\sigma^w)_i^*, \quad i \notin \alpha^w(w^*), \end{aligned}$$

The claim follows by eliminating $\lambda_\mathcal{Z}^w = 0$ and noting that $\alpha^w(w^*) = \mathcal{A}(t^*)$. □

Next, second order conditions for (I-NLP) are formulated. To this end, the Lagrangian

$$\mathcal{L}_{\text{i-abs}}(t,|z^t|,\lambda) := f(t) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(t,|z^t|) - \lambda_{\mathcal{I}}^T c_{\mathcal{I}}(t,|z^t|) + \lambda_{\mathcal{Z}}^T \left[c_{\mathcal{Z}}(t,|z^t|) - \mathcal{P}_{(\alpha^t)^c}^T \mathcal{P}_{(\alpha^t)^c}\Sigma^t|z^t|\right]$$

with $\lambda = (\lambda_{\mathcal{E}}, \lambda_{\mathcal{I}}, \lambda_{\mathcal{Z}}) \in \mathbb{R}^{m_1+m_2+s}$ and the matrix

$$U_{\text{i-abs}}(t) := \begin{bmatrix} U(t) \\ [e_i^T \Sigma^t \partial_t z^t(t) U(t)]_{i\notin\alpha^t} \end{bmatrix} \tag{3.19}$$

are needed. Here, $U(t)$ denotes a matrix which columns are a basis for $J_{\text{i-abs}}$.

Theorem 3.75 (Second Order Necessary Conditions for (I-NLP))**.** *Consider* (I-NLP) *for $d \geq 2$. Assume that $(t^*,(z^t)^*)$ is a local minimizer and that LIKQ holds at t^*. Denote by λ^* the unique Lagrange multiplier and set $\alpha^t = \alpha^t(t^*)$. Then,*

$$U_{\textit{i-abs}}(t^*)^T H_{\textit{i-abs}}(t^*,(z^t)^*,\lambda^*) U_{\textit{i-abs}}(t^*) \geq 0$$

where $H_{\textit{i-abs}}(t^*,(z^t)^*,\lambda^*) := \begin{bmatrix} I & 0 \\ 0 & P_{(\alpha^t)^c} \end{bmatrix} \begin{bmatrix} \partial_{11}\mathcal{L}_{\textit{i-abs}} & \partial_{12}\mathcal{L}_{\textit{i-abs}} \\ \partial_{21}\mathcal{L}_{\textit{i-abs}} & \partial_{22}\mathcal{L}_{\textit{i-abs}} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & P_{(\alpha^t)^c}^T \end{bmatrix}$. *Here, $\mathcal{L}_{\textit{i-abs}}$ is evaluated at $(t^*,|(z^t)^*|,\lambda^*)$.*

Proof. As in Theorem 3.74, (E-NLP) instead of (I-NLP) can be considered by Theorem 3.72. In Corollary 3.62 the second order necessary conditions for (E-NLP) have been derived:

$$\tilde{U}_t^s(x^*)^T \tilde{H}_t^s(x^*,z^*,\bar{\lambda}^*) \tilde{U}_t^s(x^*) \geq 0$$

with $\bar{\lambda}^* = (\bar{\lambda}_{\mathcal{E}}^*, \bar{\lambda}_{\mathcal{Z}}^*)$. The Hessian reads

$$\tilde{H}_t^s(x^*,z^*,\bar{\lambda}^*) = \begin{bmatrix} I & 0 \\ 0 & P_{\alpha^c} \end{bmatrix} \begin{bmatrix} \partial_{11}\tilde{\mathcal{L}}_t^s(x^*,|z^*|,\bar{\lambda}^*) & \partial_{12}\tilde{\mathcal{L}}_t^s(x^*,|z^*|,\bar{\lambda}^*) \\ \partial_{21}\tilde{\mathcal{L}}_t^s(x^*,|z^*|,\bar{\lambda}^*) & \partial_{22}\tilde{\mathcal{L}}_t^s(x^*,|z^*|,\bar{\lambda}^*) \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & P_{\alpha^c}^T \end{bmatrix},$$

where α^c is the complement of α and the Lagrangian of (E-NLP) is

$$\tilde{\mathcal{L}}_t^s(x,|z|,\bar{\lambda}) = \bar{f}(x) + \bar{\lambda}_{\mathcal{E}}^T \bar{c}_{\mathcal{E}}(x,|z|) + \bar{\lambda}_{\mathcal{Z}}^T [\bar{c}_{\mathcal{Z}}(x,|z|) - \mathcal{P}_{\alpha^c}^T \mathcal{P}_{\alpha^c}\Sigma|z|].$$

Further, the matrix $\tilde{U}_t^s(x^*)$ is given by

$$\tilde{U}_t^s(x^*) = \begin{bmatrix} U_{\text{eq}}(x^*) \\ [e_i^T \Sigma^* \partial_x z(x^*) U_{\text{eq}}(x^*)]_{i\notin\alpha} \end{bmatrix},$$

where the columns of $U_{\text{eq}}(x^*)$ are a basis for $\ker(J_{\text{eq}}(x^*))$. Using the special structure of (E-NLP), the Lagrangian can be rewritten setting $\bar{\lambda}_{\mathcal{E}} := (\lambda_{\mathcal{E}}, -\lambda_{\mathcal{I}})$, $\bar{\lambda}_{\mathcal{Z}} := (\lambda_{\mathcal{Z}}, \lambda_{\mathcal{Z}}^w)$ as

$$\begin{aligned} \mathcal{L}_{\text{e-abs}}(t,w,|z^t|,|z^w|,\bar{\lambda}) := f(t) &+ \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(t,|z^t|) - \lambda_{\mathcal{I}}^T [c_{\mathcal{I}}(t,|z^t|) - |z^w|] \\ &+ \lambda_{\mathcal{Z}}^T [c_{\mathcal{Z}}(t,|z^t|) - \mathcal{P}_{(\alpha^t)^c}^T \mathcal{P}_{(\alpha^t)^c}\Sigma^t|z^t|] \\ &+ (\lambda_{\mathcal{Z}}^w)^T [w - \mathcal{P}_{(\alpha^w)^c}^T \mathcal{P}_{(\alpha^w)^c}\Sigma^w|z^w|]. \end{aligned}$$

Then, comparing the derivatives of $\mathcal{L}_{\text{e-abs}}(t^*, w^*, |(z^t)^*|, |(z^w)^*|, \bar{\lambda}^*)$ and $\mathcal{L}_{\text{i-abs}}(t^*, |(z^t)^*|, \lambda^*)$, the Hessian $\tilde{H}_t^s(x^*, z^*\bar{\lambda}^*)$ becomes:

$$H_{\text{e-abs}} := \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & P_{(\alpha^t)^c} & 0 \\ 0 & 0 & 0 & P_{(\alpha^w)^c} \end{bmatrix} \begin{bmatrix} \partial_{11}\mathcal{L}_{\text{i-abs}} & 0 & \partial_{12}\mathcal{L}_{\text{i-abs}} & 0 \\ 0 & 0 & 0 & 0 \\ \partial_{21}\mathcal{L}_{\text{i-abs}} & 0 & \partial_{22}\mathcal{L}_{\text{i-abs}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & P^T_{(\alpha^t)^c} & 0 \\ 0 & 0 & 0 & P^T_{(\alpha^w)^c} \end{bmatrix}.$$

Here, $H_{\text{e-abs}}$ is evaluated at $(t^*, w^*, (z^t)^*, (z^w)^*, \bar{\lambda}^*)$ and all partial derivatives of $\mathcal{L}_{\text{i-abs}}$ are evaluated at $(t^*, |(z^t)^*|, \lambda^*)$. Moreover, $J_{\text{eq}}(x^*)$ can be rewritten which gives $J_{\text{e-abs}}(t^*, w^*)$ from Lemma 3.70. Thus, a basis of its nullspace is given by the columns of

$$\bar{U}(t^*, w^*) := \begin{bmatrix} U(t^*) \\ (\Sigma^w)^* \partial_t c_{\mathcal{I}} \end{bmatrix},$$

where the columns of $U(t^*)$ are a basis for the nullspace of $J_{\text{i-abs}}(t^*)$ from Definition 3.67. Using this and $\partial_w z^w(w) = I$, the matrix $\tilde{U}_t^s(x^*)$ can be rewritten as

$$U_{\text{e-abs}}(t^*, w^*) := \begin{bmatrix} U(t^*) \\ (\Sigma^w)^* \partial_t c_{\mathcal{I}} \\ [e_i^T (\Sigma^t)^* \partial_t z^t(t^*) U(t^*)]_{i \notin \alpha^t} \\ [e_i^T \partial_t c_{\mathcal{I}}]_{i \notin \alpha^w} \end{bmatrix}.$$

Finally,

$$\begin{aligned} 0 &\leq \tilde{U}_t^s(x^*)^T \tilde{H}_t^s(x^*, z^*, \bar{\lambda}^*) \tilde{U}_t^s(x^*) \\ &\leq U_{\text{e-abs}}(t^*, w^*)^T H_{\text{e-abs}}(t^*, w^*, (z^t)^*, (z^w)^*, \bar{\lambda}^*) U_{\text{e-abs}}(t^*, w^*) \\ &= U_{\text{i-abs}}(t^*)^T H_{\text{i-abs}}(t^*, (z^t)^*, \lambda^*) U_{\text{i-abs}}(t^*) \end{aligned}$$

with $U_{\text{i-abs}}(t^*)$ from (3.19) and

$$H_{\text{i-abs}}(t^*, (z^t)^*, \lambda^*) = \begin{bmatrix} I & 0 \\ 0 & P_{(\alpha^t)^c} \end{bmatrix} \begin{bmatrix} \partial_{11}\mathcal{L}_{\text{i-abs}} & \partial_{12}\mathcal{L}_{\text{i-abs}} \\ \partial_{21}\mathcal{L}_{\text{i-abs}} & \partial_{22}\mathcal{L}_{\text{i-abs}} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & P^T_{(\alpha^t)^c} \end{bmatrix}$$

where $\mathcal{L}_{\text{i-abs}}$ is evaluated at $(t^*, |(z^t)^*|, \lambda^*)$ This proves the claim. □

Theorem 3.76 (Second Order Sufficient Conditions for (I-NLP))**.** *Consider (I-NLP) for $d \geq 2$. Assume that $(t^*, (z^t)^*)$ is kink stationary for (I-NLP) with a Lagrange multiplier vector $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{I}}, \lambda^*_{\mathcal{Z}})$ that satisfies strict complementarity for $\lambda^*_{\mathcal{I}}$ and strict normal growth,*

$$[\lambda_{\mathcal{E}}^T \partial_2 c_{\mathcal{E}} - \lambda_{\mathcal{I}}^T \partial_2 c_{\mathcal{I}} + \lambda_{\mathcal{Z}}^T \partial_2 c_{\mathcal{Z}}]_i > |(\lambda_{\mathcal{Z}})_i|, \quad i \in \alpha^t(t^*).$$

Assume further that LIKQ holds at t^* and that

$$U_{\textit{i-abs}}(t^*)^T H_{\textit{i-abs}}(t^*, (z^t)^*, \lambda^*) U_{\textit{i-abs}}(t^*) > 0.$$

Then, $(t^*, (z^t)^*)$ is a strict local minimizer of (I-NLP).

Proof. As before the slack reformulation (E-NLP) of (I-NLP) is considered. The assumption of strict complementarity for $\lambda_{\mathcal{I}}^*$ and strict normal growth for (I-NLP) implies strict normal growth for (E-NLP). Moreover, the previous proof shows that the condition

$$U_{\text{i-abs}}(t^*)^T H_{\text{i-abs}}(t^*, (z^t)^*, \lambda^*) U_{\text{i-abs}}(t^*) > 0$$

is equivalent to

$$U_{\text{e-abs}}(x^*)^T H_{\text{e-abs}}(t^*, (z^t)^*, \bar{\lambda}^*) U_{\text{e-abs}}(x^*) > 0.$$

Then, Corollary 3.63 can be applied, which gives the assertion. □

3.2.2 Handling of Nonsmooth Objective Function

Now, a level-1 nonsmooth objective function is studied in more detail. As it does not interfere with inequality constraints they can be neglected to simplify notation. Thus, the problem becomes:

$$\begin{aligned} \min_{x \in D^x} \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \end{aligned} \tag{fNLP}$$

with $f \in C^d_{\text{abs}}(D^x, \mathbb{R})$ and $g \in C^d_{\text{abs}}(D^x, \mathbb{R}^m)$ for $d \geq 1$.

Direct Handling In this paragraph the direct handling of the nonsmooth objective function is looked at. Thus, f and g are rewritten in abs-normal form (2.3b) using the variables t and z^t.

This leads to the next definition.

Definition 3.77 (F-Abs-Normal NLP)**.** A nonsmooth NLP is called an *f-abs-normal NLP* if functions $c_f \in C^d(D^{t,|z^t|}, \mathbb{R})$, $c_{\mathcal{E}} \in C^d(D^{t,|z^t|}, \mathbb{R}^{m_1})$ and $c_{\mathcal{Z}} \in C^d(D^{t,|z^t|}, \mathbb{R}^{s_t})$ for $d \geq 1$ exist such that the NLP can equivalently be stated as

$$\begin{aligned} \min_{(t,z^t) \in D^{t,|z^t|}} \quad & c_f(t, |z^t|) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t, |z^t|) = 0, \\ & c_{\mathcal{Z}}(t, |z^t|) - z^t = 0, \end{aligned} \tag{F-NLP}$$

where $D^{|z^t|}$ is symmetric and $\partial_2 c_{\mathcal{Z}}(x, |z^t|)$ is strictly lower triangular. The feasible set of (F-NLP) is denoted by

$$\mathcal{F}_{\text{f}} := \{(t, z^t) \in D^{t,|z^t|} \colon c_{\mathcal{E}}(t, |z^t|) = 0,\ c_{\mathcal{Z}}(t, |z^t|) - z^t = 0\}.$$

Using the dependence $z^t = z^t(t)$ the feasible set reads

$$\mathcal{F}_{\text{f}} = \{(t, z^t(t)) \colon t \in D^t,\ c_{\mathcal{E}}(t, |z^t(t)|) = 0\}.$$

Then, LIKQ for (F-NLP) is defined such that it matches to Definition 2.34 for (unNLP) and to Definition 3.67 for (eqNLP).

Definition 3.78 (LIKQ for (F-NLP))**.** Given (F-NLP), consider $(t, z^t(t)) \in \mathcal{F}_\mathrm{f}$. One says that the *linear independence kink qualification (LIKQ)* holds at t if the matrix

$$J_f(t) := \begin{bmatrix} \partial_t c_\mathcal{E}(t, |z^t(t)|) \\ [e_i^T \partial_t z^t(t)]_{i \in \alpha^t} \end{bmatrix} \in \mathbb{R}^{(m+|\alpha^t|)\times n_t}.$$

has full row rank $m + |\alpha^t|$.

Constant Objective In this paragraph the nonsmooth objective function is replaced by a constant c. Then, an additional equality constraint $f(t) - c = 0$ occurs. This leads to:

$$\begin{aligned} \min_{(t,c)\in D^t \times \mathbb{R}} \quad & c \\ \text{s.t.} \quad & f(t) - c = 0, \\ & g(t) = 0. \end{aligned}$$

As before f and g are expressed in abs-normal form.

This leads to the next definition.

Definition 3.79 (Abs-Normal NLP with Constant Objective)**.** An abs-normal NLP posed in the following form is called an *abs-normal NLP with constant objective*:

$$\begin{aligned} \min_{(t,c,z^t)\in D^{t,c,|z^t|}} \quad & c \\ \text{s.t.} \quad & c_f(t, |z^t|) - c = 0, \\ & c_\mathcal{E}(t, |z^t|) = 0, \\ & c_\mathcal{Z}(t, |z^t|) - z^t = 0, \end{aligned} \tag{C-NLP}$$

where $D^{t,c,|z^t|} = D^t \times \mathbb{R} \times D^{|z^t|}$. The feasible set of (C-NLP) is a lifting of $\mathcal{F}_\mathrm{f}$ and is denoted by

$$\begin{aligned} \mathcal{F}_\mathrm{c} :=& \{(t, c, z^t) \in D^{t,c,|z^t|} \colon c_f(t, |z^t|) - c = 0,\ c_\mathcal{E}(t, |z^t|) = 0,\ c_\mathcal{Z}(t, |z^t|) - z^t = 0\} \\ =& \{(t, c, z^t) \colon (t, z^t) \in \mathcal{F}_\mathrm{f},\ c = c_f(t, |z^t|)\}. \end{aligned}$$

Using the dependence of z^t on t it reads

$$\begin{aligned} \mathcal{F}_\mathrm{c} &= \{(t, c, z^t(t)) \colon (t, c) \in D^t \times \mathbb{R},\ c_f(t, |z^t(t)|) - c = 0,\ c_\mathcal{E}(t, |z^t(t)|) = 0\} \\ &= \{(t, c, z^t(t)) \colon (t, c) \in D^t \times \mathbb{R},\ c = c_f(t, |z^t(t)|)\}. \end{aligned}$$

Next, LIKQ can be formulated via Definition 3.67.

Lemma 3.80. *Given* (C-NLP)*, consider* $(t, c) \in \mathcal{F}_c$*. Then, LIKQ at* $(t, c, z^t(t))$ *is full row rank of*

$$J_c(t, c) := \begin{bmatrix} \partial_t c_f(t, |z^t(t)|) & -1 \\ \partial_t c_\mathcal{E}(t, |z^t(t)|) & 0 \\ [e_i^T \partial_t z^t(t)]_{i \in \alpha^t} & 0 \end{bmatrix} \in \mathbb{R}^{(n_t+m+|\alpha^t|)\times(n_t+1)}.$$

Proof. Set $x = (t, c)$, $z = z^t$, $\bar{f}(x) = c$, $\bar{c}_{\mathcal{E}}(x, |z|) = (c_f(t, |z^t|) - c, c_{\mathcal{E}}(t, |z^t|))$ and $\bar{c}_{\mathcal{Z}}(x, |z|) = c_{\mathcal{Z}}(t, |z^t|)$. Then, (C-NLP) can be stated as

$$\begin{aligned} \min_{x,z} \quad & f(x) \\ \text{s.t.} \quad & \bar{c}_{\mathcal{E}}(x, |z|) = 0, \\ & \bar{c}_{\mathcal{Z}}(x, |z|) - z = 0. \end{aligned} \tag{$\overline{\text{C-NLP}}$}$$

Using the special structure of (C-NLP), the active Jacobian can be computed from Definition 3.66. This gives

$$\begin{aligned} \partial_x z(x) &= (I - \partial_2 \bar{c}_{\mathcal{Z}}(x, |z(x)|)\Sigma)^{-1} \partial_1 \bar{c}_{\mathcal{Z}}(x, |z(x)|) \\ &= (I - \partial_2 c_{\mathcal{Z}}(t, |z^t(t)|)\Sigma^t)^{-1} \begin{bmatrix} \partial_1 c_{\mathcal{Z}}(t, |z^t(t)|) & 0 \end{bmatrix} \\ &= \begin{bmatrix} (I - \partial_2 c_{\mathcal{Z}}(t, |z^t(t)|)\Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}}(t, |z^t(t)|) & 0 \end{bmatrix} \\ &= \begin{bmatrix} \partial_t z^t(t) & 0 \end{bmatrix} \end{aligned}$$

and further

$$\bar{J}_\alpha(x) = \begin{bmatrix} [e_i^T \partial_t z^t(t)]_{i \in \alpha^t} & 0 \end{bmatrix}$$

as well as

$$\begin{aligned} \bar{J}_{\mathcal{E}}(x) &= \partial_1 c_{\mathcal{E}}(x, |z(x)|) + \partial_2 c_{\mathcal{E}}(x, |z(x)|)\Sigma \partial_x z(x) \\ &= \begin{bmatrix} \partial_1 c_f(t, |z^t(t)|) & -1 \\ \partial_1 c_{\mathcal{E}}(t, |z^t(t)|) & 0 \end{bmatrix} + \begin{bmatrix} \partial_2 c_f(t, |z^t(t)|) \\ \partial_2 c_{\mathcal{E}}(t, |z^t(t)|) \end{bmatrix} \Sigma^t \begin{bmatrix} \partial_t z^t(t) & 0 \end{bmatrix} \\ &= \begin{bmatrix} \partial_1 c_f(t, |z^t(t)|) + \partial_2 c_f(t, |z^t(t)|)\Sigma^t \partial_t z^t(t) & -1 \\ \partial_1 c_{\mathcal{E}}(t, |z^t(t)|) + \partial_2 c_{\mathcal{E}}(t, |z^t(t)|)\Sigma^t \partial_t z^t(t) & 0 \end{bmatrix} \\ &= \begin{bmatrix} \partial_t c_f(t, |z^t(t)|) & -1 \\ \partial_t c_{\mathcal{E}}(t, |z^t(t)|) & 0 \end{bmatrix}. \end{aligned}$$

Then, $J_c(x) = \begin{bmatrix} \bar{J}_{\mathcal{E}}(x)^T & \bar{J}_\alpha(x)^T \end{bmatrix}^T$ has the form above. □

Relation of LIKQ In this paragraph the relation of LIKQ for the two different approaches is given.

Theorem 3.81. *LIKQ for (F-NLP) holds at $(t, z^t(t)) \in \mathcal{F}_f$ if and only if LIKQ for (C-NLP) holds at $(t, c, z^t(t)) \in \mathcal{F}_c$ with $c = c_f(t, z^t(t))$.*

Proof. This follows immediately by block elimination. □

The optimality conditions can be obtained using the results of the previous section and the formulation as ($\overline{\text{C-NLP}}$). As this is done analogously to the optimality conditions for (I-NLP) and not needed in the following, it is omitted here.

Chapter 4

Relations between Abs-Normal NLPs and MPECs under LIKQ

This chapter compares the theory for abs-normal NLPs to the class of Mathematical Programs with Equilibrium Constraints (MPECs). First, in section 4.1 it is shown that both problem classes can be transformed into each other. This gives rise to the definition of so-called counterpart MPECs. As LIKQ is preserved under reformulation of nonsmooth objective functions only abs-normal NLPs with a smooth objective function are considered. Thus, counterpart MPECs for (I-NLP) and (E-NLP) are formulated in section 4.2. Then, relations for constraint qualifications of linear independence type as well as for optimality conditions are compared in section 4.3 and section 4.4. In both sections the definitions from section 2.3 are first adapted to the particular counterpart MPEC. Finally, in section 4.5 special relations between constraint qualifications for the unconstrained abs-normal NLP and its counterpart MPEC are studied.

Parts of sections 4.1 and 4.5 are published in [14]. Parts of sections 4.2-4.4 can be found in [15].

4.1 Equivalence of Abs-Normal NLP and MPEC

The general level-1 nonsmooth NLP

$$\begin{aligned} \min_{x \in D^x} \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \\ & h(x) \geq 0, \end{aligned}$$

with $f \in C^d_{\text{abs}}(D^x, \mathbb{R})$, $g \in C^d_{\text{abs}}(D^x, \mathbb{R}^{m_1})$ and $h \in C^d_{\text{abs}}(D^x, \mathbb{R}^{m_2})$ for $d \geq 1$ can be formulated via the abs-normal form (2.3b) as

$$\begin{aligned} \min_{(x,z) \in D^{x,|z|}} \quad & f(x, |z|) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, |z|) = 0, \\ & c_{\mathcal{I}}(x, |z|) \geq 0, \\ & c_{\mathcal{Z}}(x, |z|) - z = 0, \end{aligned}$$

with $\partial_2 c_{\mathcal{Z}}(x, |z|)$ strictly lower triangular. Then, z can be partioned into its nonnegative part and the modulus of its nonpositive part:

$$u := [z]^+ := \max(z, 0) \quad \text{and} \quad v := [z]^- := \max(-z, 0).$$

If complementarity of these two variables holds, $|z|$ can be replaced by $u+v$ and z itself by $u-v$. This leads to the counterpart MPEC

$$\begin{aligned}
\min_{(x,u,v)\in D^{x,u,v}} \quad & f(x,u+v) \\
\text{s.t.} \quad & c_{\mathcal{E}}(x,u+v)=0, \\
& c_{\mathcal{I}}(x,u+v)\ge 0, \\
& c_{\mathcal{Z}}(x,u+v)-(u-v)=0, \\
& 0\le u\perp v\ge 0,
\end{aligned}$$

where open sets $D^u, D^v \subseteq \mathbb{R}^s$ are chosen such that $0\in D^u$, $0\in D^v$ and $u-v\in D^{|z|}$ hold.

Conversely, every MPEC can be rewritten as an abs-normal NLP. Thus, consider

$$\begin{aligned}
\min_{(x,u,v)\in D^{x,u,v}} \quad & f(x,u,v) \\
\text{s.t.} \quad & g(x,u,v)=0, \\
& h(x,u,v)\ge 0, \\
& 0\le u\perp v\ge 0,
\end{aligned}$$

with $f\in C^d(D^{x,u,v},\mathbb{R})$, $g\in C^d(D^{x,u,v},\mathbb{R}^{m_1})$ and $h\in C^d(D^{x,u,v},\mathbb{R}^{m_2})$ for $d\ge 1$. Then, the complementarity requirement $0\le u\perp v\ge 0$ is replaced by the equivalent formulation $u+v-|u-v| = 2\min(u,v)=0$. This can be rewritten in abs-normal form and gives with $y=(x,u,v)$ the counterpart abs-normal NLP

$$\begin{aligned}
\min_{(y,z)\in D^{y,|z|}} \quad & f(y) \\
\text{s.t.} \quad & g(y)=0, \\
& h(y)\ge 0, \\
& u+v-|z|=0, \\
& u-v-z=0.
\end{aligned}$$

with $D^y = D^{x,u,v}$ and $D^{|z|}$ open such that $D^{|z|}$ is symmetric with $u-v\in D^{|z|}$.

Thus it is apparent that the problem class of nonsmooth NLPs admitting an abs-normal form is equivalent to the problem class of general MPECs. In the following only abs-normal NLPs and corresponding counterpart MPEcs are considered in more detail; MPECs and corresponding counterpart abs-normal NLPs are neglected as this thesis is mainly about abs-normal NLPs.

4.2 Counterpart MPECs

In this section MPEC counterpart problems for the two formulations (I-NLP) and (E-NLP) are introduced. Then, the relation between MPEC-LICQ for both formulations is examined.

Counterpart MPEC for (I-NLP) First, the formulation (I-NLP) is recalled:

$$\begin{aligned} \min_{t,z^t} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t,|z^t|) = 0, \\ & c_{\mathcal{I}}(t,|z^t|) \geq 0, \\ & c_{\mathcal{Z}}(t,|z^t|) - z^t = 0. \end{aligned}$$

To reformulate it as an MPEC, z^t is partioned into its nonnegative part and the modulus of its nonpositive part, i.e. $u^t := [z^t]^+ := \max(z^t, 0)$ and $v^t := [z^t]^- := \max(-z^t, 0)$. Then, complementarity of these two variables has to be required as $|z^t|$ is replaced by $u^t + v^t$ and z^t itself by $u^t - v^t$.

Definition 4.1 (Counterpart MPEC of (I-NLP))**.** The *counterpart MPEC* of (I-NLP) reads

$$\begin{aligned} \min_{(t,u^t,v^t)\in D^{t,u^t,v^t}} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t,u^t+v^t) = 0, \\ & c_{\mathcal{I}}(t,u^t+v^t) \geq 0, \\ & c_{\mathcal{Z}}(t,u^t+v^t) - (u^t - v^t) = 0, \\ & 0 \leq u^t \perp v^t \geq 0, \end{aligned} \tag{I-MPEC}$$

where $0 \in D^{u^t}$ and $0 \in D^{v^t}$. The feasible set of (I-MPEC) is denoted by

$$\mathcal{F}_{\text{i-mpec}} := \left\{ (t,u^t,v^t) \in D^{t,u^t,v^t} \;\middle|\; \begin{array}{l} c_{\mathcal{E}}(t,u^t+v^t) = 0, \;\; c_{\mathcal{I}}(t,u^t+v^t) \geq 0, \\ c_{\mathcal{Z}}(t,u^t+v^t) = u^t - v^t, \;\; 0 \leq u^t \perp v^t \geq 0 \end{array} \right\}.$$

By construction of the counterpart MPEC, the following lemma holds.

Lemma 4.2. *Given* (I-NLP) *and its counterpart MPEC* (I-MPEC)*, a homeomorphism* $\phi\colon \mathcal{F}_{\textit{i-mpec}} \to \mathcal{F}_{\textit{i-abs}}$ *is defined by*

$$\phi(t,u^t,v^t) = (t, u^t - v^t) \quad \textit{and} \quad \phi^{-1}(t,z^t) = (t,[z^t]^+,[z^t]^-).$$

Now, the definition of MPEC-LICQ is specified for the structure of (I-MPEC). Note that the index sets of (I-MPEC) are denoted by $\mathcal{U}_0^t$, $\mathcal{U}_+^t$, $\mathcal{V}_0^t$, $\mathcal{V}_+^t$ and $\mathcal{D}^t$ as the variables t, u^t and v^t are used here.

Lemma 4.3 (MPEC-LICQ for (I-MPEC))**.** *Consider a feasible point* (t,u^t,v^t) *of* (I-MPEC) *with associated index sets* $\mathcal{U}_0^t$*,* $\mathcal{U}_+^t$*,* $\mathcal{V}_0^t$*,* $\mathcal{V}_+^t$ *and* $\mathcal{D}^t$*. Set* $\mathcal{A} = \mathcal{A}(t,u^t,v^t)$ *and* $c_{\mathcal{A}} = [c_i]_{i\in\mathcal{A}}$*. Then,* MPEC-LICQ *is full row rank of*

$$J_{\textit{i-mpec}}(t,u^t,v^t) := \begin{bmatrix} \partial_1 c_{\mathcal{E}} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}_+^t} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}_+^t} \\ \partial_1 c_{\mathcal{A}} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}_+^t} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}_+^t} \\ \partial_1 c_{\mathcal{Z}} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}_+^t} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}_+^t} \end{bmatrix} \in \mathbb{R}^{(m_1+|\mathcal{A}|+s_t)\times(n_t+|\mathcal{U}_+^t|+|\mathcal{V}_+^t|)}.$$

Here, all partial derivatives are evaluated at $(t, u^t + v^t)$*.*

Proof. By Definition 2.49 full row rank of the Jacobian of the tightened NLP is required. It reads

$$J_{\mathrm{TNLP}}(t,u^t,v^t) := \begin{bmatrix} \partial_1 c_{\mathcal{E}} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}^t_+} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}^t_0} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}^t_+} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}^t_0} \\ \partial_1 c_{\mathcal{A}} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}^t_+} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}^t_0} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}^t_+} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}^t_0} \\ \partial_1 c_{\mathcal{Z}} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}^t_+} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}^t_0} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}^t_+} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}^t_0} \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix},$$

where $P_{\mathcal{S}} \in \mathbb{R}^{|\mathcal{S}| \times s_t}$ denotes the projector onto the subspace defined by $\mathcal{S} \subseteq \{1, \dots, s_t\}$ and all partial derivatives are evaluated at $(t, u^t + v^t)$. Then, the two unit blocks can be exploited to state MPEC-LICQ in a more compact form. ☐

Counterpart MPEC for (E-NLP) Recall, that the slack formulation (E-NLP) reads:

$$\begin{aligned} \min_{t,w,z^t,z^w} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t, |z^t|) = 0, \\ & c_{\mathcal{I}}(t, |z^t|) - |z^w| = 0, \\ & c_{\mathcal{Z}}(t, |z^t|) - z^t = 0, \\ & w - z^w = 0. \end{aligned}$$

Using the same approach as in the preceding paragraph, the counterpart MPEC is formulated.

Definition 4.4 (Counterpart MPEC of (E-NLP))**.** The *counterpart MPEC* of (E-NLP) reads:

$$\begin{aligned} \min_{(t,w,u^t,v^t,u^w,v^w) \in D^{t,w,u^t,v^t,u^w,v^w}} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t, u^t + v^t) = 0, \\ & c_{\mathcal{I}}(t, u^t + v^t) - (u^w + v^w) = 0, \\ & c_{\mathcal{Z}}(t, u^t + v^t) - (u^t - v^t) = 0, \\ & w - (u^w - v^w) = 0, \\ & 0 \le u^t \perp v^t \ge 0, \\ & 0 \le u^w \perp v^w \ge 0, \end{aligned} \tag{E-MPEC}$$

where $0 \in D^{u^t,v^t,u^w,v^w}$. The feasible set is a lifting of $\mathcal{F}_{\text{i-mpec}}$ and denoted by

$$\begin{aligned} \mathcal{F}_{\text{e-mpec}} &:= \left\{ (t,w,u^t,v^t,u^w,v^w) \;\middle|\; \begin{array}{l} c_{\mathcal{E}}(t,u^t+v^t) = 0,\ c_{\mathcal{I}}(t,u^t+v^t) = u^w + v^w, \\ c_{\mathcal{Z}}(t,u^t+v^t) = u^t - v^t,\ w = u^w - v^w, \\ 0 \le u^t \perp v^t \ge 0,\ 0 \le u^w \perp v^w \ge 0 \end{array} \right\} \\ &= \left\{ (t,w,u^t,v^t,u^w,v^w) \;\middle|\; \begin{array}{l} (t,u^t,v^t) \in \mathcal{F}_{\text{i-mpec}},\ c_{\mathcal{I}}(t,u^t+v^t) = u^w + v^w, \\ w = u^w - v^w,\ 0 \le u^w \perp v^w \ge 0 \end{array} \right\} \end{aligned}$$

with $(t, w, u^t, v^t, u^w, v^w) \in D^{t,w,u^t,v^t,u^w,v^w}$.

Clearly, the homeomorphism between $\mathcal{F}_{\text{i-mpec}}$ and $\mathcal{F}_{\text{i-abs}}$ extends to $\mathcal{F}_{\text{e-mpec}}$ and $\mathcal{F}_{\text{e-abs}}$.

Lemma 4.5. *Given an abs-normal NLP (E-NLP) and its counterpart MPEC (E-MPEC), a homeomorphism $\bar{\phi} : \mathcal{F}_{e\text{-}mpec} \to \mathcal{F}_{e\text{-}abs}$ is defined by*

$$\begin{aligned}\bar{\phi}(t, w, u^t, v^t, u^w, v^w) &= (t, w, u^t - v^t, u^w - v^w),\\ \bar{\phi}^{-1}(t, w, z^t, z^w) &= (t, w, [z^t]^+, [z^t]^-, [z^w]^+, [z^w]^-).\end{aligned}$$

The index sets are split into components for variables t and w which gives $\mathcal{U}_+ = (\mathcal{U}_+^t, \mathcal{U}_+^w)$, $\mathcal{V}_+ = (\mathcal{V}_+^t, \mathcal{V}_+^w)$ and $\mathcal{D} = (\mathcal{D}^t, \mathcal{D}^w)$.

Lemma 4.6. *Consider a feasible point $y = (t, w, u^t, v^t, u^w, v^w)$ of (E-MPEC) with associated index sets $\mathcal{U}_+^t$, $\mathcal{V}_+^t$, $\mathcal{U}_+^w$ and $\mathcal{V}_+^w$. Then, MPEC-LICQ is full row rank of*

$$\begin{aligned}J_{e\text{-}mpec}(y) &:= \begin{bmatrix} \partial_1 c_{\mathcal{E}} & 0 & \partial_2 c_{\mathcal{E}} P_{\mathcal{U}_+^t}^T & \partial_2 c_{\mathcal{E}} P_{\mathcal{V}_+^t}^T & 0 & 0 \\ \partial_1 c_{\mathcal{I}} & 0 & \partial_2 c_{\mathcal{I}} P_{\mathcal{U}_+^t}^T & \partial_2 c_{\mathcal{I}} P_{\mathcal{V}_+^t}^T & -P_{\mathcal{U}_+^w}^T & -P_{\mathcal{V}_+^w}^T \\ \partial_1 c_{\mathcal{Z}} & 0 & [\partial_2 c_{\mathcal{Z}} - I] P_{\mathcal{U}_+^t}^T & [\partial_2 c_{\mathcal{Z}} + I] P_{\mathcal{V}_+^t}^T & 0 & 0 \\ 0 & I & 0 & 0 & -P_{\mathcal{U}_+^w}^T & +P_{\mathcal{V}_+^w}^T \end{bmatrix}\\ &\in \mathbb{R}^{(m_1+m_2+s_t+m_2)\times(n_t+m_2+|\mathcal{U}_+^t|+|\mathcal{V}_+^t|+|\mathcal{U}_+^w|+|\mathcal{V}_+^w|)},\end{aligned}$$

where all partial derivatives are evaluated at $(t, u^t + v^t)$.

Proof. Set $x = (t, w)$, $u = (u^t, u^w)$, $v = (v^t, v^w)$ as well as $\bar{f}(x) = f(t)$,

$$\bar{c}_{\mathcal{E}}(x, u + v) = \begin{pmatrix} c_{\mathcal{E}}(t, u^t + v^t) \\ c_{\mathcal{I}}(t, u^t + v^t) - (u^w + v^w) \end{pmatrix} \quad \text{and} \quad \bar{c}_{\mathcal{Z}}(x, u + v) = \begin{pmatrix} c_{\mathcal{Z}}(t, u^t + v^t) \\ w \end{pmatrix}.$$

Then, (E-MPEC) becomes

$$\begin{aligned} \min_{x,u,v} \quad & \bar{f}(x) \\ \text{s.t.} \quad & \bar{c}_{\mathcal{E}}(x, u + v) = 0, \\ & \bar{c}_{\mathcal{Z}}(x, u + v) - (u - v) = 0, \\ & 0 \le u \perp v \ge 0 \end{aligned} \tag{$\overline{\text{E-MPEC}}$}$$

and the Jacobian can be computed from Lemma 4.3 using the structure of ($\overline{\text{E-MPEC}}$). With all partial derivatives evaluated at $(t, u^t + v^t)$ this gives:

$$J_{\text{e-mpec}}(y) = \begin{bmatrix} \partial_1 c_{\mathcal{E}} & 0 & \partial_2 c_{\mathcal{E}} P_{\mathcal{U}_+^t}^T & 0 & \partial_2 c_{\mathcal{E}} P_{\mathcal{V}_+^t}^T & 0 \\ \partial_1 c_{\mathcal{I}} & 0 & \partial_2 c_{\mathcal{I}} P_{\mathcal{U}_+^t}^T & -P_{\mathcal{U}_+^w}^T & \partial_2 c_{\mathcal{I}} P_{\mathcal{V}_+^t}^T & -P_{\mathcal{V}_+^w}^T \\ \partial_1 c_{\mathcal{Z}} & 0 & [\partial_2 c_{\mathcal{Z}} - I] P_{\mathcal{U}_+^t}^T & 0 & [\partial_2 c_{\mathcal{Z}} + I] P_{\mathcal{V}_+^t}^T & 0 \\ 0 & I & 0 & -P_{\mathcal{U}_+^w}^T & 0 & +P_{\mathcal{V}_+^w}^T \end{bmatrix}.$$

The resulting matrix has the stated form, except for the last four columns belonging to variables (u^t, v^t, u^w, v^w) rather than $(u, v) = (u^t, u^w, v^t, v^w)$. ☐

Like LIKQ for (E-NLP), MPEC-LICQ for (E-MPEC) does not depend on the particular choice of w.

Relations of MPEC Constraint Qualifications In this paragraph the relation of LIKQ for the two different formulations introduced in the previous paragraphs is stated. It follows from the results in section 3.2.1 and in the *next* section. For an illustration see Fig. 4.1 below. The $2^{m_2-|\mathcal{D}^w|}$ different possible choices of w are collected in $W(t,u^t,v^t) := \{(w,u^w,v^w)\colon |w| = c_{\mathcal{I}}(t,u^t+v^t), u^w = [w]^+, v^w = [w]^-\}$.

Theorem 4.7. *MPEC-LICQ for (I-MPEC) holds at $(t,u^t,v^t) \in \mathcal{F}_{\text{i-mpec}}$ if and only if MPEC-LICQ for (E-MPEC) holds at $(t,w,u^t,u^w,v^t,v^w) \in \mathcal{F}_{\text{e-mpec}}$ for any $(w,u^w,v^w) \in W(t,u^t,v^t)$ and hence for all $(w,u^w,v^w) \in W(t,u^t,v^t)$.*

Proof. This follows directly from Theorem 3.72, Theorem 4.8 and Theorem 4.9. Alternatively, the equivalence follows directly by using

$$\mathcal{U}_+^w \cup \mathcal{V}_+^w = \mathcal{I} \setminus \mathcal{A}(t,u^t,v^t) \quad \text{and} \quad \mathcal{U}_+^w \cap \mathcal{V}_+^w = \emptyset.$$

□

4.3 Kink Qualifications and MPEC Constraint Qualifications

In this section relations between abs-normal NLPs and counterpart MPECs in both formulations are studied more closely.

Relations of (I-NLP) and (I-MPEC) Here the variables x and z instead of t and z^t are used to shorten notation because inequality slacks are not considered explicitly. Then the general abs-normal NLP (I-NLP) becomes:

$$\begin{aligned}
\min_{(x,z)\in D^{x,|z|}} \quad & f(x) \\
\text{s.t.} \quad & c_{\mathcal{E}}(x,|z|) = 0, \\
& c_{\mathcal{I}}(x,|z|) \geq 0, \\
& c_{\mathcal{Z}}(x,|z|) - z = 0.
\end{aligned}$$

The counterpart MPEC (I-MPEC) reads:

$$\begin{aligned}
\min_{(x,u,v)\in D^{x,u,v}} \quad & f(x) \\
\text{s.t.} \quad & c_{\mathcal{E}}(x,u+v) = 0, \\
& c_{\mathcal{I}}(x,u+v) \geq 0, \\
& c_{\mathcal{Z}}(x,u+v) - (u-v) = 0, \\
& 0 \leq u \perp v \geq 0.
\end{aligned}$$

The following relations of kink qualifications and MPEC constraint qualifications are obtained.

Theorem 4.8 (Equivalence of LIKQ and MPEC-LICQ)**.** *LIKQ for* (I-NLP) *holds at* $x \in \mathcal{F}_{\text{i-abs}}$ *if and only if MPEC-LICQ for* (I-MPEC) *holds at* $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\text{i-mpec}}$.

Proof. Setting $y := (x, u+v)$ and $r := m_1 + |\mathcal{A}| + s$, MPEC-LICQ for the counterpart MPEC is

$$\operatorname{rank} \begin{bmatrix} \partial_1 c_{\mathcal{E}}(y) & \partial_2 c_{\mathcal{E}}(y) P_{\mathcal{U}_+}^T & \partial_2 c_{\mathcal{E}}(y) P_{\mathcal{V}_+}^T \\ \partial_1 c_{\mathcal{A}}(y) & \partial_2 c_{\mathcal{A}}(y) P_{\mathcal{U}_+}^T & \partial_2 c_{\mathcal{A}}(y) P_{\mathcal{V}_+}^T \\ \partial_1 c_{\mathcal{Z}}(y) & [\partial_2 c_{\mathcal{Z}}(y) - I] P_{\mathcal{U}_+}^T & [\partial_2 c_{\mathcal{Z}}(y) + I] P_{\mathcal{V}_+}^T \end{bmatrix} = r.$$

By negating the second column and combining it with the third column, this is equivalent to

$$\operatorname{rank} \begin{bmatrix} \partial_1 c_{\mathcal{E}}(y) & -\partial_2 c_{\mathcal{E}}(y) \Sigma P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T \\ \partial_1 c_{\mathcal{A}}(y) & -\partial_2 c_{\mathcal{A}}(y) \Sigma P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T \\ \partial_1 c_{\mathcal{Z}}(y) & [I - \partial_2 c_{\mathcal{Z}}(y) \Sigma] P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T \end{bmatrix} = r$$

and, by non-singularity of $I - \partial_2 c_{\mathcal{Z}}(y)\Sigma$, to

$$\operatorname{rank} \begin{bmatrix} \partial_1 c_{\mathcal{E}}(y) & -\partial_2 c_{\mathcal{E}}(y) \Sigma P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T \\ \partial_1 c_{\mathcal{A}}(y) & -\partial_2 c_{\mathcal{A}}(y) \Sigma P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T \\ [I - \partial_2 c_{\mathcal{Z}}(y) \Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(y) & P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T \end{bmatrix} = r.$$

Next, the third row is used to eliminate the entries above $P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T$ to obtain

$$\operatorname{rank} \begin{bmatrix} \partial_1 c_{\mathcal{E}}(y) + \partial_2 c_{\mathcal{E}}(y) \Sigma [I - \partial_2 c_{\mathcal{Z}}(y) \Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(y) & 0 \\ \partial_1 c_{\mathcal{A}}(y) + \partial_2 c_{\mathcal{A}}(y) \Sigma [I - \partial_2 c_{\mathcal{Z}}(y) \Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(y) & 0 \\ [I - \partial_2 c_{\mathcal{Z}}(y) \Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(y) & P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T \end{bmatrix} = r,$$

which can be rewritten with $u + v = |z| = |z(x)|$ as

$$\operatorname{rank} \begin{bmatrix} \partial_x c_{\mathcal{E}}(x, |z(x)|) & 0 \\ \partial_x c_{\mathcal{A}}(x, |z(x)|) & 0 \\ \partial_x z(x) & P_{\mathcal{U}_+ \cup \mathcal{V}_+}^T \end{bmatrix} = r.$$

Finally, since $\alpha = \mathcal{D}$ is the complement of $\mathcal{U}_+ \cup \mathcal{V}_+$, this is equivalent to

$$\operatorname{rank} \begin{bmatrix} \partial_x c_{\mathcal{E}}(x, |z(x)|) \\ \partial_x c_{\mathcal{A}}(x, |z(x)|) \\ [e_i^T \partial_x z(x)]_{i \in \alpha} \end{bmatrix} = m_1 + |\mathcal{A}| + |\alpha|,$$

which is LIKQ for the abs-normal NLP. □

Relations of (E-NLP) and (E-MPEC) As the reformulation with inequality slacks is just a specialization of the general case, the same relation between LIKQ and MPEC-LICQ as in the previous paragraph holds.

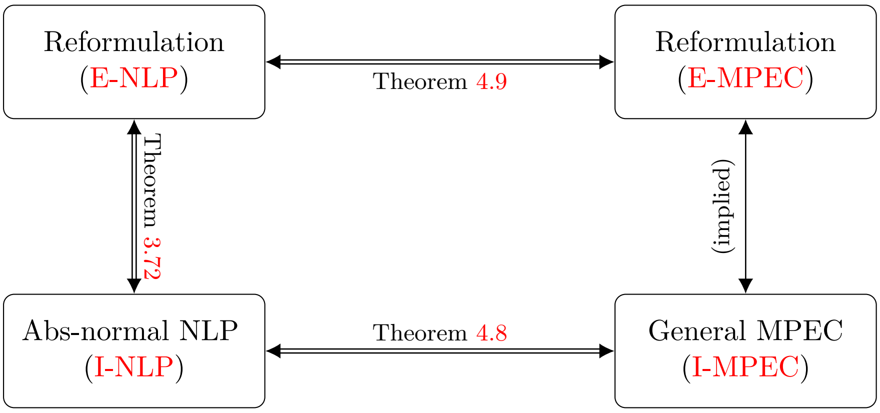

Figure 4.1: Relations between LICQ and LIKQ for different problem formulations

Theorem 4.9 (Equivalence of LIKQ and MPEC-LICQ)**.** *LIKQ for* (E-NLP) *holds at* $x \in \mathcal{F}_{i\text{-}abs}$ *if and only if MPEC-LICQ for* (E-MPEC) *holds at* $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{i\text{-}mpec}$.

Proof. In the proof of Lemma 3.70 the short notation ($\overline{\text{E-NLP}}$) for (E-NLP) was used to formulate LIKQ and similarly, in the proof of Lemma 4.6 the short notation ($\overline{\text{E-MPEC}}$) for the counterpart MPEC (E-MPEC). As ($\overline{\text{E-MPEC}}$) is the counterpart MPEC for ($\overline{\text{E-NLP}}$), Theorem 4.8 can be applied and gives the result. □

4.4 Optimality Conditions

In this section first and second order optimality conditions for (I-MPEC) under MPEC-LICQ and for (I-NLP) under LIKQ are considered and their relations are discussed. Since both regularity conditions are invariant under the slack reformulation by Theorem 3.72 and Theorem 4.7, the results hold also for (E-MPEC) and (E-NLP).

First Order Conditions In this paragraph, strong stationarity for MPECs is compared to kink stationarity for abs-normal NLPs.

First, the MPEC-Lagrangian is formulated for (I-MPEC). It reads

$$\begin{aligned}\mathcal{L}_\perp(y, \lambda, \mu) = f(t) &+ \lambda_\mathcal{E}^T c_\mathcal{E}(t, u^t + v^t) - \lambda_\mathcal{I}^T c_\mathcal{I}(t, u^t + v^t) \\ &+ \lambda_\mathcal{Z}^T [c_\mathcal{Z}(t, u^t + v^t) - (u^t - v^t)] - \mu_\mathrm{u}^T u^t - \mu_\mathrm{v}^T v^t.\end{aligned}$$

with $y = (t, u^t, v^t)$ and Lagrange multiplier vectors $\lambda = (\lambda_\mathcal{E}, \lambda_\mathcal{I}) \in \mathbb{R}^{m_1+m_2+s_t}$ and $\mu = (\mu_\mathrm{u}, \mu_\mathrm{v}) \in \mathbb{R}^{2s_t}$.

Then, the definition of strong stationarity is adapted. The result follows directly from Definition 2.61.

Lemma 4.10 (Strong Stationarity). *Consider a feasible point $y^* = (t^*, (u^t)^*, (v^t)^*)$ of (I-MPEC) with associated index sets $\mathcal{U}_+^t$, $\mathcal{V}_+^t$ and $\mathcal{D}^t$. It is strongly stationary if and only if there exist Lagrange multiplier vectors $\lambda^* = (\lambda_\mathcal{E}^*, \lambda_\mathcal{I}^*, \lambda_\mathcal{Z}^*)$ and $\mu^* = (\mu_u^*, \mu_v^*)$ such that the following conditions are satisfied:*

$$\partial_y \mathcal{L}_\perp(y^*, \lambda^*, \mu^*) = 0, \tag{4.3a}$$
$$(\mu_u^*)_i \geq 0,\ (\mu_v^*)_i \geq 0, \quad i \in \mathcal{D}^t, \tag{4.3b}$$
$$(\mu_u^*)_i = 0, \quad i \in \mathcal{U}_+^t, \tag{4.3c}$$
$$(\mu_v^*)_i = 0, \quad i \in \mathcal{V}_+^t, \tag{4.3d}$$
$$\lambda_\mathcal{I}^* \geq 0, \tag{4.3e}$$
$$(\lambda_\mathcal{I}^*)^T c_\mathcal{I}(t^*, (u^t)^* + (v^t)^*) = 0. \tag{4.3f}$$

The next theorem shows that the two stationarity concepts coincide.

Theorem 4.11 (S-Stationarity is Kink Stationarity). *A feasible point $(t^*, (z^t)^*)$ of (I-NLP) is kink stationary if and only if $(t^*, (u^t)^*, (v^t)^*) = (t^*, [z^t(t^*)]^+, [z^t(t^*)]^-)$ of (I-MPEC) is strongly stationary.*

Proof. Comparison of the stationarity conditions of (I-NLP) and (I-MPEC) shows directly that (4.3e) and (3.18d) as well as (4.3f) and (3.18e) coincide. Thus, the remaining conditions (4.3a) to (4.3d) for (I-MPEC) and (3.18a) to (3.18c) for (I-NLP) have to be checked. Condition (4.3a) of (I-MPEC), where all derivatives are evaluated at $(t^*, (u^t)^* + (v^t)^*)$, is

$$\begin{aligned}
f'(t^*) + (\lambda_\mathcal{E}^*)^T \partial_1 c_\mathcal{E} - (\lambda_\mathcal{I}^*)^T \partial_1 c_\mathcal{I} + (\lambda_\mathcal{Z}^*)^T \partial_1 c_\mathcal{Z} &= 0,\\
(\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} - (\lambda_\mathcal{I}^*)^T \partial_2 c_\mathcal{I} + (\lambda_\mathcal{Z}^*)^T [\partial_2 c_\mathcal{Z} - I] - (\mu_\mathrm{u}^*)^T &= 0,\\
(\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} - (\lambda_\mathcal{I}^*)^T \partial_2 c_\mathcal{I} + (\lambda_\mathcal{Z}^*)^T [\partial_2 c_\mathcal{Z} + I] - (\mu_\mathrm{v}^*)^T &= 0.
\end{aligned}$$

The first condition coincides with (3.18a). Combining the second and the third condition with conditions (4.3b) to (4.3d), yields

$$\begin{aligned}
\Big[(\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} - (\lambda_\mathcal{I}^*)^T \partial_2 c_\mathcal{I} + (\lambda_\mathcal{Z}^*)^T \partial_2 c_\mathcal{Z} \qquad\Big]_i &= +(\lambda_\mathcal{Z}^*)_i, \quad i \in \mathcal{U}_+^t,\\
\Big[(\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} - (\lambda_\mathcal{I}^*)^T \partial_2 c_\mathcal{I} + (\lambda_\mathcal{Z}^*)^T \partial_2 c_\mathcal{Z} \qquad\Big]_i &= -(\lambda_\mathcal{Z}^*)_i, \quad i \in \mathcal{V}_+^t,\\
\Big[(\lambda_\mathcal{E}^*)^T \partial_2 c_\mathcal{E} - (\lambda_\mathcal{I}^*)^T \partial_2 c_\mathcal{I} + (\lambda_\mathcal{Z}^*)^T [\partial_2 c_\mathcal{Z} \pm I]\Big]_i &\geq 0, \quad i \in \mathcal{D}^t.
\end{aligned}$$

These are exactly conditions (3.18b) and (3.18c) for (I-NLP) by definition of the index sets and of σ^*. ◻

As LIKQ for (I-NLP) is equivalent to MPEC-LICQ for (I-MPEC), the previous theorem provides a different perspective on Theorem 3.74 and Theorem 2.62: one can be obtained from the other directly via Theorem 4.11.

Second Order Conditions In this paragraph, second order conditions for MPECs and abs-normal NLPs are compared.

First, they are formulated for (I-MPEC). This is based on [23] (see section 2.3) but some additional assumptions on the Lagrange multiplier vectors are made. These are given in the next definition.

Definition 4.12 (MPEC-Strict Complementarity)**.** Consider a strongly stationary point $y^* = (t^*, (u^t)^*, (v^t)^*)$ of (I-MPEC) with Lagrange multiplier vectors $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{I}}, \lambda^*_{\mathcal{Z}})$ and $\mu^* = (\mu^*_{\mathrm{u}}, \mu^*_{\mathrm{v}})$. One says that *MPEC-strict complementarity* holds if $\lambda^*_i > 0$ for all $i \in \mathcal{A} = \mathcal{A}(y^*)$ as well as $(\mu^*_{\mathrm{u}})_i > 0$ and $(\mu^*_{\mathrm{v}})_i > 0$ for all $i \in \mathcal{D}^t$.

It will be shown that under MPEC-LICQ and MPEC-strict complementarity the critical cone reduces to the nullspace of the Jacobian of the tightened NLP which was introduced in the proof of Lemma 4.3 and reads

$$J_{\mathrm{TNLP}}(y^*) = \begin{bmatrix} \partial_1 c_{\mathcal{E}} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}^t_+} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}^t_0} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}^t_+} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}^t_0} \\ \partial_1 c_{\mathcal{A}} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}^t_+} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}^t_0} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}^t_+} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}^t_0} \\ \partial_1 c_{\mathcal{Z}} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}^t_+} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}^t_0} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}^t_+} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}^t_0} \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix},$$

with $y^* = (t^*, (u^t)^*, (v^t)^*)$ and associated index sets $\mathcal{U}^t_0$, $\mathcal{U}^t_+$, $\mathcal{V}^t_0$, $\mathcal{V}^t_+$ and $\mathcal{D}^t$. Here, all partial derivatives are evaluated at $(t^*, (u^t)^* + (v^t)^*)$.

To this aim, first the particular form of the nullspace of $J_{\mathrm{TNLP}}(y^*)$ is derived.

Lemma 4.13. *Consider a feasible point $y^* = (t^*, (u^t)^*, (v^t)^*)$ of (I-MPEC) with associated index sets $\mathcal{U}^t_0$, $\mathcal{U}^t_+$, $\mathcal{V}^t_0$, $\mathcal{V}^t_+$ and $\mathcal{D}^t$ and assume that MPEC-LICQ holds at y^*. Then, a basis for the nullspace of $J_{TNLP}(y^*)$ is given by the columns of the matrix*

$$U_{i\text{-}mpec}(y^*) := \begin{bmatrix} I \\ +P_{\mathcal{U}^t_+}(I - \partial_2 c_{\mathcal{Z}} \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} \\ 0 \\ -P_{\mathcal{V}^t_+}(I - \partial_2 c_{\mathcal{Z}} \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} \\ 0 \end{bmatrix} \tilde{U}(y^*), \tag{4.4}$$

where the columns of $\tilde{U}(y^*)$ are a basis for the nullspace of

$$\begin{bmatrix} \partial_1 c_{\mathcal{E}} + \partial_2 c_{\mathcal{E}} \Sigma^t (I - \partial_2 c_{\mathcal{Z}}\ \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} \\ \partial_1 c_{\mathcal{A}} + \partial_2 c_{\mathcal{A}} \Sigma^t (I - \partial_2 c_{\mathcal{Z}}\ \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} \\ [e_i^T (I - \partial_2 c_{\mathcal{Z}} \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}}]_{i \in \mathcal{D}^t} \end{bmatrix}, \tag{4.5}$$

with $\Sigma^t = \operatorname{diag}(\sigma^t)$ where $\sigma^t_i = 1$ for $i \in \mathcal{U}^t_+$, $\sigma^t_i = -1$ for $i \in \mathcal{V}^t_+$ and $\sigma^t_i = 0$ for $i \in D^t$. All partial derivatives are evaluated at $(t^*, (u^t)^* + (v^t)^*)$.

Proof. Consider the equation $0 = J_{\mathrm{TNLP}}(y^*) U(y^*)$ with $U(y^*) = [U_1\ U_2\ U_3\ U_4\ U_5]^T$. Then, exploiting the identity blocks in $J_{\mathrm{TNLP}}(y^*)$ it follows directly that $U_3 = U_5 = 0$. Thus, the equation $0 = J_{\mathrm{TNLP}}(y^*) U(y^*)$ reduces to

$$0 = J_{\text{i-mpec}}(y^*) \begin{bmatrix} U_1 \\ U_2 \\ U_4 \end{bmatrix} = \begin{bmatrix} \partial_1 c_{\mathcal{E}} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}^t_+} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}^t_+} \\ \partial_1 c_{\mathcal{A}} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}^t_+} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}^t_+} \\ \partial_1 c_{\mathcal{Z}} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}^t_+} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}^t_+} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_4 \end{bmatrix}.$$

This is equivalent to

$$0 = \begin{bmatrix} \partial_1 c_{\mathcal{E}} & -\partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}^t_+} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}^t_+} \\ \partial_1 c_{\mathcal{A}} & -\partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}^t_+} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}^t_+} \\ \partial_1 c_{\mathcal{Z}} & [I - \partial_2 c_{\mathcal{Z}}] P^T_{\mathcal{U}^t_+} & [I - (-\partial_2 c_{\mathcal{Z}})] P^T_{\mathcal{V}^t_+} \end{bmatrix} \begin{bmatrix} U_1 \\ -U_2 \\ U_4 \end{bmatrix}.$$

Combining $-U_2$ and U_4 in U_{24} and using Σ^t it can be rewritten as

$$0 = \begin{bmatrix} \partial_1 c_{\mathcal{E}} & -\partial_2 c_{\mathcal{E}} \Sigma P^T_{\mathcal{U}^t_+ \cup \mathcal{V}^t_+} \\ \partial_1 c_{\mathcal{A}} & -\partial_2 c_{\mathcal{A}} \Sigma P^T_{\mathcal{U}^t_+ \cup \mathcal{V}^t_+} \\ \partial_1 c_{\mathcal{Z}} & [I - \partial_2 c_{\mathcal{Z}} \Sigma] P^T_{\mathcal{U}^t_+ \cup \mathcal{V}^t_+} \end{bmatrix} \begin{bmatrix} U_1 \\ U_{24} \end{bmatrix}.$$

Then, the third line gives $P^T_{\mathcal{U}^t_+ \cup \mathcal{V}^t_+} U_{24} = -(I - \partial_2 c_{\mathcal{Z}} \Sigma)^{-1} \partial_1 c_{\mathcal{Z}} U_1$ since $\partial_2 c_{\mathcal{Z}}$ is strictly lower triangular. This can be multiplied by $P_{\mathcal{U}^t_+ \cup \mathcal{V}^t_+}$ and $P_{\mathcal{D}^t} = \mathcal{P}_{\mathcal{U}^t_0 \cap \mathcal{V}^t_0}$ giving

$$\begin{aligned} U_{24} &= -P_{\mathcal{U}^t_+ \cup \mathcal{V}^t_+} (I - \partial_2 c_{\mathcal{Z}} \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} U_1, \\ 0 &= [e_i^T (I - \partial_2 c_{\mathcal{Z}} \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}}]_{i \in \mathcal{D}^t} U_1 \end{aligned}$$

and inserted into the first and second equation giving

$$\begin{aligned} 0 &= \left[\partial_1 c_{\mathcal{E}} + \partial_2 c_{\mathcal{E}} \Sigma^t (I - \partial_2 c_{\mathcal{Z}} \ \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} \right] U_1, \\ 0 &= \left[\partial_1 c_{\mathcal{A}} + \partial_2 c_{\mathcal{A}} \Sigma^t (I - \partial_2 c_{\mathcal{Z}} \ \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} \right] U_1. \end{aligned}$$

Then, U_2 and U_4 read

$$\begin{aligned} U_2 &= P_{\mathcal{U}^t_+} (I - \partial_2 c_{\mathcal{Z}} \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} U_1, \\ U_4 &= -P_{\mathcal{V}^t_+} (I - \partial_2 c_{\mathcal{Z}} \Sigma^t)^{-1} \partial_1 c_{\mathcal{Z}} U_1. \end{aligned}$$

This gives the result as $U_{\text{i-mpec}}(y^*)$ and $\tilde{U}(y^*)$ are defined appropriate. □

Second order conditions can now be stated and will be proved using the following lemma and Theorem 2.70 from Scheel and Scholtes.

Lemma 4.14 (Critical Direction). *Consider a strongly stationary point $y^* = (t^*, (u^t)^*, (v^t)^*)$ of (I-MPEC) with associated index sets $\mathcal{U}^t_+$, $\mathcal{V}^t_+$ and $\mathcal{D}^t$. Then, a vector $d = (dt, du^t, dv^t) \in \mathbb{R}^{n_t} \times \mathbb{R}^{s_t} \times \mathbb{R}^{s_t}$ is a* critical direction *at y^* if and only if*

$$\min(du^t_i, dv^t_i) = 0, \quad i \in \mathcal{D}^t, \tag{4.6a}$$
$$du^t_i = 0, \quad i \in \mathcal{V}^t_+, \tag{4.6b}$$
$$dv^t_i = 0, \quad i \in \mathcal{U}^t_+, \tag{4.6c}$$
$$\partial_1 c_{\mathcal{A}} dt + \partial_2 c_{\mathcal{A}} (du^t + dv^t) \geq 0, \tag{4.6d}$$
$$\partial_1 c_{\mathcal{E}} dt + \partial_2 c_{\mathcal{E}} (du^t + dv^t) = 0, \tag{4.6e}$$
$$\partial_1 c_{\mathcal{Z}} dt + [\partial_2 c_{\mathcal{Z}} - I] du^t + [\partial_2 c_{\mathcal{Z}} + I] dv^t = 0, \tag{4.6f}$$
$$f'(t^*) dt = 0, \tag{4.6g}$$

where all constraint derivatives are evaluated at $(t^*, (u^t)^* + (v^t)^*)$.

Proof. Follows directly from Definition 2.69 using the form of (I-MPEC). $\square$

Under additional assumptions, the set of critical directions is exactly the nullspace of $J_{\mathrm{TNLP}}(y^*)$.

Lemma 4.15. *Consider a strongly stationary point $y^* = (t^*, (u^t)^*, (v^t)^*)$ of (I-MPEC) with Lagrange multiplier vectors $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{I}}, \lambda^*_{\mathcal{Z}})$ and $\mu^* = (\mu^*_{\mathrm{u}}, \mu^*_{\mathrm{v}})$ and assume that MPEC-LICQ and MPEC-strict complementarity hold. Then, the set of critical directions is* $\ker J_{TNLP}(y^*)$.

Proof. In the following all partial derivatives are evalauated at $(t^*, (u^t)^* + (v^t)^*)$. First consider a critical direction $d = (dt, du^t, dv^t)$ at the strongly stationary point y^*. Then, (4.6e) and (4.6f) imply that rows one and three of $J_{\mathrm{TNLP}}(y^*)d$ vanish and by (4.3a) and (4.6g) this gives

$$\begin{aligned}
0 &= \partial_{t,u^t,v^t}\mathcal{L}_{\perp}(y^*, \lambda^*, \mu^*)d \\
&= f'(t^*)dt + (\lambda^*_{\mathcal{E}})^T[\partial_1 c_{\mathcal{E}} dt + \partial_2 c_{\mathcal{E}}(du^t + dv^t)] \\
&\quad - (\lambda^*_{\mathcal{I}})^T[\partial_1 c_{\mathcal{I}} dt + \partial_2 c_{\mathcal{I}}(du^t + dv^t)] \\
&\quad + (\lambda^*_{\mathcal{Z}})^T[\partial_1 c_{\mathcal{Z}} dt + (\partial_2 c_{\mathcal{Z}} - I)du^t + (\partial_2 c_{\mathcal{Z}} + I)dv^t] \\
&\quad - (\mu^*_{\mathrm{u}})^T du^t - (\mu^*_{\mathrm{v}})^T dv^t \\
&= -(\lambda^*_{\mathcal{I}})^T[\partial_1 c_{\mathcal{I}} dt + \partial_2 c_{\mathcal{I}}(du^t + dv^t)] - (\mu^*_{\mathrm{u}})^T du^t - (\mu^*_{\mathrm{v}})^T dv^t.
\end{aligned}$$

With $(\lambda^*_{\mathcal{I}})^T c_{\mathcal{I}} = 0$ (4.3f), $(\mu^*_{\mathrm{u}})_i = 0$ for $i \in \mathcal{U}^t_+$ (4.3c), $(\mu^*_{\mathrm{v}})_i = 0$ for $i \in \mathcal{V}^t_+$ (4.3d) and (4.6b), (4.6c) it leads to

$$0 = (\lambda^*_{\mathcal{A}})^T[\partial_1 c_{\mathcal{A}} du^t + \partial_2 c_{\mathcal{A}}(du^t + dv^t)] + \sum_{i \in \mathcal{D}^t}[(\mu^*_{\mathrm{u}})_i du^t_i + (\mu^*_{\mathrm{v}})_i dv^t_i].$$

All factors in this sum of products are nonnegative by (4.3b), (4.3e), (4.6d) and (4.6a), which implies

$$\begin{aligned}
0 &= (\lambda^*_{\mathcal{A}})^T[\partial_1 c_{\mathcal{A}} du^t + \partial_2 c_{\mathcal{A}}(du^t + dv^t)], \\
0 &= (\mu^*_{\mathrm{u}})_i du^t_i = (\mu^*_{\mathrm{v}})_i dv^t_i, \quad i \in \mathcal{D}^t.
\end{aligned}$$

Finally, by MPEC-strict complementarity

$$\begin{aligned}
0 &= \partial_1 c_{\mathcal{A}} du^t + \partial_2 c_{\mathcal{A}}(du^t + dv^t), \\
0 &= du^t_i = dv^t_i, \quad i \in \mathcal{D}^t
\end{aligned}$$

and $du^t_i = 0$ for $i \in \mathcal{U}^t_0$ as well as $dv^t_i = 0$ for $i \in \mathcal{V}^t_0$ follow since $\mathcal{U}^t_0 = \mathcal{V}^t_+ \cup \mathcal{D}^t$ and $\mathcal{V}^t_0 = \mathcal{U}^t_+ \cup \mathcal{D}^t$. Thus, d is a nullspace vector of $J_{\mathrm{TNLP}}(y^*)$.

Conversely, given a nullspace vector $d = (dt, du^t, dv^t)$, the first three rows of $J_{\mathrm{TNLP}}(y^*)d = 0$ yield conditions (4.6e), (4.6d) and (4.6f), with equality "$= 0$" in case of (4.6d). The last two rows yield $du^t_i = 0$ for $i \in \mathcal{U}^t_0$ and $dv^t_i = 0$ for $i \in \mathcal{V}^t_0$, hence (4.6b), (4.6c) and $du^t_i = dv^t_i = 0$ for $i \in \mathcal{D}^t$ (4.6a). Moreover, it holds that $(\mu^*_{\mathrm{u}})_i = 0$ for $i \in \mathcal{U}^t_+$ (4.3c), $(\mu^*_{\mathrm{v}})_i = 0$ for $i \in \mathcal{V}^t_+$ (4.3d) and $(\lambda^*_{\mathcal{I}})_i = 0$ for $i \notin \mathcal{A}$ (4.3f), so that (4.3a) becomes (4.6g):

$$0 = \partial_{t,u^t,v^t}\mathcal{L}_{\perp}(y^*, \lambda^*, \mu^*)d = f'(t^*)dt.$$

Thus d is a critical direction. $\square$

Now Theorem 2.70 and Theorem 2.71 can be used to prove second order necessary and sufficient conditions for this setting.

Theorem 4.16 (Second Order Necessary Conditions)**.** *Consider (I-MPEC) for $d \geq 2$. Assume that $y^* = (t^*, (u^t)^*, (v^t)^*)$ is a local minimizer and that MPEC-LICQ holds. Denote by λ^* and μ^* the unique Lagrange multiplier vectors and assume further that MPEC-strict complementarity holds. Then,*

$$U_{\textit{i-mpec}}(y^*)^T H_{\textit{i-mpec}}(y^*, \lambda^*) U_{\textit{i-mpec}}(y^*) \geq 0$$

where $H_{\textit{i-mpec}}(y^*, \lambda^*) := \partial^2_{yy} \mathcal{L}_\perp(y^*, \lambda^*, \mu^*)$. (Note that $\partial^2_{yy} \mathcal{L}_\perp$ does not depend on μ^*.)

Proof. Theorem 2.70 asserts that every critical direction d satisfies $d^T H_{\text{i-mpec}}(y^*, \lambda^*) d \geq 0$ at a local minimizer y^* if MPEC-LICQ holds at y^*. Since the set of critical directions is $\ker J_{\text{TNLP}}(y^*)$ under the stronger assumptions here, the claim follows directly from Theorem 2.70. □

Remark 4.17. Here the exposition has been simplified by making the assumption of MPEC-strict complementarity so that one can directly rely on Theorem 2.70. However, the second order necessary conditions can also be proven without MPEC-strict complementarity by considering branch problems of (I-MPEC). The corresponding approach for (I-NLP) has been taken in section 3.1.4.2, such that Theorem 3.62 does not require strict complementarity.

Theorem 4.18 (Second Order Sufficient Conditions)**.** *Consider (I-MPEC) for $d \geq 2$. Assume that $y^* = (t^*, (u^t)^*, (v^t)^*)$ is strongly stationary with Lagrange multiplier vectors λ^* and μ^* satisfying MPEC-strict complementarity. Assume further that MPEC-LICQ holds and that*

$$U_{\textit{i-mpec}}(y^*)^T H_{\textit{i-mpec}}(y^*, \lambda^*) U_{\textit{i-mpec}}(y^*) > 0.$$

Then, y^* is a strict local minimizer of (I-MPEC).

Proof. In Theorem 2.71, the assertion is proven under the weaker assumption that y^* is strongly stationary and for every critical direction $d \neq 0$ there exists Lagrange multiplier vectors λ^* and μ^* such that $d^T H_{\text{i-mpec}}(y^*, \lambda^*) d > 0$. Under the additional assumptions of MPEC-LICQ and MPEC-strict complementarity, the Lagrange multiplier vectors are uniquely determined and a basis of the set of critical directions is given by the columns of the matrix $U_{\text{i-mpec}}(y^*)$, see proof of the previous lemma. Thus, the claim follows directly from Theorem 2.71. □

Next, the relation between second order optimality conditions is formulated.

Theorem 4.19. *Consider (I-NLP) for $d \geq 2$. Assume that $(t^*, (z^t)^*)$ is kink stationary with Lagrange multiplier vector λ^* such that strict complementarity and strict normal growth are satisfied. Assume further that LIKQ holds at t^*. Then,*

$$U_{\textit{i-mpec}}(y^*)^T H_{\textit{i-mpec}}(y^*, \lambda^*) U_{\textit{i-mpec}}(y^*) \geq 0 \iff U_{\textit{i-abs}}(t^*)^T H_{\textit{i-abs}}(t^*, (z^t)^*, \lambda^*) U_{\textit{i-abs}}(x^*) \geq 0,$$

where $y^* = (t^*, (u^t)^*, (v^t)^*) = (t^*, [(z^t)^*]^+, [(z^t)^*]^-)$. The equivalence holds also with strict inequalities.

Proof. Using that $u^* + v^* = |(z^t)^*|$ and $(z^t)^* = z^t(t^*)$ the matrix $U_{\text{i-mpec}}(y^*)$ in (4.4) reads

$$U_{\text{i-mpec}}(y^*) = \begin{bmatrix} I \\ +P_{\mathcal{U}_+^t}\partial_t z^t(t^*) \\ 0 \\ -P_{\mathcal{V}_+^t}\partial_t z^t(t^*) \\ 0 \end{bmatrix} \tilde{U}(y^*),$$

where the columns of $\tilde{U}(y^*)$in (4.5) are a basis for the nullspace of

$$\begin{bmatrix} \partial_t c_{\mathcal{E}}(t, |z^t(t)|) \\ \partial_t c_{\mathcal{A}}(t, |z^t(t)|) \\ [e_i^T \partial_t z^t(t)]_{i \in \mathcal{D}^t} \end{bmatrix}.$$

Thus, $\tilde{U}(y^*) = U(t^*)$ holds. The Lagrangians of (I-MPEC) and (I-NLP), respectively, are

$$\mathcal{L}_\perp(y, \lambda) = f(t) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(t, u^t + v^t) - \lambda_{\mathcal{I}}^T c_{\mathcal{I}}(t, u^t + v^t) + \lambda_{\mathcal{Z}}^T [c_{\mathcal{Z}}(t, u^t + v^t) - (u^t - v^t)],$$
$$\mathcal{L}(t, |z^t|, \lambda) = f(t) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(t, |z^t|) - \lambda_{\mathcal{I}}^T c_{\mathcal{I}}(t, |z^t|) + \lambda_{\mathcal{Z}}^T [c_{\mathcal{Z}}(t, |z^t|) - \mathcal{P}_{(\alpha^*)^c}^T \mathcal{P}_{(\alpha^*)^c} \Sigma^* |z^t|].$$

Thus, the second term in

$$U_{\text{i-mpec}}(y^*)^T H_{\text{i-mpec}}(y^*, \lambda^*) U_{\text{i-mpec}}(y^*) = U_{\text{i-mpec}}(y^*)^T \nabla_{yy} \mathcal{L}_\perp(y^*, \lambda^*) U_{\text{i-mpec}}(y^*),$$

with $(z^t)^* = (u^t)^* - (v^t)^*$ can be rewritten as

$$\begin{bmatrix} U(t^*) \\ +P_{\mathcal{U}_+^t}\partial_t z^t(t^*)U(t^*) \\ -P_{\mathcal{V}_+^t}\partial_t z^t(t^*)U(t^*) \end{bmatrix}^T \begin{bmatrix} H_{11} & H_{12}P_{\mathcal{U}_+^t}^T & H_{12}P_{\mathcal{V}_+^t}^T \\ P_{\mathcal{U}_+^t}H_{21} & P_{\mathcal{U}_+^t}H_{22}P_{\mathcal{U}_+^t}^T & P_{\mathcal{U}_+^t}H_{22}P_{\mathcal{V}_+^t}^T \\ P_{\mathcal{V}_+^t}H_{21} & P_{\mathcal{V}_+^t}H_{22}P_{\mathcal{U}_+^t}^T & P_{\mathcal{V}_+^t}H_{22}P_{\mathcal{V}_+^t}^T \end{bmatrix} \begin{bmatrix} U(t^*) \\ +P_{\mathcal{U}_+^t}\partial_t z^t(t^*)U(t^*) \\ -P_{\mathcal{V}_+^t}\partial_t z^t(t^*)U(t^*) \end{bmatrix}$$

where $H_{ij} := \partial_i \partial_j \mathcal{L}(t^*, |(z^t)^*|, \lambda^*)$. Set $\Sigma^t = \operatorname{diag}(\sigma^t)$ with $\sigma_i^t = 1$ for $i \in \mathcal{U}_+^t$,$\sigma_i^t = -1$ for $i \in \mathcal{V}_+^t$ and $\sigma_i^t = 0$ for $i \in D^t$. Then, since $\mathcal{U}_+^t \cup \mathcal{V}_+^t = (\alpha^t)^c$, the term reads

$$\begin{bmatrix} U(t^*) \\ P_{(\alpha^t)^c}\Sigma^t \partial_t z^t(t^*)U(t^*) \end{bmatrix}^T \begin{bmatrix} H_{11} & H_{12}P_{(\alpha^t)^c}^T \\ P_{(\alpha^t)^c}H_{21} & P_{(\alpha^t)^c}H_{22}P_{(\alpha^t)^c}^T \end{bmatrix} \begin{bmatrix} U(t^*) \\ P_{(\alpha^t)^c}\Sigma^t \partial_t z^t(t^*)U(t^*), \end{bmatrix}$$

which is exactly $U_{\text{i-abs}}(t^*)^T H_{\text{i-abs}}(t^*, (z^t)^*, \lambda^*) U_{\text{i-abs}}(t^*)$. $\square$

Note that the previous theorem can be used to transfer the second order conditions for (I-NLP) and (I-MPEC) into each other. This follows from the equivalence of LIKQ and MPEC-LICQ by Theorem 4.8 and from the equivalence of stationarity concepts by Theorem 4.11.

4.5 Unconstrained Abs-Normal NLP

In this section, the unconstrained abs-normal NLP (unNLP) introduced in section 2.2 is considered. It reads

$$\begin{aligned}\min_{(x,z)\in D^{x,|z|}} \quad & f(x,|z|)\\ \text{s.t.} \quad & c_{\mathcal{Z}}(x,|z|)-z=0.\end{aligned}$$

Using the same approach as in the previous section it can be rewritten as an MPEC.

Definition 4.20 (Counterpart MPEC)**.** The counterpart MPEC of (unNLP) reads

$$\begin{aligned}\min_{(x,u,v)\in D^{x,u,v}} \quad & f(x,u+v)\\ \text{s.t.} \quad & c_{\mathcal{Z}}(x,u+v)-u-v=0,\\ & 0\le u\perp v\ge 0\end{aligned} \tag{unMPEC}$$

with $0\in D^u$ and $0\in D^v$. The feasible set is denoted by

$$\mathcal{F}_{\text{un}} := \{(x,u,v)\in D^{x,u,v} \colon c_{\mathcal{Z}}(x,u+v)-u-v=0,\ 0\le u\perp v\ge 0\}.$$

To compare KQs and MPEC-CQs, the latter are formulated for (unMPEC).

Lemma 4.21 (MPEC Constraint Qualifications)**.** *Consider a feasible point (x,u,v) of (unMPEC). Then, MPEC-LICQ is full row rank of*

$$\Big[\partial_1 c_{\mathcal{Z}}(x,u,v) \quad [\partial_2 c_{\mathcal{Z}}(x,u+v)-I]P_{\mathcal{U}_+}^T \quad [\partial_2 c_{\mathcal{Z}}(x,u+v)+I]P_{\mathcal{V}_+}^T\Big] \in \mathbb{R}^{s\times(n+|\mathcal{U}_+|+|\mathcal{V}_+|)} \tag{4.7}$$

and MPEC-MFCQ *is that the linear system*

$$\partial_1 c_{\mathcal{Z}}(x,u,v)d_x + [\partial_2 c_{\mathcal{Z}}(x,u+v)-I]P_{\mathcal{U}_+}^T d_u + [\partial_2 c_{\mathcal{Z}}(x,u+v)+I]P_{\mathcal{V}_+}^T d_v = 0 \tag{4.8}$$

admits a solution $d=(d_x,d_u,d_v)\in\mathbb{R}^n\times\mathbb{R}^{|\mathcal{U}_+|}\times\mathbb{R}^{|\mathcal{V}_+|}$ and additionally (4.7) holds.

Proof. MPEC-LICQ at (x,u,v) requires

$$\operatorname{rank}[\partial_{x,u,v}\left(c_{\mathcal{Z}}(x,u,v)-(u-v)\right)] = s$$

This is equivalent to the condition (4.7) by exploiting the structure as in Lemma 4.3. Similarly, (4.8) at (x,u,v) requires

$$\operatorname{rank}[\partial_{x,u,v}\left(c_{\mathcal{Z}}(x,u,v)-(u-v)\right)] = s$$

and the existence of a vector $(d_{\text{x}},d_{\text{u}},d_{\text{v}})\in\mathbb{R}^n\times\mathbb{R}^s\times\mathbb{R}^s$ that satisfies

$$\begin{aligned}\partial_x c_{\mathcal{Z}}(x,u,v)d_{\text{x}} + \partial_u\left(c_{\mathcal{Z}}(x,u,v)-(u-v)\right)d_{\text{u}} + \partial_v\left(c_{\mathcal{Z}}(x,u,v)-(u-v)\right)d_{\text{v}} &= 0,\\ P_{\mathcal{U}_0}d_{\text{u}} &= 0,\\ P_{\mathcal{V}_0}d_{\text{v}} &= 0.\end{aligned}$$

These are equivalent to the conditions (4.7) and (4.8) above. $\square$

Note that MPEC-LICQ and MPEC-MFCQ coincide in this setting as the system (4.8) has always the solution $d = 0$.

The equivalence between LIKQ for (unNLP) and MPEC-LICQ for (unMPEC) follows directly from Theorem 4.8.

Corollary 4.22 (Equivalence of LIKQ and MPEC-LICQ)**.** *LIKQ for* (unNLP) *holds at* $x \in D^x$ *if and only if MPEC-LICQ holds at* $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{un}$.

The structure of (unMPEC), in particular that no inequalities occur, has some remarkable consequences on relations between MPEC-CQs and KQs. As MPEC-MFCQ coincides with MPEC-LICQ which in turn is equivalent to LIKQ and stronger than MFKQ, MPEC-MFCQ is stronger than MFKQ. For an illustration see figure 4.2.

Next, one would be interested in learning of a kink qualification that is equivalent to MPEC-MFCQ such that it holds in absence of LIKQ. This is done in the next chapter and for this inequality constraints are needed.

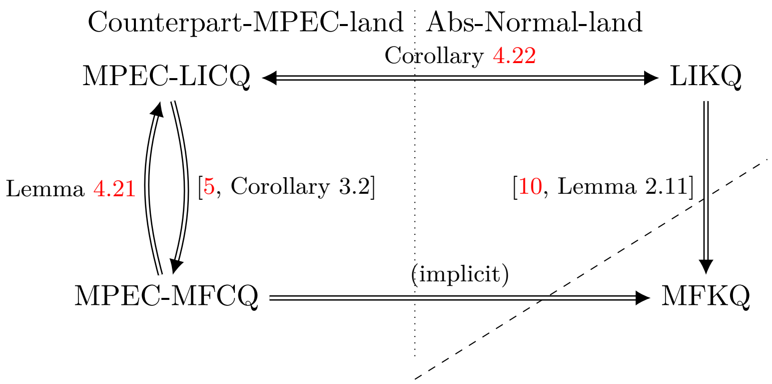

Figure 4.2: Relations between CQs and KQs for the unconstrained abs-normal NLP

Chapter 5

Optimality Conditions for Abs-Normal NLPs under Weaker KQs using MPECs

This chapter introduces weaker kink qualifactions and corresponding stationarity concepts for abs-normal NLPs using weaker constraint qualifactions and corresponding stationarity concepts for MPECs. First, kink qualifications are defined for abs-normal NLPs with a smooth objective function in section 5.1. Then, corresponding constraint qualifications for counterpart MPECs are formulated in section 5.2. In both cases this is done for the direct handling of inequalities and for the slack reformulation. Then, relations between these kink and constraint qualifications as well as the stationarity concepts are discussed in section 5.3 and section 5.6. Next, in section 5.5 it is considered if the new kink qualifications are preserved under reformulation of nonsmooth objective functions. Finally, section 5.6 deals with the weaker kink qualifications and stationarity concepts in the unconstrained case.

Part of sections 5.1-5.4 can be found in [15, 16] and parts of section 5.6 are published in [14].

5.1 Abs-Normal NLPs

In this section weaker kink qualifications than LIKQ are formulated for the problem formulations (I-NLP) and (E-NLP). Then, relations between them are considered.

Direct Handling In this paragraph the direct handling of inequalities is considered and kink qualifications are formulated. Thus, recall the problem formulation (I-NLP):

$$\begin{aligned} \min_{t,z^t} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t,|z^t|) = 0, \\ & c_{\mathcal{I}}(t,|z^t|) \geq 0, \\ & c_{\mathcal{Z}}(t,|z^t|) - z^t = 0. \end{aligned}$$

Definition 5.1 (Interior Direction Kink Qualification (IDKQ))**.** Given (I-NLP), consider $(t, z^t(t)) \in \mathcal{F}_{\text{i-abs}}$. One says that the *interior direction kink qualification (IDKQ)* holds at t if

$$\begin{bmatrix} J_{\mathcal{E}}(t) \\ J_{\alpha^t}(t) \end{bmatrix} = \begin{bmatrix} \partial_t c_{\mathcal{E}}(t,|z^t(t)|) \\ [e_i^T \partial_t z^t(t)]_{i\in\alpha^t} \end{bmatrix} \in \mathbb{R}^{(m_1+|\alpha^t|)\times n_t}$$

has full row rank $m_1 + |\alpha^t|$ and if there exists a vector $d \in \mathbb{R}^{n_t}$ such that

$$J_{\mathcal{E}}(t)d = 0, \quad J_{\alpha^t}(t)d = 0 \quad \text{and} \quad J_{\mathcal{A}}(t)d > 0.$$

For the general abs-normal NLP (I-NLP) considered here, IDKQ actually generalizes MFCQ from the smooth case and corresponds to MPEC-MFCQ, as it will be shown below. The canonical name MFKQ cannot be used, however, since Griewank and Walther have already defined MFKQ as a different weakening of LIKQ in [10] (see section 2.2.5). As other possible names like “Abs-normal MFKQ” or “Constrained MFKQ” could produce confusion rather than clarification the descriptive name “Interior Direction KQ” is suggested.

The following example from [23] (converted from MPEC form to abs-normal NLP form) shows that IDKQ is weaker than LIKQ in the presence of inequality constraints.

Example 5.2 (IDKQ is weaker than LIKQ)**.** Consider the problem

$$\begin{aligned} \min_{t \in \mathbb{R}^3, z^t \in \mathbb{R}} \quad & t_1 + t_2 - t_3 \\ \text{s.t.} \quad & t_1 + t_2 - |z^t| = 0, \\ & 4t_1 - t_3 \geq 0, \\ & 4t_2 - t_3 \geq 0, \\ & t_1 - t_2 - z^t = 0, \end{aligned}$$

with solution $t^* = (0, 0, 0)$ and $z^t(t^*) = 0$. Thus, the switching is localized and $(\alpha^t)^* = \{1\}$ as well as $(\sigma^t)^* = 0$ hold. Then, the Jacobians can be computed and read

$$\begin{aligned} J_{\mathcal{A}}(t^*) &= \partial_1 c_{\mathcal{I}}(t^*, |z^t(t^*)|) = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 4 & -1 \end{bmatrix}, \\ J_{\mathcal{E}}(t^*) &= \partial_1 c_{\mathcal{E}}(t^*, |z^t(t^*)|) = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}, \\ J_{\alpha^t}(t^*) = \partial_t z(t^*) &= \partial_1 c_{\mathcal{Z}}(t^*, |z^t(t^*)|) = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}. \end{aligned}$$

Here, LIKQ is not satisfied since the Jacobian $J_{\text{i-abs}}(t^*) = [J_{\mathcal{E}}(t^*)^T \; J_{\mathcal{A}}(t^*) \; J_{\alpha^t}(t^*)^T]^T$ cannot have full row rank by dimension. But IDKQ is satisfied as $[J_{\mathcal{E}}(t^*)^T \; J_{\alpha^t}(t^*)^T]^T$ has full row rank and $J_{\mathcal{E}}(t^*)d = J_{\alpha^t}(t^*)d = 0,\; J_{\mathcal{A}}(t^*)d = 1 > 0$ with $d = (0, 0, -1)^T$.

With the goal of considering kink qualifications in the spirit of Abadie and Guignard, the tangential cone and the abs-normal-linearized cone are defined. Here, (t, z^t) instead of $(t, z^t(t))$ is written to shorten notation. Just in the definitions of the kink qualifications $z^t(t)$ is used to emphasize the dependence of t.

Definition 5.3 (Tangential Cone and Abs-Normal-Linearized Cone for (I-NLP))**.** Consider a feasible point (t, z^t) of (I-NLP). The *tangential cone* to $\mathcal{F}_{\text{i-abs}}$ at (t, z^t) is

$$\mathcal{T}_{\text{i-abs}}(t, z^t) := \left\{ (\delta t, \delta z^t) \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\; \mathcal{F}_{\text{i-abs}} \ni (t_k, z_k^t) \to (t, z^t): \\ \tau_k^{-1}(t_k - t, z_k^t - z^t) \to (\delta t, \delta z^t) \end{array} \right\}.$$

With $\delta\zeta_i := |\delta z_i^t|$ if $i \in \alpha^t(t)$ and $\delta\zeta_i := \sigma_i^t(t)\delta z_i^t$ if $i \notin \alpha^t(t)$, the *abs-normal-linearized cone* is

$$\mathcal{T}_{\text{i-abs}}^{\text{lin}}(t, z^t) := \left\{ (\delta t, \delta z^t) \;\middle|\; \begin{array}{c} \partial_1 c_{\mathcal{E}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{E}}(t, |z^t|)\delta\zeta = 0, \\ \partial_1 c_{\mathcal{A}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{A}}(t, |z^t|)\delta\zeta \geq 0, \\ \partial_1 c_{\mathcal{Z}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{Z}}(t, |z^t|)\delta\zeta = \delta z^t \end{array} \right\}.$$

Note that in $\mathcal{T}_{\text{i-abs}}$ the relation $\delta z^t = z(\delta t)$ holds due to the dependence $z^t = z^t(t)$ and the continuity of $c_{\mathcal{Z}}$. Whereas, such a relation does not hold in $\mathcal{T}_{\text{i-abs}}^{\text{lin}}$.

The proof that the tangential cone is a subset of the abs-normal-linearized cone follows the idea for MPECs presented in subsection 2.3.2 and originally given in [5]. First, the definition of the smooth branch NLPs for (I-NLP) with their standard tangential cones and linearized cones is needed.

Definition 5.4 (Branch NLPs for (I-NLP))**.** Consider a feasible point $(\hat{t}, \hat{z}^t)$ of (I-NLP). Choose $\sigma^t \in \{-1, 1\}^{s_t}$ with $\sigma^t \succeq \sigma^t(\hat{t})$ and set $\Sigma^t = \operatorname{diag}(\sigma^t)$. The branch problem $\text{NLP}(\Sigma^t)$ is defined as

$$\begin{aligned} \min_{(t,z^t) \in D^{t,|z^t|}} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t, \Sigma^t z^t) = 0, \\ & c_{\mathcal{I}}(t, \Sigma^t z^t) \geq 0, \\ & c_{\mathcal{Z}}(t, \Sigma^t z^t) - z^t = 0, \\ & \Sigma^t z^t \geq 0. \end{aligned} \tag{NLP(Σ^t)}$$

The feasible set of (NLP(Σ^t)), which always contains $(\hat{t}, \hat{z}^t)$, is denoted by

$$\mathcal{F}_{\Sigma^t} := \left\{ (t, z^t) \in D^{t,|z^t|} \;\middle|\; \begin{array}{c} c_{\mathcal{E}}(t, \Sigma^t z^t) = 0, \;\; c_{\mathcal{I}}(t, \Sigma^t z^t) \geq 0, \\ c_{\mathcal{Z}}(t, \Sigma^t z^t) - z^t = 0, \;\; \Sigma^t z^t \geq 0 \end{array} \right\}$$

Note that in contrast to chapter 3 z^t instead of only z_+^t is considered. Nevertheless, this does not change the branch problem as for $i \notin \alpha^t(\hat{t})$ additional $\sigma_i^t z_i^t \geq 0$ is required with identical σ_i^t for all branch problems.

Lemma 5.5 (Tangential Cone and Linearized Cone for (NLP(Σ^t)))**.** *Given* (NLP(Σ^t)), *consider a feasible point* (t, z^t)*. The tangential cone to* $\mathcal{F}_{\Sigma^t}$ *at* (t, z^t) *reads*

$$\mathcal{T}_{\Sigma^t}(t, z^t) := \left\{ (\delta t, \delta z^t) \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0, \;\; \mathcal{F}_{\Sigma^t} \ni (t_k, z_k^t) \to (t, z^t): \\ \tau_k^{-1}(t_k - t, z_k^t - z^t) \to (\delta t, \delta z^t) \end{array} \right\}.$$

The linearized cone reads

$$\mathcal{T}_{\Sigma^t}^{lin}(t, z^t) := \left\{ (\delta t, \delta z^t) \;\middle|\; \begin{array}{r} \partial_1 c_{\mathcal{E}}(t, \Sigma^t z^t)\delta t + \partial_2 c_{\mathcal{E}}(t, \Sigma^t z^t)\Sigma^t \delta z^t = 0, \\ \partial_1 c_{\mathcal{A}}(t, \Sigma^t z^t)\delta t + \partial_2 c_{\mathcal{A}}(t, \Sigma^t z^t)\Sigma^t \delta z^t \geq 0, \\ \partial_1 c_{\mathcal{Z}}(t, \Sigma^t z^t)\delta t + \partial_2 c_{\mathcal{Z}}(t, \Sigma^t z^t)\Sigma^t \delta z^t = \delta z^t, \\ \sigma_i^t \delta z_i^t \geq 0, \;\; i \in \alpha^t(t) \end{array} \right\}.$$

Proof. Every branch problem (NLP(Σ^t)) is smooth and thus Definition 2.14 can be applied. $\square$

Remark 5.6. Observe that $|z^t| = \Sigma^t z^t$ in Definitions 5.4 and 5.5. Thus, for every Σ^t the relations $\mathcal{F}_{\Sigma^t} \subseteq \mathcal{F}_{\text{i-abs}}$, $\mathcal{T}_{\Sigma^t}(t, z^t) \subseteq \mathcal{T}_{\text{i-abs}}(t, z^t)$ and $\mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$ hold.

Lemma 5.7. *Consider a feasible point* $(\hat{t}, \hat{z}^t)$ *of* (I-NLP) *with associated branch problems* (NLP(Σ^t)). *Then, the following decompositions of the tangential cone and of the abs-normal-linearized cone of* (I-NLP) *hold:*

$$\mathcal{T}_{i\text{-}abs}(\hat{t}, \hat{z}^t) = \bigcup_{\Sigma^t} \mathcal{T}_{\Sigma^t}(\hat{t}, \hat{z}^t) \quad \textit{and} \quad \mathcal{T}^{lin}_{i\text{-}abs}(\hat{t}, \hat{z}^t) = \bigcup_{\Sigma^t} \mathcal{T}^{lin}_{\Sigma^t}(\hat{t}, \hat{z}^t).$$

Proof. First, the tangential cones are considered and it is shown that a neighborhood $\mathcal{N}$ of $(\hat{t}, \hat{z}^t)$ exists such that

$$\mathcal{F}_{\text{i-abs}} \cap \mathcal{N} = \bigcup_{\Sigma^t} (\mathcal{F}_{\Sigma^t} \cap \mathcal{N}).$$

The inclusion $\supseteq$ holds for every neighborhood $\mathcal{N}$ since $\mathcal{F}_{\Sigma^t} \subseteq \mathcal{F}_{\text{i-abs}}$ for all Σ^t. To show the inclusion $\subseteq$ consider an index $i \notin \alpha^t(\hat{t})$. Then, by continuity, $\epsilon_i > 0$ exists with $\sigma_i(t) = \sigma_i^t(t) \in \{-1, +1\}$ for all $t \in B_{\epsilon_i}(\hat{t})$. Now set $\epsilon := \min_{i \notin \alpha^t(\hat{t})} \epsilon_i$, $\mathcal{N} := B_\epsilon \times \mathbb{R}^{n_t}$ and consider $(t, z^t) \in \mathcal{N} \cap \mathcal{F}_{\text{i-abs}}$. With the choice $\sigma_i^t = \sigma_i^t(t)$ for $i \notin \alpha^t(t)$ and $\sigma_i^t = 1$ for $i \in \alpha^t(t)$ one finds $\Sigma^t = \operatorname{diag}(\sigma^t)$ such that $(t, z^t) \in \mathcal{N} \cap \mathcal{F}_{\Sigma^t}$ since $\alpha^t(t) \subseteq \alpha^t(\hat{t})$. Thus,

$$\mathcal{F}_{\text{i-abs}} \cap \mathcal{N} = \bigcup_{\Sigma^t} (\mathcal{F}_{\Sigma^t} \cap \mathcal{N}).$$

Now, let $\mathcal{T}(\hat{t}, \hat{z}^t; \mathcal{F})$ generically denote the tangential cone to $\mathcal{F}$ at $(\hat{t}, \hat{z}^t)$. Then,

$$\begin{aligned} \mathcal{T}_{\text{i-abs}}(\hat{t}, \hat{z}^t) = \mathcal{T}(\hat{t}, \hat{z}^t; \mathcal{F}_{\text{i-abs}}) &= \mathcal{T}(\hat{t}, \hat{z}^t; \mathcal{F}_{\text{i-abs}} \cap \mathcal{N}) = \mathcal{T}(\hat{t}, \hat{z}^t; \textstyle\bigcup_{\Sigma^t} (\mathcal{F}_{\Sigma^t} \cap \mathcal{N})) \\ &= \bigcup_{\Sigma^t} \mathcal{T}(\hat{t}, \hat{z}^t; \mathcal{F}_{\Sigma^t} \cap \mathcal{N}) = \bigcup_{\Sigma^t} \mathcal{T}(\hat{t}, \hat{z}^t; \mathcal{F}_{\Sigma^t}) = \bigcup_{\Sigma^t} \mathcal{T}_{\Sigma^t}(\hat{t}, \hat{z}^t). \end{aligned}$$

Here the fourth equality holds since the number of branch problems is finite. The decomposition of $\mathcal{T}^{\text{lin}}_{\text{i-abs}}$ follows directly by comparing definitions of $\mathcal{T}^{\text{lin}}_{\text{i-abs}}$ and $\mathcal{T}^{\text{lin}}_{\Sigma^t}$. $\square$

Lemma 5.8. *Let* (t, z^t) *be feasible for* (I-NLP). *Then,*

$$\mathcal{T}_{i\text{-}abs}(t, z^t) \subseteq \mathcal{T}^{lin}_{i\text{-}abs}(t, z^t) \quad \textit{and} \quad \mathcal{T}_{i\text{-}abs}(t, z^t)^* \supseteq \mathcal{T}^{lin}_{i\text{-}abs}(t, z^t)^*.$$

Proof. The branch NLPs are smooth, hence the inclusion $\mathcal{T}_{\Sigma^t}(t, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t)$ holds by standard NLP theory. Then, the first result follows directly from Lemma 5.7 and the second result follows by dualization. $\square$

In general, the reverse inclusions do not hold. This leads to the following definitions.

Definition 5.9 (Abadie's Kink Qualification (AKQ) for (I-NLP))**.** Consider a feasible point $(t, z^t(t))$ of (I-NLP). One says that *Abadie's Kink Qualification (AKQ)* holds for (I-NLP) at t if $\mathcal{T}_{\text{i-abs}}(t, z^t(t)) = \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t(t))$.

Definition 5.10 (Guignard's Kink Qualification (GKQ) for (I-NLP))**.** Consider a feasible point $(t, z^t(t))$ of (I-NLP). One says that *Guignard's Kink Qualification (GKQ)* holds for (I-NLP) at t if $\mathcal{T}_{\text{i-abs}}(t, z^t(t))^* = \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t(t))^*$.

The decomposition in Lemma 5.7 leads to the next two results. As the branch problems are smooth, ACQ and GCQ as introduced in subsection 2.1.4 are considered. Thus, they are given at a tuple $(t, z^t(t))$ in contrast to AKQ and GKQ which are given at t.

Theorem 5.11 (ACQ for all (NLP(Σ^t)) implies AKQ for (I-NLP))**.** *Consider a feasible point $(t, z^t(t))$ of (I-NLP) with associated branch problems (NLP(Σ^t)). Then, AKQ holds for (I-NLP) at t if ACQ holds for all (NLP(Σ^t)) at $(t, z^t(t))$.*

Proof. This follows directly from Lemma 5.7. □

Theorem 5.12 (GCQ for all (NLP(Σ^t)) implies GKQ for (I-NLP))**.** *Consider a feasible point $(t, z^t(t))$ of (I-NLP) with associated branch problems (NLP(Σ^t)). Then, GKQ holds for (I-NLP) at t if GCQ holds for all (NLP(Σ^t)) at $(t, z^t(t))$.*

Proof. This follows directly from Lemma 5.7 because it gives the relations

$$\mathcal{T}_{\text{i-abs}}(t, z^t(t))^* = \bigcap_{\Sigma^t} \mathcal{T}_{\Sigma^t}(t, z^t(t))^* \quad \text{and} \quad \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t(t))^* = \bigcap_{\Sigma^t} \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t(t))^*$$

by dualization. □

Inequality Slacks In this paragraph the kink qualifications which were introduced in the previous paragraph are formulated for (E-NLP), which is

$$\begin{aligned} \min_{t,w,z^t,z^w} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t, |z^t|) = 0, \\ & c_{\mathcal{I}}(t, |z^t|) - |z^w| = 0, \\ & c_{\mathcal{Z}}(t, |z^t|) - z^t = 0, \\ & w - z^w = 0. \end{aligned}$$

To this end, its formulation as ($\overline{\text{E-NLP}}$) is needed:

$$\begin{aligned} \min_{x,z} \quad & f(x) \\ \text{s.t.} \quad & \bar{c}_{\mathcal{E}}(x, |z|) = 0, \\ & \bar{c}_{\mathcal{Z}}(x, |z|) - z = 0, \end{aligned}$$

with $x = (t, w)$, $z = (z^t, z^w)$, $\bar{f}(x) = f(t)$, $\bar{c}_{\mathcal{E}}(x, |z|) = (c_{\mathcal{E}}(t, |z^t|), c_{\mathcal{I}}(t, |z^t|) - |z^w|)$ and $\bar{c}_{\mathcal{Z}}(x, |z|) = (c_{\mathcal{Z}}(t, |z^t|), w)$. Thus, (E-NLP) can be seen as a special case of (I-NLP) and the following set of lemmas follow from the results in the previous section.

Lemma 5.13 (IDKQ for (E-NLP))**.** *Given (E-NLP), consider $(t, w, z^t(t), z^w(w)) \in \mathcal{F}_{e\text{-}abs}$. Then, IDKQ at (t, w) is LIKQ at (t, w).*

Proof. By construction, (E-NLP) and thus ($\overline{\text{E-NLP}}$) contain no inequalities. Hence, definition 5.1 applied to ($\overline{\text{E-NLP}}$) reduces to full row rank of $J_{\text{e-abs}}(x) = [\bar{J}_{\mathcal{E}}(x)^T \ \bar{J}_{\alpha}(x)^T]^T$ which is exactly LIKQ for (E-NLP) (see Lemma 3.70). □

This lemma stands in contrast to the standard reformulation of smooth NLP inequalities as equalities with nonnegative slacks where the validity of LICQ and MFCQ are both unaffected.

Next, Abadie's and Guignard's KQ for (E-NLP) will be defined. For that, the tangential and the abs-normal-linearized cone are needed which will be formulated for (E-NLP) in the next lemma.

Lemma 5.14 (Tangential Cone and Abs-Normal-Linearized Cone for (E-NLP))**.** *The tangential cone to $\mathcal{F}_{\text{e-abs}}$ at (t, w, z^t, z^w) reads*

$$\mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w) = \left\{ \delta \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\text{e-abs}} \ni (t_k, w_k, z_k^t, z_k^w) \to (t, w, z^t, z^w): \\ \tau_k^{-1}(t_k - t, w_k - w, z_k^t - z^t) \to (\delta t, \delta w, \delta z^t),\ \delta z^w = \delta w \end{array} \right\}$$

with $\delta = (\delta t, \delta w, \delta z^t, \delta z^w) \in \mathbb{R}^{n_t+m_2+s_t+m_2}$ and the abs-normal-linearized cone reads

$$\mathcal{T}^{lin}_{\text{e-abs}}(t, w, z^t, z^w) = \left\{ \delta \;\middle|\; \begin{array}{l} \partial_1 c_{\mathcal{I}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{I}}(t, |z^t|)\delta\zeta = \delta\omega, \\ (\delta t, \delta z^t) \in \mathcal{T}^{lin}_{\text{i-abs}}(t, z^t),\ \delta z^w = \delta w \end{array} \right\},$$

with $\delta = (\delta t, \delta w, \delta z^t, \delta z^w) \in \mathbb{R}^{n_t+m_2+s_t+m_2}$ and

$$\delta\zeta_i = \begin{cases} \sigma_i(t)\delta z_i^t, & i \notin \alpha^t(t), \\ |\delta z_i^t|, & i \in \alpha^t(t), \end{cases} \qquad \delta\omega_i = \begin{cases} \sigma_i(w)\delta z_i^w, & i \notin \alpha^w(w), \\ |\delta z_i^w|, & i \in \alpha^w(w), \end{cases}$$

where $\alpha(t, w) = (\alpha^t(t), \alpha^w(w))$.

Proof. Definition 5.3 can be applied to (E-NLP) and gives using its definition

$$\mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w) = \left\{ \delta \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\text{e-abs}} \ni (t_k, w_k, z_k^t, z_k^w) \to (t, w, z^t, z^w): \\ \tau_k^{-1}(t_k - t, w_k - w, z_k^t - z^t, z_k^w - z^w) \to (\delta t, \delta w, \delta z^t, \delta z^w) \end{array} \right\}$$

as well as

$$\mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t, z^w) = \left\{ \delta \;\middle|\; \begin{array}{r} \partial_1 c_{\mathcal{E}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{E}}(t, |z^t|)\delta\zeta = 0, \\ \partial_1 c_{\mathcal{I}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{I}}(t, |z^t|)\delta\zeta = \delta\omega, \\ \partial_1 c_{\mathcal{Z}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{Z}}(t, |z^t|)\delta\zeta = \delta z^t, \\ \delta w = \delta z^w \end{array} \right\}.$$

Using, that $w_k = z_k^w$ and $w = z^w$ by definition of $\mathcal{F}_{\text{e-abs}}$ and the definition of $\mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$ gives the presentation above. □

Lemma 5.15 (Branch NLPs for (E-NLP))**.** *Given a feasible point $(\hat{t}, \hat{w}, \hat{z}^t, \hat{z}^w)$ of (E-NLP). Choose $\sigma^t \in \{-1, 1\}^{s_t}$ with $\sigma^t \succeq \sigma^t(\hat{t})$, $\sigma^w \in \{-1, 1\}^{m_2}$ with $\sigma^w \succeq \sigma^t(\hat{w})$ and set $\Sigma^t = \operatorname{diag}(\sigma^t)$, $\Sigma^w = \operatorname{diag}(\sigma^w)$. The branch problem NLP($\Sigma^{t,w}$) for $\Sigma^{t,w} := \operatorname{diag}(\sigma^t, \sigma^w)$ reads*

$$\begin{aligned} \min_{(t, w, z^t, z^w) \in D^{t, w, |z^t|, |z^w|}} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t, \Sigma^t z^t) = 0, \\ & c_{\mathcal{I}}(t, \Sigma^t z^t) - \Sigma^w z^w = 0, \\ & c_{\mathcal{Z}}(t, \Sigma^t z^t) - z^t = 0, \\ & w - z^w = 0, \\ & \Sigma^t z^t \geq 0, \\ & \Sigma^w z^w \geq 0. \end{aligned} \tag{NLP($\Sigma^{t,w}$)}$$

The feasible set of (NLP($\Sigma^{t,w}$)), *which always contains* $(\hat{t}, \hat{w}, \hat{z}^t, \hat{z}^w)$ *and is a lifting of* $\mathcal{F}_{\Sigma^t}$, *is denoted by*

$$\mathcal{F}_{\Sigma^{t,w}} = \left\{ (t, w, z^t, z^w) \in D^{t,w,|z^t|,|z^w|} \;\middle|\; \begin{array}{l} (t, z^t) \in \mathcal{F}_{\Sigma^t}, c_{\mathcal{I}}(t, \Sigma^t z^t) - \Sigma^w z^w = 0, \\ w - z^w = 0,\ \Sigma^w z^w \geq 0 \end{array} \right\}.$$

Proof. Definition 5.4 can be applied to (E-NLP) and gives the branch problem above with feasible set

$$\mathcal{F}_{\Sigma^{t,w}} = \left\{ (t, w, z^t, z^w) \in D^{t,w,|z^t|,|z^w|} \;\middle|\; \begin{array}{l} c_{\mathcal{E}}(t, \Sigma^t z^t) = 0,\ c_{\mathcal{I}}(t, \Sigma^t z^t) - \Sigma^w z^w = 0, \\ c_{\mathcal{Z}}(t, \Sigma^t z^t) - z^t = 0,\ w - z^w = 0, \\ \Sigma^t z^t \geq 0,\ \Sigma^w z^w \geq 0 \end{array} \right\}.$$

Using the definition of $\mathcal{F}_{\Sigma^t}$ it has the form stated above. $\square$

Lemma 5.16 (Tangential and Linearized Cone for (NLP($\Sigma^{t,w}$))). *Given* (NLP($\Sigma^{t,w}$)), *consider a feasible point* (t, w, z^t, z^w). *The tangential cone to* $\mathcal{F}_{\Sigma^{t,w}}$ *at* (t, w, z^t, z^w) *reads*

$$\mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w) = \left\{ \delta \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\Sigma^{t,w}} \ni (t_k, w_k, z^t_k, z^w_k) \to (t, w, z^t, z^w): \\ \tau_k^{-1}(t_k - t, w_k - w, z^t_k - z^t) \to (\delta t, \delta w, \delta z^t),\ \delta z^w = \delta w \end{array} \right\}.$$

The linearized cone reads

$$\mathcal{T}^{lin}_{\Sigma^{t,w}}(t, w, z^t, z^w) = \left\{ \delta \;\middle|\; \begin{array}{l} \partial_1 c_{\mathcal{I}} \delta t + \partial_2 c_{\mathcal{I}} \Sigma^t \delta z^t - \Sigma^w \delta z^w = 0,\ \delta z^w = \delta w, \\ (\delta t, \delta z^t) \in \mathcal{T}^{lin}_{\Sigma^t}(t, z^t),\ \sigma^w_i \delta z^w_i \geq 0,\ i \in \alpha^w(w) \end{array} \right\}.$$

Here, $\delta = (\delta t, \delta w, \delta z^t, \delta z^w) \in \mathbb{R}^{n_t + m_2 + s_t + m_2}$ *and all partial derivatives are evaluated at* $(t, \Sigma^t z^t)$.

Proof. Definition 5.5 can be applied to (E-NLP). Then, the cones read

$$\mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w) = \left\{ \delta \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\Sigma^{t,w}} \ni (t_k, w_k, z^t_k, z^w_k) \to (t, w, z^t, z^w): \\ \tau_k^{-1}(t_k - t, w_k - w, z^t_k - z^t), z^w_k - z^w \to (\delta t, \delta w, \delta z^t, \delta z^w) \end{array} \right\}$$

and

$$\mathcal{T}^{\mathrm{lin}}_{\Sigma^{t,w}}(t, w, z^t, z^w) = \left\{ \delta \;\middle|\; \begin{array}{r} \partial_1 c_{\mathcal{E}} \delta t + \partial_2 c_{\mathcal{E}} \Sigma^t \delta z^t = 0, \\ \partial_1 c_{\mathcal{I}} \delta t + \partial_2 c_{\mathcal{I}} \Sigma^t \delta z^t - \Sigma^w \delta z^w = 0, \\ \partial_1 c_{\mathcal{Z}} \delta t + \partial_2 c_{\mathcal{Z}} \Sigma^t \delta z^t = \delta z^t, \\ \delta w = \delta z^w, \\ \sigma^t_i \delta z^t_i \geq 0,\ i \in \alpha^t(t),\ \sigma^w_i \delta z^w_i \geq 0,\ i \in \alpha^w(w) \end{array} \right\}.$$

Then, $w_k = z^w_k$ and $w = z^w$ hold by definition of $\mathcal{F}_{\Sigma^{t,w}}$ and the definition of $\mathcal{T}^{\mathrm{lin}}_{\Sigma^t}(t, z^t)$ can be inserted. This leads to the stated form of the two cones. $\square$

Moreover, the following decompositions are obtained by applying Lemma 5.7 to (E-NLP):

$$\mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w) = \bigcup_{\Sigma^{t,w}} \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w), \quad \mathcal{T}^{\mathrm{lin}}_{\text{e-abs}}(t, w, z^t, z^w) = \bigcup_{\Sigma^{t,w}} \mathcal{T}^{\mathrm{lin}}_{\Sigma^{t,w}}(t, w, z^t, z^w).$$

As before, the tangential cone is a subset of the linearized cone and the reverse inclusion holds for the dual cones:

$$\mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w) \subseteq \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t, z^w), \quad \mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w)^* \supseteq \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t, z^w)^*.$$

This follows directly by applying Lemma 5.8 to (E-NLP).

Again, equality does not hold in general. Thus, Abadie's Kink Qualification (AKQ) and Guignard's Kink Qualification (GKQ) for (E-NLP) are considered.

Lemma 5.17 (AKQ for (E-NLP))**.** *Given (E-NLP), consider* $(t, w, z^t(t), z^w(w)) \in \mathcal{F}_{\text{e-abs}}$*. Then, AKQ for (E-NLP) at* (t, w) *is* $\mathcal{T}_{\text{e-abs}}(t, w, z^t(t), z^w(w)) = \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t(t), z^w(w))$*.*

Proof. This follows directly by applying Definition 5.9 to (E-NLP). □

Lemma 5.18 (GKQ for (E-NLP))**.** *Given (E-NLP), consider* $(t, w, z^t(t), z^w(w)) \in \mathcal{F}_{\text{e-abs}}$*. Then, GKQ for (E-NLP) at* (t, w) *is* $\mathcal{T}_{\text{e-abs}}(t, w, z^t(t), z^w(w))^* = \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t(t), z^w(w))^*$*.*

Proof. This follows directly by applying Definition 5.10 to (E-NLP). □

The possible slack values $w \in W(t) := \{w \in \mathbb{R}^{m_2} : |w| = c_{\mathcal{I}}(t, |z^t(t)|)\}$ just differ by the signs of components w_i for $i \notin \mathcal{A}(t)$. Thus, homeomorphism between the cones for different choices of slack values exist as is shown in the next lemma.

Lemma 5.19. *Given (E-NLP), consider* $(t, w, z^t(t), z^w(w)) \in \mathcal{F}_{\text{e-abs}}$ *and* $\tilde{w} \in W(t)$, $\tilde{w} \neq w$*. Define* $\chi \colon \mathcal{T}_{\text{e-abs}}(t, w, z^t(t), z^w(w)) \to \mathcal{T}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))$, $\chi \colon \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t(t), z^w(w)) \to \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))$ *as*

$$\chi(\delta t, \delta w, \delta z^t, \delta z^w) = (\delta t, \Sigma \delta w, \delta z^t, \Sigma \delta z^w) \text{ and } \chi^{-1}(\delta t, \tilde{\delta} w, \delta z^t, \tilde{\delta} z^w) = (\delta t, \Sigma \tilde{\delta} w, \delta z^t, \Sigma \tilde{\delta} z^w),$$

where $\Sigma = \operatorname{diag}(\sigma)$ *with* $\sigma_i = -1$ *if* $w_i \neq \tilde{w}_i$ *and* $\sigma_i = 1$ *if* $w_i = \tilde{w}_i$*. Then, both functions* χ *are homeomorphisms.*

Proof. Set $\tilde{w}_k = \Sigma w_k$ and $\tilde{z}^w_k = \Sigma z^w_k$ for a given vector $(\delta t, \delta w, \delta z^t, \delta z^w) = \lim \tau_k^{-1}(t_k - t, w_k - w, z^t_k - \hat{z}^t, z^t_w - \hat{z}^w) \in \mathcal{T}_{\text{e-abs}}(t, w, z^t(t), z^w(w))$ to obtain

$$\lim \frac{\tilde{w}_k - \tilde{w}}{\tau_k} = \Sigma \lim \frac{w_k - w}{\tau_k} = \Sigma \delta w, \; \lim \frac{\tilde{z}^w_k - \tilde{z}^w}{\tau_k} = \Sigma \lim \frac{z^w_k - z^w}{\tau_k} = \Sigma \delta z^w.$$

This implies that $\chi(\delta) = (\delta t, \Sigma \delta w, \delta z^t, \Sigma \delta z^w) \in \mathcal{T}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))$. Conversely, the same argument holds with interchanged variables. Further, χ and χ^{-1} are continuous by definition and inverse to each other by definition of Σ.

Now, let $\delta = (\delta t, \delta w, \delta z^t, \delta z^w) \in \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, \hat{z}^t, \hat{z}^w)$, then $\chi(\delta) = (\delta t, \Sigma \delta w, \delta z^t, \Sigma \delta z^w) \in \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, \hat{z}^t, \tilde{z}^w)$ as $\alpha^w(w) = \alpha^w(\tilde{w})$, $\sigma_i(w)\delta z^w_i = \sigma_i(\tilde{w})\sigma_i \delta z^w_i$ for $i \notin \alpha^w(\tilde{w})$ and $|\delta z^w_i| = |\sigma_i \delta z^w_i|$ for $i \in \alpha^w(\tilde{w})$. Again, the same argument holds with interchanged variables. This gives the converse direction. □

Using these homeomorphisms, one can prove next that neither AKQ nor GKQ depend on the particular choice of w. Thus, both conditions above are well-defined for (E-NLP).

Lemma 5.20. *Consider a feasible point* $(t, w, z^t(t), z^w(w))$ *of* (E-NLP)*. AKQ holds at* $(t, \tilde{w})$ *for arbitrary* $\tilde{w} \in W(t)$*,* $\tilde{w} \neq w$ *if AKQ holds at* (t, w)*.*

Proof. One has to show

$$\mathcal{T}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w})) = \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))$$

This follows directly from the homeomorphism χ in Lemma 5.19. □

Lemma 5.21. *Consider a feasible point* $(t, w, z^t(t), z^w(w))$ *of* (E-NLP)*. GKQ holds at* $(t, \tilde{w})$ *for arbitrary* $\tilde{w} \in W(t)$*,* $\tilde{w} \neq w$ *if GKQ holds at* (t, w)*.*

Proof. As $\mathcal{T}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))^* \supseteq \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))^*$ is always satisfied, it is left to show that

$$\mathcal{T}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))^* \subseteq \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))^*.$$

Consider $\tilde{\omega} = (\omega t, \tilde{\omega} w, \omega z^t, \tilde{\omega} z^w) \in \mathcal{T}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))^*$, i.e. $\tilde{\omega}^T \tilde{\delta} \geq 0$ for all $\tilde{\delta} \in \mathcal{T}_{\text{e-abs}}(t, \tilde{w}, z^t(\tilde{t}), z^w(\tilde{w}))^*$. Then, set $\omega = (\omega t, \Sigma \tilde{\omega} w, \omega z^t, \Sigma \tilde{\omega} w)$ and thus $\omega^T \delta = \omega^T \chi^{-1}(\tilde{\delta}) = \tilde{\omega} \tilde{\delta} \geq 0$ for all $\delta \in \mathcal{T}_{\text{e-abs}}(t, w, z^t(t), z^w(w))$. This is $\omega \in \mathcal{T}_{\text{e-abs}}(t, w, z^t(t), z^w(w))^*$ and further $\omega \in \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t(t), z^w(w))^*$ by the assumption. Hence, $\tilde{\omega}^T \tilde{\delta} = \omega^T \delta \geq 0$ for all $\mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))$ which means $\omega \in \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, \tilde{w}, z^t(t), z^w(\tilde{w}))^*$. □

As before, AKQ or GKQ are implied if ACQ or GCQ hold for all branch problems.

Theorem 5.22 (ACQ for all (NLP($\Sigma^{t,w}$)) implies AKQ for (E-NLP))**.** *Consider a feasible point* $(t, w, z^t(t), z^w(w))$ *of* (E-NLP) *with associated branch problems* (NLP($\Sigma^{t,w}$))*. Then, AKQ holds for* (E-NLP) *at* (t, w) *if ACQ holds for all* (NLP($\Sigma^{t,w}$)) *at* $(t, w, z^t(t), z^w(w))$*.*

Proof. This follows directly by applying Theorem 5.11 to ($\overline{\text{E-NLP}}$). □

Theorem 5.23 (GCQ for all (NLP($\Sigma^{t,w}$)) implies GKQ for (E-NLP))**.** *Consider a feasible point* $(t, w, z^t(t), z^w(w))$ *of* (E-NLP) *with associated branch problems* (NLP($\Sigma^{t,w}$))*. Then, GKQ holds for* (E-NLP) *at* (t, w) *if GCQ holds for all* (NLP($\Sigma^{t,w}$)) *at* $(t, w, z^t(t), z^w(w))$*.*

Proof. This follows directly by applying Theorem 5.12 to ($\overline{\text{E-NLP}}$). □

Relations of Kink Qualifications In this paragraph the previous defined kink qualifications for both formulations will be compared.

Theorem 5.24. *IDKQ for* (I-NLP) *holds at* $(t, z^t(t)) \in \mathcal{F}_{i\text{-}abs}$ *if IDKQ for* (E-NLP) *holds at* $(t, w, z^t(t), z^w(w)) \in \mathcal{F}_{e\text{-}abs}$ *for any (and hence all)* $w \in W(t)$*. The converse is not true.*

Proof. Since (E-NLP) has no inequalities, the concepts of IDKQ and LIKQ coincide. LIKQ for (E-NLP) is equivalent to LIKQ for (I-NLP) by Theorem 3.72 and LIKQ for (I-NLP) implies IDKQ for (I-NLP). The converse does not hold since LIKQ for (I-NLP) is stronger then IDKQ as was shown in Example 5.2. □

Theorem 5.25. *AKQ for* (I-NLP) *holds at* $(t, z^t(t)) \in \mathcal{F}_{i\text{-}abs}$ *if and only if AKQ for* (E-NLP) *holds at* $(t, w, z^t(t), z^w(w)) \in \mathcal{F}_{e\text{-}abs}$ *for any (and hence all)* $w \in W(t)$*.*

Proof. As $\mathcal{T}_{\text{i-abs}}(t, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$ and $\mathcal{T}_{\text{e-abs}}(t, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, z^t)$ always hold, one just need to prove

$$\mathcal{T}_{\text{i-abs}}(t, z^t) \supseteq \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t) \iff \mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w) \supseteq \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t, z^w).$$

The implication "$\Rightarrow$" is shown first. Let $\delta = (\delta t, \delta w, \delta z^t, \delta z^w) \in \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t, z^w)$. Then, $\tilde{\delta} = (\delta t, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$ and by the assumption $\tilde{\delta} = \mathcal{T}_{\text{i-abs}}(t, z^t)$ holds. Hence, there exist sequences $(t_k, z^t_k) \in \mathcal{F}_{\text{i-abs}}$ and $\tau_k \searrow 0$ with $(t_k, z^t_k) \to (t, z^t)$ and $\tau_k^{-1}(t_k - t, z^t_k - z^t) \to (\delta t, \delta z^t)$. Now, define

$$\Sigma^w = \operatorname{diag}(\sigma) \quad \text{with} \quad \sigma_i = \begin{cases} \sigma^w(w_i), & i \notin \alpha^w(w), \\ \operatorname{sign}(\delta z^w_i), & i \in \alpha^w(w) \end{cases}$$

and set $z^w_k := w_k := \Sigma^w c_{\mathcal{I}}(t_k, |z^t_k|)$. Then, $z^w = w = \Sigma^w c_{\mathcal{I}}(t, |z^t|)$ holds which leads to

$$\begin{aligned} z^w_k - z^w &= \Sigma^w[c_{\mathcal{I}}(t_k, |z^t_k|) - c_{\mathcal{I}}(t, |z^t|)] \\ &= \Sigma^w[\partial_1 c_{\mathcal{I}}(t, |z^t|)(t_k - t) + \partial_2 c_{\mathcal{I}}(t, |z^t|)(|z^t_k| - |z^t|) + o(\|(t_k - t, |z^t_k| - |z^t|)\|)]. \end{aligned}$$

Further, for k large enough $|z^t_k| - |z^t| = \Sigma^t_k z^t_k - \Sigma^t z^t$ holds using $\Sigma^t_k = \operatorname{diag}(\sigma^t_k)$ with $\sigma^t_k = \sigma(t_k)$ and $\Sigma^t = \operatorname{diag}(\sigma^t)$ with $\sigma^t = \sigma(t)$. Then, for $z^t_i \neq 0$

$$\tau_k^{-1}(|(z^t_k)_i| - |z^t_i|) = \tau_k^{-1}\sigma^t_i((z^t_k)_i - z^t_i) \to \sigma^t_i \delta z^t_i$$

can be obtained. For $z^t_i = 0$ one has $\tau_k^{-1}(z^t_k)_i \to \delta z^t_i$ and hence

$$\tau_k^{-1}(|(z^t_k)_i| - |z^t_i|) = \tau_k^{-1}|(z^t_k)_i| \to |\delta z^t_i|.$$

Thus, $\tau_k^{-1}(|(z^t_k)| - |z^t|) \to \delta\zeta$ holds and in total

$$\tau_k^{-1}(z^w_k - z^w) \to \Sigma^w[\partial_1 c_{\mathcal{I}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{I}}(t, |z^t|)\delta\zeta] = \Sigma^w \delta\zeta = \delta z^w.$$

Additionally, one obtains $\tau_k^{-1}(w_k - w) \to \delta w$ and finally $d \in \mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w)$. To prove the implication "$\Leftarrow$", consider $\delta = (\delta t, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$. Define

$$\Sigma^w = \operatorname{diag}(\sigma) \quad \text{with} \quad \sigma_i = \begin{cases} \pm 1, & i \in \mathcal{A}(t), \\ \operatorname{sign}([\partial_1 c_{\mathcal{I}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{I}}(t, |z^t|)\delta\zeta]_i), & i \notin \mathcal{A}(t) \end{cases}$$

and set $\delta w = \delta z^w = \Sigma^w[\partial_1 c_{\mathcal{I}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{I}}(t, |z^t|)\delta\zeta]$. Then, $\tilde{\delta} = (\delta t, \delta w, \delta z^t, \delta z^w) \in \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t, z^w)$ for $w = z^w = \Sigma^w c_{\mathcal{I}}(t, |z^t|)$. By assumption, $\tilde{\delta} \in \mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w)$ holds and this directly implies $\delta = (\delta t, \delta z^t) \in \mathcal{T}_{\text{i-abs}}(t, z^t)$. □

Theorem 5.26. *GKQ for (I-NLP) holds at $(t, z^t(t)) \in \mathcal{F}_{\textit{i-abs}}$ if GKQ for (E-NLP) holds at $(t, w, z^t(t), z^w(w)) \in \mathcal{F}_{\textit{e-abs}}$ for any (and hence all) $w \in W(t)$.*

Proof. The inclusion $\mathcal{T}_{\text{i-abs}}(t, z^t)^* \supseteq \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)^*$ is always satisfied. Thus, it is left to show that

$$\mathcal{T}_{\text{i-abs}}(t, z^t)^* \subseteq \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)^*.$$

Let $\omega = (\omega t, \omega z^t) \in \mathcal{T}_{\text{i-abs}}(t, z^t)^*$, i.e. $\omega^T\delta \geq 0$ for all $\delta = (\delta t, \delta z^t) \in \mathcal{T}_{\text{i-abs}}(t, z^t)$. Then, set $\tilde{\omega} = (\omega t, 0, \omega z^t, 0)$ and obtain $\tilde{\omega}^T\tilde{\delta} = \omega^T\delta \geq 0$ for all $\tilde{\delta} \in \mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w)$ where $w \in W(t)$ is arbitrary. By assumption, then $\tilde{\omega}^T\tilde{\delta} \geq 0$ for all $\tilde{\delta} \in \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t, z^w)$ holds. This implies $\omega^T\delta = \tilde{\omega}^T\tilde{\delta} \geq 0$ for all $\delta \in \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$. □

It is an open question if the converse holds. Next, the branch problems are considered and relations for ACQ and GCQ for all branch problems are obtained. Here, the sign informations will be exploited to show equivalence for GCQ for all branch problems.

Theorem 5.27. *ACQ for (NLP(Σ^t)) holds at $(t, z^t(t)) \in \mathcal{F}_{\Sigma^t}$ if and only if ACQ for (NLP($\Sigma^{t,w}$)) holds at $(t, w, z^t(t), z^t(w)) \in \mathcal{F}_{\Sigma^{t,w}}$ for any (and hence all) $w \in W(t)$.*

Proof. This follows as in the proof in Theorem 5.25. The inclusions $\mathcal{T}_{\Sigma^t}(t, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t)$ and $\mathcal{T}_{\Sigma^{t,w}}(t, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\Sigma^{t,w}}(t, z^t)$ are always satisfied. Thus, one just needs to prove

$$\mathcal{T}_{\Sigma^t}(t, z^t) \supseteq \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t) \iff \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w) \supseteq \mathcal{T}^{\text{lin}}_{\Sigma^{t,w}}(t, w, z^t, z^w).$$

First, the implication "$\Rightarrow$" is shown. Let $\delta = (\delta t, \delta w, \delta z^t, \delta z^w) \in \mathcal{T}^{\text{lin}}_{\Sigma^{t,w}}(t, w, z^t, z^w)$. Then, $\tilde{\delta} = (\delta t, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t)$ and by the assumption $\tilde{\delta} \in \mathcal{T}_{\Sigma^t}(t, z^t)$. Thus, sequences $(t_k, z^t_k) \in \mathcal{F}_{\text{i-abs}}$ and $\tau_k \searrow 0$ exist such that $(t_k, z^t_k) \to (t, z^t)$ and $\tau_k^{-1}(t_k - t, z^t_k - z^t) \to (\delta t, \delta z^t)$. Then, define $z^w_k := w_k := \Sigma^w c_{\mathcal{I}}(t_k, \Sigma^t z^t_k)$ and obtain with $z^w = w = \Sigma^w c_{\mathcal{I}}(t, \Sigma^t z^t)$

$$\begin{aligned} z^w_k - z^w &= \Sigma^w (c_{\mathcal{I}}(t_k, \Sigma^t z^t_k) - c_{\mathcal{I}}(t, \Sigma^t z^t)) \\ &= \Sigma^w [\partial_1 c_{\mathcal{I}}(t, \Sigma^t z^t)(t_k - t) + \partial_2 c_{\mathcal{I}}(t, \Sigma^t z^t)\Sigma^t (z^t_k - z^t) + o(\|(t_k - t, \Sigma^t (z^t_k - z^t))\|)]. \end{aligned}$$

Thus,

$$\tau_k^{-1}(z^w_k - z^w) \to \Sigma^w [\partial_1 c_{\mathcal{I}}(t, \Sigma^t z^t)(t_k - t) + \partial_2 c_{\mathcal{I}}(t, \Sigma^t z^t)\Sigma^t (z^t_k - z^t)] = \delta z^w$$

and $\tau_k^{-1}(w_k - w) \to \delta w$. This yields $d \in \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w)$. Second, the implication "$\Leftarrow$" is proven. Let $\delta = (\delta t, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$ and set $\tilde{\delta} = (\delta t, \delta w, \delta z^t, \delta z^w)$ with

$$\delta w := \delta z^w := \Sigma^w [\partial_1 c_{\mathcal{I}}(t, \Sigma^t z^t)\delta t + \partial_2 c_{\mathcal{I}}(t, \Sigma^t z^t)\delta \zeta].$$

Then, $\tilde{\delta} \in \mathcal{T}^{\text{lin}}_{\Sigma^{t,w}}(t, w, z^t, z^w)$ for $w = z^w = \Sigma^w c_{\mathcal{I}}(t, \Sigma^t z^t)$ and the assumption yields $\tilde{\delta} \in \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w)$. By construction, this directly implies $\delta = (\delta t, \delta z^t) \in \mathcal{T}_{\Sigma^t}(t, z^t)$. ☐

Theorem 5.28. *GCQ for (NLP(Σ^t)) holds at $(t, z^t(t)) \in \mathcal{F}_{\Sigma^t}$ if and only if GCQ for (NLP($\Sigma^{t,w}$)) holds at $(t, w, z^t(t), z^t(w)) \in \mathcal{F}_{\Sigma^{t,w}}$ for any (and hence all) $w \in W(t)$.*

Proof. The inclusions $\mathcal{T}_{\Sigma^t}(t, z^t)^* \supseteq \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t)^*$ and $\mathcal{T}_{\Sigma^{t,w}}(t, z^t)^* \supseteq \mathcal{T}^{\text{lin}}_{\Sigma^{t,w}}(t, z^t)^*$ are always satisfied. Thus, one just need to prove

$$\mathcal{T}_{\Sigma^t}(t, z^t)^* \subseteq \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t)^* \iff \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w)^* \subseteq \mathcal{T}^{\text{lin}}_{\Sigma^{t,w}}(t, w, z^t, z^w)^*.$$

It is started with the implication "$\Rightarrow$". Let $\omega = (\omega t, \omega w, \omega z^t, \omega z^w) \in \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w)^*$, i.e. $\omega^T \delta \geq 0$ for all $\delta = (\delta t, \delta w, \delta z^t, \delta z^w) \in \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w)$. Set

$$\tilde{\omega} = (\tilde{\omega} t, \tilde{\omega} z^t) = (\omega t, \omega z^t) + (\omega w + \omega z^w)\Sigma^w (\partial_1 c_{\mathcal{I}}(t, \Sigma^t z^t), \partial_2 c_{\mathcal{I}}(t, \Sigma^t z^t)\Sigma^t).$$

Then, $\tilde{\omega}^T \tilde{\delta} = \omega^T \delta \geq 0$ for all $\delta = (\delta t, \delta z^t) \in \mathcal{T}_{\Sigma^t}(t, z^t)$ and thus $\tilde{\omega} \in \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t)$. Then, $\omega^T \delta \geq 0$ for all $\delta = (\delta t, \delta w, \delta z^t \delta z^w) \in \mathcal{T}^{\text{lin}}_{\Sigma^{t,w}}(t, w, z^t, z^w)$ as $\omega^T \delta = \tilde{\omega}^T \tilde{\delta}$ holds. The proof of the reverse implication "$\Leftarrow$" follows as in the proof of Theorem 5.26. Let $\omega = (\omega t, \omega z^t) \in \mathcal{T}_{\Sigma^t}(t, z^t)^*$, i.e. $\omega^T \delta \geq 0$ for all $\delta = (\delta t, \delta z^t) \in \mathcal{T}_{\Sigma^t}(t, z^t)$. Then, with $\tilde{\omega} = (\omega t, 0, \omega z^t, 0)$ one obtains $\tilde{\omega}^T \tilde{\delta} = \omega^T \delta \geq 0$ for all $\tilde{\delta} \in \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w)$ and $w \in W(t)$ is arbitrary. Thus, $\tilde{\omega} \in \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w)$ and by assumption $\tilde{\omega} \in \mathcal{T}^{\text{lin}}_{\Sigma^{t,w}}(t, w, z^t, z^w)$. This directly implies $\omega^T \delta \geq 0$ for all $\delta \in \mathcal{T}^{\text{lin}}_{\Sigma^t}(t, z^t)$. ☐

5.2 Counterpart MPECs

In this section the definitions of MPEC-MFCQ, MPEC-ACQ and MPEC-GCQ are formulated for (I-MPEC) and (E-MPEC). Then, relations between them are shown.

Counterpart MPEC for (I-NLP) In this paragraph the MPEC variants of MFCQ, ACQ and GCQ are formulated for (I-MPEC), which is recalled here:

$$\begin{aligned}\min_{t,u^t,v^t} \quad & f(t)\\ \text{s.t.} \quad & c_{\mathcal{E}}(t, u^t + v^t) = 0,\\ & c_{\mathcal{I}}(t, u^t + v^t) \geq 0,\\ & c_{\mathcal{Z}}(t, u^t + v^t) - (u^t - v^t) = 0,\\ & 0 \leq u^t \perp v^t \geq 0.\end{aligned}$$

Lemma 5.29 (MPEC-MFCQ for (I-MPEC))**.** *Given a feasible point (t, u^t, v^t) of* (I-MPEC) *with associated index sets $\mathcal{U}_0^t$, $\mathcal{U}_+^t$, $\mathcal{V}_0^t$, $\mathcal{V}_+^t$ and $\mathcal{D}^t$. Then,* MPEC-MFCQ *at (t, u^t, v^t) is full row rank of*

$$J_{\mathcal{E}}(t, u^t, v^t) := \begin{bmatrix} \partial_1 c_{\mathcal{E}} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}_+^t} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}_+^t} \\ \partial_1 c_{\mathcal{Z}} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}_+^t} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}_+^t} \end{bmatrix} \in \mathbb{R}^{(m_1+s_t)\times(n_t+|\mathcal{U}_+^t|+|\mathcal{V}_+^t|)}$$

and existence of a vector $d \in \mathbb{R}^{n_t+|\mathcal{U}_+^t|+|\mathcal{V}_+^t|}$ such that $J_{\mathcal{E}}(t, u^t, v^t)d = 0$ and $J_{\mathcal{A}}(t, u^t, v^t)d > 0$ with

$$J_{\mathcal{A}}(t, u^t, v^t) := \begin{bmatrix} \partial_1 c_{\mathcal{A}} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}_+^t} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}_+^t} \end{bmatrix}.$$

Here, all partial derivatives are evaluated at $(t, u^t + v^t)$.

Proof. By Definition 2.49 MPEC-MFCQ is MFCQ for the tightened NLP. Thus, full row rank of the matrix

$$\tilde{J}_{\mathcal{E}}(t, u^t, v^t) = \begin{bmatrix} \partial_1 c_{\mathcal{E}} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}_+^t} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{U}_0^t} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}_+^t} & \partial_2 c_{\mathcal{E}} P^T_{\mathcal{V}_0^t} \\ \partial_1 c_{\mathcal{Z}} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}_+^t} & [\partial_2 c_{\mathcal{Z}} - I] P^T_{\mathcal{U}_0^t} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}_+^t} & [\partial_2 c_{\mathcal{Z}} + I] P^T_{\mathcal{V}_0^t} \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}$$

is required. Exploiting the two unit blocks gives the rank condition stated above. Moreover, the existence of a vector $d \in \mathbb{R}^{n_t+2s}$ is required such that $\tilde{J}_{\mathcal{E}} d = 0$ and $\tilde{J}_{\mathcal{A}} d > 0$ with

$$\tilde{J}_{\mathcal{A}}(t, u^t, v^t) = \begin{bmatrix} \partial_1 c_{\mathcal{A}} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}_+^t} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{U}_0^t} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}_+^t} & \partial_2 c_{\mathcal{A}} P^T_{\mathcal{V}_0^t} \end{bmatrix}.$$

The condition $\tilde{J}_{\mathcal{E}} d = 0$ implies directly that $d_i = 0$ for $i \in \mathcal{U}_0^t \cup \mathcal{V}_0^t$ and thus d has to satisfy the conditions stated above. □

As with LIKQ and IDKQ for (I-NLP), MPEC-MFCQ is weaker than MPEC-LICQ for the counterpart MPEC of (I-NLP). This can be seen by rewriting Example 5.2 as the counterpart MPEC and checking the above conditions.

Example 5.30 (MPEC-MFCQ is weaker than MPEC-LICQ)**.** With $z^t = u^t - v^t$ and $|z^t| = u^t + v^t$ the problem reads

$$\begin{aligned} \min_{t\in\mathbb{R}^3, u^t\in\mathbb{R}, v^t\in\mathbb{R}} \quad & t_1 + t_2 - t_3 \\ \text{s.t.} \quad & t_1 + t_2 - (u^t + v^t) = 0, \\ & 4t_1 - t_3 \geq 0, \\ & 4t_2 - t_3 \geq 0, \\ & t_1 - t_2 - (u^t - v^t) = 0, \\ & 0 \leq u^t \perp v^t \geq 0 \end{aligned}$$

with solution $t^* = (0,0,0)$ and $(u^t)^* = (v^t)^* = 0$. Thus, $\mathcal{D}(y^*) = \{1\}$ for $y^* = (t^*, (u^t)^*, (v^t)^*)$ and the Jacobians read

$$J_{\mathcal{E}}(y^*) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}, \quad J_{\mathcal{A}}(y^*) = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 4 & -1 \end{bmatrix}.$$

Here, MPEC-LICQ is not satisfied since the Jacobian $J_{\text{i-mpec}}(y^*) = [J_{\mathcal{E}}(y^*)^T,\ J_{\mathcal{A}}(y^*)^T]^T$ cannot have full row rank by dimension. But MPEC-MFCQ is satisfied as $J_{\mathcal{E}}(y^*)$ has full row rank and $J_{\mathcal{E}}(y^*)d = 0,\ J_{\mathcal{A}}(y^*)d = 1 > 0$ with $d = (0,0,-1)$.

In order to define MPEC-CQs in the spirit of Abadie and Guignard, the tangential cone, the complementarity cone and the MPEC-linearized cone are introduced.

Lemma 5.31 (Tangential Cone and MPEC-Linearized Cone for (I-MPEC))**.** *Consider a feasible point (t, u^t, v^t) of (I-MPEC) with associated index sets $\mathcal{U}^t_+$, $\mathcal{V}^t_+$ and $\mathcal{D}^t$. The* tangential cone *to $\mathcal{F}_{\text{i-mpec}}$ at (t, u^t, v^t) reads*

$$\mathcal{T}_{\text{i-mpec}}(t, u^t, v^t) := \left\{ \begin{pmatrix} \delta t \\ \delta u^t \\ \delta v^t \end{pmatrix} \in \mathbb{R}^{n_t + 2s_t} \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\text{i-mpec}} \ni (t_k, u^t_k, v^t_k) \to (t, u^t, v^t): \\ \tau_k^{-1}(t_k - t, u^t_k - u^t, v^t_k - v^t) \to (\delta t, \delta u^t, \delta v^t) \end{array} \right\}.$$

Setting $\mathcal{A} = \mathcal{A}(t, u^t, v^t)$ and $c_{\mathcal{A}} = [c_i]_{i\in\mathcal{A}}$, the MPEC-linearized cone *reads*

$$\mathcal{T}^{lin}_{\text{i-mpec}}(t, u^t, v^t) := \left\{ \begin{pmatrix} \delta t \\ \delta u^t \\ \delta v^t \end{pmatrix} \in \mathbb{R}^{n_t + 2s_t} \;\middle|\; \begin{array}{r} \partial_1 c_{\mathcal{E}} \delta t + \partial_2 c_{\mathcal{E}} (\delta u^t + \delta v^t) = 0, \\ \partial_1 c_{\mathcal{A}} \delta t + \partial_2 c_{\mathcal{A}} (\delta u^t + \delta v^t) \geq 0, \\ \partial_1 c_{\mathcal{Z}} \delta t + \partial_2 c_{\mathcal{Z}} (\delta u^t + \delta v^t) = \delta u^t - \delta v^t, \\ (\delta u^t, \delta v^t) \in \mathcal{T}_\perp(u^t, v^t) \end{array} \right\}$$

with complementarity cone

$$\mathcal{T}_\perp(u^t, v^t) = \left\{ \begin{pmatrix} \delta u^t \\ \delta v^t \end{pmatrix} \in \mathbb{R}^{2s_t} \;\middle|\; \begin{array}{r} \delta v^t_i = 0,\ i \in \mathcal{U}^t_+, \\ \delta u^t_i = 0,\ i \in \mathcal{V}^t_+, \\ 0 \leq \delta u^t_i \perp \delta v^t_i \geq 0,\ i \in \mathcal{D}^t \end{array} \right\}.$$

Here, all partial derivatives are evaluated at $(t, u^t + v^t)$.

Proof. This follows directly from Definition 2.50. $\square$

Lemma 5.32. *Given (I-NLP) with counterpart MPEC (I-MPEC). Consider* $(t, z^t) \in \mathcal{F}_{\text{i-abs}}$ *with* $\sigma^t = \sigma^t(t)$ *and* $(t, u^t, v^t) = \phi^{-1}(t, z^t) \in \mathcal{F}_{\text{i-mpec}}$ *with associated index sets* $\mathcal{U}_+^t$, $\mathcal{V}_+^t$ *and* $\mathcal{D}^t$. *Define* $\psi\colon \mathcal{T}_{\text{i-mpec}}(t, u^t, v^t) \to \mathcal{T}_{\text{i-abs}}(t, z^t)$ *and* $\psi\colon \mathcal{T}^{lin}_{\text{i-mpec}}(t, u^t, v^t) \to \mathcal{T}^{lin}_{\text{i-abs}}(t, z^t)$ *as*

$$\psi(\delta t, \delta u^t, \delta v^t) = (\delta t, \delta u^t - \delta v^t) \quad \textit{and} \quad \psi^{-1}(\delta t, \delta z^t) = (\delta t, \langle \delta z^t \rangle^+, \langle \delta z^t \rangle^-).$$

Here, $\langle \delta z^t \rangle^+, \langle \delta z^t \rangle^-$ *map* δz^t *into the complementarity cone via*

$$\langle \delta z_i^t \rangle^+ = \left\{ \begin{array}{ll} +\delta z_i^t, & i \in \mathcal{U}_+^t \ (\sigma_i^t = +1) \\ 0, & i \in \mathcal{V}_+^t \ (\sigma_i^t = -1) \\ [\delta z_i^t]^+, & i \in \mathcal{D}^t \ (\sigma_i^t = 0) \end{array} \right\} \quad \textit{and} \quad \langle \delta z_i^t \rangle^- = \left\{ \begin{array}{ll} 0, & i \in \mathcal{U}_+^t \ (\sigma_i^t = +1) \\ -\delta z_i^t, & i \in \mathcal{V}_+^t \ (\sigma_i^t = -1) \\ [\delta z_i^t]^-, & i \in \mathcal{D}^t \ (\sigma_i^t = 0) \end{array} \right\}.$$

Then, both functions ψ *are homeomorphisms.*

Proof. First, consider $\psi\colon \mathcal{T}_{\text{i-mpec}}(t, u^t, v^t) \to \mathcal{T}_{\text{i-abs}}(t, z^t)$: Given a vector $(\delta t, \delta u^t, \delta v^t) = \lim \tau_k^{-1}(t_k - t, u_k^t - u^t, v_k^t - v^t) \in \mathcal{T}_{\text{i-mpec}}(t, u^t, v^t)$, set $(t_k, z_k^t) = \phi(t_k, u_k^t, v_k^t) = (t_k, u_k^t - v_k^t) \in \mathcal{F}_{\text{i-abs}}$ to obtain

$$\lim \frac{z_k^t - z^t}{\tau_k} = \lim \frac{(u_k^t - u^t) - (v_k^t - v^t)}{\tau_k} = \delta u^t - \delta v^t \implies (\delta t, \delta u^t - \delta v^t) \in \mathcal{T}_{\text{i-abs}}(t, z^t).$$

Conversely, given a vector $(\delta t, \delta z^t) = \lim \tau_k^{-1}(t_k - t, z_k^t - z^t) \in \mathcal{T}_{\text{i-abs}}(t, z^t)$, define $(t_k, u_k^t, v_k^t) = \phi^{-1}(t_k, z_k^t) = (t_k, [z_k^t]^+, [z_k^t]^-) \in \mathcal{F}_{\text{i-mpec}}$. Then, $\tau_k^{-1}((u_k - u) - (v_k - v)) \to \langle \delta z^t \rangle^+ - \langle \delta z^t \rangle^-$ holds. Thus, it remains to show $\tau_k^{-1}(u_k - u, v_k - v) \to (\langle \delta z^t \rangle^+, \langle \delta z^t \rangle^-)$ which is done componentwise:

- $i \in \mathcal{U}_+^t$: $v_i^t = 0$ holds by feasibility and $\langle \delta z^t \rangle^- = 0$ by definition. Thus, $(u_k^t)_i > 0$ holds for k large enough and by complementarity $(v_k^t)_i = 0$ holds. Then, $\tau_k^{-1}((u_k^t)_i - u_i^t) \to \langle \delta z^t \rangle_i^+$ follows.
- $i \in \mathcal{V}_+^t$: $\tau_k^{-1}((v_k^t)_i - v_i^t) \to \langle \delta z^t \rangle_i^-$ follows as in the previous case.
- $i \in \mathcal{D}^t$ and $\langle \delta z^t \rangle_i^+ > 0$: $\langle \delta z^t \rangle_i^- = 0$ holds by complementarity and so $\tau_k^{-1}((u_i^t)_k - (v_i^t)_k) \to \langle \delta z^t \rangle_i^+$. Then, $\tau_k^{-1}(u_i^t)_k \to \langle \delta z^t \rangle_i^+$ and $\tau_k^{-1}(v_i^t)_k \to 0$ because of sign constraints.
- $i \in \mathcal{D}^t$ and $\langle \delta z^t \rangle_i^- > 0$: $\tau_k^{-1}(u_i^t)_k \to 0$ and $\tau_k^{-1}(v_i^t)_k \to \langle \delta z^t \rangle_i^-$ follow as in the previous case.
- $i \in \mathcal{D}^t$ and $\langle \delta z^t \rangle_i^+ = \langle \delta z^t \rangle_i^- = 0$: Then, $\tau_k^{-1}((u_i^t)_k - (v_i^t)_k) \to 0$ holds. Because of sign constraints and complementarity, this can only hold if $\tau_k^{-1}(u_i^t)_k \to 0$, $\tau_k^{-1}(v_i^t)_k \to 0$.

Altogether, this implies

$$\lim \frac{(t_k - t, u_k^t - u^t, v_k^t - v^t)}{\tau_k} = (\delta t, \langle \delta z^t \rangle^+, \langle \delta z^t \rangle^-) \in \mathcal{T}_{\text{i-mpec}}(t, u^t, v^t).$$

By construction, ψ and ψ^{-1} are both continuous and inverse to each other.
Second, consider $\psi\colon \mathcal{T}^{\text{lin}}_{\text{i-mpec}}(t, u^t, v^t) \to \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$: Given $(\delta t, \delta u^t, \delta v^t) \in \mathcal{T}^{\text{lin}}_{\text{i-mpec}}(t, u^t, v^t)$, the vectors $\delta z^t = \delta u^t - \delta v^t$ and $\delta\zeta = \delta u^t + \delta v^t$ satisfy

$$\delta z^t_i = \left\{\begin{array}{ll} \delta u^t_i, & i \in \mathcal{U}^t_+ \\ -\delta v^t_i, & i \in \mathcal{V}^t_+ \\ \delta u^t_i - \delta v^t_i, & i \in \mathcal{D}^t \end{array}\right\}, \quad \delta\zeta_i = \left\{\begin{array}{ll} \delta u^t_i, & i \in \mathcal{U}^t_+ \\ \delta v^t_i, & i \in \mathcal{V}^t_+ \\ \delta u^t_i + \delta v^t_i, & i \in \mathcal{D}^t \end{array}\right\} = \left\{\begin{array}{ll} \sigma^t_i \delta z^t_i, & \sigma^t_i = +1 \\ \sigma^t_i \delta z^t_i, & \sigma^t_i = -1 \\ |\delta z^t_i|, & \sigma^t_i = 0 \end{array}\right\}.$$

Thus, $(\delta t, \delta z^t) = \psi(\delta t, \delta u^t, \delta v^t) \in \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$. Conversely, consider $(\delta t, \langle \delta z^t\rangle^+, \langle \delta z^t\rangle^-) = \psi^{-1}(\delta t, \delta z^t)$ for $(\delta t, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\text{i-abs}}(t, z^t)$. Then, $0 \le \langle \delta z^t\rangle^+ \perp \langle \delta z^t\rangle^- \ge 0$, $\delta z^t = \langle \delta z^t\rangle^+ - \langle \delta z^t\rangle^-$ hold by construction and $\delta\zeta = \langle \delta z^t\rangle^+ + \langle \delta z^t\rangle^-$ as

$$\delta\zeta_i = \left\{\begin{array}{ll} \sigma_i \delta z^t_i, & \sigma^t_i = +1 \\ \sigma^t_i \delta z^t_i, & \sigma^t_i = -1 \\ |\delta z^t_i|, & \sigma^t_i = 0 \end{array}\right\} = \left\{\begin{array}{ll} \delta z^t_i, & i \in \mathcal{U}^t_+ \\ -\delta z^t_i, & i \in \mathcal{V}^t_+ \\ \langle \delta z^t\rangle^+_i + \langle \delta z^t\rangle^-_i, & i \in \mathcal{D}^t \end{array}\right\} = \langle \delta z^t\rangle^+_i + \langle \delta z^t\rangle^-_i.$$

Thus, $(\delta t, \langle \delta z^t\rangle^+, \langle \delta z^t\rangle^-) \in \mathcal{T}^{\text{lin}}_{\text{i-mpec}}(t, u^t, v^t)$. Again, ψ and ψ^{-1} are both continuous and inverse to each other by construction. □

Lemma 5.33 (Branch NLPs for (I-MPEC))**.** *Given a feasible point $(\hat t, \hat u^t, \hat v^t)$ of (I-MPEC) with associated index sets $\mathcal{U}^t_+$, $\mathcal{V}^t_+$ and $\mathcal{D}^t$ and choose $\mathcal{P}^t \subseteq \mathcal{D}^t$. The branch problem NLP($\mathcal{P}^t$) reads*

$$\begin{aligned} \min_{(t,u^t,v^t)\in D^{t,u^t,v^t}} \quad & f(t) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t, u^t + v^t) = 0, \\ & c_{\mathcal{I}}(t, u^t + v^t) \ge 0, \\ & c_{\mathcal{Z}}(t, u^t + v^t) - (u^t - v^t) = 0, \\ & 0 = u^t_i,\ 0 \le v^t_i,\ i \in \mathcal{V}^t_+ \cup \mathcal{P}^t, \\ & 0 \le u^t_i,\ 0 = v^t_i,\ i \in \mathcal{U}^t_+ \cup \bar{\mathcal{P}}^t. \end{aligned} \tag{NLP($\mathcal{P}^t$)}$$

Here, $\bar{\mathcal{P}}^t$ denotes the complement of $\mathcal{P}^t$ in $\mathcal{D}^t$. The feasible set of (NLP($\mathcal{P}^t$)), which always contains $(\hat t, \hat u^t, \hat v^t)$, is denoted by

$$\mathcal{F}_{\mathcal{P}^t} := \left\{ (t, u^t, v^t) \;\middle|\; \begin{array}{l} c_{\mathcal{E}}(t, u^t + v^t) = 0,\ c_{\mathcal{I}}(t, u^t + v^t) \ge 0, \\ c_{\mathcal{Z}}(t, u^t + v^t) - (u^t - v^t) = 0, \\ 0 = u^t_i \textit{ and } 0 \le v^t_i \textit{ for } i \in \mathcal{V}^t_+ \cup \mathcal{P}^t, \\ 0 \le u^t_i \textit{ and } 0 = v^t_i \textit{ for } i \in \mathcal{U}^t_+ \cup \bar{\mathcal{P}}^t \end{array} \right\}.$$

Proof. This follows directly from Definition 2.52. □

Lemma 5.34. *Given (NLP(Σ^t)) and (NLP($\mathcal{P}^t$)) with $\mathcal{P}^t = \{i \in \alpha^t(\hat t)\colon \sigma^t_i = -1\}$. Define $\phi_{\mathcal{P}^t} := \phi|_{\mathcal{F}_{\mathcal{P}^t}}$ and $\phi^{-1}_{\mathcal{P}^t} := \phi^{-1}|_{\mathcal{F}_{\Sigma^t}}$. Then,*

$$\phi_{\mathcal{P}^t}\colon \mathcal{F}_{\mathcal{P}^t} \to \mathcal{F}_{\Sigma^t} \quad \textit{with} \quad \phi^{-1}_{\mathcal{P}^t}\colon \mathcal{F}_{\Sigma^t} \to \mathcal{F}_{\mathcal{P}^t}$$

is a homeomorphism.

Proof. This follows directly from Lemma 4.2 by the definition of $\mathcal{P}^t$. □

Lemma 5.35 (Tangential Cone and Linearized Cone for (NLP($\mathcal{P}^t$)))**.** *Consider a feasible point* (t, u^t, v^t) *of* (NLP($\mathcal{P}^t$))*. The* tangential cone *to* $\mathcal{F}_{\mathcal{P}^t}$ *at* (t, u^t, v^t) *reads*

$$\mathcal{T}_{\mathcal{P}^t}(t, u^t, v^t) := \left\{ (\delta t, \delta u^t, \delta v^t) \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\mathcal{P}^t} \ni (t_k, u_k^t, v_k^t) \to (t, u^t, v^t): \\ \tau_k^{-1}(t_k - t, u_k^t - u^t, v_k^t - v^t) \to (\delta t, \delta u^t, \delta v^t) \end{array} \right\}$$

and the linearized cone *reads*

$$\mathcal{T}_{\mathcal{P}^t}^{lin}(t, u^t, v^t) := \left\{ \begin{pmatrix} \delta t \\ \delta u^t \\ \delta v^t \end{pmatrix} \;\middle|\; \begin{array}{l} \partial_1 c_{\mathcal{E}} \delta t + \partial_2 c_{\mathcal{E}}(\delta u^t + \delta v^t) = 0, \\ \partial_1 c_{\mathcal{A}} \delta t + \partial_2 c_{\mathcal{A}}(\delta u^t + \delta v^t) \geq 0, \\ \partial_1 c_{\mathcal{Z}} \delta t + \partial_2 c_{\mathcal{Z}}(\delta u^t + \delta v^t) = \delta u^t - \delta v^t, \\ 0 = \delta u_i^t \text{ for } i \in \mathcal{V}_+^t \cup \mathcal{P}^t,\ \ 0 = \delta v_i^t \text{ for } i \in \mathcal{U}_+^t \cup \bar{\mathcal{P}}^t, \\ 0 \leq \delta u_i^t \text{ for } i \in \bar{\mathcal{P}}^t,\ \ 0 \leq \delta v_i^t \text{ for } i \in \mathcal{P}^t \end{array} \right\}.$$

Here, all partial derivatives are evaluated at $(t, u^t + v^t)$*.*

Proof. This follows directly from Definition 2.53. □

Lemma 5.36. *Given* (NLP(Σ^t)) *and* (NLP($\mathcal{P}^t$)) *with* $\mathcal{P}^t = \{i \in \alpha^t(\hat{t})\colon \sigma_i^t = -1\}$*. Consider* $(t, z^t) \in \mathcal{F}_{\Sigma^t}$ *and* $(t, u^t, v^t) = \phi_{\mathcal{P}^t}^{-1}(t, z^t)$*. Define* $\psi_{\mathcal{P}^t} := \psi|_{\mathcal{T}_{\mathcal{P}^t}}$*,* $\psi_{\mathcal{P}^t}^{-1} := \psi^{-1}|_{\mathcal{T}_{\Sigma^t}}$ *and* $\psi_{\mathcal{P}^t} := \psi|_{\mathcal{T}_{\mathcal{P}^t}^{lin}}$*,* $\psi_{\mathcal{P}^t}^{-1} := \psi^{-1}|_{\mathcal{T}_{\Sigma^t}^{lin}}$*. Then,*

$$\psi_{\mathcal{P}}\colon \mathcal{T}_{\mathcal{P}^t}(t, u^t, v^t) \to \mathcal{T}_{\Sigma^t}(t, \hat{z}^t) \quad \textit{and} \quad \psi_{\mathcal{P}}\colon \mathcal{T}_{\mathcal{P}^t}^{lin}(t, \hat{u}^t, \hat{v}^t) \to \mathcal{T}_{\Sigma^t}^{lin}(t, \hat{z}^t)$$

are homeomorphisms.

Proof. The first claim follows from Lemma 5.32 as $\phi|_{\mathcal{P}^t} : \mathcal{F}_{\mathcal{P}^t} \to \mathcal{F}_{\Sigma^t}$ is a homeomorphism (see Lemma 5.34). Also, the second claim follows from Lemma 5.32 as

$$\mathcal{P}^t = \{i \in \mathcal{D}^t\colon \sigma_i^t = -1\} \quad \text{and} \quad \bar{\mathcal{P}}^t = \{i \in \mathcal{D}^t\colon \sigma_i^t = +1\}$$

since $\alpha^t(\hat{t}) = \mathcal{D}^t$. □

Using this notation MPEC-ACQ and MPEC-GCQ read as follows.

Lemma 5.37 (MPEC-ACQ for (I-MPEC))**.** *Given a feasible point* (t, u^t, v^t) *of* (I-MPEC)*. Then, MPEC-ACQ for* (I-MPEC) *at* (t, u^t, v^t) *reads* $\mathcal{T}_{\text{i-mpec}}(t, u^t, v^t) = \mathcal{T}_{\text{i-mpec}}^{lin}(t, u^t, v^t)$*.*

Lemma 5.38 (MPEC-GCQ for (I-MPEC))**.** *Given a feasible point* (t, u^t, v^t) *of* (I-MPEC)*. Then, MPEC-GCQ for* (I-MPEC) *at* (t, u^t, v^t) *reads* $\mathcal{T}_{\text{i-mpec}}(t, u^t, v^t)^* = \mathcal{T}_{\text{i-mpec}}^{lin}(t, u^t, v^t)^*$*.*

Both MPEC-CQs are implied if the corresponding constraint qualifications hold for all branch problems.

Theorem 5.39 (ACQ for all (NLP($\mathcal{P}^t$)) implies MPEC-ACQ for (I-MPEC))**.** *Consider a feasible point* (t, u^t, v^t) *of* (I-MPEC) *with associated branch problems* (NLP($\mathcal{P}^t$))*. Then, MPEC-ACQ holds for* (I-MPEC) *at* (t, u^t, v^t) *if ACQ holds for all* (NLP($\mathcal{P}^t$)) *at* (t, u^t, v^t)*.*

Proof. This follows directly from Theorem 2.58. □

Theorem 5.40 (GCQ for all (NLP($\mathcal{P}^t$)) implies MPEC-GCQ for (I-MPEC))**.** *Consider a feasible point* (t, u^t, v^t) *of* (I-MPEC) *with associated branch problems* (NLP($\mathcal{P}^t$))*. Then, MPEC-GCQ holds for* (I-MPEC) *at* (t, u^t, v^t) *if GCQ holds for all* (NLP($\mathcal{P}^t$)) *at* (t, u^t, v^t)*.*

Proof. This follows directly from Theorem 2.59. □

Counterpart MPEC for (E-NLP) In this paragraph MPEC-MFCQ, MPEC-ACQ and MPEC-GCQ are stated for (E-MPEC), which reads

$$\begin{aligned}
\min_{t,w,u^t,v^t,u^w,v^w} \quad & f(t) \\
\text{s.t.} \quad & c_{\mathcal{E}}(t, u^t + v^t) = 0, \\
& c_{\mathcal{I}}(t, u^t + v^t) - (u^w + v^w) = 0, \\
& c_{\mathcal{Z}}(t, u^t + v^t) - (u^t - v^t) = 0, \\
& w - (u^w - v^w) = 0, \\
& 0 \le u^t \perp v^t \ge 0, \\
& 0 \le u^w \perp v^w \ge 0.
\end{aligned}$$

For this, the formulation as ($\overline{\text{E-MPEC}}$) is used and recalled here:

$$\begin{aligned}
\min_{x,u,v} \quad & f(x) \\
\text{s.t.} \quad & \bar{c}_{\mathcal{E}}(x, u + v) = 0, \\
& \bar{c}_{\mathcal{Z}}(x, u + v) - (u - v) = 0, \\
& 0 \le u \perp v \ge 0,
\end{aligned}$$

where $x = (t, w)$, $u = (u^t, u^w)$, $v = (v^t, v^w)$ as well as $\bar{f}(x) = f(t)$, $\bar{c}_{\mathcal{E}}(x, u + v) = (c_{\mathcal{E}}(t, u^t + v^t), c_{\mathcal{I}}(t, u^t + v^t) - (u^w + v^w))$ and $\bar{c}_{\mathcal{Z}}(x, u + v) = (c_{\mathcal{Z}}(t, u^t + v^t), w)$.

Hence, problem (E-MPEC) is a special case of (I-MPEC) and the next lemmas can be obtained from the corresponding definitions and lemmas for (I-MPEC).

As in the abs-normal case, constraint qualifications of Mangasarian Fromovitz type are not preserved under slack formulation. Thus, MPEC-MFCQ coincides with MPEC-LICQ for (E-MPEC).

Lemma 5.41 (MPEC-MFCQ for (E-MPEC))**.** *Given* (E-MPEC)*, consider a feasible point* $y = (t, w, u^t, v^t, u^w, v^w)$*. Then, MPEC-MFCQ at* y *is MPEC-LICQ at* y*.*

Proof. By construction neither (E-MPEC) nor ($\overline{\text{E-MPEC}}$) contain inequalities. Thus, just full row rank of $J_{\text{e-mpec}}(x)$ is required by Definition 2.49 which is exactly the definition of MPEC-LICQ for (E-MPEC). □

To define MPEC-ACQ and MPEC-GCQ the tangential and abs-normal linearized cone are needed.

Lemma 5.42 (Tangential Cone and MPEC-Linearized Cone for (E-MPEC))**.** *Consider a feasible point* $y = (t, w, u^t, v^t, u^w, v^w)$ *of* (E-MPEC)*. The* tangential cone *to* $\mathcal{F}_{\text{e-mpec}}$ *at* y *reads*

$$\mathcal{T}_{\text{e-mpec}}(y) = \left\{ \delta \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\text{e-mpec}} \ni y_k = (t_k, w_k, u_k^t, v_k^t, u_k^w, v_k^w) \to y\colon \\ \tau_k^{-1}(y_k - y) \to \delta = (\delta t, \delta w, \delta u^t, \delta v^t, \delta u^w, \delta v^w) \end{array} \right\}.$$

The MPEC-linearized cone *reads*

$$\mathcal{T}^{\text{lin}}_{\text{e-mpec}}(y) = \left\{ \delta \;\middle|\; \begin{array}{l} \partial_1 c_{\mathcal{I}} \delta t + \partial_2 c_{\mathcal{I}} (\delta u^t + \delta v^t) = \delta u^w + \delta v^w,\ \ \delta w = \delta u^w - \delta v^w, \\ (\delta t, \delta u^t \delta v^t) \in \mathcal{T}^{\text{lin}}_{\text{i-mpec}}(t, u^t, v^t),\ \ (\delta u^w, \delta v^w) \in \mathcal{T}_{\perp}(u^w, v^w) \end{array} \right\}.$$

Here, $\delta = (\delta t, \delta w, \delta u^t, \delta v^t, \delta u^w, \delta v^w) \in \mathbb{R}^{n_t + m_2 + 2s_t + 2m_2}$ *and all partial derivatives are evaluated at* $(t, u^t + v^t)$*.*

Proof. Definition 5.31 can be applied to (E-MPEC) and gives

$$\mathcal{T}_{\text{e-mpec}}(y) = \left\{ \delta \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\text{e-mpec}} \ni y_k = (t_k, w_k, u_k^t, v_k^t, u_k^w, v_k^w) \to y\colon \\ \tau_k^{-1}(y_k - y) \to \delta = (\delta t, \delta w, \delta u^t, \delta v^t, \delta u^w, \delta v^w) \end{array} \right\}$$

and

$$\mathcal{T}^{\text{lin}}_{\text{e-mpec}}(y) = \left\{ \delta \;\middle|\; \begin{array}{r} \partial_1 c_{\mathcal{E}} \delta t + \partial_2 c_{\mathcal{E}} (\delta u^t + \delta v^t) = 0, \\ \partial_1 c_{\mathcal{I}} \delta t + \partial_2 c_{\mathcal{I}} (\delta u^t + \delta v^t) = \delta u^w + \delta v^w, \\ \partial_1 c_{\mathcal{Z}} \delta t + \partial_2 c_{\mathcal{Z}} (\delta u^t + \delta v^t) = \delta u^t - \delta v^t, \\ \delta w = \delta u^w - \delta v^w, \\ (\delta u^t, \delta v^t) \in \mathcal{T}_{\perp}(u^t, v^t), \\ (\delta u^w, \delta v^w) \in \mathcal{T}_{\perp}(u^w, v^w) \end{array} \right\}.$$

Using the definition of $\mathcal{T}^{\text{lin}}_{\text{i-mpec}}(t, u^t, v^t)$ leads to the presentation above. □

Lemma 5.43. *Given* (E-NLP) *with counterpart* (E-MPEC)*. Consider* $(t, w, z^t, z^w) \in \mathcal{F}_{\text{e-abs}}$ *and* $(t, w, u^t, v^t, u^w, v^w) = \bar{\phi}^{-1}(t, w, z^t, z^w) \in \mathcal{F}_{\text{e-mpec}}$*. Define*

$$\bar{\psi}\colon \mathcal{T}_{\text{e-mpec}}(t, w, u^t, v^t, u^w, v^w) \to \mathcal{T}_{\text{e-abs}}(t, w, z^t, z^w),$$
$$\bar{\psi}\colon \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(t, w, u^t, v^t, u^w, v^w) \to \mathcal{T}^{\text{lin}}_{\text{e-abs}}(t, w, z^t, z^w),$$

as

$$\bar{\psi}(\delta t, \delta w, \delta u^t, \delta v^t, \delta u^w, \delta v^w) = (\delta t, \delta w, \delta u^t - \delta v^t, \delta u^w - \delta v^w),$$
$$\bar{\psi}^{-1}(\delta t, \delta w, \delta z^t, \delta z^w) = (\delta t, \delta w, \langle \delta z^t \rangle^+, \langle \delta z^t \rangle^-, \langle \delta z^w \rangle^+, \langle \delta z^w \rangle^-).$$

Then, both functions $\bar{\psi}$ *are homeomorphisms.*

Proof. This follows directly from Lemma 5.32. □

Lemma 5.44 (Branch NLPs for (E-MPEC)). *Given a feasible point $\hat{y} = (\hat{t}, \hat{w}, \hat{u}^t, \hat{v}^t, \hat{u}^w, \hat{v}^w)$ of (E-MPEC) with associated index sets $\mathcal{U}_+^t$, $\mathcal{V}_+^t$, $\mathcal{D}_+^t$, $\mathcal{U}_+^w$, $\mathcal{V}_+^w$ and $\mathcal{D}_+^w$. Choose $\mathcal{P}^t \subseteq \mathcal{D}^t$ as well as $\mathcal{P}^w \subseteq \mathcal{D}^w$. The branch problem NLP($\mathcal{P}^{t,w}$) is defined as*

$$\begin{aligned}
\min_{(t,w,u^t,v^t,u^w,v^w)\in D^{t,w,u^t,v^t,u^w,v^w}} \quad & f(t) \\
s.t. \quad & c_{\mathcal{E}}(t, u^t + v^t) = 0, \\
& c_{\mathcal{I}}(t, u^t + v^t) - (u^w + v^w) = 0, \\
& c_{\mathcal{Z}}(t, u^t + v^t) - (u^t - v^t) = 0, \\
& w - (u^w - v^w) = 0, \\
& 0 = u_i^t,\ 0 \le v_i^t,\ i \in \mathcal{V}_+^t \cup \mathcal{P}^t, \\
& 0 \le u_i^t,\ 0 = v_i^t,\ i \in \mathcal{U}_+^t \cup \bar{\mathcal{P}}^t, \\
& 0 = u_i^w,\ 0 \le v_i^w,\ i \in \mathcal{V}_+^w \cup \mathcal{P}^w, \\
& 0 \le u_i^w,\ 0 = v_i^w,\ i \in \mathcal{U}_+^w \cup \bar{\mathcal{P}}^w.
\end{aligned} \tag{NLP($\mathcal{P}^{t,w}$)}$$

The feasible set of (NLP($\mathcal{P}^{t,w}$)), which always contains $\hat{y}$, is a lifting of $\mathcal{F}_{\mathcal{P}^t}$ and is denoted by

$$\mathcal{F}_{\mathcal{P}^{t,w}} := \left\{ (t, w, u^t, v^t, u^w, v^w) \;\middle|\; \begin{array}{l} (t, u^t, v^t) \in \mathcal{F}_{\mathcal{P}^t},\ w - (u^w - v^w) = 0, \\ c_{\mathcal{I}}(t, u^t + v^t) - (u^w + v^w) = 0, \\ 0 = u_i^w,\ 0 \le v_i^w,\ i \in \mathcal{V}_+^w \cup \mathcal{P}^w, \\ 0 \le u_i^w,\ 0 = v_i^w,\ i \in \mathcal{U}_+^w \cup \bar{\mathcal{P}}^w \end{array} \right\}.$$

Proof. Definition 5.33 applied to (E-MPEC) gives the branch problem above with feasible set

$$\mathcal{F}_{\mathcal{P}^{t,w}} = \left\{ (t, w, u^t, v^t, u^w, v^w) \;\middle|\; \begin{array}{l} c_{\mathcal{E}}(t, u^t + v^t) = 0,\ c_{\mathcal{I}}(t, u^t + v^t) - (u^w - v^w) = 0, \\ c_{\mathcal{Z}}(t, u^t + v^t) - (u^t - v^t) = 0,\ w - (u^w - v^w) = 0, \\ 0 = u_i^t,\ 0 \le v_i^t,\ i \in \mathcal{V}_+^t \cup \mathcal{P}^t, \\ 0 \le u_i^t,\ 0 = v_i^t,\ i \in \mathcal{U}_+^t \cup \bar{\mathcal{P}}^t, \\ 0 = u_i^w,\ 0 \le v_i^w,\ i \in \mathcal{V}_+^w \cup \mathcal{P}^w, \\ 0 \le u_i^w,\ 0 = v_i^w,\ i \in \mathcal{U}_+^w \cup \bar{\mathcal{P}}^w \end{array} \right\}.$$

Using the definition of $\mathcal{F}_{\mathcal{P}^t}$ gives the stated form above. □

Lemma 5.45. *Given (NLP($\Sigma^{t,w}$)) and (NLP($\mathcal{P}^{t,w}$)) with $\mathcal{P}^t = \{i \in \alpha^t(\hat{t})\colon \sigma_i^t = -1\}$ and $\mathcal{P}^w = \{i \in \alpha^w(\hat{w})\colon \sigma_i^w = -1\}$. Define $\bar{\phi}_{\mathcal{P}^{t,w}} := \bar{\phi}|_{\mathcal{F}_{\mathcal{P}^{t,w}}}$ and $\bar{\phi}_{\mathcal{P}^{t,w}}^{-1} := \bar{\phi}^{-1}|_{\mathcal{F}_{\Sigma^{t,w}}}$. Then,*

$$\bar{\phi}_{\mathcal{P}^{t,w}} \colon \mathcal{F}_{\mathcal{P}^{t,w}} \to \mathcal{F}_{\Sigma^{t,w}} \quad \textit{with} \quad \bar{\phi}_{\mathcal{P}^{t,w}}^{-1} \colon \mathcal{F}_{\Sigma^{t,w}} \to \mathcal{F}_{\mathcal{P}^{t,w}}$$

is a homeomorphism.

Proof. This follows directly from Lemma 4.5 by the definition of $\mathcal{P}^t$ and $\mathcal{P}^w$. □

Lemma 5.46 (Tangential Cone and Linearized Cone for (NLP($\mathcal{P}^{t,w}$))**.** *Consider a feasible point* $y = (t, w, u^t, v^t, u^w, v^w)$ *of* (NLP($\mathcal{P}^{t,w}$))*. The* tangential cone *to* $\mathcal{F}_{\mathcal{P}^{t,w}}$ *at* y *is*

$$\mathcal{T}_{\mathcal{P}^{t,w}}(y) = \left\{ \delta \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\mathcal{P}^{t,w}} \ni (t_k, w_k, u_k^t, v_k^t, u_k^w, v_k^w) \to (t, w, u^t, v^t, u^w, v^w)\colon \\ \tau_k^{-1}(t_k - t, w_k - w, u_k^t - u^t, v_k^t - v^t, u_k^w - u^w, v_k^w - v^w) \to \delta \end{array} \right\},$$

where $\delta = (\delta t, \delta w, \delta u^t, \delta v^t, \delta u^w, \delta v^w)$*. The* linearized cone *is*

$$\mathcal{T}^{lin}_{\mathcal{P}^{t,w}}(y) = \left\{ \delta \;\middle|\; \begin{array}{l} (\delta t, \delta u^t, \delta v^t) \in \mathcal{T}^{lin}_{\mathcal{P}^t}, \\ \partial_1 c_{\mathcal{I}} \delta t + \partial_2 c_{\mathcal{I}}(\delta u^t + \delta v^t) = (\delta u^w + \delta v^w),\ \ \delta w = \delta u^w + \delta v^w, \\ 0 = \delta u_i^w \text{ for } i \in \mathcal{V}_+^w \cup \mathcal{P}^w,\ \ 0 = \delta v_i^w \text{ for } i \in \mathcal{U}_+^w \cup \bar{\mathcal{P}}^w, \\ 0 \le \delta u_i^w \text{ for } i \in \bar{\mathcal{P}}^w,\ \ 0 \le \delta v_i^w \text{ for } i \in \mathcal{P}^w \end{array} \right\}.$$

Here, all partial derivatives are evaluated at $(t, u^t + v^t)$*.*

Proof. Definition 5.35 applied to ($\overline{\text{E-MPEC}}$) gives

$$\mathcal{T}_{\mathcal{P}^{t,w}}(y) = \left\{ \delta \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\mathcal{P}^{t,w}} \ni (t_k, w_k, u_k^t, v_k^t, u_k^w, v_k^w) \to (t, w, u^t, v^t, u^w, v^w)\colon \\ \tau_k^{-1}(t_k - t, w_k - w, u_k^t - u^t, v_k^t - v^t, u_k^w - u^w, v_k^w - v^w) \to \delta \end{array} \right\}$$

and

$$\mathcal{T}^{\text{lin}}_{\mathcal{P}^{t,w}}(y) = \left\{ \delta \;\middle|\; \begin{array}{r} \partial_1 c_{\mathcal{E}} \delta t + \partial_2 c_{\mathcal{E}}(\delta u^t + \delta v^t) = 0, \\ \partial_1 c_{\mathcal{I}} \delta t + \partial_2 c_{\mathcal{I}}(\delta u^t + \delta v^t) = \delta u^w + \delta v^w, \\ \partial_1 c_{\mathcal{Z}} \delta t + \partial_2 c_{\mathcal{Z}}(\delta u^t + \delta v^t) = \delta u^t - \delta v^t, \\ \delta w = \delta u^w - \delta v^w, \\ 0 = \delta u_i^t \text{ for } i \in \mathcal{V}_+^t \cup \mathcal{P}^t,\ \ 0 = \delta v_i^t \text{ for } i \in \mathcal{U}_+^t \cup \bar{\mathcal{P}}^t, \\ 0 \le \delta u_i^t \text{ for } i \in \bar{\mathcal{P}}^t,\ \ 0 \le \delta v_i^t \text{ for } i \in \mathcal{P}^t \\ 0 = \delta u_i^w \text{ for } i \in \mathcal{V}_+^w \cup \mathcal{P}^w,\ \ 0 = \delta v_i^w \text{ for } i \in \mathcal{U}_+^w \cup \bar{\mathcal{P}}^w, \\ 0 \le \delta u_i^w \text{ for } i \in \bar{\mathcal{P}}^w,\ \ 0 \le \delta v_i^w \text{ for } i \in \mathcal{P}^w \end{array} \right\}.$$

Using the definition of $\mathcal{T}^{\text{lin}}_{\mathcal{P}^t}$ gives the result. ☐

Lemma 5.47. *Given* (NLP($\Sigma^{t,w}$)) *and* (NLP($\mathcal{P}^{t,w}$)) *with* $\mathcal{P}^t = \{i \in \alpha^t(\hat{t})\colon \sigma_i^t = -1\}$ *and* $\mathcal{P}^w = \{i \in \alpha^t(\hat{t})\colon \sigma_i^w = -1\}$*. Consider* $(t, w, z^t, z^w) \in \mathcal{F}_{e\text{-}abs}$ *and* $(t, w, u^t, v^t, u^w, v^w) = \bar{\phi}^{-1}(t, w, z^t, z^w)$*. Define* $\bar{\psi}_{\mathcal{P}^{t,w}} := \bar{\psi}|_{\mathcal{T}_{\mathcal{P}^{t,w}}}$*,* $\bar{\psi}^{-1}_{\mathcal{P}^{t,w}} := \bar{\psi}|_{\mathcal{T}_{\Sigma^{t,w}}}$ *and* $\bar{\psi}_{\mathcal{P}^{t,w}} := \bar{\psi}|_{\mathcal{T}^{lin}_{\mathcal{P}^{t,w}}}$*,* $\bar{\psi}^{-1}_{\mathcal{P}^{t,w}} := \bar{\psi}|_{\mathcal{T}^{lin}_{\Sigma^{t,w}}}$*. Then,*

$$\begin{aligned} \bar{\psi}_{\mathcal{P}^{t,w}} &\colon \mathcal{T}_{\mathcal{P}^{t,w}}(t, w, u^t, v^t, u^w, v^w) \to \mathcal{T}_{\Sigma^{t,w}}(t, w, z^t, z^w), \\ \bar{\psi}_{\mathcal{P}^{t,w}} &\colon \mathcal{T}^{lin}_{\mathcal{P}^{t,w}}(t, w, u^t, v^t, u^w, v^w) \to \mathcal{T}^{lin}_{\Sigma^{t,w}}(t, w, z^t, z^w) \end{aligned}$$

are homeomorphisms.

Proof. This follows directly from Lemma 5.36. ☐

With these cones MPEC-ACQ and MPEC-GCQ read as follows.

Lemma 5.48 (MPEC-ACQ for (E-MPEC))**.** *Given a feasible point $y = (t, w, u^t, v^t, u^w, v^w)$ of* (E-MPEC)*. Then, MPEC-ACQ at y reads $\mathcal{T}_{\text{e-mpec}}(y) = \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(y)$.*

Proof. This follows by applying Lemma 5.37 to (E-MPEC). □

Lemma 5.49 (MPEC-GCQ for (E-MPEC))**.** *Given a feasible point $y = (t, w, u^t, v^t, u^w, v^w)$ of* (E-MPEC)*, Then, MPEC-GCQ at y reads $\mathcal{T}_{\text{e-mpec}}(y)^* = \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(y)^*$.*

Proof. This follows by applying Lemma 5.38 to (E-MPEC). □

Due to symmetry, the above equality of cones (respectively dual cones) holds for all elements $(w, \hat{u}^w, \hat{v}^w) \in W(t, \hat{u}^t, \hat{v}^t) = \{(w, u^w, v^w) \colon |w| = c_{\mathcal{I}}(t, u^t + v^t), u^w = [w]^+, v^w = [w]^-\}$ if it holds for any element. Thus, both conditions are well-defined for (E-MPEC). This is proven formally next and to begin with, homeomorphisms between the cones are defined.

Lemma 5.50. *Given* (E-MPEC)*, consider a feasible point $y = (t, w, u^t, v^t, u^w, v^w)$ and $(\tilde{w}, \tilde{u}^w, \tilde{v}^w) \in W(t, u^t, v^t)$, $(\tilde{w}, \tilde{u}^w, \tilde{v}^w) \neq (w, u^t, v^t)$. Set $\tilde{y} = (t, \tilde{w}, u^t, v^t, \tilde{u}^w, \tilde{v}^w)$ and define $\rho \colon \mathcal{T}_{\text{e-mpec}}(y) \to \mathcal{T}_{\text{e-mpec}}(\tilde{y})$ and $\rho \colon \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(y) \to \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(\tilde{y})$ as*

$$\begin{aligned} \rho(\delta t, \delta w, \delta u^t, \delta v^t, \delta u^w, \delta v^w) &= (\delta t, \Sigma \delta w, \delta u^t, \delta v^t, \Sigma \delta u^w, \Sigma \delta v^w), \\ \rho^{-1}(\delta t, \tilde{\delta} w, \delta u^t, \delta v^t, \tilde{\delta} u^w, \tilde{\delta} v^w) &= (\delta t, \Sigma \tilde{\delta} w, \delta u^t, \delta v^t, \Sigma \tilde{\delta} u^w, \Sigma \tilde{\delta} v^w), \end{aligned}$$

where $\Sigma = \operatorname{diag}(\sigma)$ with $\sigma_i = -1$ if $w_i \neq \tilde{w}_i$ and $\sigma_i = 1$ if $w_i = \tilde{w}_i$. Then, both functions ρ are homeomorphisms.

Proof. This follows directly as $\rho = \psi^{-1} \circ \chi \circ \psi$ and $\rho^{-1} = \psi^{-1} \circ \chi^{-1} \circ \psi$. □

Now, these homeomorphisms will be used to show that MPEC-ACQ and MPEC-GCQ are independent of the particular choice of w. Thus, both conditions above are well-defined for (E-MPEC).

Lemma 5.51. *Consider a feasible point $y = (t, w, u^t, v^t, u^w, v^w)$ of* (E-MPEC)*. MPEC-ACQ holds at $\tilde{y} = (t, \tilde{w}, u^t, v^t, \tilde{u}^w, \tilde{v}^w)$ for arbitrary $(\tilde{w}, \tilde{u}^w, \tilde{v}^w) \in W(t, u^t, v^t)$, $(\tilde{w}, \tilde{u}^w, \tilde{v}^w) \neq (w, u^w, v^w)$ if MPEC-ACQ holds at y.*

Proof. One has to show

$$\mathcal{T}_{\text{e-abs}}(\tilde{y}) = \mathcal{T}^{\text{lin}}_{\text{e-abs}}(y).$$

This follows directly from the homeomorphism ρ in Lemma 5.50. □

Lemma 5.52. *Consider a feasible point $y = (t, w, u^t, v^t, u^w, v^w)$ of* (E-MPEC)*. MPEC-GCQ holds at $\tilde{y} = (t, \tilde{w}, u^t, v^t, \tilde{u}^w, \tilde{v}^w)$ for arbitrary $(\tilde{w}, \tilde{u}^w, \tilde{v}^w) \in W(t, u^t, v^t)$, $(\tilde{w}, \tilde{u}^w, \tilde{v}^w) \neq (w, u^w, v^w)$ if MPEC-GCQ holds at y.*

Proof. As $\mathcal{T}_{\text{e-mpec}}(\tilde{y})^* \supseteq \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(\tilde{y})^*$ is always satisfied, it is left to show that

$$\mathcal{T}_{\text{e-mpec}}(\tilde{y})^* \subseteq \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(\tilde{y})^*.$$

Consider $\tilde{\omega} = (\omega t, \tilde{\omega} w, \omega u^t, \omega v^t, \tilde{\omega} u^w, \tilde{\omega} v^w) \in \mathcal{T}_{\text{e-abs}}(\tilde{y})^*$, i.e. $\tilde{\omega}^T \tilde{\delta} \geq 0$ for all $\tilde{\delta} \in \mathcal{T}_{\text{e-mpec}}(\tilde{y})^*$. Then, set $\omega = (\omega t, \Sigma \tilde{\omega} w, \omega u^t, \omega v^t, \Sigma \tilde{\omega} u^w, \Sigma \tilde{\omega} v^w)$ and thus $\omega^T \delta = \omega^T \rho^{-1}(\tilde{\delta}) = \tilde{\omega} \tilde{\delta} \geq 0$ for all $\delta \in \mathcal{T}_{\text{e-mpec}}(y)$. This is $\omega \in \mathcal{T}_{\text{e-mpec}}(y)^*$ and further $\omega \in \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(y)^*$ by the assumption. Hence, $\tilde{\omega}^T \tilde{\delta} = \omega^T \delta \geq 0$ for all $\mathcal{T}^{\text{lin}}_{\text{e-mpec}}(\tilde{y})$ which means $\omega \in \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(\tilde{y})^*$. □

As before, the MPEC-CQs are implied if they are satisfied for all branch problems.

Theorem 5.53 (ACQ for all (NLP($\mathcal{P}^{t,w}$)) implies MPEC-ACQ for (E-MPEC))**.** *Consider a feasible point $y = (t, w, u^t, v^t, u^w, v^w)$ of (E-MPEC) with associated branch problems (NLP($\mathcal{P}^{t,w}$)). Then, MPEC-ACQ holds for (E-MPEC) at y if ACQ holds for all (NLP($\mathcal{P}^{t,w}$)) at y.*

Proof. This follows from Theorem 5.39. □

Theorem 5.54 (GCQ for all (NLP($\mathcal{P}^{t,w}$)) implies MPEC-GCQ for (E-MPEC))**.** *Consider a feasible point $y = (t, w, u^t, v^t, u^w, v^w)$ of (E-MPEC) with associated branch problems (NLP($\mathcal{P}^{t,w}$)). Then, MPEC-GCQ holds for (E-MPEC) at y if GCQ holds for all (NLP($\mathcal{P}^{t,w}$)) at y.*

Proof. This follows from Theorem 5.40. □

Relations of MPEC Constraint Qualifications In this paragraph relations between constraint qualifications for the two different formulations (I-MPEC) and (E-MPEC) are proven. Some relations follow from the results in the previous section and in the two following sections. For an illustration, see Fig. 5.1 and Fig. 5.2.

Theorem 5.55. *MPEC-MFCQ for (I-MPEC) holds at $(t, u^t, v^t) \in \mathcal{F}_{\textit{i-mpec}}$ if MPEC-MFCQ for (E-MPEC) holds at $(t, w, u^t, u^w, v^t, v^w) \in \mathcal{F}_{\textit{e-mpec}}$ for any (and hence all) $(w, u^w, v^w) \in W(t, u^t, v^t)$. The converse is not true.*

Proof. This follows directly from Theorem 5.24, Theorem 5.60 and Corollary 5.65. □

Theorem 5.56. *MPEC-ACQ for (I-MPEC) holds at $(t, u^t, v^t) \in \mathcal{F}_{\textit{i-mpec}}$ if and only if MPEC-ACQ for (E-MPEC) holds at $(t, w, u^t, u^w, v^t, v^w) \in \mathcal{F}_{\textit{e-mpec}}$ for any (and hence all) $(w, u^w, v^w) \in W(t, u^t, v^t)$.*

Proof. This follows immediately from Theorem 5.25, Theorem 5.61 and Theorem 5.66. □

Theorem 5.57. *MPEC-GCQ for (I-MPEC) holds at $(t, u^t, v^t) \in \mathcal{F}_{\textit{i-mpec}}$ if MPEC-GCQ for (E-MPEC) holds at $(t, w, u^t, v^t, u^w, v^w) \in \mathcal{F}_{\textit{e-mpec}}$ for any (and hence all) $(w, u^w, v^w) \in W(t, u^t, v^t)$.*

Proof. The inclusion $\mathcal{T}_{\text{i-mpec}}(t, u^t, v^t)^* \supseteq \mathcal{T}^{\text{lin}}_{\text{i-mpec}}(t, u^t, v^t)^*$ always holds. Thus, it is left to show that

$$\mathcal{T}_{\text{i-mpec}}(t, u^t, v^t)^* \subseteq \mathcal{T}^{\text{lin}}_{\text{i-mpec}}(t, u^t, v^t)^*.$$

Let $\omega = (\omega t, \omega u^t, \omega v^t) \in \mathcal{T}_{\text{i-mpec}}(t, u^t, v^t)^*$, i.e. $\omega^T \delta \geq 0$ for all $\delta \in \mathcal{T}_{\text{i-mpec}}(t, u^t, v^t)$. Then, define $\tilde{\omega} = (\omega t, 0, \omega u^t, \omega v^t, 0, 0)$ and obtain $\tilde{\omega}^T \tilde{\delta} = \omega^T \delta \geq 0$ for all $\tilde{\delta} \in \mathcal{T}_{\text{e-mpec}}(t, w, u^t, v^t, u^w, v^w)$ where $w \in W(t)$ is arbitrary. By assumption, $\tilde{\omega}^T \tilde{\delta} \geq 0$ for all $\tilde{\delta} \in \mathcal{T}^{\text{lin}}_{\text{e-mpec}}(t, w, u^t, v^t, u^w, v^w)$ holds, which implies $\omega^T \delta = \tilde{\omega}^T \tilde{\delta} \geq 0$ for all $\delta \in \mathcal{T}^{\text{lin}}_{\text{i-mpec}}(t, u^t, v^t)$. □

It is still an open question if the converse holds. Nevertheless, when moving to the branch problems, equivalence for ACQ and GCQ holds.

Theorem 5.58. *ACQ for (NLP($\mathcal{P}^t$)) holds at $(t, u^t, v^t) \in \mathcal{F}_{\mathcal{P}^t}$ if and only if ACQ for (NLP($\mathcal{P}^{t,w}$)) holds at $(t, w, u^t, v^t, u^w, v^w) \in \mathcal{F}_{\mathcal{P}^{t,w}}$ for any $(w, u^w, v^w) \in W(t, u^t, v^t)$.*

Proof. This follows immediately from Theorem 5.27, Theorem 5.63 and Theorem 5.68. □

Theorem 5.59. *GCQ for (NLP($\mathcal{P}^t$)) holds at $(t, u^t, v^t) \in \mathcal{F}_{\mathcal{P}^t}$ if and only if GCQ for (NLP($\mathcal{P}^{t,w}$)) holds at $(t, w, u^t, v^t, u^w, v^w) \in \mathcal{F}_{\mathcal{P}^{t,w}}$ for any $(w, u^w, v^w) \in W(t, u^t, v^t)$.*

Proof. This follows immediately from Theorem 5.28, Theorem 5.64 and Theorem 5.69. □

5.3 Kink Qualifications and MPEC Constraint Qualifications

In this section a closer look is taken at relations between abs-normal NLPs and counterpart MPECs in both formulations.

Relations of (I-NLP) and (I-MPEC) First, relations between KQs for (I-NLP) and MPEC-CQs for (I-MPEC) of Mangasarian Fromowitz, Abadie and Guignard type are considered. In the following the variables x and z instead of t and z^t are used. Thus, the abs-normal NLP (I-NLP) reads:

$$\begin{aligned} \min_{x,z} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, |z|) = 0, \\ & c_{\mathcal{I}}(x, |z|) \geq 0, \\ & c_{\mathcal{Z}}(x, |z|) - z = 0. \end{aligned}$$

The counterpart MPEC (I-MPEC) becomes:

$$\begin{aligned} \min_{x,u,v} \quad & f(x) \\ \text{s.t.} \quad & c_{\mathcal{E}}(x, u + v) = 0, \\ & c_{\mathcal{I}}(x, u + v) \geq 0, \\ & c_{\mathcal{Z}}(x, u + v) - (u - v) = 0, \\ & 0 \leq u \perp v \geq 0. \end{aligned}$$

Then, the subsequent relations of kink qualifications and MPEC constraint qualifications can be shown.

Theorem 5.60 (Equivalence of IDKQ and MPEC-MFCQ)**.** *IDKQ for (I-NLP) holds at $(x, z(x)) \in \mathcal{F}_{\text{i-abs}}$ if and only if MPEC-MFCQ for (I-MPEC) holds at $(x, u, v) \in \mathcal{F}_{\text{i-mpec}}$ with $(x, u, v) = (x, [z(x)]^+, [z(x)]^-)$.*

Proof. Using the short notation $y := (x, u + v)$, MPEC-MFCQ for the counterpart MPEC is

1. full row rank of

$$\begin{bmatrix} \partial_1 c_{\mathcal{E}}(y) & \partial_2 c_{\mathcal{E}}(y) P_{\mathcal{U}_+}^T & \partial_2 c_{\mathcal{E}}(y) P_{\mathcal{V}_+}^T \\ \partial_1 c_{\mathcal{Z}}(y) & [\partial_2 c_{\mathcal{Z}}(y) - I] P_{\mathcal{U}_+}^T & [\partial_2 c_{\mathcal{Z}}(y) + I] P_{\mathcal{V}_+}^T \end{bmatrix} \in \mathbb{R}^{(m_1+s)\times(n+|\mathcal{U}_+ \cup \mathcal{V}_+|)}.$$

As in the proof of Theorem 4.8, this is seen to be full row rank of

$$\begin{bmatrix} \partial_x c_{\mathcal{E}}(x, |z(x)|) \\ [e_i^T \partial_x z(x)]_{i\in\alpha} \end{bmatrix} \in \mathbb{R}^{(m_1+|\alpha|)\times n}.$$

2. the existence of a vector $d = (d_{\mathrm{x}}, d_{\mathrm{u}}, d_{\mathrm{v}}) \in \mathbb{R}^{n+|\mathcal{U}_+\cup\mathcal{V}_+|}$ such that

$$\begin{bmatrix} \partial_1 c_{\mathcal{E}}(y) & \partial_2 c_{\mathcal{E}}(y) P_{\mathcal{U}_+}^T & \partial_2 c_{\mathcal{E}}(y) P_{\mathcal{V}_+}^T \\ \partial_1 c_{\mathcal{Z}}(y) & [\partial_2 c_{\mathcal{Z}}(y) - I] P_{\mathcal{U}_+}^T & [\partial_2 c_{\mathcal{Z}}(y) + I] P_{\mathcal{V}_+}^T \end{bmatrix} d = 0,$$
$$\begin{bmatrix} \partial_1 c_{\mathcal{A}}(y) & \partial_2 c_{\mathcal{A}}(y) P_{\mathcal{U}_+}^T & \partial_2 c_{\mathcal{A}}(y) P_{\mathcal{V}_+}^T \end{bmatrix} d > 0.$$

Now, d_u and $-d_v$ are combined to $d_{uv} \in \mathbb{R}^{|\mathcal{U}_+\cup\mathcal{V}_+|}$. Then, this is equivalent to

$$\begin{aligned} \partial_1 c_{\mathcal{E}}(y) d_x + \partial_2 c_{\mathcal{E}}(y) \Sigma P_{\mathcal{U}_+\cup\mathcal{V}_+}^T d_{uv} &= 0, \\ \partial_1 c_{\mathcal{Z}}(y) d_x - [I - \partial_2 c_{\mathcal{Z}}(y)\Sigma] P_{\mathcal{U}_+\cup\mathcal{V}_+}^T d_{uv} &= 0, \\ \partial_1 c_{\mathcal{A}}(y) d_x + \partial_2 c_{\mathcal{A}}(y) \Sigma P_{\mathcal{U}_+\cup\mathcal{V}_+}^T d_{uv} &> 0 \end{aligned}$$

for $\Sigma = \operatorname{diag}(\sigma)$ with $\sigma = \sigma(x)$. The second condition can be written as

$$[I - \partial_2 c_{\mathcal{Z}}(y)\Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(y) d_x = P_{\mathcal{U}_+\cup\mathcal{V}_+}^T d_{uv}. \tag{5.1}$$

Multiplying this by $P_{\mathcal{D}}^T = P_\alpha^T$ and using $u + v = |z(x)|$ yield

$$\left[e_i^T [I - \partial_2 c_{\mathcal{Z}}(y)\Sigma]^{-1} \partial_1 c_{\mathcal{Z}}(y)\right]_{i\in\alpha} d_x = [e_i^T \partial_x z(x)]_{i\in\alpha} d_x = 0.$$

Substituting the right-hand side of (5.1) into the first and third condition with $u+v = |z(x)|$, finally gives

$$\begin{aligned} \partial_x c_{\mathcal{E}}(x, |z(x)|) d_x &= 0, \\ [e_i^T \partial_x z(x)]_{i\in\alpha} d_x &= 0, \\ \partial_x c_{\mathcal{A}}(x, |z(x)|) d_x &> 0, \end{aligned}$$

which is IDKQ for the abs-normal NLP. □

Theorem 5.61 (Equivalence AKQ and MPEC-ACQ). *AKQ for (I-NLP) holds at $(x, z(x)) \in \mathcal{F}_{\text{i-abs}}$ if and only if MPEC-ACQ for (I-MPEC) holds at $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\text{i-mpec}}$.*

Proof. One has to show

$$\mathcal{T}_{\text{i-abs}}(x, z) = \mathcal{T}_{\text{i-abs}}^{\text{lin}}(x, z) \iff \mathcal{T}_{\text{i-mpec}}(x, u, v) = \mathcal{T}_{\text{i-mpec}}^{\text{lin}}(x, u, v).$$

This is obvious from the homeomorphisms ψ in Lemma 5.32. □

Theorem 5.62 (MPEC-GCQ implies GKQ). *GKQ for (I-NLP) holds at $(x, z(x)) \in \mathcal{F}_{\text{i-abs}}$ if MPEC-GCQ for (I-MPEC) holds at $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\text{i-mpec}}$.*

Proof. The inclusion $\mathcal{T}^{\text{lin}}_{\text{i-abs}}(x,z)^* \subseteq \mathcal{T}_{\text{i-abs}}(x,z)^*$ always holds by Lemma 5.8. Thus, one just has to show

$$\mathcal{T}_{\text{i-abs}}(x,z)^* \subseteq \mathcal{T}^{\text{lin}}_{\text{i-abs}}(x,z)^*.$$

Consider $\omega = (\omega x, \omega z) \in \mathcal{T}_{\text{i-abs}}(x,z)^*$, i.e. $\omega^T\delta \geq 0$ for all $\delta = (\delta x, \delta z) \in \mathcal{T}_{\text{i-abs}}(x,z)$, and set $\tilde{\omega} = (\omega x, \omega z, -\omega z)$. Then,

$$\tilde{\omega}^T\psi^{-1}(\delta) = \omega x^T\delta x + \omega z^T\langle\delta z\rangle^+ - \omega z^T\langle\delta z\rangle^- = \omega x^T\delta x + \omega z^T\delta z = \omega^T\delta \geq 0.$$

holds for every $\delta \in \mathcal{T}_{\text{i-abs}}(x,z)$. This means $\tilde{\omega} \in \mathcal{T}_{\text{i-mpec}}(x,u,v)^*$ and hence, by assumption, $\tilde{\omega} \in \mathcal{T}^{\text{lin}}_{\text{i-mpec}}(x,u,v)^*$. Thus, $\omega^T\delta = \tilde{\omega}^T\psi^{-1}(\delta) \geq 0$ for every $\delta \in \mathcal{T}^{\text{lin}}_{\text{i-abs}}(x,z)$, which means $\omega \in \mathcal{T}^{\text{lin}}_{\text{i-abs}}(x,z)^*$. □

It is an open question if the converse holds. Once again, moving to branch problems allows to exploit additional sign information.

Theorem 5.63 (Equivalence of ACQ for (NLP(Σ^t)) and ACQ for (NLP($\mathcal{P}^t$))). *ACQ for (NLP(Σ^t)) holds at $(x, z(x)) \in \mathcal{F}_{\Sigma^t}$ if and only if ACQ for the corresponding (NLP($\mathcal{P}^t$)) holds at $(x,u,v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\mathcal{P}^t}$.*

Proof. One just need to show

$$\mathcal{T}_{\Sigma^t}(x,z) = \mathcal{T}^{\text{lin}}_{\Sigma^t}(x,z) \iff \mathcal{T}_{\mathcal{P}^t}(x,u,v) = \mathcal{T}^{\text{lin}}_{\mathcal{P}^t}(x,u,v).$$

This is obvious from the homeomorphisms $\psi_{\mathcal{P}}$ in Lemma 5.36. □

Theorem 5.64 (Equivalence of GCQ for (NLP(Σ^t)) and GCQ for (NLP($\mathcal{P}^t$))). *GCQ for (NLP(Σ^t)) holds at $(x, z(x)) \in \mathcal{F}_{\Sigma^t}$ if and only if GCQ for the corresponding (NLP($\mathcal{P}^t$)) holds at $(x,u,v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\mathcal{P}^t}$.*

Proof. The inclusions $\mathcal{T}^{\text{lin}}_{\mathcal{P}^t}(x,u,v)^* \subseteq \mathcal{T}_{\mathcal{P}^t}(x,u,v)^*$ and $\mathcal{T}^{\text{lin}}_{\Sigma^t}(x,z)^* \subseteq \mathcal{T}_{\Sigma^t}(x,z)^*$ hold always. Thus, one just has to show

$$\mathcal{T}_{\Sigma^t}(x,z)^* \supseteq \mathcal{T}^{\text{lin}}_{\Sigma^t}(x,z)^* \iff \mathcal{T}_{\mathcal{P}^t}(x,u,v)^* \supseteq \mathcal{T}^{\text{lin}}_{\mathcal{P}^t}(x,u,v)^*.$$

First, the implication "⇒" is shown. Consider $\omega = (\omega x, \omega u, \omega v) \in \mathcal{T}_{\mathcal{P}^t}(x,u,v)^*$, i.e. $\omega^T\delta \geq 0$ for all $\delta = (\delta x, \delta u, \delta v) \in \mathcal{T}_{\mathcal{P}^t}(x,u,v)$. Set $\tilde{\omega} = (\omega x, \omega z)$ with

$$\omega z_i = \begin{cases} +\omega u_i, & i \in \mathcal{U}_+ \cup \bar{\mathcal{P}}, \\ -\omega v_i, & i \in \mathcal{V}_+ \cup \mathcal{P}. \end{cases}$$

This leads to

$$\tilde{\omega}^T\psi_{\mathcal{P}}(\delta) = \omega x^T\delta x + \omega z^T(\delta u - \delta v) = \omega x^T\delta x + \omega u^T\delta u + \omega v^T\delta v = \omega^T\delta \geq 0$$

for every $\delta \in \mathcal{T}_{\mathcal{P}^t}(x,u,v)$, i.e. $\tilde{\omega} \in \mathcal{T}_{\Sigma^t}(x,z)^*$. Then, the assumption yields $\tilde{\omega} \in \mathcal{T}^{\text{lin}}_{\Sigma^t}(x,z)^*$. As $\omega^T\delta = \tilde{\omega}^T\psi_{\mathcal{P}}(\delta) \geq 0$ is satisfied for every $\delta \in \mathcal{T}^{\text{lin}}_{\mathcal{P}^t}(x,u,v)$, $\omega \in \mathcal{T}^{\text{lin}}_{\mathcal{P}^t}(x,u,v)^*$ follows. The reverse implication can be proven as in Theorem 5.62. Thus, consider $\omega = (\omega x, \omega z) \in \mathcal{T}_{\Sigma^t}(x,z)^*$. Then, $\omega^T\delta \geq 0$ for all $\delta = (\delta x, \delta z) \in \mathcal{T}_{\Sigma^t}(x,z)$ and with $\tilde{\omega} = (\omega x, \omega z, -\omega z)$ this implies

$$\tilde{\omega}^T\psi_{\mathcal{P}}^{-1}(\delta) = \omega x^T\delta x + \omega z^T\langle\delta z\rangle^+ - \omega z^T\langle\delta z\rangle^- = \omega x^T\delta x + \omega z^T\delta z = \omega^T\delta \geq 0$$

for every $\delta \in \mathcal{T}_{\Sigma^t}(x,z)$, i.e. $\tilde{\omega} \in \mathcal{T}_{\mathcal{P}^t}(x,u,v)^*$. Then, $\tilde{\omega} \in \mathcal{T}^{\text{lin}}_{\mathcal{P}^t}(x,u,v)^*$ by the assumption and $\omega^T\delta = \tilde{\omega}^T\psi_{\mathcal{P}}^{-1}(\delta) \geq 0$ holds for every $\delta \in \mathcal{T}^{\text{lin}}_{\Sigma^t}(x,z)$. This proves $\omega \in \mathcal{T}^{\text{lin}}_{\Sigma^t}(x,z)^*$. □

Relations of (E-NLP) and (E-MPEC) Now, the KQs and MPEC-CQs of Mangasarian Fromovitz, Abadie and Guignard type are compared for the slack formulations. As both slack formulations can be seen as special cases of the direct handling, the results of the previous paragraph hold here.

However, the equivalence of IDKQ and MPEC-MFCQ follows directly as it holds for LIKQ and MPEC-MFCQ (see Theorem 4.9) and LIKQ and IDKQ as well as MPEC-LICQ and MPEC-MFCQ coincide in the purely equality constrained setting.

Corollary 5.65 (Equivalence of IDKQ and MPEC-MFCQ). *IDKQ for* (E-NLP) *holds at* $(x, z(x)) \in \mathcal{F}_{\text{e-abs}}$ *if and only if MPEC-MFCQ for* (E-MPEC) *holds at* $(x, u, v) \in \mathcal{F}_{\text{e-mpec}}$ *with* $(x, u, v) = (x, [z(x)]^+, [z(x)]^-)$.

But to prove relations for Abadie and Guignard type regularity assumptions, the formulations $(\overline{\text{E-NLP}})$ and $(\overline{\text{E-MPEC}})$ and the results of the previous paragraph will be used.

Theorem 5.66 (Equivalence of AKQ for (E-NLP) and MPEC-ACQ for (E-MPEC)). *AKQ for* (E-NLP) *holds at* $x \in \mathcal{F}_{\text{e-abs}}$ *if and only if MPEC-ACQ for* (E-MPEC) *holds at* $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\text{e-mpec}}$.

Proof. The short notation $(\overline{\text{E-NLP}})$ was used to formulate AKQ in Lemma 5.17 and, similarly, the short notation $(\overline{\text{E-MPEC}})$ to formulate MPEC-ACQ Lemma 5.48. Then, Theorem 5.61 can be applied and gives the result as $(\overline{\text{E-MPEC}})$ is the counterpart MPEC for $(\overline{\text{E-NLP}})$. □

Theorem 5.67 (MPEC-GCQ for (E-MPEC) implies GKQ for (E-NLP)). *GKQ for* (E-NLP) *holds at* $x \in \mathcal{F}_{\text{e-abs}}$ *if MPEC-GCQ for* (E-MPEC) *holds at* $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\text{e-mpec}}$.

Proof. The short notation $(\overline{\text{E-NLP}})$ was used to formulate GKQ in Lemma 5.18 and, similarly, the short notation $(\overline{\text{E-MPEC}})$ to formulate MPEC-GCQ in Lemma 5.49. Then, Theorem 5.62 can be applied and gives the result as $(\overline{\text{E-MPEC}})$ is the counterpart MPEC for $(\overline{\text{E-NLP}})$. □

As before, it is still an open question if the converse holds.

Theorem 5.68 (Equivalence of ACQ for (NLP($\Sigma^{t,w}$)) and ACQ for (NLP($\mathcal{P}^{t,w}$))). *ACQ for* (NLP($\Sigma^{t,w}$)) *holds at* $(x, z(x)) \in \mathcal{F}_{\Sigma^{t,w}}$ *if and only if ACQ for the corresponding* (NLP($\mathcal{P}^{t,w}$)) *holds at* $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\mathcal{P}^{t,w}}$.

Proof. This follows from Theorem 5.63 as $(\overline{\text{E-MPEC}})$ is the counterpart MPEC for $(\overline{\text{E-NLP}})$ and these were used to state ACQ for (NLP($\Sigma^{t,w}$)) and (NLP($\mathcal{P}^{t,w}$)). □

Theorem 5.69 (Equivalence of GCQ for (NLP($\Sigma^{t,w}$)) and GCQ for (NLP($\mathcal{P}^{t,w}$))). *GCQ for* (NLP($\Sigma^{t,w}$)) *holds at* $(x, z(x)) \in \mathcal{F}_{\Sigma^{t,w}}$ *if and only if GCQ for the corresponding* (NLP($\mathcal{P}^{t,w}$)) *holds at* $(x, u, v) = (x, [z(x)]^+, [z(x)]^-) \in \mathcal{F}_{\mathcal{P}^{t,w}}$.

Proof. This follows from Theorem 5.64 as $(\overline{\text{E-MPEC}})$ is the counterpart MPEC for $(\overline{\text{E-NLP}})$ and these were used to state GCQ for (NLP($\Sigma^{t,w}$)) and (NLP($\mathcal{P}^{t,w}$)). □

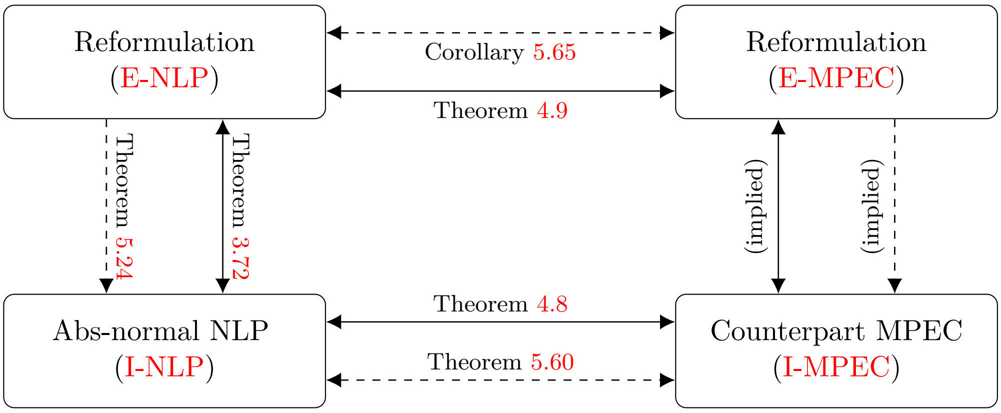

Figure 5.1: Solid arrows: relations between LIKQ and MPEC-LICQ; dashed arrows: relations between IDKQ and MPEC-MFCQ.

5.4 Optimality Conditions

In this section, Mordukhovich stationarity and Bouligand stationarity are defined for the abs-normal NLP and compared to the definitions of M-stationarity and B-stationarity for MPECs.

Mordukhovich Stationarity In this paragraph a closer look is taken at M-stationarity which is a necessary optimality condition for MPECs under MPEC-GCQ. But as the converse of Theorem 5.62 cannot be proven so far, MPEC-ACQ will be considered instead.

First, the definition of M-stationarity is stated for (I-MPEC).

Lemma 5.70 (M-Stationarity for (I-MPEC))**.** *Given* (I-MPEC)*, consider a feasible point* (x^*, u^*, v^*) *with associated index sets* $\mathcal{U}_+$, $\mathcal{V}_+$ *and* $\mathcal{D}$*. It is an* M-stationary *point if and only if there exist Lagrange multiplier vectors* $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{I}}, \lambda^*_{\mathcal{Z}})$ *and* $\mu^* = (\mu^*_u, \mu^*_v)$ *such that the following conditions are satisfied:*

$$\partial_{x,u,v}\mathcal{L}_\perp(x^*, u^*, v^*, \lambda^*, \mu^*) = 0, \tag{5.2a}$$
$$((\mu^*_u)_i > 0,\ (\mu^*_v)_i > 0)\ \vee\ (\mu^*_u)_i(\mu^*_v)_i = 0,\ \ i \in \mathcal{D} \tag{5.2b}$$
$$(\mu^*_u)_i = 0,\ \ i \in \mathcal{U}_+, \tag{5.2c}$$
$$(\mu^*_v)_i = 0,\ \ i \in \mathcal{V}_+, \tag{5.2d}$$
$$(\lambda^*_{\mathcal{I}}) \ge 0, \tag{5.2e}$$
$$(\lambda^*_{\mathcal{I}})^T c_{\mathcal{I}}(x^*, u^*, v^*) = 0. \tag{5.2f}$$

Recall that $\mathcal{L}_\perp$ *is the MPEC-Lagrangian function*

$$\begin{aligned}\mathcal{L}_\perp(x, u, v, \lambda, \mu) = f(x) &+ \lambda^T_{\mathcal{E}} c_{\mathcal{E}}(x, u + v) - \lambda^T_{\mathcal{I}} c_{\mathcal{I}}(x, u + v) \\ &+ \lambda^T_{\mathcal{Z}}[c_{\mathcal{Z}}(x, u + v) - (u - v)] - \mu^T_u u - \mu^T_v v.\end{aligned}$$

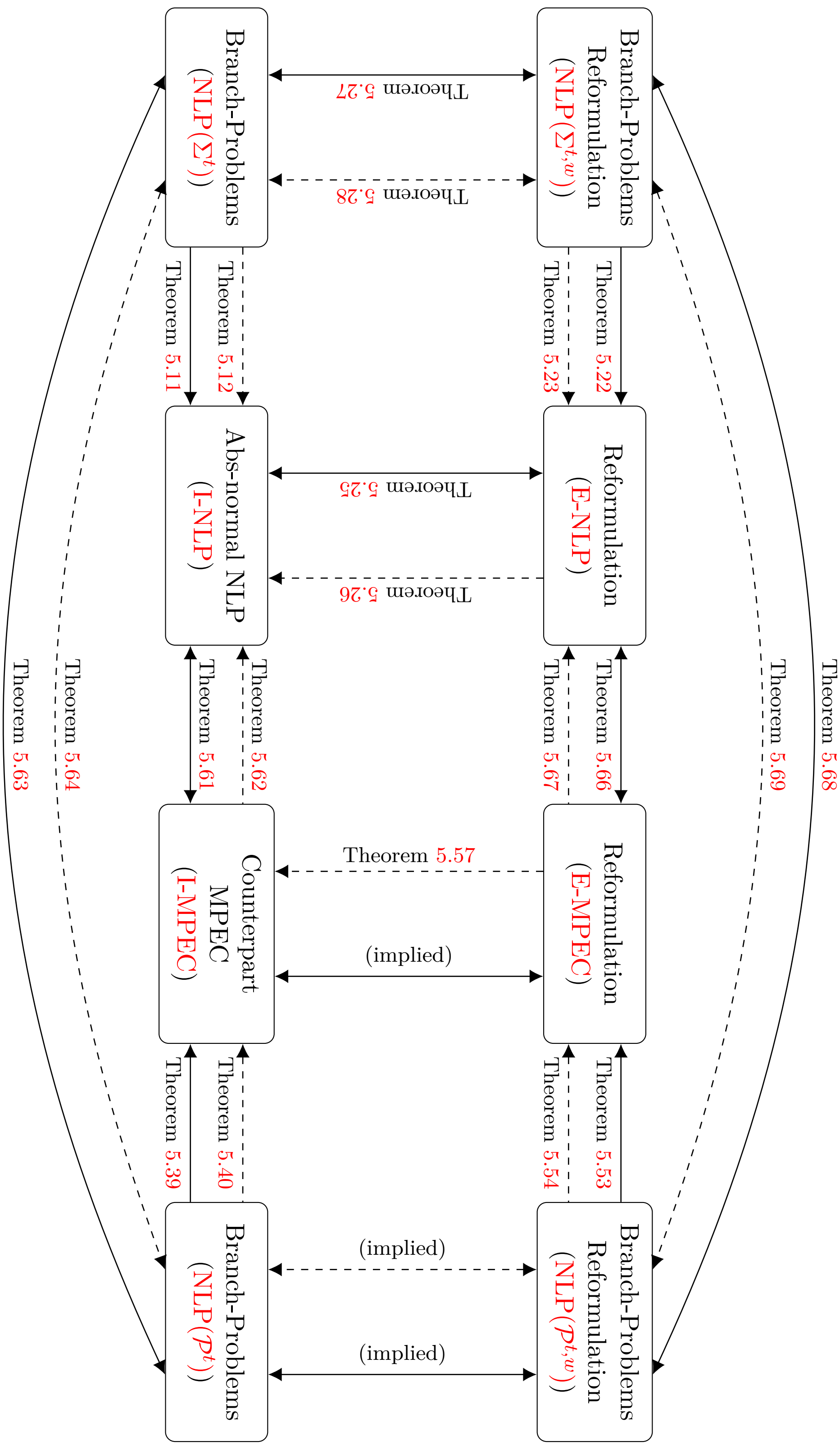

Figure 5.2: Solid arrows: relations between AKQ and MPEC-ACQ; dashed arrows: relations between GKQ and MPEC-GCQ. Note that in Theorem 5.26, Theorem 5.57, Theorem 5.62 and Theorem 5.67 only one-sided implications are proven and it is open whether the reverse implications hold.

Proof. This follows directly from Definition 2.63. □

Then, M-stationarity for (I-NLP) is defined such that it is equivalent to M-Stationarity for (I-MPEC).

Definition 5.71 (M-Stationarity for (I-NLP))**.** Consider a feasible point (x^*, z^*) of (I-NLP). It is an *M-stationary* point if there exist a Lagrange multiplier vector $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{I}}, \lambda^*_{\mathcal{Z}})$ such that the following conditions are satisfied:

$$f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 c_{\mathcal{E}} - (\lambda^*_{\mathcal{I}})^T \partial_1 c_{\mathcal{I}} + (\lambda^*_{\mathcal{Z}})^T \partial_1 c_{\mathcal{Z}} = 0, \tag{5.3a}$$

$$[(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}} - (\lambda^*_{\mathcal{I}})^T \partial_2 c_{\mathcal{I}} + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}]_i = (\lambda^*_{\mathcal{Z}})_i \sigma^*_i, \; i \notin \alpha(x^*), \tag{5.3b}$$

$$(\mu^-_i)(\mu^+_i) = 0 \;\vee\; [(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}} - (\lambda^*_{\mathcal{I}})^T \partial_2 c_{\mathcal{I}} + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}]_i > |(\lambda^*_{\mathcal{Z}})_i|, \; i \in \alpha(x^*), \tag{5.3c}$$

$$(\lambda^*_{\mathcal{I}}) \geq 0, \tag{5.3d}$$

$$(\lambda^*_{\mathcal{I}})^T c_{\mathcal{I}} = 0. \tag{5.3e}$$

Here the notation

$$\mu^+_i := \left[(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}} - (\lambda^*_{\mathcal{I}})^T \partial_2 c_{\mathcal{I}} + (\lambda^*_{\mathcal{Z}})^T [\partial_2 c_{\mathcal{Z}} - I]\right]_i,$$
$$\mu^-_i := \left[(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}} - (\lambda^*_{\mathcal{I}})^T \partial_2 c_{\mathcal{I}} + (\lambda^*_{\mathcal{Z}})^T [\partial_2 c_{\mathcal{Z}} + I]\right]_i,$$

is used and the constraints and the partial derivatives are evaluated at $(x^*, |z^*|)$.

Theorem 5.72 (M-Stationarity for (I-MPEC) is M-Stationarity for (I-NLP))**.** *A feasible point (x^*, z^*) of (I-NLP) is M-stationary if and only if $(x^*, u^*, v^*) = (x^*, [z^*]^+, [z^*]^-)$ of (I-MPEC) is M-stationary.*

Proof. For indices that satisfy the first condition in (5.2b), the equivalence with the second condition in (5.3c) was shown in Theorem 4.11. Thus, one just need to consider the alternative conditions. For (I-MPEC) the relations

$$\left[(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}} - (\lambda^*_{\mathcal{I}})^T \partial_2 c_{\mathcal{I}} + (\lambda^*_{\mathcal{Z}})^T [\partial_2 c_{\mathcal{Z}} - I]\right]_i = (\mu^*_{\mathrm{u}})_i, \; i \in \mathcal{D},$$
$$\left[(\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}} - (\lambda^*_{\mathcal{I}})^T \partial_2 c_{\mathcal{I}} + (\lambda^*_{\mathcal{Z}})^T [\partial_2 c_{\mathcal{Z}} + I]\right]_i = (\mu^*_{\mathrm{v}})_i, \; i \in \mathcal{D},$$

hold which was also shown in Theorem 4.11. These are exactly the definitions of μ^+_i and μ^-_i in the definition of M-Stationarity for (I-NLP). □

Finally, it is shown that M-Stationarity for (I-NLP) is a necessary condition under AKQ.

Theorem 5.73 (Minimizers and M-Stationarity for (I-NLP))**.** *Assume that (x^*, z^*) is a local minimizer of (I-NLP) and that AKQ holds at x^*. Then, (x^*, z^*) is M-stationary for (I-NLP).*

Proof. First, note that (x^*, z^*) is a local minimizer of (I-NLP) if and only if $(x^*, u^*, v^*) = (x^*, [z^*]^+, [z^*]^-)$ is a local minimizer of (I-MPEC). Further, MPEC-ACQ holds by Theorem 5.61. Now, Theorem 2.64 implies that (x^*, u^*, v^*) is M-stationary for (I-MPEC) and thus Theorem 5.72 implies that (x^*, z^*) is M-stationary for (I-NLP). □

Linearized Bouligand Stationarity In this paragraph abs-normal-linearized Bouligand stationarity for (I-NLP) is defined and compared to MPEC-linearized Bouligand stationarity which is a necessary optimality condition under MPEC-GCQ. As equivalence between MPEC-GCQ and GKQ cannot be shown so far, GCQs for the branch problems (NLP(Σ^t)) and (NLP($\mathcal{P}^t$)) are considered instead. For these CQs equivalence was shown in Theorem 5.62.

To begin with, MPEC-linearized B-stationarity for (I-MPEC) is formulated.

Lemma 5.74 (MPEC-linearized B-Stationarity for (I-MPEC))**.** *Consider a feasible point (x^*, u^*, v^*) of (I-MPEC) with associated index sets $\mathcal{U}_+$, $\mathcal{V}_+$ and $\mathcal{D}$. It is a* B-stationary *point if and only if it is a stationary point of all branch problems (NLP($\mathcal{P}^t$)) for $\mathcal{P}^t = \mathcal{P} \subseteq \mathcal{D}$.*

Proof. This follows directly from Lemma (2.67). $\square$

Next, abs-normal-linearized B-Stationarity for (I-NLP) is defined.

Definition 5.75 (Abs-Normal-Linearized B-Stationarity for (I-NLP))**.** Consider a feasible point (x^*, z^*) of (I-NLP). It is an *abs-normal-linearized B-stationary* point if it is a stationary point of all branch problems (NLP(Σ^t)) for $\Sigma^t = \operatorname{diag}(\sigma)$ with $\sigma \succeq \sigma(x)$.

Both linearized B-stationarity concepts are equivalent. This is shown in the following theorem.

Theorem 5.76 (MPEC-linearized B-stationarity for (I-MPEC) is abs-normal-linearized B-stationarity for (I-NLP))**.** *A feasible point (x^*, z^*) of (I-NLP) is abs-normal-linearized B-stationary if and only if $(x^*, u^*, v^*) = (x^*, [z^*]^+, [z^*]^-)$ of (I-MPEC) is MPEC-linearized B-stationary.*

Proof. Consider (NLP(Σ^t)) for $\Sigma^t = \Sigma = \operatorname{diag}(\sigma)$ with $\sigma \succeq \sigma(x^*)$ and (NLP($\mathcal{P}^t$)) with $\mathcal{P}^t = \mathcal{P} = \{i \in \alpha(x^*) : \sigma_i = -1\}$. Then, $f'(x)\delta x \geq 0$ holds for all $(\delta x, \delta z) \in \mathcal{T}_{\Sigma^t}(x^*, z^*)$ if and only if it holds for all $\delta = (\delta x, \delta u, \delta v) \in \mathcal{T}_{\mathcal{P}^t}(x^*, u^*, v^*)$ by Lemma 5.36. Thus, the result follows as these conditions are equivalent to the definitions of stationary point for the branch problems by Theorem 2.17. $\square$

Then, it can be proven that abs-normal-linearized B-Stationarity is a necessary optimality condition under GCQ for all (NLP(Σ^t)).

Theorem 5.77 (Minimizers and abs-normal-linearized B-Stationarity for (I-NLP))**.** *Assume that (x^*, z^*) is a local minimizer of (I-NLP) and that GCQ holds at x^* for all (NLP(Σ^t)) with $\Sigma^t = \operatorname{diag}(\sigma)$, $\sigma = \sigma(x^*)$. Then, (x^*, z^*) is abs-normal-linearized B-stationary for (I-NLP).*

Proof. The point (x^*, z^*) is a local minimizer of (I-NLP) if and only if $(x^*, u^*, v^*) = (x^*, [z^*]^+, [z^*]^-)$ is a local minimizer of (I-MPEC). Moreover, GCQ for all (NLP(Σ^t)) and GCQ for all (NLP($\mathcal{P}^t$)) are equivalent by Theorem 5.64. Thus, (x^*, u^*, v^*) is MPEC-linearized B-stationary by Corollary 2.68 and finally (x^*, z^*) is abs-normal-linearized B-stationary by Theorem 5.76. $\square$

5.5 Handling of Nonsmooth Objective Function

This section checks if the previous weaker kink qualification are preserved under the reformulation of a nonsmooth objective function. To this end, the kink qualifactions are stated for the two formulations (F-NLP) and (C-NLP). Note that IDKQ is not considered here as it coincides with LIKQ in both settings. This is due to the fact that both formulations do not contain any inequalities.

Direct Handling First, recall the formulation (F-NLP), where the objective is handled directly:

$$\begin{aligned} \min_{t,z^t} \quad & c_f(t, |z^t|) \\ \text{s.t.} \quad & c_{\mathcal{E}}(t, |z^t|) = 0, \\ & c_{\mathcal{Z}}(t, |z^t|) - z^t = 0, \end{aligned}$$

with feasible set

$$\begin{aligned} \mathcal{F}_{\mathrm{f}} =& \{(t, z^t) \colon c_{\mathcal{E}}(t, |z^t|) = 0,\ c_{\mathcal{Z}}(t, |z^t|) - z^t = 0\} \\ =& \{(t, z^t(t)) \colon t \in D^t,\ c_{\mathcal{E}}(t, |z^t|) = 0\}. \end{aligned}$$

Next, the cones in Abadie's and Guignard's kink qualifications are formulated for (F-NLP).

Definition 5.78 (Tangential Cone and Abs-Normal-Linearized Cone for (F-NLP))**.** Consider a feasible point (t, z^t) of (F-NLP). The *tangential cone* to $\mathcal{F}_{\mathrm{f}}$ at (t, z^t) is

$$\mathcal{T}_{\text{f-abs}}(t, z^t) := \left\{ (\delta t, \delta z^t) \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{\mathrm{f}} \ni (t_k, z_k^t) \to (t, z^t) \colon \\ \tau_k^{-1}(t_k - t, z_k^t - z^t) \to (\delta t, \delta z^t) \end{array} \right\}.$$

With $\delta\zeta_i = |\delta z_i^t|$ if $i \in \alpha^t(t)$ and $\delta\zeta_i = \sigma_i(t)\delta z_i^t$ if $i \notin \alpha^t(t)$, the *abs-normal-linearized cone* is

$$\mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t) := \left\{ (\delta t, \delta z^t) \;\middle|\; \begin{array}{l} \partial_1 c_{\mathcal{E}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{E}}(t, |z^t|)\delta\zeta = 0, \\ \partial_1 c_{\mathcal{Z}}(t, |z^t|)\delta t + \partial_2 c_{\mathcal{Z}}(t, |z^t|)\delta\zeta = \delta z^t \end{array} \right\}.$$

Note that

$$\mathcal{T}_{\text{f-abs}}(t, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t) \quad \text{and} \quad \mathcal{T}_{\text{f-abs}}(t, z^t)^* \supseteq \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t)^*.$$

This follows from Lemma 5.8 as the definitions of $\mathcal{F}_{\mathrm{f}}$ and thus $\mathcal{T}_{\text{f-abs}}(t, z^t)$ as well as of $\mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t)$ do not depend on the objective function.

As the converses do not hold, AKQ and GKQ will be defined next.

Definition 5.79 (AKQ for (F-NLP))**.** Consider a feasible point t of (F-NLP). One says that *Abadie's kink qualification (AKQ)* holds for (F-NLP) at t if $\mathcal{T}_{\text{f-abs}}(t, z^t(t)) = \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t(t))$.

Definition 5.80 (GKQ for (F-NLP))**.** Consider a feasible point t of (F-NLP). One says that *Guginard's kink qualification (GKQ)* holds for (F-NLP) at t reads $\mathcal{T}_{\text{f-abs}}(t, z^t(t))^* = \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t(t))^*$.

Constant Objective Introducing an additional variable c the reformulation in abs-normal form (C-NLP) reads:

$$\begin{aligned}\min_{t,c,z^t} \quad & c \\ \text{s.t.} \quad & c_f(t,|z^t|) - c = 0, \\ & c_{\mathcal{E}}(t,|z^t|) = 0, \\ & c_{\mathcal{Z}}(t,|z^t|) - z^t = 0,\end{aligned}$$

with feasible set

$$\begin{aligned}\mathcal{F}_{\mathrm{c}} &= \{(t,c,z^t)\colon c_f(t,|z^t|) - c = 0,\ c_{\mathcal{E}}(t,|z^t|) = 0,\ c_{\mathcal{Z}}(t,|z^t|) - z^t = 0\} \\ &= \{(t,c,z^t)\colon (t,z^t) \in \mathcal{F}_{\mathrm{f}},\ c = c_f(t,|z^t|)\}.\end{aligned}$$

To obtain the kink qualifications for (C-NLP), the definitions of subsection 5.1 will be applied to the reformulation ($\overline{\text{C-NLP}}$):

$$\begin{aligned}\min_{x,z} \quad & f(x) \\ \text{s.t.} \quad & \bar{c}_{\mathcal{E}}(x,|z|) = 0, \\ & \bar{c}_{\mathcal{Z}}(x,|z|) - z = 0,\end{aligned}$$

with $x = (t,c)$, $z = z^t$, $\bar{f}(x) = c$, $\bar{c}_{\mathcal{E}}(x,|z|) = (c_f(t,|z^t|) - c, c_{\mathcal{E}}(t,|z^t|))$ and $\bar{c}_{\mathcal{Z}}(x,|z|) = c_{\mathcal{Z}}(t,|z^t|)$.

Lemma 5.81 (Tangential Cone and Abs-Normal-Linearized Cone for (C-NLP))**.** *Consider a feasible point (t,c,z^t) of (C-NLP). The tangential cone to $\mathcal{F}_c$ at (t,c,z^t) reads*

$$\mathcal{T}_{\textit{c-abs}}(t,c,z^t) := \left\{ (\delta t, \delta c, \delta z^t) \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_c \ni (t_k, c_k, z_k^t) \to (t,c,z^t)\colon \\ \tau_k^{-1}(t_k - t, c_k - c, z_k^t - z^t) \to (\delta t, \delta c, \delta z^t) \end{array} \right\}.$$

With $\delta\zeta_i = |\delta z_i^t|$ if $i \in \alpha^t(t)$ and $\delta\zeta_i = \sigma_i(t)\delta z_i^t$ if $i \notin \alpha^t(t)$, the abs-normal-linearized cone reads

$$\mathcal{T}^{\textit{lin}}_{\textit{c-abs}}(t,c,z^t) := \left\{ (\delta t, \delta c, \delta z^t) \;\middle|\; \begin{array}{r} \partial_1 c_f(t,|z^t|)\delta t + \partial_2 c_f(t,|z^t|)\delta\zeta = \delta c, \\ (\delta t, \delta z^t) \in \mathcal{T}^{\textit{lin}}_{\textit{f-abs}}(t,z^t) \end{array} \right\}.$$

Proof. Definition 5.3 applied to ($\overline{\text{C-NLP}}$) gives $\mathcal{T}_{\text{c-abs}}(t,c,z^t)$ above and

$$\mathcal{T}^{\text{lin}}_{\text{c-abs}}(t,c,z^t) = \left\{ (\delta t, \delta c, \delta z^t) \;\middle|\; \begin{array}{r} \partial_1 c_f(t,|z^t|)\delta t - \delta c + \partial_2 c_f(t,|z^t|)\delta\zeta = 0, \\ \partial_1 c_{\mathcal{E}}(t,|z^t|)\delta t + \partial_2 c_{\mathcal{E}}(t,|z^t|)\delta\zeta = 0, \\ \partial_1 c_{\mathcal{Z}}(t,|z^t|)\delta t + \partial_2 c_{\mathcal{Z}}(t,|z^t|)\delta\zeta = \delta z^t \end{array} \right\}.$$

Then, the form of $\mathcal{T}^{\text{lin}}_{\text{f-abs}}(t,z^t)$ gives the presentation above. □

Here, the relations

$$\mathcal{T}_{\text{c-abs}}(t,c,z^t) \subseteq \mathcal{T}^{\text{lin}}_{\text{c-abs}}(t,c,z^t) \quad \text{and} \quad \mathcal{T}_{\text{c-abs}}(t,c,z^t)^* \supseteq \mathcal{T}^{\text{lin}}_{\text{c-abs}}(t,c,z^t)^*$$

follows from Lemma 5.8 applied to the formulation ($\overline{\text{C-NLP}}$).

Again, the converse relations do not hold in general. Thus, AKQ and GKQ are defined appropiately.

Lemma 5.82 (AKQ for (C-NLP)). *Consider a feasible point (t, c) of (C-NLP). Then, AKQ for (C-NLP) at (t, c) reads $\mathcal{T}_{c\text{-}abs}(t, c, z^t(t)) = \mathcal{T}^{lin}_{c\text{-}abs}(t, c, z^t(t))$.*

Proof. This follows directly by applying Definition 5.9 to ($\overline{\text{C-NLP}}$). □

Lemma 5.83 (GKQ for (C-NLP)). *Consider a feasible point (t, c) of (C-NLP). Then, GKQ holds for (C-NLP) at (t, c) if $\mathcal{T}_{c\text{-}abs}(t, z^t(t))^* = \mathcal{T}^{lin}_{c\text{-}abs}(t, z^t(t))^*$.*

Proof. This follows directly by applying Definition 5.10 to ($\overline{\text{C-NLP}}$). □

Relations Between Kink Qualifications Next, relations between AKQ and GKQ for both formulations are proven.

Theorem 5.84. *AKQ for (F-NLP) holds at $t \in \mathcal{F}_f$ if and only if AKQ for (C-NLP) holds at $(t, c) \in \mathcal{F}_c$.*

Proof. As $\mathcal{T}_{\text{f-abs}}(t, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t)$ and $\mathcal{T}_{\text{c-abs}}(t, w, z^t) \subseteq \mathcal{T}^{\text{lin}}_{\text{c-abs}}(t, w, z^t)$ always hold, just

$$\mathcal{T}_{\text{f-abs}}(t, z^t) \supseteq \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t) \iff \mathcal{T}_{\text{c-abs}}(t, w, z^t, z^w) \supseteq \mathcal{T}^{\text{lin}}_{\text{c-abs}}(t, w, z^t, z^w)$$

need to be proven. First, the implication "⇒" is proven. Let $\delta = (\delta t, \delta c, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\text{c-abs}}(t, c, z^t)$. Then, $\tilde{\delta} = (\delta t, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t)$ and by assumption $\tilde{\delta} \in \mathcal{T}_{\text{f-abs}}(t, z^t)$. Hence, there exist sequences $(t_k, z^t_k) \in \mathcal{F}_{\text{i-abs}}$ and $\tau_k \searrow 0$ with $(t_k, z^t_k) \to (t, z^t)$ and $\tau_k^{-1}(t_k - t, z^t_k - z^t) \to (\delta t, \delta z^t)$. Set $c_k := c_f(t_k, |z^t_k|)$ and thus

$$c_k - c = \partial_1 c_f(t, |z^t|)(t_k - t) + \partial_2 c_f(t, |z^t|)(|z^t_k| - |z^t|) + o(\|(t_k - t, |z^t_k| - |z^t|)\|)]$$

by Taylor. As in Theorem 5.25 it follows that

$$\tau_k^{-1}(c_k - c) \to \partial_1 c_f(t, |z^t|)\delta t + \partial_2 c_f(t, |z^t|)\delta\zeta = \delta c$$

and thus $d \in \mathcal{T}_{\text{c-abs}}(t, c, z^t)$.
To prove the implication "⇐", consider $\delta = (\delta t, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t)$ and set

$$\delta c = \partial_1 c_f(t, |z^t|)\delta t + \partial_2 c_f(t, |z^t|)\delta\zeta.$$

Then, $\tilde{\delta} = (\delta t, \delta c, \delta z^t) \in \mathcal{T}^{\text{lin}}_{\text{c-abs}}(t, c, z^t)$ and by the assumption $\tilde{\delta} \in \mathcal{T}_{\text{c-abs}}(t, w, z^t)$ holds. This directly implies $\delta = (\delta t, \delta z^t) \in \mathcal{T}_{\text{f-abs}}(t, z^t)$. □

Theorem 5.85. *GKQ for (F-NLP) holds at $t \in \mathcal{F}_f$ if GKQ for (C-NLP) holds at $(t, c) \in \mathcal{F}_c$.*

Proof. The inclusion $\mathcal{T}_{\text{f-abs}}(t, z^t)^* \supseteq \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t)^*$ always holds. Thus, it is left to prove

$$\mathcal{T}_{\text{f-abs}}(t, z^t)^* \subseteq \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t)^*.$$

Let $\omega = (\omega t, \omega z^t) \in \mathcal{T}_{\text{f-abs}}(t, z^t)^*$, i.e. $\omega^T\delta \geq 0$ for all $\delta = (\delta t, \delta z^t) \in \mathcal{T}_{\text{f-abs}}(t, z^t)$. Then, one obtains $\tilde{\omega}^T\tilde{\delta} = \omega^T\delta \geq 0$ for all $\tilde{\delta} \in \mathcal{T}_{\text{c-abs}}(t, c, z^t)$ with the choice $\tilde{\omega} = (\omega t, 0, \omega z^t)$. By assumption, $\tilde{\omega} \in \mathcal{T}_{\text{c-abs}}(t, c, z^t) = \mathcal{T}^{\text{lin}}_{\text{c-abs}}(t, c, z^t)$, which gives directly $\omega^T\delta \geq 0$ for all $\delta \in \mathcal{T}^{\text{lin}}_{\text{f-abs}}(t, z^t)$. □

As in section 5.3 one can define branch problems for (F-NLP). Then, sign information can be exploited to prove equivalence between ACQ resp. GCQ for all branch problems of (F-NLP) and (C-NLP). But this is omitted here as it is just technical without additional insights.

Further, these results can be used to transfer the stationarity concepts of section 5.4 to (F-NLP). But as in section 3.2.2, this is omitted here as it is just technical.

5.6 Unconstrained Abs-Normal NLP

This section revisits the unconstrained abs-normal NLP (unNLP)

$$\begin{aligned} \min_{x,z} \quad & f(x,|z|) \\ \text{s.t.} \quad & c_{\mathcal{Z}}(x,|z|) - z = 0, \end{aligned}$$

with feasible set $\mathcal{F}_{\text{un}} = \{(x,z)\colon c_{\mathcal{Z}}(x,|z|) - z = 0\}$. In this setting AKQ and GKQ hold without prerequisites and thus every local minimizer of (unNLP) is M- and abs-normal-linearized B-stationary. These results are proven in the following.

Kink Qualifications First, AKQ and GKQ are formulated for (unNLP).

Lemma 5.86 (Tangential and abs-normal-linearized cone)**.** *Given (unNLP), consider a feasible point $(x,z(x))$. The tangential cone to $\mathcal{F}_{un}$ at (x,z) reads*

$$\mathcal{T}_{un}(x,z) = \left\{ (\delta x, \delta z) \;\middle|\; \begin{array}{l} \exists \tau_k \searrow 0,\ \mathcal{F}_{un} \ni (x_k,z_k) \to (x,z)\colon \\ \tau_k^{-1}(x_k - x, z_k - z) \to (\delta x, \delta z) \end{array} \right\}.$$

With $\delta\zeta_i := |\delta z_i|$ if $i \in \alpha(x)$ and $\delta\zeta_i := \sigma_i(x)\delta z_i$ if $i \notin \alpha(x)$, the abs-normal-linearized cone reads

$$\mathcal{T}_{un}^{lin}(x,z) = \left\{ (\delta x, \delta z) \;\middle|\; \partial_1 c_{\mathcal{Z}}(x,|z|)\delta t + \partial_2 c_{\mathcal{Z}}(x,|z|)\delta\zeta = \delta z \right\}.$$

Proof. This follows directly from Definition 5.78. □

Lemma 5.87 (Kink Qualifications)**.** *Given (unNLP), consider $(x,z(x)) \in \mathcal{F}_{un}$. Then, AKQ reads*

$$\mathcal{T}_{un}(x,z(x)) = \mathcal{T}_{un}^{lin}(x,z(x))$$

and GKQ reads

$$\mathcal{T}_{un}(x,z(x))^* = \mathcal{T}_{un}^{lin}(x,z(x))^*.$$

Proof. This follows directly from Definition 5.79 and Definition 5.80. □

The next theorem shows that AKQ is always satisified. Key is the absence of constraints besides of the switching feasibility in (unNLP).

Theorem 5.88 (AKQ holds)**.** *Given (unNLP), consider $(x,z(x)) \in \mathcal{F}_{un}$. Then, AKQ holds at x.*

Proof. To show AKQ, i.e. $\mathcal{T}_{\text{un}}(x,z(x)) = \mathcal{T}_{\text{un}}^{\text{lin}}(x,z(x))$, it suffices to show

$$\mathcal{T}_{\text{un}}(x,z(x)) \supseteq \mathcal{T}_{\text{un}}^{\text{lin}}(x,z(x))$$

as the reverse inclusion always holds. Thus, let $\delta = (\delta x, \delta z) \in \mathcal{T}_{\text{un}}^{\text{lin}}$. Then,

$$\delta z = (I - \partial_2 c_{\mathcal{Z}}(x,|z|)\Sigma)^{-1}\partial_1 c_{\mathcal{Z}}(x,|z|)\delta x$$

for all $\tilde{\Sigma} = \operatorname{diag}(\tilde{\sigma})$ with $\tilde{\sigma}_i = \sigma_i(x)$ for $i \notin \alpha(x)$ and $\tilde{\sigma}_i \in \{-1,1\}$ such that $\tilde{\sigma}_i\delta z_i \geq 0$ for $i \in \alpha(x)$. Further, there exist a neighborhood B of (x,z) which can be decomposed via $B = \bigcup_{\hat{\sigma} \geq \sigma(x)} (\mathcal{F}_{\hat{\Sigma}} \cap B)$ with $\mathcal{F}_{\hat{\Sigma}} = \{c_{\mathcal{Z}}(x, \hat{\Sigma}z) - z = 0,\ \hat{\Sigma}z \geq 0\}$. Then, there exists a null sequence τ_k such that $(x_k, z_k) := (x + \tau_k \delta x, z(x_k)) \in \mathcal{F}_{\hat{\Sigma}}$ for $\hat{\Sigma} = \operatorname{diag}(\hat{\sigma})$ with $\hat{\sigma}_i = \hat{\sigma}(x_i)$ for $i \notin \alpha(x)$ and $\hat{\sigma}_i \in \{-1, 1\}$ for $i \in \alpha(x)$. By construction and continuity of $c_{\mathcal{Z}}$, $(x_k, z_k) \to (x, z)$ holds. Further, $|z| = \hat{\Sigma}z$ as $\hat{\sigma} \succeq \sigma(x)$ and thus $z = c_{\mathcal{Z}}(x, \hat{\Sigma}z)$. Using this, Taylor gives

$$\begin{aligned} z_k - z &= \partial_1 c_{\mathcal{Z}}(x, \hat{\Sigma}z)(x_k - x) + \partial_2 c_{\mathcal{Z}}(x, \hat{\Sigma}z)\hat{\Sigma}(z_k - z) + o(\|(x_k - x, \hat{\Sigma}(z_k - z))\|) \\ &= \partial_1 c_{\mathcal{Z}}(x, \hat{\Sigma}z)\delta x + \partial_2 c_{\mathcal{Z}}(x, \hat{\Sigma}z)\hat{\Sigma}(z_k - z) + o(\|(x_k - x, \hat{\Sigma}(z_k - z))\|) \end{aligned}$$

which means

$$\frac{z_k - z}{\tau^k} \to (I - \partial_2 c_{\mathcal{Z}}(x, \hat{\Sigma}z)\hat{\Sigma})^{-1}\partial_1 c_{\mathcal{Z}}(x, \hat{\Sigma}z)\delta x = (I - \partial_2 c_{\mathcal{Z}}(x, |z|)\hat{\Sigma})^{-1}\partial_1 c_{\mathcal{Z}}(x, |z|)\delta x.$$

For $i \in \alpha(x)$ this implies

$$\hat{\sigma}_i \left[(I - \partial_2 c_{\mathcal{Z}}(x, |z|)\hat{\Sigma})^{-1}\partial_1 c_{\mathcal{Z}}(x, |z|)\delta x\right]_i \geq 0$$

as $\hat{\sigma}_i \tau_k^{-1}((z_k)_i - z_i) = \hat{\sigma}_i \tau_k^{-1}(z_k)_i \geq 0$. Altogether,

$$\frac{z_k - z}{\tau^k} \to (I - \partial_2 c_{\mathcal{Z}}(x, |z|)\hat{\Sigma})^{-1}\partial_1 c_{\mathcal{Z}}(x, |z|)\delta x = \delta z$$

and hence $(\delta t, \delta z^t) \in \mathcal{T}_{\text{un}}(x, z(x))$. $\square$

Thus, equality holds also for the dual cones.

Corollary 5.89 (GKQ holds)**.** *Given* (unNLP)*, consider* $(x, z(x)) \in \mathcal{F}_{un}$*. Then, GKQ holds at* x*.*

Optimality Conditions The previous results imply that every local minimizer of (unNLP) is M- and abs-normal-linearized B-stationary. Before this is stated formally and proven, the stationarity concepts are formulated for (unNLP).

Definition 5.90 (M-Stationarity)**.** Consider a feasible point $(t, z^t(t))$ of (unNLP). It is an *M-stationary* point if there exist a Lagrange multiplier vector $\lambda^*_{\mathcal{Z}}$ such that the following conditions are satisfied:

$$\begin{aligned} \partial_1 c_f(t^*) + (\lambda^*_{\mathcal{Z}})^T \partial_1 c_{\mathcal{Z}} &= 0, \\ [\partial_2 c_f + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}]_i &= (\lambda^*_{\mathcal{Z}})_i \sigma^*_i,\ i \notin \alpha(t^*), \\ (\mu_i^-)(\mu_i^+) = 0 \ \vee \ [\partial_2 c_f + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}]_i &> |(\lambda^*_{\mathcal{Z}})_i|,\ i \in \alpha(t^*). \end{aligned}$$

Here, the notation

$$\mu_i^+ := \left[\partial_2 c_f + (\lambda^*_{\mathcal{Z}})^T[\partial_2 c_{\mathcal{Z}} - I]\right]_i \quad \text{and} \quad \mu_i^- := \left[\partial_2 c_f + (\lambda^*_{\mathcal{Z}})^T[\partial_2 c_{\mathcal{Z}} + I]\right]_i$$

is used and the constraints and the partial derivatives are evaluated at $(t^*, |(z^t)^*|)$.

Definition 5.91 (Branch NLPs for (unNLP))**.** Consider a feasible point $(\hat{t}, \hat{z}^t)$ of (unNLP). Choose $\sigma^t \in \{-1, 1\}^{s_t}$ with $\sigma^t \succeq \sigma^t(\hat{t})$ and set $\Sigma^t = \operatorname{diag}(\sigma^t)$. The branch problem $\mathrm{NLP}(\Sigma^t_{\mathrm{un}})$ is defined as

$$\begin{aligned} \min_{t,z^t} \quad & c_f(t, \Sigma^t z^t) \\ \text{s.t.} \quad & c_{\mathcal{Z}}(t, \Sigma^t z^t) = z^t, \\ & \Sigma^t z^t \geq 0. \end{aligned} \tag{$\mathrm{NLP}(\Sigma^t_{\mathrm{un}})$}$$

Definition 5.92 (Abs-normal-linearized B-Stationarity for (unNLP))**.** Consider a feasible point $(t^*, (z^t)^*)$ of (unNLP). It is an *abs-normal-linearized B-stationary* point if it is a stationary point of all branch problems ($\mathrm{NLP}(\Sigma^t_{\mathrm{un}})$).

Theorem 5.93 (Minimizers and M-Stationarity for (unNLP))**.** *Assume that $(t^*, (z^t)^*)$ is a local minimizer of* (unNLP)*. Then, it is M-stationary for* (unNLP)*.*

Proof. Using a reformulation as in the previous subsection, (unNLP) can be written as

$$\begin{aligned} \min_{x,z} \quad & \bar{f}(x) \\ \text{s.t.} \quad & \bar{c}_{\mathcal{E}}(x, |z|) = 0, \\ & \bar{c}_{\mathcal{Z}}(x, |z|) - z = 0, \end{aligned}$$

with $x = (t, c)$, $z = z^t$, $\bar{f}(x) = c$, $\bar{c}_{\mathcal{E}}(x, |z|) = c_f(t, |z^t|) - c$ and $\bar{c}_{\mathcal{Z}}(x, |z|) = c_{\mathcal{Z}}(t, |z^t|)$. AKQ for this problem holds by Theorem 5.84 and thus Theorem 5.73 gives the conditions

$$\begin{aligned} \bar{f}'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 \bar{c}_{\mathcal{E}} + (\lambda^*_{\mathcal{Z}})^T \partial_1 \bar{c}_{\mathcal{Z}} &= 0, \\ [(\lambda^*_{\mathcal{E}})^T \partial_2 \bar{c}_{\mathcal{E}} + (\lambda^*_{\mathcal{Z}})^T \partial_2 \bar{c}_{\mathcal{Z}}]_i &= (\lambda^*_{\mathcal{Z}})_i \sigma^*_i, \ i \notin \alpha(x^*), \\ (\mu^-_i)(\mu^+_i) = 0 \ \vee \ [(\lambda^*_{\mathcal{E}})^T \partial_2 \bar{c}_{\mathcal{E}} + (\lambda^*_{\mathcal{Z}})^T \partial_2 \bar{c}_{\mathcal{Z}}]_i &> |((\lambda^*_{\mathcal{Z}}))_i|, \ i \in \alpha(x^*), \end{aligned}$$

with $\mu^+_i := \Big[(\lambda^*_{\mathcal{E}})^T \partial_2 \bar{c}_{\mathcal{E}} + (\lambda^*_{\mathcal{Z}})^T [\partial_2 \bar{c}_{\mathcal{Z}} - I]\Big]_i$ and $\mu^-_i := \Big[(\lambda^*_{\mathcal{E}})^T \partial_2 \bar{c}_{\mathcal{E}} + (\lambda^*_{\mathcal{Z}})^T [\partial_2 \bar{c}_{\mathcal{Z}} + I]\Big]_i$. Then, using the definitions this yields

$$\begin{aligned} (\lambda^*_{\mathcal{E}})^T \partial_1 c_f + (\lambda^*_{\mathcal{Z}})^T \partial_1 c_{\mathcal{Z}} &= 0, \\ 1 - (\lambda^*_{\mathcal{E}})^T &= 0, \\ [(\lambda^*_{\mathcal{E}})^T \partial_2 c_f + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}]_i &= (\lambda^*_{\mathcal{Z}})_i \sigma^*_i, \ i \notin \alpha(x^*), \\ (\mu^-_i)(\mu^+_i) = 0 \ \vee \ [(\lambda^*_{\mathcal{E}})^T \partial_2 c_f + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}]_i &> |(\lambda^*_{\mathcal{Z}})_i|, \ i \in \alpha(x^*), \end{aligned}$$

with $\mu^+_i := \Big[(\lambda^*_{\mathcal{E}})^T \partial_2 c_f + (\lambda^*_{\mathcal{Z}})^T [\partial_2 c_{\mathcal{Z}} - I]\Big]_i$ and $\mu^-_i := \Big[(\lambda^*_{\mathcal{E}})^T \partial_2 c_f + (\lambda^*_{\mathcal{Z}})^T [\partial_2 c_{\mathcal{Z}} + I]\Big]_i$. Inserting $\lambda^*_{\mathcal{E}} = 1$ gives the result. □

Theorem 5.94 (Minimizers and abs-normal-linearized B-Stationarity for (unNLP))**.** *Assume that (t, z^t) is a local minimizer of* (unNLP)*. Then, it is abs-normal-linearized B-stationary for* (unNLP)*.*

Proof. Using the same reformulation as in the previous proof, AKQ implies GKQ for this formulation. Then, Theorem 5.77 gives that (t, z^t) is a stationary point for every branch problem of the form

$$\begin{aligned} \min_{x,z} \quad & \bar{f}(x) \\ \text{s.t.} \quad & \bar{c}_{\mathcal{E}}(x, \Sigma z) = 0, \\ & \bar{c}_{\mathcal{Z}}(x, \Sigma z) = z, \\ & \Sigma z \geq 0. \end{aligned}$$

Inserting the definitions leads to the formulation

$$\begin{aligned} \min_{t,z^t} \quad & c \\ \text{s.t.} \quad & c_f(t, \Sigma^t z^t) - c = 0, \\ & c_{\mathcal{Z}}(t, \Sigma^t z^t) = z^t, \\ & \Sigma^t z^t \geq 0. \end{aligned}$$

Finally, this is equivalent to the definition of abs-normal-linearized B-stationary for (unNLP). □

Remark 5.95. In [11], Griewank and Walther have presented a stationarity concept that holds without any kink qualification. Indeed, this concept is precisely abs-normal-linearized Bouligand stationarity: it requires the conditions of Definition 5.75 specialized to (unNLP).

Chapter 6

Conclusion and Outlook

This thesis provides optimality conditions for abs-normal NLPs.

To begin with, a straightforward extension of first and second order optimality conditions for unconstrained abs-normal NLPs and for classical smooth NLPs to the general case of abs-normal NLPs is given. Here, the fundamental regularity assumption to prove these optimality conditions is LIKQ which is a generalization of LIKQ for unconstrained abs-normal NLPs and LICQ for smooth NLPs. At first, abs-normal NLPs without inequality constraints are considered. Thus, any nonsmoothness as well as the distinction of active and inactive constraints are captured by the switching variables of the equality constraints. Afterwards, these results are extended to additional inequality constraints by rewriting them via absolute value slacks as equality constraints. Here, the key is that LIKQ is preserved under the reformulation and so the previous results can be used. Further, abs-normal NLPs with a nonsmooth objective function are reformulated replacing it by an additional variable and adding an additional equality constraint. Again, LIKQ is preserved under this reformulation and so optimality conditions can also be transfered.

Next, it is shown that abs-normal NLPs are essentially the same problem class as MPECs. In particular, any abs-normal NLP can be reformulated as an MPEC and vice versa. Then, equivalence between LIKQ and MPEC-LICQ for both formulations of inequality constraints is proven and corresponding first and second order conditions are compared. It turns out that they are equivalent except for some technical assumptions. Moreover, unconstrained abs-normal NLPs are considered and it is shown that MFKQ is weaker than MPEC-MFCQ.

After that weaker constraint qualifications for MPECs are used to obtain corresponding concepts for abs-normal NLPs. In particular, the concepts of Mangasarian Fromovitz, Abadie and Guignard type are considered. It turns out that the reformulation with absolute value slacks which was useful to simplify derivations under LIKQ, does not preserve IDKQ as no sign condition of slack variables has to hold. Then, it is shown that constraint qualifications of Abadie type are preserved, whereas for Guignard type one can only prove some implications. Here, one subtle drawback is the non-uniqueness of slack variables as they occur inside an absolute value and no sign condition is required. Thus, branch formulations of general abs-normal NLPs and counterpart MPECs are introduced to exploit additional sign conditions and then, constraint qualifications of Abadie and Guignard type are preserved. Nevertheless it is still an open question if GKQ and MPEC-GCQ are equivalent for both formulations of inequality constraints. Next, M-stationarity and linearized-B-stationarity are defined for abs-normal NLPs and first order conditions are proven using the corresponding concepts for MPECs. Moreover, it is shown that and how the extension of these results to nonsmooth objective functions is possible. Finally, it turns out that AKQ and thus GKQ are always satisfied in the case of unconstrained abs-normal NLPs. Hence, every local minimizer is M- and abs-normal-linearized B-stationary in this special case.

The aim of this thesis is to extend optimality conditions to general abs-normal NLPs and to develop a deeper understanding of the relation to MPECs and their associated theoretical concepts. Thus, a natural next step would be to extend the algorithm SALMIN for unconstrained abs-normal NLPs to general ones and to implement a correlating solver. A related but also slightly different next step would be to combine advantanges of abs-normal NLPs and MPECs to obtain a new solution algorithm. Here, the abs-normal side provides the possibility of computations via algorithmic differentiation and the MPEC side a variety of existing sophisticated solution algorithms. Nevertheless, it would be also a possible next step to extend SALMIN first and combine advantages after that.

Bibliography

[1] H. Amann and J. Escher. *Analysis II*. Birkhäuser Verlag, second corrected edition, 2006.

[2] D.P. Bertsekas. *Nonlinear Programming*. Athena Scientific, second edition, 2003.

[3] F. H. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley & Sons, 1983.

[4] S. Fiege. *Minimization of Lipschitzian piecewise smooth objective functions*. Dissertation, Universität Paderborn, 2017.

[5] M.L. Flegel. *Constraint Qualifications and Stationarity Concepts for Mathematical Programs with Equilibrium Constraints*. Dissertation, Universität Würzburg, 2005.

[6] C. Geiger and C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer Verlag, 2002.

[7] A. Griewank. On stable piecewise linearization and generalized algorithmic differentiation. *Optimization Methods and Software*, 28(6):1139–1178, 2013.

[8] A. Griewank and A. Walther. First and second order optimality conditions for piecewise smooth objective functions. *Optimization Methods and Software*, 31(5):904–930, 2016.

[9] A. Griewank and A. Walther. Finite convergence of an active signature method to local minima of piecewise linear functions. *Optimization Methods and Software*, 2018. published online https://doi.org/10.1080/10556788.2018.1546856.

[10] A. Griewank and A. Walther. Characterizing and testing subdifferential regularity for piecewise smooth objective functions. *SIAM Journal on Optimization*, 29(2):1473–1501, November 2019.

[11] A. Griewank and A. Walther. Relaxing kink qualifications and proving convergence rates in piecewise smooth optimization. *SIAM Journal on Optimization*, 29(1):262–289, January 2019.

[12] Andreas Griewank. Personal conversation, Juni 2018.

[13] L. C. Hegerhorst-Schultchen, C. Kirches, and M. C. Steinbach. Comparing abs-normal NLPs to MPECs. *Proceedings in Applied Mathematics and Mechanics*, 2019. published online https://doi.org/10.1002/pamm.201900263.

[14] L. C. Hegerhorst-Schultchen, C. Kirches, and M. C. Steinbach. On the relation between MPECs and optimization problems in abs-normal form. *Optimization Methods and Software*, 2019. published online https://doi.org/10.1080/10556788.2019.1588268.

[15] L. C. Hegerhorst-Schultchen, C. Kirches, and M. C. Steinbach. Relations between abs-normal NLPs and MPECs under strong constraint qualifications. 2019. preprint available online http://www.optimization-online.org/DB_HTML/2019/07/7302.html.

[16] L. C. Hegerhorst-Schultchen, C. Kirches, and M. C. Steinbach. Relations between abs-normal NLPs and MPECs under weak constraint qualifications. 2019. preprint available online http://www.optimization-online.org/DB_HTML/2019/08/7337.html.

[17] L. C. Hegerhorst-Schultchen and M. C. Steinbach. On first and second order optimality conditions for abs-normal NLP. *Optimization*, 2019. published online https://doi.org/10.1080/02331934.2019.1626386.

[18] T. Koch, B. Hiller, M. Pfetsch, and L. Schewe. *Evaluating Gas Network Capacities*. MOS-SIAM Series on Optimization. 2015.

[19] Z. Luo, J. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, 1996.

[20] B. Mordukhovich. *Variational Analysis and Generalized Differentiation I*. Springer-Verlag Berlin Heidelberg, 2006.

[21] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, New York, NY, USA, second edition, 2006.

[22] R. T. Rockafellar. *Convex Analysis*. Princton University Press, 1970.

[23] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, February 2000.

[24] A. Schwartz. Mathematical programs with complementarity constraints and related problems. *Course Notes, Graduate School CE, Technische Universität Darmstadt*, 2018. (available from https://github.com/alexandrabschwartz/Winterschool2018/blob/master/LectureNotes.pdf).

[25] 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.

Curriculum Vitae

Persönliche Angaben

Name	Lisa Christine Hegerhorst-Schultchen, geb. Hegerhorst
Geburtsdatum	22. Dezember 1990
Geburtsort	Stadthagen

Akademischer und Beruflicher Werdegang

06/2010	Abitur am Ratsgymnasium Stadthagen
10/2010–09/2013	Studium Bachelor of Science im Studiengang Mathematik an der Leibniz Universität Hannover
09/2013	Bachelorarbeit: *Funktionalanalytische Behandlung der nichtlinearen Schrödingergleichung*, Institut für Angewandte Mathematik, Prof. Dr. Escher
10/2013–09/2015	Studium Master of Science im Studiengang Mathematik an der Leibniz Universität Hannover
09/2015	Masterarbeit: *Theory and Numerics of the Gauss-Newton Method*, Institut für Angewandte Mathematik, Prof. Dr. Steinbach
seit 10/2015	Wissenschaftliche Mitarbeiterin und Promotionsstudentin am Institut für Angewandte Mathematik, AG Algorithmische Optimierung (Prof. Dr. Steinbach), Leibniz Universität Hannover

Publikationen

L. C. Hegerhorst-Schultchen, C. Kirches, and M. C. Steinbach. On the relation between MPECs and optimization problems in abs-normal form. *Optimization Methods and Software*, 2019. published online https://doi.org/10.1080/10556788.2019.1588268

L. C. Hegerhorst-Schultchen and M. C. Steinbach. On first and second order optimality conditions for abs-normal NLP. *Optimization*, 2019. published online https://doi.org/10.1080/02331934.2019.1626386

L. C. Hegerhorst-Schultchen, C. Kirches, and M. C. Steinbach. Comparing abs-normal NLPs to MPECs. *Proceedings in Applied Mathematics and Mechanics*, 2019. published online https://doi.org/10.1002/pamm.201900263