ON FIRST AND SECOND ORDER OPTIMALITY CONDITIONS FOR ABS-NORMAL NLP

L. C. HEGERHORST-SCHULTCHEN AND M. C. STEINBACH

Abstract. Structured nonsmoothness is widely present in practical optimization. A particularly attractive class of nonsmooth problems, both from a theoretical and from an algorithmic perspective, are optimization problems in so-called abs-normal form as developed by Griewank and Walther. Here we generalize their theory for the unconstrained case to nonsmooth NLPs with equality and inequality constraints in abs-normal form, obtaining similar necessary and sufficient conditions of first and second order that are directly based on classical Karush-Kuhn-Tucker (KKT) theory for smooth NLPs. Several small examples illustrate the theoretical results. We also give some brief remarks on the intimate relationship of abs-normal NLPs with MPECs.

1. Introduction

Nonsmoothness is a common phenomenon in practical optimization problems from various areas. In finite dimensions, nonsmooth problems are often a consequence of low-dimensional modeling and can be recast in smooth form by introducing additional variables and constraints. This is well-known, for instance, in case of the ℓ_1 and ℓ_∞ norms, and for expectations and certain risk measures over finitely many scenarios in two-stage and multi-stage stochastic optimization. However, genuine nonsmoothness remains in many cases, and suitable structural assumptions can lead to useful theoretical results and algorithmically tractable problem classes with lower complexity than common mixed integer reformulations. Griewank and Walther have developed a particularly attractive class of unconstrained finite-dimensional nonlinear optimization problems where the nonsmoothness is caused by possibly nested occurrences of the absolute value function, which also covers minimum and maximum of finitely many arguments [1, 2]. These problems can be cast in so-called abs-normal form with theoretical properties that directly generalize standard KKT theory, and they are tractable by sophisticated algorithms with guaranteed convergence based on piecewise linearizations and using algorithmic differentiation techniques [3, 4]. There is a substantial body of literature on other areas of nonsmooth optimization in finite and infinite dimensions, such as MPECs, variational inequalities, and many more. We do not address any of them here since our focus is entirely on the abs-normal problems mentioned above.

This paper generalizes the unconstrained theory of Griewank and Walther to abs-normal NLPs with nonsmooth equality and inequality constraints. In particular, we extend the linear independence kink qualification (LIKQ) and derive first and second order necessary and sufficient optimality conditions for the abs-normal NLP. As in the unconstrained case, this is achieved by applying standard KKT theory (see, e.g., [6]) to suitable trunk and branch problems that satisfy the LICQ whenever the abs-normal NLP satisfies the LIKQ.

The remaining material is structured as follows. In Sect. 2 we introduce the abs-normal NLP along with the required basic concepts of Griewank and Walther, such as level-1 smoothness and abs-normal form. In Sect. 3 we introduce further basic concepts like signatures and active switching variables to state the linear independence kink qualification (LIKQ) as a fundamental regularity condition. Necessary and sufficient optimality conditions of first and second order for the abs-normal NLP are first derived with so-called localized switching (where all switching variables are active) in Sect. 4, then for the more intricate non-localized case in Sect. 5. Finally we give some

Date: March 6, 2019.

2010 *Mathematics Subject Classification.* 49J52, 90C30, 90C46.

Key words and phrases. Nonsmooth NLP, abs-normal form, linear independence kink qualification, first and second order necessary and sufficient conditions.

conclusions in Sect. 6. Throughout the paper our notation and technical setting differs slightly from Griewank and Walther, mostly due to the more general problem formulation.
Notation. We denote by ∂_i the partial derivative with respect to the i-th argument and by ∂_y the derivative with respect to the variable y. For $\sigma, \tilde{\sigma} \in \{-1, 0, 1\}^s$ we use the partial order $\sigma \geq \tilde{\sigma} \iff \sigma_i \tilde{\sigma}_i \geq \tilde{\sigma}_i^2$ for $i = 1, \ldots, s$. Thus, σ_i is arbitrary if $\tilde{\sigma}_i = 0$ and $\sigma_i = \tilde{\sigma}_i$ otherwise.

2. Problem Formulation

Here we introduce basic classes of nonsmooth functions and the nonsmooth NLP of interest. We also show that it suffices to consider a smooth objective with equality constraints in abs-normal form, which simplifies the subsequent presentation considerably.

2.1. **Abs-Normal Form.**

We consider nonsmooth functions where any nonsmoothness is caused by (possibly nested) occurrences of the absolute value function. This leads to the class of level-1 nonsmooth functions.

Definition 1 (Level-1 Nonsmooth Function)**.** Let D be an open subset of $\mathbb{R}^n$. We say that a function $\varphi \colon D \to \mathbb{R}$ is *level-1 nonsmooth* if it can be expressed as a composition of smooth functions and the absolute value function. We say that a function $g \colon D \to \mathbb{R}^m$ is level-1 nonsmooth if all component functions $g_i \colon D \to \mathbb{R}$ are level-1 nonsmooth.

Every level-1 nonsmooth function $g \colon D^x \to \mathbb{R}^m$ can be reformulated as the system

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

where $c_{\mathcal{E}} \colon D^x \times D^{|z|} \to \mathbb{R}^m$ and $c_{\mathcal{Z}} \colon D^x \times D^{|z|} \to \mathbb{R}^s$ are smooth and $D^{|z|} \subset \mathbb{R}^s$ is open with $0 \in D^{|z|}$ (otherwise $|\cdot|$ would be redundant). From now on we write $D^{x,|z|}$ for $D^x \times D^{|z|}$.

To derive this formulation, we replace from left to right all arguments of absolute value evaluations by variables z_i, $i = 1, \ldots, s$. Moreover, we replace from inside to outside if nested absolute value evaluations occur and use already defined variables z_i if arguments repeat. For details see [1, 2]. Note that 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 $i < j$.

The above construction leads to the following definition.

Definition 2 (Abs-Normal Form)**.** We say that a function $g \colon D^x \to \mathbb{R}^m$ is given in *abs-normal form* if $c_{\mathcal{E}} \colon D^{x,|z|} \to \mathbb{R}^m$ and $c_{\mathcal{Z}} \colon D^{x,|z|} \to \mathbb{R}^s$ exist such that $D^{|z|} \subset \mathbb{R}^s$ is open with $0 \in D^{|z|}$ and

$$g(x) = c_{\mathcal{E}}(x, |z|), \tag{1a}$$

$$z = c_{\mathcal{Z}}(x, |z|) \quad \text{with } \partial_2 c_{\mathcal{Z}}(x, |z|) \text{ strictly lower triangular.} \tag{1b}$$

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

Note that, although z is implicitly defined by (1b), the switching variables z_j can be computed in turn from x and z_i with $i < j$ since the partial derivative $\partial_2 c_{\mathcal{Z}}(x, |z|)$ is strictly lower triangular. In the following we sometimes write $z(x)$ to denote the dependence on x explicitly.

Definition 3 (Class $C^d_{\text{abs}}(D)$)**.** Let D^x be an open subset of $\mathbb{R}^n$ and $d \in \mathbb{N} \cup \{\infty\}$. We denote by $C^d_{\text{abs}}(D^x, \mathbb{R}^m)$ the set of functions $g \colon D^x \to \mathbb{R}^m$ in abs-normal form (1) with $c_{\mathcal{E}} \in C^d(D^{x,|z|}, \mathbb{R}^m)$ and $c_{\mathcal{Z}} \in C^d(D^{x,|z|}, \mathbb{R}^s)$.

Example 4. The maximum and minimum functions are level-1 nonsmooth. They can easily be reformulated in abs-normal form and are elements of $C^\infty_{\text{abs}}(\mathbb{R}^2)$:

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

Here and in the following we deviate slightly from Griewank and Walther [2, 3, 4] by considering *open* domains D^x and $D^{|z|}$. We prefer this setting since we will need continuous derivatives and the implicit function theorem. Besides, we extend the definitions of the case $m = 1$ from [2, 3, 4] to functions $g \colon D^x \to \mathbb{R}^m$ in the spirit of Griewank's m-dimensional abs-normal form [1].

2.2. **Abs-Normal NLP.** The subject of this paper is the level-1 nonsmooth equality-constrained NLP

$$\min_{x} \quad f(x) \quad \text{s.t.} \quad g(x) = 0, \tag{2}$$

where $D^x \subset \mathbb{R}^n$ is open, $f \in C^d(D^x, \mathbb{R})$, and $g \in C^d_{\text{abs}}(D^x, \mathbb{R}^m)$. Without loss of generality we restrict ourselves to the case with a smooth objective function and without (level-1 nonsmooth) inequality constraints. This is possible because the general level-1 nonsmooth NLP

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

with $f \in C^d_{\text{abs}}(D^x, \mathbb{R})$, $g \in C^d_{\text{abs}}(D^x, \mathbb{R}^{m_1})$, and $h \in C^d_{\text{abs}}(D^x, \mathbb{R}^{m_2})$, can be reformulated as

$$\begin{aligned} \min_{x,v,w} \quad v \quad \text{s.t.} \quad & f(x) - v = 0, \\ & g(x) = 0, \\ & h(x) - |w| = 0. \end{aligned}$$

Here we have $v \in \mathbb{R}$, 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 $h(x) - |w| \in C^d_{\text{abs}}(D^x \times \mathbb{R}^{m_2}, \mathbb{R}^{m_2})$. The conversion of inequalities $h(x) \geq 0$ to equalities $h(x) - |w| = 0$ has been suggested by Griewank in a personal discussion.

To obtain the desired abs-normal NLP associated with NLP (2), the nonsmooth constraint $g(x) = 0$ is stated in abs-normal form:

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

Definition 5 (Abs-Normal NLP)**.** Let D^x be an open subset of $\mathbb{R}^n$. We call a nonsmooth optimization problem an *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 $c_{\mathcal{Z}} \in C^d(D^{x,|z|}, \mathbb{R}^s)$ with $d \geq 1$ exist such that the NLP can equivalently be stated as

$$\begin{aligned} \min_{x,z} \quad f(x) \quad \text{s.t.} \quad & c_{\mathcal{E}}(x, |z|) = 0, \\ & c_{\mathcal{Z}}(x, |z|) - z = 0, \end{aligned} \tag{3}$$

where $0 \in D^{|z|}$ and $\partial_2 c_{\mathcal{Z}}(x, |z|)$ is strictly lower triangular. The feasible set of (3) is denoted

$$\Omega := \{(x, z) \colon c_{\mathcal{E}}(x, |z|) = 0, c_{\mathcal{Z}}(x, |z|) = z\}.$$

Clearly x^* is a local minimizer of the nonsmooth NLP (2) if and only if $(x^*, z^*) = (x^*, z(x^*))$ is a local minimizer of the abs-normal NLP (3). In the following we will therefore only consider the abs-normal NLP (3). It provides a standard form for any level-1 nonsmooth NLP.

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

Example 6. Consider the level-1 nonsmooth NLP

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

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:

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

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

3. LIKQ

This section introduces the linear independence kink qualification (LIKQ) as the fundamental nonsmooth regularity condition. First we consider the solvability of the nonlinear switching system $z = c_{\mathcal{Z}}(x, |z|)$. To this end, we use the reformulation $|z_i| = \text{sign}(z_i) z_i$, which leads to the definition of a signature.

Definition 7 (Signature of z)**.** For $x \in D^x$ we define the *signature* $\sigma(x)$ and the associated *signature matrix* $\Sigma(x)$ as

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

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

Observe that the absolute value can be expressed as $|z| = \hat{\Sigma} z$ where $\hat{\sigma}_i = \operatorname{sign}(z_i)$ 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 $\hat{\sigma} \geq \sigma(x)$. 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 $z_i = 0$. A straightforward application of the implicit function theorem to the switching system for z then leads to the next lemma.

Lemma 8. *Assume that $c_{\mathcal{Z}} \in C^d(D^{x,|z|}, \mathbb{R}^s)$ for $d \geq 1$ and fix $\tilde{x} \in D^x$. Let $\tilde{z} = z(\tilde{x})$, $\tilde{\sigma} = \sigma(\tilde{x})$, and $\Sigma = \operatorname{diag}(\sigma)$ for arbitrary $\sigma \geq \tilde{\sigma}$. Then, the switching system $z = c_{\mathcal{Z}}(x, \Sigma z)$ has a locally unique solution $z^\sigma(x)$ which is d times continuously differentiable with Jacobian*

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

Clearly we have $z^\sigma(\tilde{x}) = z(\tilde{x})$ and $\Sigma z^\sigma(\tilde{x}) = |z(\tilde{x})|$. This implies

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

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

Definition 9 (Active Switching Variables)**.** A switching variable z_i is called *active* if $z_i(x) = 0$. 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 we have $|\alpha(x)|$ active switching variables and $|\sigma(x)| := s - |\alpha(x)|$ inactive ones.

We are now ready to extend the definitions of active Jacobian and LIKQ of Griewank and Walther [2] to the abs-normal NLP. This requires the constraints Jacobian $J_{\mathcal{E}}(x)$.

Definition 10 (Active Jacobian)**.** Let $c_{\mathcal{E}} \in C^1(D^{x,|z|}, \mathbb{R}^m)$ and $c_{\mathcal{Z}} \in C^1(D^{x,|z|}, \mathbb{R}^s)$. For $x \in D^x$ let $\alpha = \alpha(x)$, $\sigma = \sigma(x)$, and $\Sigma = \operatorname{diag}(\sigma)$. The *active Jacobian* of the abs-Normal NLP (3) is

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

It consists of the *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*

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

Similar to the smooth case, optimality conditions for the abs-normal NLP (3) 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 11 (Linear Independence Kink Qualification, LIKQ)**.** Given $c_{\mathcal{E}} \in C^1(D^{x,|z|}, \mathbb{R}^m)$ and $c_{\mathcal{Z}} \in C^1(D^{x,|z|}, \mathbb{R}^s)$, consider $x \in D^x$. We say that the *linear independence kink qualification (LIKQ)* holds at x if

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

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

In the following we denote by $U(x) \in \mathbb{R}^{n \times [n - (m+|\alpha|)]}$ a matrix whose columns are a basis of $\ker(J(x))$, i.e., $\ker(J(x)) = \operatorname{im}(U(x))$.

Let us return to Example 6 and check whether LIKQ holds at the solution.

Example 12 (LIKQ for Example 6)**.** Recall the abs-normal NLP

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

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

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

This matrix has full row rank. Hence, LIKQ is satisfied at $x^* = (0,0,0)$ with $z^* = 0$. The nullspace of $J(x^*)$ is spanned by $U(x^*) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$.

4. Localized Case

As in the unconstrained case [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.

Definition 13 (Localization)**.** Let $c_{\mathcal{E}} \in C^d(D^{x,|z|}, \mathbb{R}^m)$, $c_{\mathcal{Z}} \in C^d(D^{x,|z|}, \mathbb{R}^s)$, and consider $x \in D^x$. We say that the switching is *localized* at x if $z(x) = 0$ and *non-localized* otherwise.

To derive optimality conditions, we first consider the localized case. Given a point of interest x^*, we thus assume in this section that we have $z(x^*) = 0$ with active switching set $\alpha(x^*) = \{1, \dots, s\}$ of cardinality $|\alpha(x^*)| = s$.

4.1. **Trunk Problem.** Substituting $z = 0$ into the abs-normal NLP (3) yields the so-called localized trunk problem.

Definition 14 (Localized Trunk Problem)**.** The *localized trunk problem* of (3) is

$$\min_x \quad f(x) \quad \text{s.t.} \quad c_{\mathcal{E}}(x,0) = 0, \quad c_{\mathcal{Z}}(x,0) = 0. \tag{4}$$

Its feasible set is denoted $\Omega_t := \{(x,0) \colon c_{\mathcal{E}}(x,0) = 0, c_{\mathcal{Z}}(x,0) = 0\}$.

The following lemma is immediately clear by construction.

Lemma 15. *If $(x^*,0)$ is a local minimizer of the abs-normal NLP (3), then x^* is a local minimizer of the localized trunk problem (4).*

Moreover, the trunk problem is obviously smooth and we can apply standard theory 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 16. *Let $c_{\mathcal{E}} \in C^1(D^{x,|z|}, \mathbb{R}^m)$ and $c_{\mathcal{Z}} \in C^1(D^{x,|z|}, \mathbb{R}^s)$. Then, LIKQ at x^* with $z(x^*) = 0$ is LICQ at x^* for the localized trunk problem.*

Proof. With $\sigma = \operatorname{sign}(z(x^*)) = 0$ and $\Sigma = \operatorname{diag}(\sigma) = 0$, Definition 10 yields

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

This proves the claim. □

Note that a nullspace matrix of $J_t(x^*)$ is $U_t(x^*) = U(x^*) \in \mathbb{R}^{n\times[n-(m+s)]}$. First and second order necessary conditions of the localized trunk problem are now readily stated in terms of the Lagrangian

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

Theorem 17 (First Order Necessary Conditions)**.** *Let $f \in C^1(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^1(D^{x,|z|}, \mathbb{R}^m)$, and $c_{\mathcal{Z}} \in C^1(D^{x,|z|}, \mathbb{R}^s)$. Assume that $(x^*, 0)$ is a local minimizer of the abs-normal NLP (3) and that LIKQ holds at x^*. Then, there exists $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{Z}})$ such that the following conditions are satisfied:*

$$\begin{aligned} f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_1 c_{\mathcal{Z}}(x^*, 0) &= 0 \quad \textit{(tangential stationarity)},\\ c_{\mathcal{E}}(x^*, 0) &= 0 \quad \textit{(primal feasibility)},\\ c_{\mathcal{Z}}(x^*, 0) &= 0 \quad \textit{(switching feasibility)}. \end{aligned} \tag{5}$$

The Lagrange multipliers are unique since LICQ holds by Lemma 16.

Remark 18. In contrast to standard theory for smooth NLPs, these conditions are *not* sufficient in the linear case since the converse of Lemma 15 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 19 (Second Order Necessary Condition)**.** *Let $f \in C^2(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^2(D^{x,|z|}, \mathbb{R}^m)$, and $c_{\mathcal{Z}} \in C^2(D^{x,|z|}, \mathbb{R}^s)$. Assume that $(x^*, 0)$ is a local minimizer of the abs-normal NLP (3) 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 $H_t(x^*, \lambda^*) = \partial^2_{xx} \mathcal{L}_t(x^*, \lambda^*) \in \mathbb{R}^{n \times n}$.

Again we return to Example 6 to illustrate the trunk problem and its necessary conditions.

Example 20 (Trunk problem of Example 6)**.** With $z = 0$ in Example 12 one obtains the localized trunk problem

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

LICQ holds by Lemma 16 since LIKQ is satisfied at $x^* = (0, 0, 0)$ as was shown in Example 12. The Lagrangian reads

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

and 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

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

and the matrix $U(x^*) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ from Example 12:

$$U(x^*)^T H(x^*, \lambda^*) U(x^*) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0\\ 0\\ 1 \end{bmatrix} = 2 \geq 0.$$

Thus, second order necessary conditions of Theorem 19 hold at the solution x^* with $z^* = 0$.

4.2. **Branch Problems.** One cannot derive sufficient conditions for the abs-normal NLP using the localized trunk problem alone since the converse of Lemma 15 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 21 (Localized Branch Problems)**.** The *localized branch problem* associated with the definite signature $\sigma \in \{-1, 1\}^s$ and $\Sigma = \operatorname{diag}(\sigma)$ reads

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

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

$$\min_{x,\bar{z}} \quad f(x) \quad \text{s.t.} \quad \begin{aligned} & c_{\mathcal{E}}(x, \bar{z}) = 0, \\ & c_{\mathcal{Z}}(x, \bar{z}) - \Sigma \bar{z} = 0, \\ & \bar{z} \geq 0. \end{aligned} \tag{6}$$

The feasible set is

$$\begin{aligned} \Omega_\Sigma &:= \{(x, z) \colon c_{\mathcal{E}}(x, \Sigma z) = 0, c_{\mathcal{Z}}(x, \Sigma z) = z, \Sigma z \geq 0\} \\ &= \{(x, \bar{z}) \colon c_{\mathcal{E}}(x, \bar{z}) = 0, c_{\mathcal{Z}}(x, \bar{z}) = \Sigma \bar{z}, \bar{z} \geq 0\}. \end{aligned}$$

By construction, we have the following inclusions of feasible sets:

$$\Omega_t \subset \Omega_\Sigma \subset \Omega \quad \text{for all } \sigma \in \{-1, 1\}^s.$$

Moreover, it is easily seen that the branch problems coincide in the localized trunk problem and provide a decomposition of the feasible set of the abs-normal NLP:

$$\bigcap_{\sigma \in \{-1,1\}^s} \Omega_\Sigma = \Omega_t \quad \text{and} \quad \bigcup_{\sigma \in \{-1,1\}^s} \Omega_\Sigma = \Omega.$$

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

Lemma 22. *The pair $(x^*, 0)$ is a local minimizer of the abs-normal NLP (3) if and only if it is a local minimizer of the localized branch problem (6) for every definite $\Sigma = \operatorname{diag}(\sigma)$.*

As every branch problem is smooth, we can again apply standard NLP theory 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 immediately yields the following implication.

Lemma 23. *If LICQ at x^* holds for the localized trunk problem, then LICQ at $(x^*, 0)$ holds for all branch problems.*

Thus, due to Lemma 16, LICQ for all branch problems is implied by LIKQ at x^* with $z(x^*) = 0$. A nullspace matrix of $J_\Sigma(x^*)$ is given by $U_\Sigma(x^*) = \begin{bmatrix} U(x^*)^T & 0 \end{bmatrix}^T \in \mathbb{R}^{(n+s)\times([n-(m+s)]}$, as a comparison of $J_\Sigma(x^*)$ with $J_t(x^*)$ shows. The Lagrangians of the branch problems read

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

Theorem 24 (First Order Necessary Conditions)**.** *Let $f \in C^1(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^1(D^{x,|z|}, \mathbb{R}^m)$, and $c_{\mathcal{Z}} \in C^1(D^{x,|z|}, \mathbb{R}^s)$. Assume that $(x^*, 0)$ is a local minimizer of the abs-normal NLP (3) and that*

LIKQ holds at x^*. *Then, there exists* $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{Z}})$ *such that the following conditions are satisfied for every definite* $\Sigma = \operatorname{diag}(\sigma)$*:*

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

Proof. Given $\sigma \in \{-1, 1\}^s$, applying the first order necessary conditions of smooth NLPs yields the existence of $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{Z}})$ and μ^* such that

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

The first, third, and fourth condition are precisely the necessary conditions of the trunk problem from Theorem 17. Thus, the Lagrange multiplier vectors λ^* are unique and are 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:

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

Fortunately, since the multipliers 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 satisfies $\Sigma \lambda^*_{\mathcal{Z}} = |\lambda^*_{\mathcal{Z}}|$. 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 25. *The equivalence just stated immediately implies:*

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

As in the unconstrained case, we thus get rid of the combinatorial complexity of the first order necessary conditions stated in Theorem 24. We obtain a single set of conditions for *all* 2^s branch problems.

Corollary 26 (First Order Necessary Conditions)**.** *Let* $f \in C^1(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^1(D^{x,|z|}, \mathbb{R}^m)$, *and* $c_{\mathcal{Z}} \in C^1(D^{x,|z|}, \mathbb{R}^s)$. *Assume that* $(x^*, 0)$ *is a local minimizer of the abs-normal NLP* (3) *and that LIKQ holds at* x^*. *Then, there exists* $\lambda^* = (\lambda^*_{\mathcal{E}}, \lambda^*_{\mathcal{Z}})$ *such that the following conditions are satisfied:*

$$\begin{aligned} f'(x^*) + (\lambda^*_{\mathcal{E}})^T \partial_1 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_1 c_{\mathcal{Z}}(x^*, 0) &= 0 && \textit{(tangential stationarity)},\\ (\lambda^*_{\mathcal{E}})^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda^*_{\mathcal{Z}})^T \partial_2 c_{\mathcal{Z}}(x^*, 0) &\geq |\lambda^*_{\mathcal{Z}}|^T && \textit{(normal growth)},\\ c_{\mathcal{E}}(x^*, 0) &= 0 && \textit{(primal feasibility)},\\ c_{\mathcal{Z}}(x^*, 0) &= 0 && \textit{(switching feasibility)}. \end{aligned} \tag{7}$$

Observe that the normal growth condition provides a considerable strengthening of the otherwise identical first order conditions of the trunk problem derived in Theorem 17. Later we will show that, as in the smooth case, (7) provides sufficient first order conditions under special assumptions, possibly requiring normal growth in strict form. Strict normal growth in (7) is also needed for the following second order sufficient conditions.

Theorem 27 (Second Order Sufficient Conditions). *Let $f \in C^2(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^2(D^{x,|z|}, \mathbb{R}^m)$, and $c_{\mathcal{Z}} \in C^2(D^{x,|z|}, \mathbb{R}^s)$. Assume that $(x^*, 0)$ is feasible for the abs-normal NLP (3) and that LIKQ holds at x^*. Assume further that a Lagrange multiplier vector λ^* exists such that the first order necessary conditions (7) 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

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

Then, $(x^*, 0)$ is a strict local minimizer of the abs-normal NLP (3).

Proof. The strict normal growth condition implies $\mu^* > 0$ for all branch problems. Thus, we have strict complementarity in every branch problem, and by LIKQ we have unique Lagrange multipliers. To apply standard second order conditions for smooth NLPs we have to show 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^*: we have

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

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

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

and the assertion follows by standard theory for the smooth case. □

Again we have a closer look at Example 6.

Example 28. With $\bar{z} = \Sigma z$ for definite $\Sigma = \operatorname{diag}(\sigma)$ we obtain the following branch problems:

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

Here, LICQ holds at $x^* = (0, 0, 0)$ by Lemma 23 since LIKQ holds, see Example 12. Applying Theorem 24 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, we have strict normal growth

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

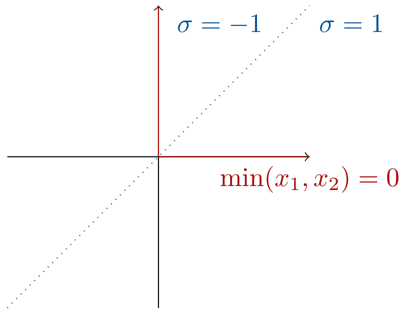

FIGURE 1. Illustration of Example 30.

and the reduced Hessian is positive definite,

$$U(x^*)^T H(x^*, \lambda^*) U(x^*) = 2 > 0.$$

Thus, we can apply Theorem 27, saying that $x^* = (0, 0, 0)$ with $z^* = 0$ is a strict local minimizer of the abs-normal NLP.

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 an immediate consequence of Corollary 26 due to the convexitiy of all branch problems.

Theorem 29 (First Order Necessary and Sufficient Conditions for Linear Functions)**.** *Given $f \in C^1(D^x, \mathbb{R})$, $c_\mathcal{E} \in C^1(D^{x,|z|}, \mathbb{R}^m)$, and $c_\mathcal{Z} \in C^1(D^{x,|z|}, \mathbb{R}^s)$, assume that f, $c_\mathcal{E}$, and $c_\mathcal{Z}$ are linear and that LIKQ holds at x^*. Then, $(x^*, 0)$ is a local minimizer of the abs-normal NLP (3) if and only if there exists $\lambda^* = (\lambda^*_\mathcal{E}, \lambda^*_\mathcal{Z})$ such that the following conditions are satisfied:*

$$\begin{aligned} f'(x^*) + (\lambda^*_\mathcal{E})^T \partial_1 c_\mathcal{E}(x^*, 0) + (\lambda^*_\mathcal{Z})^T \partial_1 c_\mathcal{Z}(x^*, 0) &= 0 && \textit{(tangential stationarity)},\\ (\lambda^*_\mathcal{E})^T \partial_2 c_\mathcal{E}(x^*, 0) + (\lambda^*_\mathcal{Z})^T \partial_2 c_\mathcal{Z}(x^*, 0) &\geq |\lambda^*_\mathcal{Z}|^T && \textit{(normal growth)},\\ c_\mathcal{E}(x^*, 0) &= 0 && \textit{(primal feasibility)},\\ c_\mathcal{Z}(x^*, 0) &= 0 && \textit{(switching feasibility)}. \end{aligned} \tag{8}$$

In contrast to the smooth case we actually need some regularity here. It is provided by the LIKQ which ensures that the Lagrange multipliers of all branch problems coincide.

Next, we have a closer look at a very simple example.

Example 30. Consider the problem depicted in Fig. 1:

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

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

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

Clearly, the switching is localized at $z^* = z(x^*) = 0$, and LIKQ holds at x^* by full row rank of

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

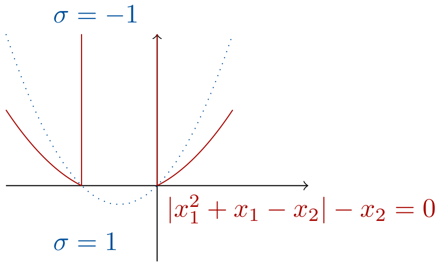

FIGURE 2. Illustration of Example 32.

Since all functions are linear, the assumptions of Theorem 29 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 $U(x^*)$ is empty and the second order conditions become trivial.

Theorem 31 (First Order Sufficient Conditions for $s + m = n$)**.** *Let $f \in C^2(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^2(D^{x,|z|}, \mathbb{R}^m)$, and $c_{\mathcal{Z}} \in C^2(D^{x,|z|}, \mathbb{R}^s)$. Assume that $s + m = n$ and that LIKQ holds at x^*. Then, $(x^*, 0)$ is a strict local minimizer of the abs-normal NLP (3) if there exists $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_{\mathcal{Z}}^*)$ such that the following conditions are satisfied:*

$$\begin{aligned} f'(x^*) + (\lambda_{\mathcal{E}}^*)^T \partial_1 c_{\mathcal{E}}(x^*, 0) + (\lambda_{\mathcal{Z}}^*)^T \partial_1 c_{\mathcal{Z}}(x^*, 0) &= 0 && \textit{(tangential stationarity)}, \\ (\lambda_{\mathcal{E}}^*)^T \partial_2 c_{\mathcal{E}}(x^*, 0) + (\lambda_{\mathcal{Z}}^*)^T \partial_2 c_{\mathcal{Z}}(x^*, 0) &> |\lambda_{\mathcal{Z}}^*|^T && \textit{(strict normal growth)}, \\ c_{\mathcal{E}}(x^*, 0) &= 0 && \textit{(primal feasibility)}, \\ c_{\mathcal{Z}}(x^*, 0) &= 0 && \textit{(switching feasibility)}. \end{aligned} \tag{9}$$

Next we look at a slightly more complicated example.

Example 32. Consider the abs-normal NLP depicted in Fig. 2:

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

The feasible set consists of two components and we have two local minimizers, $(0,0)$ and $(-1,0)$. We consider the global solution $x^* = (0,0)$. The abs-normal NLP reads

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

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

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

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 31 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 $\lambda_{\mathcal{Z}}^* = 2$.

5. General Non-Localized Case

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

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

By construction we have $z_+(x^*) > 0$, and z_+ keeps 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 we set

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

Then, we have $z_0(x^*) = 0$ but no additional information on the sign of z_0 in the neighborhood B. This leads to the definition

$$\sigma_0(x) := \operatorname{sign}(z_0(x)) \in \{-1,0,1\}^{|\alpha^*|},$$

and we set $z_+^* := z_+(x^*)$, $z_0^* := z_0(x^*) = 0$, $\bar{z}_0^* := \bar{z}_0(x^*) = 0$, and $\sigma_0^* := \sigma_0(x^*) = 0$. To partition the switching constraints, we divide the domain of $|z|$ via $D^{|z|} = D^{z_0} \times \Sigma_+^* D^{z_+}$ with $\Sigma_+^* = \operatorname{diag}(\sigma_+^*)$. Using the shorthand notation $D^{x,z_0,z_+} := D^x \times D^{z_0} \times D^{z_+}$, we set

$$\begin{aligned} c_+ &:= (\sigma_i^* e_i^T c_{\mathcal{Z}})_{i\notin\alpha^*} \in C^d(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|}),\\ c_0 &:= (e_i^T c_{\mathcal{Z}})_{i\in\alpha^*} \in C^d(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|}). \end{aligned}$$

Definition 33 (Split Abs-Normal NLP)**.** The *split abs-normal NLP* reads

$$\min_{x,z_0,z_+} \quad f(x) \quad \text{s.t.} \quad \begin{aligned} & c_{\mathcal{E}}(x,|z_0|,z_+) = 0,\\ & c_0(x,|z_0|,z_+) - z_0 = 0,\\ & c_+(x,|z_0|,z_+) - z_+ = 0. \end{aligned} \tag{10}$$

Its feasible set is

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

By construction, $(x^*,z^*) = (x^*,z(x^*))$ is a local minimizer of the abs-normal NLP (3) if and only if $(x^*,0,z_+^*) = (x^*,0,z_+(x^*))$ is a local minimizer of the split abs-normal NLP (10). Moreover, we can eliminate z_+ to derive a reduced formulation of the abs-normal NLP (3) by means of the implicit function theorem.

Lemma 34. *Assume that $c_+ \in C^d(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|})$ for some $d \geq 1$. Then, 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*

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

Note that we have $z_+(x^*, 0) = z_+(x^*)$. Substitution of this implicit function into the abs-normal NLP (3) yields the reduced abs-normal NLP.

Definition 35. The *reduced abs-normal NLP* reads

$$\min_{x,z_0} \quad f(x) \quad \text{s.t.} \quad \begin{aligned} &c_\mathcal{E}(x, |z_0|, z_+(x, |z_0|)) = 0, \\ &c_0(x, |z_0|, z_+(x, |z_0|)) - z_0 = 0. \end{aligned} \tag{11}$$

Its feasible set is

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

Clearly, $(x^*, 0, z_+^*)$ with $z_+^* = z_+(x^*, 0)$ is a local minimizer of the abs-normal NLP (10) if and only if $(x^*, 0)$ is a local minimizer of the reduced abs-normal NLP (11).

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 we will again study trunk and branch problems to derive optimality conditions for the split abs-normal NLP and for the reduced abs-normal NLP, yielding equivalent but different formulations. It turns out that some conditions of the split abs-normal NLP and others of the reduced abs-normal NLP are more convenient. We will combine the convenient formulations in the third subsection as optimality conditions for the non-localized abs-normal NLP.

5.1. **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.

5.1.1. *Reduced Trunk Problem.*

Setting $z_0 = 0$ in the reduced abs-normal NLP (11) yields the reduced trunk problem.

Definition 36 (Reduced Trunk Problem)**.** The *reduced trunk problem* reads

$$\min_{x} \quad f(x) \quad \text{s.t.} \quad \begin{aligned} &c_\mathcal{E}(x, 0, z_+(x, 0)) = 0, \\ &c_0(x, 0, z_+(x, 0)) = 0. \end{aligned} \tag{12}$$

Its feasible set is $\Omega_t^r := \{(x, 0) \colon c_\mathcal{E}(x, 0, z_+(x, 0)) = 0, c_0(x, 0, z_+(x, 0)) = 0\}$.

As in the localized case, x^* has to be a local minimizer of the reduced trunk problem (12) if $(x^*, 0)$ is a local minimizer of the reduced abs-normal NLP (11). Again, the trunk problem is smooth and we show 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 37. *Let $c_\mathcal{E} \in C^1(D^{x,z_0,z_+}, \mathbb{R}^m)$ and $c_0 \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$. Then, LIKQ at x^* is LICQ at x^* for the reduced trunk problem.*

Proof. Let $\Sigma^* = \operatorname{diag}(\sigma^*)$ and split the variables of Lemma 8 (with $\tilde{x} = x^*$) in a neighborhood B of x^*:

$$z = \begin{pmatrix} z_0 \\ \Sigma_+^* z_+ \end{pmatrix}, \quad c_\mathcal{Z} = \begin{pmatrix} c_0 \\ \Sigma_+^* c_+ \end{pmatrix}, \qquad \Sigma^* = \begin{bmatrix} \Sigma_0^* & 0 \\ 0 & \Sigma_+^* \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_+^* \end{bmatrix}.$$

Then we obtain with $z_0(x^*) = 0$ the Jacobian of the implicit function,

$$\partial_x z(x^*) = \begin{bmatrix} I & \partial_3 c_0(x^*, 0, z_+(x^*))[I - \partial_3 c_+(x^*, 0, z_+(x^*))]^{-1} \\ 0 & \Sigma_+^*[I - \partial_3 c_+(x^*, 0, z_+(x^*))]^{-1} \end{bmatrix} \begin{bmatrix} \partial_1 c_0(x^*, 0, z_+(x^*)) \\ \partial_1 c_+(x^*, 0, z_+(x^*)) \end{bmatrix}.$$

Using this in Definition 10, one obtains

$$\begin{aligned} J_\mathcal{E}(x^*) &= \partial_1 c_\mathcal{E}(x^*, 0, z_+(x^*)) + \partial_3 c_\mathcal{E}(x^*, 0, z_+(x^*))[I - \partial_3 c_+(x^*, 0, z_+(x^*))]^{-1} \partial_1 c_+(x^*, 0, z_+(x^*)), \\ J_\alpha(x^*) &= \partial_1 c_0(x^*, 0, z_+(x^*)) + \partial_3 c_0(x^*, 0, z_+(x^*))[I - \partial_3 c_+(x^*, 0, z_+(x^*))]^{-1} \partial_1 c_+(x^*, 0, z_+(x^*)), \end{aligned}$$

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

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

Thus, $J_t^r(x^*)$ has the form above. □

A nullspace matrix of $J_t^r(x^*)$ is $U_t^r(x^*) = U(x^*) \in \mathbb{R}^{n\times[n-(m+|\alpha^*|)]}$, and again we obtain first and second order necessary conditions by standard NLP theory for the smooth case. The Lagrangian is

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

Theorem 38 (First Order Necessary Conditions)**.** *Let $f \in C^1(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^1(D^{x,z_0,z_+}, \mathbb{R}^m)$, and $c_0 \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$. Assume that $(x^*, 0)$ is a local minimizer of the reduced abs-normal NLP (11) and that LIKQ holds at x^*. Then, there exists $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_0^*)$ such that the following conditions are satisfied, where $c_{\mathcal{E}}, c_0, \partial_x c_{\mathcal{E}}, \partial_x c_0$ are evaluated at $(x^*, 0, z_+(x^*, 0))$:*

$$\begin{aligned} f'(x^*) + (\lambda_{\mathcal{E}}^*)^T \partial_x c_{\mathcal{E}} + (\lambda_0^*)^T \partial_x c_0 &= 0 \quad \textit{(tangential stationarity)}, \\ c_{\mathcal{E}} &= 0 \quad \textit{(primal feasibility)}, \\ c_0 &= 0 \quad \textit{(switching feasibility)}. \end{aligned} \tag{13}$$

The Lagrange multipliers λ^* are unique.

Theorem 39 (Second Order Necessary Condition)**.** *Let $f \in C^2(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^2(D^{x,z_0,z_+}, \mathbb{R}^m)$, and $c_0 \in C^2(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$. Assume that $(x^*, 0)$ is a local minimizer of the reduced abs-normal NLP (11) and that LIKQ holds at x^*. Denote by λ^* the unique Lagrange multiplier vector. Then,*

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

with $H_t^r(x, \lambda) = \partial_{xx}^2 \mathcal{L}_t^r(x, \lambda)$.

5.1.2. *Reduced branch problems.* To obtain sufficient conditions, we now consider the $2^{|\alpha^*|}$ branch problems.

Definition 40 (Reduced Branch Problems)**.** The *reduced branch problem* associated with the definite signature $\sigma_0 \in \{-1, 1\}^{|\alpha^*|}$ and $\Sigma_0 = \operatorname{diag}(\sigma_0)$ reads

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

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

$$\begin{aligned} \min_{x,\bar{z}_0} \quad f(x) \quad \text{s.t.} \quad & c_{\mathcal{E}}(x, \bar{z}_0, z_+(x, \bar{z}_0)) = 0, \\ & c_0(x, \bar{z}_0, z_+(x, \bar{z}_0)) - \Sigma_0 \bar{z}_0 = 0, \\ & \bar{z}_0 \geq 0. \end{aligned} \tag{14}$$

Its feasible set is

$$\Omega_{\Sigma_0}^r := \{(x, z_0) \colon c_{\mathcal{E}}(x, \Sigma_0 z_0, z_+(x, \Sigma_0 z_0)) = 0, c_0(x, \Sigma_0 z_0, z_+(x, \Sigma_0 z_0)) = z_0, \Sigma_0 z_0 \geq 0\}.$$

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

Lemma 41. *The point $(x^*, 0)$ is a local minimizer of the reduced abs-normal NLP (11) if and only if $(x^*, 0)$ is a local minimizer of the reduced branch problem (14) for every definite $\Sigma_0 = \operatorname{diag}(\sigma_0)$.*

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

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

A nullspace matrix of $J^r_{\Sigma_0}(x^*)$ is $U^r_{\Sigma_0}(x^*) = \begin{bmatrix} U^r_t(x^*)^T & 0 \end{bmatrix}^T \in \mathbb{R}^{(n+|\alpha^*|)\times[n-(m+|\alpha^*|)]}$. With the Lagrangian

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

we obtain the first order conditions in the following theorem. Note that we state them directly with a single normal growth condition rather than $2^{|\alpha^*|}$ of them: as in the localized case it suffices to consider one branch with $\Sigma_0 \lambda_0^* = |\lambda_0^*|$.

Theorem 42 (First Order Necessary Conditions)**.** *Let $f \in C^1(D^x, \mathbb{R})$, $c_\mathcal{E} \in C^1(D^{x,z_0,z_+}, \mathbb{R}^m)$, and $c_0 \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$. Assume that $(x^*, 0)$ is a local minimizer of the reduced abs-normal NLP (11) and that LIKQ holds at x^*. Then, there exists $\lambda^* = (\lambda^*_\mathcal{E}, \lambda^*_0)$ such that the following conditions are satisfied, where $c_\mathcal{E}, c_0, \partial_x c_\mathcal{E}, \partial_x c_0, \partial_{\bar{z}_0} c_\mathcal{E}, \partial_{\bar{z}_0} c_0$ are evaluated at $(x^*, 0, z_+(x^*, 0))$:*

$$\begin{aligned} f'(x^*) + (\lambda^*_\mathcal{E})^T \partial_x c_\mathcal{E} + (\lambda^*_0)^T \partial_x c_0 &= 0 && \textit{(tangential stationarity)}, \\ (\lambda^*_\mathcal{E})^T \partial_{\bar{z}_0} c_\mathcal{E} + (\lambda^*_0)^T \partial_{\bar{z}_0} c_0 &\geq |\lambda^*_0|^T && \textit{(normal growth)}, \\ c_\mathcal{E} &= 0 && \textit{(primal feasibility)}, \\ c_0 &= 0 && \textit{(switching feasibility)}. \end{aligned} \tag{15}$$

Proof. As in the localized case. □

Finally we obtain second order sufficient conditions.

Theorem 43 (Second Order Sufficient Conditions)**.** *Let $f \in C^2(D^x, \mathbb{R})$, $c_\mathcal{E} \in C^2(D^{x,z_0,z_+}, \mathbb{R}^m)$, and $c_0 \in C^2(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$. Assume that $(x^*, 0)$ is feasible for the reduced abs-normal NLP (11) and that LIKQ holds at x^*. Assume further that a Lagrange multiplier vector λ^* exists such that the first order necessary conditions (15) 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

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

Then, $(x^*, 0)$ is a strict local minimizer of the reduced abs-normal NLP (11).

Proof. As in the localized case. □

5.2. **Optimality Conditions of Split Abs-Normal NLP.** Here we avoid the implicit function of the reduced form 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 smaller nullspace of the active Jacobian.

5.2.1. *Split trunk problem.* The split trunk problem is obtained from the split abs-normal NLP by setting $z_0 = 0$.

Definition 44 (Split Trunk Problem)**.** The *split trunk problem* reads

$$\min_{x, z_+} \quad f(x) \quad \text{s.t.} \quad \begin{aligned} c_\mathcal{E}(x, 0, z_+) &= 0, \\ c_0(x, 0, z_+) &= 0, \\ c_+(x, 0, z_+) - z_+ &= 0. \end{aligned} \tag{16}$$

Its feasible set is $\Omega^s_t := \{(x, 0, z_+) \colon c_\mathcal{E}(x, 0, z_+) = 0, c_0(x, 0, z_+) = 0, c_+(x, 0, z_+) = z_+\}$.

If $(x^*, 0, z^*_+)$ is a local minimizer of the split abs-normal NLP (10), then (x^*, z^*_+) is a local minimizer of the split trunk problem (16). As before, we use the trunk problem to obtain necessary conditions. Here, LICQ at (x^*, z^*_+) means full row rank of

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

This is equivalent to LIKQ at x^*.

Lemma 45. *Let $c_{\mathcal{E}} \in C^1(D^{x,z_0,z_+}, \mathbb{R}^m)$, $c_0 \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$, and $c_+ \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|})$. 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 37 yields

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

where the second equality holds since $\partial_3 c_+(x^*, 0, z_+^*)$ is strictly lower triangular. □

Here we denote by $U_t^s(x^*, z_+^*) \in \mathbb{R}^{(n+|\sigma^*|)\times[n-(m+|\alpha^*|)]}$ a nullspace matrix of $J_t^s(x^*, z_+^*)$. Using once again the theory for the smooth case, we obtain the following necessary conditions.

Theorem 46 (First Order Necessary Conditions)**.** *Let $f \in C^1(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^1(D^{x,z_0,z_+}, \mathbb{R}^m)$, $c_0 \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$, and $c_+ \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|})$. Assume that $(x^*, 0, z_+^*)$ is a local minimizer of problem (10) and that LIKQ holds at x^*. Then, there exists $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_0^*, \lambda_+^*)$ such that the following conditions are satisfied, where the constraints $c_{\mathcal{E}}, c_0, c_+$ and all their partial derivatives are evaluated at $(x^*, 0, z_+^*)$:*

$$\begin{aligned} f'(x^*) + (\lambda_{\mathcal{E}}^*)^T \partial_1 c_{\mathcal{E}} + (\lambda_0^*)^T \partial_1 c_0 + (\lambda_+^*)^T \partial_1 c_+ &= 0, \\ (\lambda_{\mathcal{E}}^*)^T \partial_3 c_{\mathcal{E}} + (\lambda_0^*)^T \partial_3 c_0 + (\lambda_+^*)^T [\partial_3 c_+ - I] &= 0 \quad \textit{(tangential stationarity)}, \\ c_{\mathcal{E}} &= 0 \quad \textit{(primal feasibility)}, \\ c_0 &= 0, \\ c_+ - z_+^* &= 0 \quad \textit{(switching feasibility)}. \end{aligned} \tag{17}$$

Theorem 47 (Second Order Necessary Condition)**.** *Let $f \in C^2(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^2(D^{x,z_0,z_+}, \mathbb{R}^m)$, $c_0 \in C^2(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$, and $c_+ \in C^2(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|})$. Assume that $(x^*, 0, z_+^*)$ is a local minimizer of the split abs-normal NLP (10) and that LIKQ holds at x^*. Denote by λ^* the unique Lagrange multiplier. Then,*

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

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

As we cannot obtain sufficient conditions by using the split trunk problem, we again study branch problems in the neighborhood of the local minimizer.

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

Definition 48 (Split Branch Problem)**.** The *split branch problem* associated with the definite signature $\sigma_0 \in \{-1, 1\}^{|\alpha^*|}$ and $\Sigma_0 = \operatorname{diag}(\sigma_0)$ reads:

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

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

$$\begin{aligned} \min_{x,\bar{z}_0,z_+} \quad f(x) \quad \text{s.t.} \quad & c_{\mathcal{E}}(x, \bar{z}_0, z_+) = 0, \\ & c_0(x, \bar{z}_0, z_+) - \Sigma_0 \bar{z}_0 = 0, \\ & c_+(x, \bar{z}_0, z_+) - z_+ = 0, \\ & \bar{z}_0 \geq 0. \end{aligned} \tag{18}$$

Its feasible set is

$$\Omega_{\Sigma_0}^s := \{(x, z_0, z_+) : c_{\mathcal{E}}(x, \Sigma_0 z_0, z_+) = 0, c_0(x, \Sigma_0 z_0, z_+) = z_0, c_+(x, \Sigma_0 z_0, z_+) = z_+, \Sigma_0 z_0 \geq 0\}.$$

Once again, the relation between feasible sets implies a relation between minimizers, and we use the smooth branch problems to derive sufficient conditions.

Lemma 49. *The point x^* is a local minimizer of problem (3) if and only if $(x^*, 0, z_+^*)$ is a local minimizer of problem (18) for every definite $\Sigma_0 = \operatorname{diag}(\sigma_0)$.*

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

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

We use the partitioning $U_t^s(x^*, z_+^*) = \begin{bmatrix} U_1(x^*, z_+^*)^T & U_3(x^*, z_+^*)^T \end{bmatrix}^T$. Then, a nullspace matrix of $J^s_{\Sigma_0}(x^*, z_+^*)$ is $U^s_{\Sigma_0}(x^*, z_+^*) = \begin{bmatrix} U_1(x^*, z_+^*)^T & 0 & U_3(x^*, z_+^*)^T \end{bmatrix}^T \in \mathbb{R}^{(n+|\sigma^*|)\times[n-(m+|\alpha^*|)]}$. To state optimality conditions, we need the Lagrangian:

$$\begin{aligned} \mathcal{L}^s_{\Sigma_0}(x, \bar{z}_0, z_+, \lambda_{\mathcal{E}}, \lambda_0, \lambda_+, \mu_0) = f(x) + \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x, \bar{z}_0, z_+) + \lambda_0^T [c_0(x, \bar{z}_0, z_+) - \Sigma_0 \bar{z}_0] \\ + \lambda_+^T [c_+(x, \bar{z}_0, z_+) - z_+] - \mu_0^T \bar{z}_0. \end{aligned}$$

The first order necessary conditions then read as follows.

Theorem 50 (First Order Necessary Conditions)**.** *Let $f \in C^1(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^1(D^{x,z_0,z_+}, \mathbb{R}^m)$, $c_0 \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$, and $c_+ \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|})$. Assume that $(x^*, 0, z_+^*)$ is a local minimizer of the split abs-normal NLP (10) and that LIKQ holds at x^*. Then, there exists $\lambda^* = (\lambda_{\mathcal{E}}^*, \lambda_0^*, \lambda_+^*)$ such that the following conditions are satisfied, where the constraints $c_{\mathcal{E}}, c_0, c_+$ and all their partial derivatives are evaluated at $(x^*, 0, z_+^*)$:*

$$\begin{aligned} f'(x^*) + (\lambda_{\mathcal{E}}^*)^T \partial_1 c_{\mathcal{E}} + (\lambda_0^*)^T \partial_1 c_0 + (\lambda_+^*)^T \partial_1 c_+ = 0, & \\ (\lambda_{\mathcal{E}}^*)^T \partial_3 c_{\mathcal{E}} + (\lambda_0^*)^T \partial_3 c_0 + (\lambda_+^*)^T [\partial_3 c_+ - I] = 0 & \quad \textit{(tangential stationarity)}, \\ (\lambda_{\mathcal{E}}^*)^T \partial_2 c_{\mathcal{E}} + (\lambda_0^*)^T \partial_2 c_0 + (\lambda_+^*)^T \partial_2 c_+ \ge |\lambda_0^*|^T & \quad \textit{(normal growth)}, \\ c_{\mathcal{E}} = 0 & \quad \textit{(primal feasibility)}, \\ c_0 = 0, & \\ c_+ - z_+^* = 0 & \quad \textit{(switching feasibility)}. \end{aligned} \tag{19}$$

Proof. As in the localized case. ◻

Under the additional assumption of strict normal growth, we obtain a second order sufficient condition.

Theorem 51 (Second Order Sufficient Conditions)**.** *Let $f \in C^2(D^x, \mathbb{R})$, $c_{\mathcal{E}} \in C^2(D^{x,z_0,z_+}, \mathbb{R}^m)$, $c_0 \in C^2(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$, and $c_+ \in C^2(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|})$. Assume that $(x^*, 0, z_+^*)$ is feasible for the split abs-normal NLP (10) and that LIKQ holds at x^*. Assume further that a Lagrange multiplier vector λ^* exists such that the first order necessary conditions (19) 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

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

Then, $(x^*, 0, z_+^*)$ is a strict local minimizer of the split abs-normal NLP (10).

Proof. As in the localized case. ◻

5.3. **Combined Optimality Conditions.** In the two preceding sections we have derived optimality conditions for the non-localized abs-normal NLP in split form and in reduced form. Now we select a suitable combination of these conditions and extend them to a complete set of necessary or sufficient first and second order conditions. Observe first that the common Lagrange multipliers $(\lambda_{\mathcal{E}}^*, \lambda_0^*, \lambda_+^*)$ are identical in both versions, which was to be expected and is easily verified. There are also direct correspondences between stationary points and between the normal growth conditions, which we now state formally without proof.

Lemma 52. *The point* $(x^*, \lambda_\mathcal{E}^*, \lambda_0^*, \lambda_+^*)$ *satisfies the first order necessary conditions of Theorem 50 for the split abs-normal NLP* (10) *if and only if the point* $(x^*, \lambda_\mathcal{E}^*, \lambda_0^*)$ *satisfies the first order necessary conditions of Theorem 42 for the reduced abs-normal NLP* (11)*.*

Lemma 53. *The point* $(x^*, \lambda_\mathcal{E}^*, \lambda_0^*, \lambda_+^*)$ *satisfies the strict normal growth condition for the split abs-normal NLP* (10) *if and only if the point* $(x^*, \lambda_\mathcal{E}^*, \lambda_0^*)$ *satisfies the strict normal growth condition for the reduced abs-normal NLP* (11)*.*

Now observe 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. Hence, we choose the conditions of the split abs-normal NLP in Theorem 50 as the more convenient ones. Moreover, we have again sufficient first order conditions if all functions are linear or if the number of active switchings plus the number of constraints is equal to the dimension, i.e., $|\alpha^*| + m = n$. We state this formally without proof.

Theorem 54 (First Order Necessary and Sufficient Conditions for Linear Functions)**.** *Given* $f \in C^1(D^x, \mathbb{R})$, $c_\mathcal{E} \in C^1(D^{x,z_0,z_+}, \mathbb{R}^m)$, $c_0 \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$, *and* $c_+ \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|})$, *assume that* f, $c_\mathcal{E}$, *and* c_0 *are linear and that LIKQ holds at* x^**. Then,* $(x^*, 0)$ *is a local minimizer of the split abs-norml NLP* (10) *if and only if there exists* $\lambda^* = (\lambda_\mathcal{E}^*, \lambda_0^*, \lambda_+^*)$ *such that the conditions* (19) *hold.*

Theorem 55 (First Order Sufficient Conditions for $|\alpha^*| + m = n$)**.** *Given* $f \in C^2(D^x, \mathbb{R})$, $c_\mathcal{E} \in C^1(D^{x,z_0,z_+}, \mathbb{R}^m)$, $c_0 \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\alpha^*|})$, *and* $c_+ \in C^1(D^{x,z_0,z_+}, \mathbb{R}^{|\sigma^*|})$, *assume that we have* $|\alpha^*| + m = n$ *and that LIKQ holds at* x^**. Then,* $(x^*, 0)$ *is a strict local minimizer of the split abs-normal NLP* (10) *if there exists* $\lambda^* = (\lambda_\mathcal{E}^*, \lambda_0^*, \lambda_+^*)$ *such that the conditions* (19) *hold with strict normal growth.*

Observe finally that the second order conditions in split form involve the complicated nullspace matrix $U_t^s(x^*, z_+^*)$ whereas the corresponding conditions in reduced form involve the basic nullspace matrix $U(x^*)$ of $J(x^*)$. Therefore, we select the second order necessary and sufficient conditions of the reduced abs-normal NLP as the more convenient ones. They are stated in Theorem 39 and in Theorem 43, respectively.

6. Conclusion and Outlook

We have provided a straightforward extension of first and second order KKT theory for unconstrained abs-normal problems and for classical smooth NLPs to the general case of nonsmooth NLPs in abs-normal form. For technical simplification we have chosen the equivalent subclass of abs-normal NLPs with smooth objective and without inequality constraints, so that any nonsmoothness as well as the distinction of active and inactive constraints are captured by the switching variables of the equality constraints. In [5] we have shown that abs-normal NLPs as considered here are essentially the same problem class as MPECs. For unconstrained abs-normal problems we have also given a detailed comparison of several strong and weak MPEC constraint qualifications with LIKQ (and MFKQ) of Griewank and Walther, where it turns out that the abs-normal form provides some degree of built-in regularity. Corresponding relations between general abs-normal NLPs and general MPECs are a subject of ongoing research.

References

[1] A. Griewank. On stable piecewise linearization and generalized algorithmic differentiation. *Optimization Methods and Software*, 28(6):1139–1178, 2013.

[2] 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.

[3] A. Griewank and A. Walther. Characterizing and testing subdifferential regularity for piecewise smooth objective functions. Technical Report 6314, Optimization Online, November 2017. (Available from http://www.optimization-online.org/DB_HTML/2017/11/6314.html).

[4] 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.

[5] L. C. Hegerhorst-Schultchen, C. Kirches, and M. C. Steinbach. On the relation between MPECs and optimization problems in abs-normal form. *Accepted for publication in Optimization Methods and Software*, 2019. (Preprint available from http://www.optimization-online.org/DB_FILE/2018/06/6676.pdf).

[6] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, New York, NY, USA, second edition, 2006.

L. C. Hegerhorst-Schultchen, Institut für Angewandte Mathematik, Leibniz Universität Hannover, Welfengarten 1, 30167 Hannover, Germany
Email address: hegerhorst@ifam.uni-hannoverde

M. C. Steinbach Institut für Angewandte Mathematik, Leibniz Universität Hannover, Welfengarten 1, 30167 Hannover, Germany
Email address: mcs@ifam.uni-hannover.de