

# Relaxation schemes for mathematical programs with switching constraints

Christian Kanzow\*    Patrick Mehlitz†    Daniel Steck‡

September 6, 2018

Switching-constrained optimization problems form a difficult class of mathematical programs since their feasible set is almost disconnected while standard constraint qualifications are likely to fail at several feasible points. That is why the application of standard methods from nonlinear programming does not seem to be promising in order to solve such problems. In this paper, we adapt the relaxation method from Kanzow and Schwartz (SIAM J. Optim., 23(2):770–798, 2013) for the numerical treatment of mathematical programs with complementarity constraints to the setting of switching-constrained optimization. It is shown that the proposed method computes M-stationary points under mild assumptions. Furthermore, we comment on other possible relaxation approaches which can be used to tackle mathematical programs with switching constraints. As it turns out, adapted versions of Scholtes’ global relaxation scheme as well as the relaxation scheme of Steffensen and Ulbrich only find W-stationary points of switching-constrained optimization problems in general. Some computational experiments visualize the performance of the proposed relaxation method.

**Keywords:** Constraint qualifications, Mathematical program with switching constraints, Relaxation methods, Global convergence.

**MSC:** 65K05, 90C30, 90C33.

---

\*University of Würzburg, Institute of Mathematics, 97074 Würzburg, Germany, [kanzow@mathematik.uni-wuerzburg.de](mailto:kanzow@mathematik.uni-wuerzburg.de), <https://www.mathematik.uni-wuerzburg.de/~kanzow/>, ORCID: 0000-0003-2897-2509

†Brandenburgische Technische Universität Cottbus–Senftenberg, Institute of Mathematics, 03046 Cottbus, Germany, [mehlitz@b-tu.de](mailto:mehlitz@b-tu.de), ORCID: 0000-0002-9355-850X

‡University of Würzburg, Institute of Mathematics, 97074 Würzburg, Germany, [daniel.steck@mathematik.uni-wuerzburg.de](mailto:daniel.steck@mathematik.uni-wuerzburg.de), <https://www.mathematik.uni-wuerzburg.de/~steck/>, ORCID: 0000-0001-5335-5428

# 1 Introduction

This paper is dedicated to so-called *mathematical programs with switching constraints*, MPSCs for short. These are optimization problems of the form

$$\begin{aligned} f(x) &\rightarrow \min \\ g_i(x) &\leq 0, & i \in \mathcal{M}, \\ h_j(x) &= 0, & j \in \mathcal{P}, \\ G_l(x)H_l(x) &= 0, & l \in \mathcal{Q}, \end{aligned} \tag{MPSC}$$

where  $\mathcal{M} := \{1, \dots, m\}$ ,  $\mathcal{P} := \{1, \dots, p\}$ ,  $\mathcal{Q} := \{1, \dots, q\}$  are index sets and the functions  $f, g_i, h_j, G_l, H_l: \mathbb{R}^n \rightarrow \mathbb{R}$  are continuously differentiable for all  $i \in \mathcal{M}$ ,  $j \in \mathcal{P}$ , and  $l \in \mathcal{Q}$ . For brevity,  $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $h: \mathbb{R}^n \rightarrow \mathbb{R}^p$ ,  $G: \mathbb{R}^n \rightarrow \mathbb{R}^q$ , and  $H: \mathbb{R}^n \rightarrow \mathbb{R}^q$  are the mappings which possess the component functions  $g_i$  ( $i \in \mathcal{M}$ ),  $h_j$  ( $j \in \mathcal{P}$ ),  $G_l$  ( $l \in \mathcal{Q}$ ), and  $H_l$  ( $l \in \mathcal{Q}$ ), respectively. The last  $q$  constraints in (MPSC) force  $G_l(x)$  or  $H_l(x)$  to be zero for all  $l \in \mathcal{Q}$ , which gives rise to the terminology “switching constraints”.

Switching structures appear frequently in the context of optimal control, see Clason et al. [2017], Gugat [2008], Hante and Sager [2013], Liberzon [2003], Seidman [2013], Wang and Yan [2015], Zuazua [2011], and the references therein, or as a reformulation of so-called *either-or constraints*, see [Mehlitz, 2018, Section 7]. Naturally, (MPSC) is related to other problem classes from disjunctive programming such as *mathematical programs with complementarity constraints*, MPCCs for short, see Luo et al. [1996], Outrata et al. [1998], or *mathematical programs with vanishing constraints*, MPVCs for short, see Achtziger and Kanzow [2008], Hoheisel and Kanzow [2008]. Indeed, similarly to MPCCs and MPVCs, standard constraint qualifications are likely to be violated at the feasible points of (MPSC). Recently, stationarity conditions and constraint qualifications for (MPSC) were introduced in Mehlitz [2018].

Here, we focus on the computational treatment of (MPSC). Clearly, standard methods from nonlinear programming may run into difficulties when applied to (MPSC) due to two reasons: first, the feasible set of (MPSC) is likely to be disconnected or at least *almost* disconnected. Secondly, standard regularity conditions like the Mangasarian–Fromovitz constraint qualification, MFCQ for short, are likely to fail at the feasible points of (MPSC) under mild assumptions, see [Mehlitz, 2018, Lemma 4.1]. Similar issues appear in the context of MPCCs and MPVCs where several different relaxation schemes were introduced to overcome these shortcomings, see Hoheisel et al. [2012, 2013] and the references therein. Basically, the idea is to relax the *irregular* constraints using a relaxation parameter such that the resulting surrogate problems are (regular) standard nonlinear problems which can be tackled by common methods. The relaxation parameter is then iteratively reduced to zero and, in each iteration, a Karush–Kuhn–Tucker (KKT) point of the surrogate problem is computed. Ideally, the resulting sequence possesses a limit point and, under some problem-tailored constraint qualification, this point satisfies a suitable stationarity condition. Furthermore, it is desirable that the relaxed problems satisfy standard constraint qualifications in a neighborhood of the limit point under reasonable assumptions.

In this paper, we show that the relaxation scheme from [Kanzow and Schwartz \[2013\]](#), which was designed for the numerical investigation of MPCCs, can be adapted for the computational treatment of (MPSC). Particularly, it will be shown that the modified method can be used to find M-stationary points of (MPSC). By means of examples it is demonstrated that other relaxation methods which are well known from the theory of MPCCs only yield W-stationary points of (MPSC). We present the results of some numerical experiments which show the performance of the proposed method.

The remaining part of the paper is structured as follows. In [Section 2](#), we describe the general notation used throughout the paper and recall some fundamental theory on nonlinear programming and switching constraints. [Section 3](#) is dedicated to the main relaxation approach and contains various properties of the resulting algorithm as well as convergence results. In [Section 4](#), we describe how some common regularization methods from MPCCs can be carried over to the switching-constrained setting, and discuss the convergence properties of the resulting algorithms. [Section 5](#) contains some numerical applications, including either-or constraints, switching-constrained optimal control, and semi-continuous optimization problems arising in portfolio optimization. We conclude the paper with some final remarks in [Section 6](#).

## 2 Notation and preliminaries

### 2.1 Basic notation

The subsequently introduced tools of variational analysis can be found in [Rockafellar and Wets \[1998\]](#).

For a nonempty set  $A \subset \mathbb{R}^n$ , we call

$$A^\circ := \{y \in \mathbb{R}^n \mid \forall x \in A: x \cdot y \leq 0\}$$

the *polar cone* of  $A$ . Here,  $x \cdot y$  denotes the Euclidean inner product of the two vectors  $x, y \in \mathbb{R}^n$ . It is well known that  $A^\circ$  is a nonempty, closed, convex cone. For any two sets  $B_1, B_2 \subset \mathbb{R}^n$ , the polarization rule  $(B_1 \cup B_2)^\circ = B_1^\circ \cap B_2^\circ$  holds by definition. The polar of a polyhedral cone can be characterized by means of, e.g., Motzkin's theorem of alternatives. Note that we interpret the relations  $\leq$  and  $\geq$  for vectors componentwise.

**Lemma 2.1.** *For matrices  $\mathbf{C} \in \mathbb{R}^{m \times n}$  and  $\mathbf{D} \in \mathbb{R}^{p \times n}$ , let  $\mathcal{K} \subset \mathbb{R}^n$  be the polyhedral cone*

$$\mathcal{K} := \{d \in \mathbb{R}^n \mid \mathbf{C}d \leq 0, \mathbf{D}d = 0\}.$$

*Then  $\mathcal{K}^\circ = \{\mathbf{C}^\top \lambda + \mathbf{D}^\top \rho \mid \lambda \in \mathbb{R}^m, \lambda \geq 0, \rho \in \mathbb{R}^p\}$ .*

Let  $A \subset \mathbb{R}^n$  be a nonempty set and  $\bar{x} \in A$ . Then the closed cone

$$\mathcal{T}_A(\bar{x}) := \{d \in \mathbb{R}^n \mid \exists \{x_k\}_{k \in \mathbb{N}} \subset A \exists \{\tau_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+: x_k \rightarrow \bar{x}, \tau_k \rightarrow 0, (x_k - \bar{x})/\tau_k \rightarrow d\}$$

is called *tangent* or *Bouligand cone* to  $A$  at  $\bar{x}$ . Here,  $\mathbb{R}_+ := \{r \in \mathbb{R} \mid r > 0\}$  denotes the set of all positive reals.

The union  $\{v^1, \dots, v^r\} \cup \{w^1, \dots, w^s\}$  of sets  $\{v^1, \dots, v^r\}, \{w^1, \dots, w^s\} \subset \mathbb{R}^n$  is called *positive-linearly dependent* if there exist vectors  $\alpha \in \mathbb{R}^r$ ,  $\alpha \geq 0$ , and  $\beta \in \mathbb{R}^s$  which do not vanish at the same time such that

$$0 = \sum_{i=1}^r \alpha_i v^i + \sum_{j=1}^s \beta_j w^j.$$

Otherwise,  $\{v^1, \dots, v^r\} \cup \{w^1, \dots, w^s\}$  is called *positive-linearly independent*. Clearly, if the set  $\{v^1, \dots, v^r\}$  is empty, then the above definitions reduce to linear dependence and independence, respectively. The following lemma will be useful in this paper; its proof is similar to that of [Qi and Wei, 2000, Proposition 2.2] and therefore omitted.

**Lemma 2.2.** *Let  $\{v^1, \dots, v^r\}, \{w^1, \dots, w^s\} \subset \mathbb{R}^n$  be given sets whose union  $\{v^1, \dots, v^r\} \cup \{w^1, \dots, w^s\}$  is positive-linearly independent. Then there exists  $\varepsilon > 0$  such that, for all vectors  $\tilde{v}^1, \dots, \tilde{v}^r, \tilde{w}^1, \dots, \tilde{w}^s \in \{z \in \mathbb{R}^n \mid \|z\|_2 \leq \varepsilon\}$ , the union  $\{v^1 + \tilde{v}^1, \dots, v^r + \tilde{v}^r\} \cup \{w^1 + \tilde{w}^1, \dots, w^s + \tilde{w}^s\}$  is positive-linearly independent.*

For some vector  $z \in \mathbb{R}^n$  and an index set  $I \subset \{1, \dots, n\}$ ,  $z_I \in \mathbb{R}^{|I|}$  denotes the vector which results from  $z$  by deleting all  $z_i$  with  $i \in \{1, \dots, n\} \setminus I$ . Finally, let us mention that  $\text{supp } z := \{i \in \{1, \dots, n\} \mid z_i \neq 0\}$  is called the *support* of the vector  $z \in \mathbb{R}^n$ .

## 2.2 Standard nonlinear programs

Here, we recall some fundamental constraint qualifications from standard nonlinear programming, see, e.g., Bazarraa et al. [1993]. Therefore, we consider the nonlinear program

$$\begin{aligned} f(x) &\rightarrow \min \\ g_i(x) &\leq 0, & i \in \mathcal{M}, \\ h_j(x) &= 0, & j \in \mathcal{P}, \end{aligned} \tag{NLP}$$

i.e., we forget about the switching constraints in (MPSC) for a moment.

Let  $\tilde{X} \subset \mathbb{R}^n$  denote the feasible set of (NLP) and fix some point  $\bar{x} \in \tilde{X}$ . Then

$$I^g(\bar{x}) := \{i \in \mathcal{M} \mid g_i(\bar{x}) = 0\}$$

is called the index set of active inequality constraints at  $\bar{x}$ . Furthermore, the set

$$\mathcal{L}_{\tilde{X}}(\bar{x}) := \left\{ d \in \mathbb{R}^n \mid \begin{array}{l} \nabla g_i(\bar{x}) \cdot d \leq 0 \quad i \in I^g(\bar{x}) \\ \nabla h_j(\bar{x}) \cdot d = 0 \quad j \in \mathcal{P} \end{array} \right\}$$

is called the linearization cone to  $\tilde{X}$  at  $\bar{x}$ . Obviously,  $\mathcal{L}_{\tilde{X}}(\bar{x})$  is a polyhedral cone, and thus closed and convex. It is well known that  $\mathcal{T}_{\tilde{X}}(\bar{x}) \subset \mathcal{L}_{\tilde{X}}(\bar{x})$  is always satisfied. The converse inclusion generally only holds under some constraint qualification.

In the definition below, we recall several standard constraint qualifications which are applicable to (NLP).

**Definition 2.3.** Let  $\bar{x} \in \mathbb{R}^n$  be a feasible point of (NLP). Then  $\bar{x}$  is said to satisfy the

- (a) linear independence constraint qualification (LICQ) if the following vectors are linearly independent:

$$\{\nabla g_i(\bar{x}) \mid i \in I^g(\bar{x})\} \cup \{\nabla h_j(\bar{x}) \mid j \in \mathcal{P}\}. \quad (1)$$

- (b) Mangasarian–Fromovitz constraint qualification (MFCQ) if the vectors in (1) are positive-linearly independent.

- (c) constant positive linear dependence condition (CPLD) if, for any sets  $I \subset I^g(\bar{x})$  and  $J \subset \mathcal{P}$  such that the gradients

$$\{\nabla g_i(\bar{x}) \mid i \in I\} \cup \{\nabla h_j(\bar{x}) \mid j \in J\}$$

are positive-linearly dependent, there exists a neighborhood  $U \subset \mathbb{R}^n$  of  $\bar{x}$  such that the gradients

$$\{\nabla g_i(x) \mid i \in I\} \cup \{\nabla h_j(x) \mid j \in J\}$$

are linearly dependent for all  $x \in U$ .

- (d) Abadie constraint qualification (ACQ) if  $\mathcal{T}_{\bar{X}}(\bar{x}) = \mathcal{L}_{\bar{X}}(\bar{x})$ .

- (e) Guignard constraint qualification (GCQ) if  $\mathcal{T}_{\bar{X}}(\bar{x})^\circ = \mathcal{L}_{\bar{X}}(\bar{x})^\circ$ .

Note that the following relations hold between the constraint qualifications from Definition 2.3:

$$\text{LICQ} \implies \text{MFCQ} \implies \text{CPLD} \implies \text{ACQ} \implies \text{GCQ},$$

see [Hoheisel et al., 2013, Section 2.1] for some additional information.

It is well known that the validity of GCQ at some local minimizer  $\bar{x} \in \mathbb{R}^n$  of (NLP) implies that the KKT conditions

$$\begin{aligned} 0 &= \nabla f(\bar{x}) + \sum_{i \in I^g(\bar{x})} \lambda_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(\bar{x}), \\ \forall i \in I^g(\bar{x}): \lambda_i &\geq 0 \end{aligned}$$

provide a necessary optimality condition. Thus, the same holds for the stronger constraint qualifications ACQ, CPLD, MFCQ, and LICQ.

### 2.3 Mathematical programs with switching constraints

The statements of this section are taken from Mehlitz [2018]. Let  $X \subset \mathbb{R}^n$  denote the feasible set of (MPSC) and fix a point  $\bar{x} \in X$ . Then the index sets

$$\begin{aligned} I^G(\bar{x}) &:= \{l \in \mathcal{Q} \mid G_l(\bar{x}) = 0, H_l(\bar{x}) \neq 0\}, \\ I^H(\bar{x}) &:= \{l \in \mathcal{Q} \mid G_l(\bar{x}) \neq 0, H_l(\bar{x}) = 0\}, \\ I^{GH}(\bar{x}) &:= \{l \in \mathcal{Q} \mid G_l(\bar{x}) = 0, H_l(\bar{x}) = 0\} \end{aligned}$$

form a disjoint partition of  $\mathcal{Q}$ . It is easily seen that MFCQ (and thus LICQ) cannot hold for (MPSC) at  $\bar{x}$  if  $I^{GH}(\bar{x}) \neq \emptyset$ . Taking a look at the associated linearization cone

$$\mathcal{L}_X(\bar{x}) = \left\{ d \in \mathbb{R}^n \left| \begin{array}{ll} \nabla g_i(\bar{x}) \cdot d \leq 0 & i \in I^g(\bar{x}) \\ \nabla h_j(\bar{x}) \cdot d = 0 & j \in \mathcal{P} \\ \nabla G_k(\bar{x}) \cdot d = 0 & k \in I^G(\bar{x}) \\ \nabla H_k(\bar{x}) \cdot d = 0 & k \in I^H(\bar{x}) \end{array} \right. \right\},$$

which is always convex, one can imagine that ACQ is likely to fail as well if  $I^{GH}(\bar{x}) \neq \emptyset$  since, in the latter situation,  $\mathcal{T}_X(\bar{x})$  might be nonconvex. Note that GCQ may hold for (MPSC) even in the aforementioned context.

Due to the inherent lack of regularity, stationarity conditions for (MPSC) which are weaker than the associated KKT conditions were introduced.

**Definition 2.4.** A feasible point  $\bar{x} \in X$  of (MPSC) is called

- (a) weakly stationary (W-stationary) if there exist multipliers  $\lambda_i$  ( $i \in I^g(\bar{x})$ ),  $\rho_j$  ( $j \in \mathcal{P}$ ),  $\mu_l$  ( $l \in \mathcal{Q}$ ), and  $\nu_l$  ( $l \in \mathcal{Q}$ ) which solve the following system:

$$\begin{aligned} 0 = \nabla f(\bar{x}) + \sum_{i \in I^g(\bar{x})} \lambda_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(\bar{x}) \\ + \sum_{l \in \mathcal{Q}} [\mu_l \nabla G_l(\bar{x}) + \nu_l \nabla H_l(\bar{x})], \end{aligned}$$

$$\forall i \in I^g(\bar{x}): \lambda_i \geq 0,$$

$$\forall l \in I^H(\bar{x}): \mu_l = 0,$$

$$\forall l \in I^G(\bar{x}): \nu_l = 0.$$

- (b) Mordukhovich-stationary (M-stationary) if it is W-stationary and the associated multipliers additionally satisfy

$$\forall l \in I^{GH}(\bar{x}): \mu_l \nu_l = 0.$$

- (c) strongly stationary (S-stationary) if it is W-stationary while the associated multipliers additionally satisfy

$$\forall l \in I^{GH}(\bar{x}): \mu_l = 0 \wedge \nu_l = 0.$$

Clearly, the following implications hold:

$$\text{S-stationarity} \implies \text{M-stationarity} \implies \text{W-stationarity}.$$

Moreover, the KKT conditions of (MPSC) are equivalent to the S-stationarity conditions from Definition 2.4. One may check Figure 1 for a geometric interpretation of the Lagrange multipliers associated with the switching conditions from  $I^{GH}(\bar{x})$ . !DS

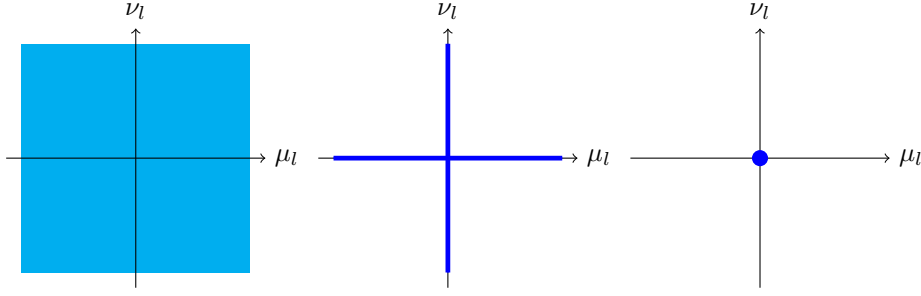


Figure 1: Geometric illustrations of weak, M-, and S-stationarity for an index  $l \in I^{GH}(\bar{x})$ .

In order to ensure that one of these stationarity notions plays the role of a necessary optimality condition for (MPSC), suitable problem-tailored constraint qualifications need to be valid. For the definition of such conditions, the following so-called *tightened nonlinear problem* is of interest:

$$\begin{aligned}
 f(x) &\rightarrow \min \\
 g_i(x) &\leq 0, & i \in \mathcal{M}, \\
 h_j(x) &= 0, & j \in \mathcal{P}, \\
 G_l(x) &= 0, & l \in I^G(\bar{x}) \cup I^{GH}(\bar{x}), \\
 H_l(x) &= 0, & l \in I^H(\bar{x}) \cup I^{GH}(\bar{x}).
 \end{aligned} \tag{TNLP}$$

Note that (TNLP) is a standard nonlinear program.

**Definition 2.5.** Let  $\bar{x} \in X$  be a feasible point of (MPSC). Then MPSC-LICQ (MPSC-MFCQ) is said to hold for (MPSC) at  $\bar{x}$  if LICQ (MFCQ) holds for (TNLP) at  $\bar{x}$ , i.e., if the vectors

$$\begin{aligned}
 &\{\nabla g_i(\bar{x}) \mid i \in I^g(\bar{x})\} \cup \left[ \{\nabla h_j(\bar{x}) \mid j \in \mathcal{P}\} \right. \\
 &\quad \cup \{\nabla G_l(\bar{x}) \mid l \in I^G(\bar{x}) \cup I^{GH}(\bar{x})\} \\
 &\quad \left. \cup \{\nabla H_l(\bar{x}) \mid l \in I^H(\bar{x}) \cup I^{GH}(\bar{x})\} \right]
 \end{aligned}$$

are linearly independent (positive-linearly independent).

It is obvious that MPSC-LICQ is stronger than MPSC-MFCQ. Furthermore, MPSC-LICQ implies standard GCQ for (MPSC) at the reference point.

In this paper, we will use another MPSC-tailored constraint qualification called MPSC-NNAMCQ, where NNAMCQ stands for the *No Nonzero Abnormal Multiplier Constraint Qualification* which has been introduced to investigate optimization problems whose feasible sets are preimages of closed but not necessarily convex sets under continuously differentiable mappings, see [Rockafellar and Wets, 1998, Section 6.D] and Ye and Ye [1997]. Clearly, (MPSC) belongs to this problem class as well if one reformulates the switching constraints as

$$(G_l(x), H_l(x)) \in C, \quad l \in \mathcal{Q},$$

with  $C := \{(a, b) \in \mathbb{R}^2 \mid ab = 0\}$ .

**Definition 2.6.** Let  $\bar{x} \in X$  be a feasible point of (MPSC). Then MPSC-NNAMCQ is said to hold for (MPSC) at  $\bar{x}$  if the following condition is valid:

$$\left. \begin{aligned} 0 &= \sum_{i \in I^g(\bar{x})} \lambda_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(\bar{x}) \\ &\quad + \sum_{l \in \mathcal{Q}} [\mu_l \nabla G_l(\bar{x}) + \nu_l \nabla H_l(\bar{x})] \\ \forall i \in I^g(\bar{x}): \lambda_i &\geq 0, \\ \forall l \in I^H(\bar{x}): \mu_l &= 0, \\ \forall l \in I^G(\bar{x}): \nu_l &= 0, \\ \forall l \in I^{GH}(\bar{x}): \mu_l \nu_l &= 0 \end{aligned} \right\} \implies \lambda = 0, \rho = 0, \mu = 0, \nu = 0.$$

It is easy to check that MPSC-NNAMCQ is implied by MPSC-MFCQ. Note that a local minimizer of (MPSC) where MPSC-LICQ holds is an S-stationary point. Furthermore, one can easily check that the associated multipliers which solve the system of S-stationarity are uniquely determined. Under MPSC-MFCQ (and, thus, MPSC-NNAMCQ), a local minimizer of (MPSC) is, in general, only M-stationary. Finally, there exist several problem-tailored constraint qualifications for (MPSC) which are weaker than MPSC-MFCQ but also imply that local solutions are M-stationary, see Mehlitz [2018].

!DS

### 3 The relaxation scheme and its convergence properties

#### 3.1 On the relaxation scheme

For our relaxation approach, we will make use of the function  $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$  defined below:

$$\forall (a, b) \in \mathbb{R}^2: \quad \varphi(a, b) := \begin{cases} ab & \text{if } a + b \geq 0, \\ -\frac{1}{2}(a^2 + b^2) & \text{if } a + b < 0. \end{cases}$$

The function  $\varphi$  was introduced in Kanzow and Schwartz [2013] to study a relaxation method for the numerical solution of MPCCs. In the following lemma, which parallels [Kanzow and Schwartz, 2013, Lemma 3.1], some properties of  $\varphi$  are summarized.

**Lemma 3.1.** (a) The function  $\varphi$  is an NCP-function, i.e. it satisfies

$$\forall (a, b) \in \mathbb{R}^2: \quad \varphi(a, b) = 0 \iff a \geq 0 \wedge b \geq 0 \wedge ab = 0.$$

(b) The function  $\varphi$  is continuously differentiable and satisfies

$$\forall (a, b) \in \mathbb{R}^2: \quad \nabla \varphi(a, b) = \begin{cases} \begin{pmatrix} b \\ a \end{pmatrix} & \text{if } a + b \geq 0, \\ \begin{pmatrix} -a \\ -b \end{pmatrix} & \text{if } a + b < 0. \end{cases}$$



For some parameter  $t \geq 0$  as well as indices  $s \in \mathcal{S} := \{1, 2, 3, 4\}$  and  $l \in \mathcal{Q}$ , we define functions  $\Phi_l^s(\cdot; t): \mathbb{R}^n \rightarrow \mathbb{R}$  via

$$\begin{aligned}\Phi_l^1(x; t) &:= \varphi(G_l(x) - t, H_l(x) - t), & \Phi_l^2(x; t) &:= \varphi(-G_l(x) - t, H_l(x) - t), \\ \Phi_l^3(x; t) &:= \varphi(-G_l(x) - t, -H_l(x) - t), & \Phi_l^4(x; t) &:= \varphi(G_l(x) - t, -H_l(x) - t)\end{aligned}$$

for any  $x \in \mathbb{R}^n$ . Now, we are in position to introduce the surrogate problem of our interest:

$$\begin{aligned}f(x) &\rightarrow \min \\ g_i(x) &\leq 0 & i \in \mathcal{M} \\ h_j(x) &= 0 & j \in \mathcal{P} \\ \Phi_l^s(x; t) &\leq 0 & s \in \mathcal{S}, l \in \mathcal{Q}.\end{aligned}\tag{P(t)}$$

The feasible set of (P(t)) will be denoted by  $X(t)$ . Figure 2 provides an illustration of  $X(t)$ .

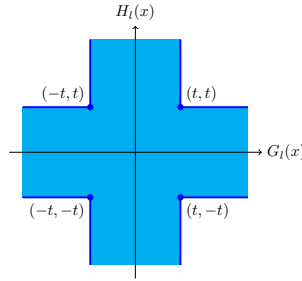


Figure 2: Geometric interpretation of the relaxed feasible set  $X(t)$ .

Let  $x \in X(t)$  be a feasible point of (P(t)) for some fixed parameter  $t > 0$ . Later on, it will be beneficial to work with the index sets defined below:

$$\begin{aligned}I_{t,1}^{00}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) = t, H_l(x) = t\}, & I_{t,2}^{00}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) = -t, H_l(x) = t\}, \\ I_{t,1}^{0+}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) = t, H_l(x) > t\}, & I_{t,2}^{0+}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) = -t, H_l(x) > t\}, \\ I_{t,1}^{+0}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) > t, H_l(x) = t\}, & I_{t,2}^{-0}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) < -t, H_l(x) = t\}, \\ I_{t,3}^{00}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) = -t, H_l(x) = -t\}, & I_{t,4}^{00}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) = t, H_l(x) = -t\}, \\ I_{t,3}^{0-}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) = -t, H_l(x) < -t\}, & I_{t,4}^{0-}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) = t, H_l(x) < -t\}, \\ I_{t,3}^{-0}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) < -t, H_l(x) = -t\}, & I_{t,4}^{+0}(x) &:= \{l \in \mathcal{Q} \mid G_l(x) > t, H_l(x) = -t\}.\end{aligned}$$

Note that all these sets are pairwise disjoint. The index sets  $I_{t,1}^{00}(x)$ ,  $I_{t,1}^{0+}(x)$ , and  $I_{t,1}^{+0}(x)$  !DS subsume the three possible cases where the constraints  $\Phi_l^1(x; t) \leq 0$  ( $l \in \mathcal{Q}$ ) are active. Similarly, the other index sets cover those indices where the constraints  $\Phi_l^s(x; t) \leq 0$  ( $l \in \mathcal{Q}$ ,  $s \in \mathcal{S}$ ) are active. It follows that an index  $l \in \mathcal{Q}$  which does not belong to any of the above sets is inactive for (P(t)) and can therefore be disregarded (locally).

For brevity, we set

$$I_t^{00}(x) := \bigcup_{s \in \mathcal{S}} I_{t,s}^{00}(x),$$

$$\begin{aligned}
I_t^{0\pm}(x) &:= I_{t,1}^{0+}(x) \cup I_{t,2}^{0+}(x) \cup I_{t,3}^{0-}(x) \cup I_{t,4}^{0-}(x), \\
I_t^{\pm 0}(x) &:= I_{t,1}^{+0}(x) \cup I_{t,2}^{-0}(x) \cup I_{t,3}^{-0}(x) \cup I_{t,4}^{+0}(x).
\end{aligned}$$

Thus, we collect all indices in  $I_t^{0\pm}(x)$  where  $G_l(x) \in \{-t, t\}$  holds while  $|H_l(x)| > t$  is valid. Similarly,  $I_t^{\pm 0}(x)$  comprises all indices where  $H_l(x) \in \{-t, t\}$  and  $|G_l(x)| > t$  hold true. The set  $I_t^{00}(x)$  contains all those indices where  $G_l(x), H_l(x) \in \{-t, t\}$  is valid. In order to address any of the four quadrants separately, we will exploit

$$\begin{aligned}
I_{t,1}^0(x) &:= I_{t,1}^{00}(x) \cup I_{t,1}^{0+}(x) \cup I_{t,1}^{+0}(x), & I_{t,2}^0(x) &:= I_{t,2}^{00}(x) \cup I_{t,2}^{0+}(x) \cup I_{t,2}^{-0}(x), \\
I_{t,3}^0(x) &:= I_{t,3}^{00}(x) \cup I_{t,3}^{0-}(x) \cup I_{t,3}^{-0}(x), & I_{t,4}^0(x) &:= I_{t,4}^{00}(x) \cup I_{t,4}^{0-}(x) \cup I_{t,4}^{+0}(x),
\end{aligned}$$

i.e.,  $I_{t,s}^0(x)$  collects all indices  $l \in \mathcal{Q}$  where  $\Phi_l^s(x; t)$  is active. !DS

If  $\bar{x} \in X$  is feasible to (MPSC),  $x$  lies in a sufficiently small neighborhood of  $\bar{x}$ , and  $t \geq 0$  is sufficiently small, then the following inclusions follow from the continuity of all appearing functions:

$$\begin{aligned}
I_t^{00}(x) \cup I_t^{0\pm}(x) &\subset I^G(\bar{x}) \cup I^{GH}(\bar{x}), \\
I_t^{\pm 0}(x) \cup I_t^{00}(x) &\subset I^H(\bar{x}) \cup I^{GH}(\bar{x}).
\end{aligned} \tag{2}$$

In the lemma below, we present explicit formulas for the gradients of  $\Phi_l^s(\cdot; t)$  with  $l \in \mathcal{Q}$  and  $s \in \mathcal{S}$ . They can be derived exploiting Lemma 3.1 as well as the chain rule. !DS

**Lemma 3.2.** *For  $x \in \mathbb{R}^n$ ,  $t > 0$ , and  $l \in \mathcal{Q}$ , the following formulas are valid:*

$$\begin{aligned}
\nabla_x \Phi_l^1(x; t) &= \begin{cases} (H_l(x) - t)\nabla G_l(x) + (G_l(x) - t)\nabla H_l(x) & \text{if } G_l(x) + H_l(x) \geq 2t, \\ -(G_l(x) - t)\nabla G_l(x) - (H_l(x) - t)\nabla H_l(x) & \text{if } G_l(x) + H_l(x) < 2t, \end{cases} \\
\nabla_x \Phi_l^2(x; t) &= \begin{cases} (t - H_l(x))\nabla G_l(x) - (G_l(x) + t)\nabla H_l(x) & \text{if } -G_l(x) + H_l(x) \geq 2t, \\ -(G_l(x) + t)\nabla G_l(x) - (H_l(x) - t)\nabla H_l(x) & \text{if } -G_l(x) + H_l(x) < 2t, \end{cases} \\
\nabla_x \Phi_l^3(x; t) &= \begin{cases} (H_l(x) + t)\nabla G_l(x) + (G_l(x) + t)\nabla H_l(x) & \text{if } -G_l(x) - H_l(x) \geq 2t, \\ -(G_l(x) + t)\nabla G_l(x) - (H_l(x) + t)\nabla H_l(x) & \text{if } -G_l(x) - H_l(x) < 2t, \end{cases} \\
\nabla_x \Phi_l^4(x; t) &= \begin{cases} -(H_l(x) + t)\nabla G_l(x) + (t - G_l(x))\nabla H_l(x) & \text{if } G_l(x) - H_l(x) \geq 2t, \\ -(G_l(x) - t)\nabla G_l(x) - (H_l(x) + t)\nabla H_l(x) & \text{if } G_l(x) - H_l(x) < 2t. \end{cases}
\end{aligned}$$

Particularly, we have

$$\begin{aligned}
\forall l \in I_{t,1}^0(x): \quad \nabla_x \Phi_l^1(x; t) &= \begin{cases} (G_l(x) - t)\nabla H_l(x) & \text{if } l \in I_{t,1}^{+0}(x), \\ (H_l(x) - t)\nabla G_l(x) & \text{if } l \in I_{t,1}^{0+}(x), \\ 0 & \text{if } l \in I_{t,1}^{00}(x), \end{cases} \\
\forall l \in I_{t,2}^0(x): \quad \nabla_x \Phi_l^2(x; t) &= \begin{cases} -(G_l(x) + t)\nabla H_l(x) & \text{if } l \in I_{t,2}^{-0}(x), \\ (t - H_l(x))\nabla G_l(x) & \text{if } l \in I_{t,2}^{0+}(x), \\ 0 & \text{if } l \in I_{t,2}^{00}(x), \end{cases}
\end{aligned}$$

$$\begin{aligned} \forall l \in I_{t,3}^0(x): \quad \nabla_x \Phi_l^3(x; t) &= \begin{cases} (G_l(x) + t)\nabla H_l(x) & \text{if } l \in I_{t,3}^{-0}(x), \\ (H_l(x) + t)\nabla G_l(x) & \text{if } l \in I_{t,3}^{0-}(x), \\ 0 & \text{if } l \in I_{t,3}^{00}(x), \end{cases} \\ \forall l \in I_{t,4}^0(x): \quad \nabla_x \Phi_l^4(x; t) &= \begin{cases} (t - G_l(x))\nabla H_l(x) & \text{if } l \in I_{t,4}^{+0}(x), \\ -(H_l(x) + t)\nabla G_l(x) & \text{if } l \in I_{t,4}^{0-}(x), \\ 0 & \text{if } l \in I_{t,4}^{00}(x). \end{cases} \end{aligned}$$

As a corollary of the above lemma, we obtain an explicit formula for the linearization cone associated with  $(\mathbf{P}(t))$ .

**Corollary 3.3.** *Fix  $t > 0$  and a feasible point  $x \in X(t)$  of  $(\mathbf{P}(t))$ . Then the following formula is valid:*

$$\mathcal{L}_{X(t)}(x) = \left\{ d \in \mathbb{R}^n \left| \begin{array}{l} \nabla g_i(x) \cdot d \leq 0 \quad i \in I^g(x) \\ \nabla h_j(x) \cdot d \leq 0 \quad j \in \mathcal{P} \\ \nabla G_l(x) \cdot d \leq 0 \quad l \in I_{t,1}^{0+}(x) \cup I_{t,4}^{0-}(x) \\ \nabla G_l(x) \cdot d \geq 0 \quad l \in I_{t,2}^{0+}(x) \cup I_{t,3}^{0-}(x) \\ \nabla H_l(x) \cdot d \leq 0 \quad l \in I_{t,1}^{+0}(x) \cup I_{t,2}^{-0}(x) \\ \nabla H_l(x) \cdot d \geq 0 \quad l \in I_{t,3}^{-0}(x) \cup I_{t,4}^{+0}(x) \end{array} \right. \right\}.$$

The upcoming lemma justifies that  $(\mathbf{P}(t))$  is indeed a *relaxation* of  $(\mathbf{MPSC})$ . Its proof parallels the one of [Kanzow and Schwartz, 2013, Lemma 3.2].

**Lemma 3.4.** *The following statements hold:*

- (P1)  $X(0) = X$ ,
- (P2)  $0 \leq t_1 \leq t_2 \implies X(t_1) \subset X(t_2)$ ,
- (P3)  $\bigcap_{t>0} X(t) = X$ .

Now, we are in position to characterize a conceptual method for the numerical solution of  $(\mathbf{MPSC})$ : First, a sequence  $\{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+$  of positive relaxation parameters is chosen which converges to zero. Next, one solves the surrogate problem  $(\mathbf{P}(t_k))$  via standard methods. If one computes a (local) minimizer of one of these surrogate problems which is feasible to  $(\mathbf{MPSC})$ , then it is already a (local) minimizer of  $(\mathbf{MPSC})$ . In general, it will be only possible to compute KKT points of the surrogate problem. However, if such a sequence of KKT points converges to some point  $\bar{x} \in \mathbb{R}^n$ , then this point must be feasible to  $(\mathbf{MPSC})$  by construction. Furthermore, it will be shown in Section 3.2 that whenever MPSC-NNAMCQ is valid at  $\bar{x}$ , then it is an M-stationary point of  $(\mathbf{MPSC})$ .

### 3.2 Convergence properties

In this section, we analyze the theoretical properties of our relaxation scheme. In order to do so, we fix some standing assumptions below.

**Assumption 3.5.** Let  $\{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+$  be a sequence of positive relaxation parameters converging to zero. For each  $k \in \mathbb{N}$ , let  $x_k \in X(t_k)$  be a KKT point of  $(\mathbf{P}(t_k))$ . We assume that  $\{x_k\}_{k \in \mathbb{N}}$  converges to  $\bar{x} \in \mathbb{R}^n$ . Note that  $\bar{x} \in X$  holds by Lemma 3.4.

First, we will show that whenever MPSC-NNAMCQ is valid at  $\bar{x}$ , then it is an M-stationary point of  $(\mathbf{MPSC})$ . Second, it will be demonstrated that MPSC-LICQ at  $\bar{x}$  implies that GCQ holds for the surrogate problem  $(\mathbf{P}(t))$  at all feasible points from a sufficiently small neighborhood of  $\bar{x}$  and sufficiently small  $t > 0$ . This property ensures that local minima of the surrogate problem  $(\mathbf{P}(t))$  which are located near  $\bar{x}$  are in fact KKT points. This way, it is shown that Assumption 3.5 is reasonable.

**Theorem 3.6.** Let Assumption 3.5 be valid. Suppose that MPSC-NNAMCQ holds at  $\bar{x}$ . Then  $\bar{x}$  is an M-stationary point of  $(\mathbf{MPSC})$ .

*Proof.* Since  $x_k$  is a KKT point of  $(\mathbf{P}(t_k))$ , there exist multipliers  $\lambda^k \in \mathbb{R}^m$ ,  $\rho^k \in \mathbb{R}^p$ , and  $\alpha^k, \beta^k, \gamma^k, \delta^k \in \mathbb{R}^q$  which solve the following system:

$$0 = \nabla f(x_k) + \sum_{i \in I^g(x_k)} \lambda_i^k \nabla g_i(x_k) + \sum_{j \in \mathcal{P}} \rho_j^k \nabla h_j(x_k) \\ + \sum_{l \in \mathcal{Q}} \left[ \alpha_l^k \nabla_x \Phi_l^1(x_k; t_k) + \beta_l^k \nabla_x \Phi_l^2(x_k; t_k) + \gamma_l^k \nabla_x \Phi_l^3(x_k; t_k) + \delta_l^k \nabla_x \Phi_l^4(x_k; t_k) \right],$$

$$\forall i \in I^g(x_k): \lambda_i^k \geq 0; \forall i \in \mathcal{M} \setminus I^g(x_k): \lambda_i^k = 0,$$

$$\forall l \in I_{t_k,1}^0(x_k): \alpha_l^k \geq 0; \forall l \in \mathcal{Q} \setminus I_{t_k,1}^0(x_k): \alpha_l^k = 0,$$

$$\forall l \in I_{t_k,2}^0(x_k): \beta_l^k \geq 0; \forall l \in \mathcal{Q} \setminus I_{t_k,2}^0(x_k): \beta_l^k = 0,$$

$$\forall l \in I_{t_k,3}^0(x_k): \gamma_l^k \geq 0; \forall l \in \mathcal{Q} \setminus I_{t_k,3}^0(x_k): \gamma_l^k = 0,$$

$$\forall l \in I_{t_k,4}^0(x_k): \delta_l^k \geq 0; \forall l \in \mathcal{Q} \setminus I_{t_k,4}^0(x_k): \delta_l^k = 0.$$

Next, let us define new multipliers  $\alpha_G^k, \alpha_H^k, \beta_G^k, \beta_H^k, \gamma_G^k, \gamma_H^k, \delta_G^k, \delta_H^k \in \mathbb{R}^q$  as stated below for all  $l \in \mathcal{Q}$ :

$$\alpha_{G,l}^k := \begin{cases} \alpha_l^k (H_l(x_k) - t_k) & l \in I_{t_k,1}^{0+}(x_k), \\ 0 & \text{otherwise,} \end{cases} \quad \alpha_{H,l}^k := \begin{cases} \alpha_l^k (G_l(x_k) - t_k) & l \in I_{t_k,1}^{+0}(x_k), \\ 0 & \text{otherwise,} \end{cases}$$

$$\beta_{G,l}^k := \begin{cases} \beta_l^k (t_k - H_l(x_k)) & l \in I_{t_k,2}^{0+}(x_k), \\ 0 & \text{otherwise,} \end{cases} \quad \beta_{H,l}^k := \begin{cases} \beta_l^k (-G_l(x_k) - t_k) & l \in I_{t_k,2}^{-0}(x_k), \\ 0 & \text{otherwise,} \end{cases}$$

$$\gamma_{G,l}^k := \begin{cases} \gamma_l^k (H_l(x_k) + t_k) & l \in I_{t_k,3}^{0-}(x_k), \\ 0 & \text{otherwise,} \end{cases} \quad \gamma_{H,l}^k := \begin{cases} \gamma_l^k (G_l(x_k) + t_k) & l \in I_{t_k,3}^{-0}(x_k), \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_{G,l}^k := \begin{cases} \delta_l^k (-H_l(x_k) - t_k) & l \in I_{t_k,4}^{0-}(x_k), \\ 0 & \text{otherwise,} \end{cases} \quad \delta_{H,l}^k := \begin{cases} \delta_l^k (t_k - G_l(x_k)) & l \in I_{t_k,4}^{+0}(x_k), \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, we set  $\mu^k := \alpha_G^k + \beta_G^k + \gamma_G^k + \delta_G^k$  and  $\nu^k := \alpha_H^k + \beta_H^k + \gamma_H^k + \delta_H^k$ . By definition, we have  $\text{supp } \mu^k \subset I_{t_k}^{0\pm}(x_k)$  as well as  $\text{supp } \nu^k \subset I_{t_k}^{\pm 0}(x_k)$ . Thus, for sufficiently large

$k \in \mathbb{N}$ , (2) yields

$$\text{supp } \mu^k \subset I^G(\bar{x}) \cup I^{GH}(\bar{x}), \quad \text{supp } \nu^k \subset I^H(\bar{x}) \cup I^{GH}(\bar{x}).$$

Using Lemma 3.2,

$$0 = \nabla f(x_k) + \sum_{i \in I^g(x_k)} \lambda_i^k \nabla g_i(x_k) + \sum_{j \in \mathcal{P}} \rho_j^k \nabla h_j(x_k) + \sum_{l \in \mathcal{Q}} \left[ \mu_l^k \nabla G_l(x_k) + \nu_l^k \nabla H_l(x_k) \right] \quad (3)$$

is obtained.

Next, it will be shown that the sequence  $\{(\lambda^k, \rho^k, \mu^k, \nu^k)\}_{k \in \mathbb{N}}$  is bounded. Assuming the contrary, we define

$$\forall k \in \mathbb{N}: \quad (\tilde{\lambda}^k, \tilde{\rho}^k, \tilde{\mu}^k, \tilde{\nu}^k) := \frac{(\lambda^k, \rho^k, \mu^k, \nu^k)}{\|(\lambda^k, \rho^k, \mu^k, \nu^k)\|_2}.$$

Clearly,  $\{(\tilde{\lambda}^k, \tilde{\rho}^k, \tilde{\mu}^k, \tilde{\nu}^k)\}_{k \in \mathbb{N}}$  is bounded and, thus, possesses a converging subsequence (without relabeling) with nonvanishing limit  $(\tilde{\lambda}, \tilde{\rho}, \tilde{\mu}, \tilde{\nu})$ . The continuity of  $g$  yields  $\text{supp } \tilde{\lambda} \subset I^g(\bar{x})$ . The above considerations yield

$$\text{supp } \tilde{\mu} \subset I^G(\bar{x}) \cup I^{GH}(\bar{x}), \quad \text{supp } \tilde{\nu} \subset I^H(\bar{x}) \cup I^{GH}(\bar{x}).$$

Choose  $l \in I^{GH}(\bar{x})$  arbitrarily. If  $\tilde{\mu}_l \neq 0$  holds true, then  $\tilde{\mu}_l^k \neq 0$  must be valid for sufficiently large  $k \in \mathbb{N}$ . By definition of  $\mu_l^k$ ,  $l \in I_{t_k}^{0\pm}(x_k)$  follows. Since  $I_{t_k}^{0\pm}(x_k)$  and  $I_{t_k}^{\pm 0}(x_k)$  are disjoint,  $\nu_l^k = 0$  holds for sufficiently large  $k \in \mathbb{N}$ . This yields  $\tilde{\nu}_l^k = 0$  for sufficiently large  $k \in \mathbb{N}$ , i.e.  $\tilde{\nu}_l = 0$  is obtained. This shows

$$\forall l \in I^{GH}(\bar{x}): \quad \tilde{\mu}_l \tilde{\nu}_l = 0.$$

Dividing (3) by  $\|(\lambda^k, \rho^k, \mu^k, \nu^k)\|_2$ , taking the limit  $k \rightarrow \infty$ , respecting the continuous differentiability of  $f, g, h, G$ , as well as  $H$ , and invoking the above arguments, we obtain

$$\begin{aligned} 0 &= \sum_{i \in I^g(\bar{x})} \tilde{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{P}} \tilde{\rho}_j \nabla h_j(\bar{x}) + \sum_{l \in \mathcal{Q}} \left[ \tilde{\mu}_l \nabla G_l(\bar{x}) + \tilde{\nu}_l \nabla H_l(\bar{x}) \right], \\ \forall i \in I^g(\bar{x}): \quad &\tilde{\lambda}_i \geq 0; \quad \forall i \in \mathcal{M} \setminus I^g(\bar{x}): \quad \tilde{\lambda}_i = 0, \\ \forall l \in I^H(\bar{x}): \quad &\tilde{\mu}_l = 0, \\ \forall l \in I^G(\bar{x}): \quad &\tilde{\nu}_l = 0, \\ \forall l \in I^{GH}(\bar{x}): \quad &\tilde{\mu}_l \tilde{\nu}_l = 0. \end{aligned}$$

Due to the fact that  $(\tilde{\lambda}, \tilde{\rho}, \tilde{\mu}, \tilde{\nu})$  does not vanish, this is a contradiction to the validity of MPSC-NNAMCQ. Thus,  $\{(\lambda^k, \rho^k, \mu^k, \nu^k)\}_{k \in \mathbb{N}}$  is bounded.

We assume w.l.o.g. that  $\{(\lambda^k, \rho^k, \mu^k, \nu^k)\}_{k \in \mathbb{N}}$  converges to  $(\lambda, \rho, \mu, \nu)$  (otherwise, we choose an appropriate subsequence). Reprising the above arguments, we have

$$\begin{aligned} \forall i \in I^g(\bar{x}): \quad &\lambda_i \geq 0; \quad \forall i \in \mathcal{M} \setminus I^g(\bar{x}): \quad \lambda_i = 0, \\ \forall l \in I^H(\bar{x}): \quad &\mu_l = 0, \\ \forall l \in I^G(\bar{x}): \quad &\nu_l = 0, \\ \forall l \in I^{GH}(\bar{x}): \quad &\mu_l \nu_l = 0. \end{aligned}$$

Taking the limit in (3) and respecting the continuous differentiability of all appearing mappings, we obtain

$$0 = \nabla f(\bar{x}) + \sum_{i \in I^g(\bar{x})} \lambda_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(\bar{x}) + \sum_{l \in \mathcal{Q}} [\mu_l \nabla G_l(\bar{x}) + \nu_l \nabla H_l(\bar{x})].$$

This shows that  $\bar{x}$  is an M-stationary point of (MPSC).  $\square$

In order to show that the validity of MPSC-LICQ at the limit point  $\bar{x}$  ensures that GCQ holds at all feasible points of (P(t)) sufficiently close to  $\bar{x}$  where  $t$  is sufficiently small, we need to study the variational geometry of the sets  $X(t)$  in some more detail.

Fix  $t > 0$  and some point  $\tilde{x} \in X(t)$ . For an arbitrary index set  $I \subset I_t^{00}(\tilde{x})$ , we consider the subsequent program:

$$\begin{aligned} f(x) &\rightarrow \min \\ g_i(x) &\leq 0, & i \in \mathcal{M}, \\ h_j(x) &= 0, & j \in \mathcal{P}, \\ -t \leq G_l(x) \leq t, & & l \in I_t^{0\pm}(\tilde{x}) \cup I, \\ -t \leq H_l(x) \leq t, & & l \in I_t^{\pm 0}(\tilde{x}) \cup (I_t^{00}(\tilde{x}) \setminus I), \\ \Phi_l^s(x; t) &\leq 0, & l \in \mathcal{Q} \setminus I_{t,s}^0(\tilde{x}), s \in \mathcal{S}. \end{aligned} \tag{P}(t, \tilde{x}, I) \quad \text{!DS}$$

The feasible set of (P(t,  $\tilde{x}$ , I)) will be denoted by  $X(t, \tilde{x}, I)$ . Clearly,  $\tilde{x}$  is a feasible point of (P(t,  $\tilde{x}$ , I)) for arbitrary  $I \subset I_t^{00}(\tilde{x})$ . Furthermore,  $X(t, \tilde{x}, I) \subset X(t)$  is valid for any choice of  $I \subset I_t^{00}(\tilde{x})$ .  $\text{!DS}$

**Lemma 3.7.** For fixed  $t > 0$  and  $\tilde{x} \in X(t)$ , we have

$$\mathcal{T}_{X(t)}(\tilde{x}) = \bigcup_{I \subset I_t^{00}(\tilde{x})} \mathcal{T}_{X(t, \tilde{x}, I)}(\tilde{x}).$$

*Proof.* We show both inclusions separately.

“ $\subset$ ” Fix an arbitrary direction  $d \in \mathcal{T}_{X(t)}(\tilde{x})$ . Then we find sequences  $\{y_k\}_{k \in \mathbb{N}} \subset X(t)$  and  $\{\tau_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+$  such that  $y_k \rightarrow \tilde{x}$ ,  $\tau_k \downarrow 0$ , and  $(y_k - \tilde{x})/\tau_k \rightarrow d$  as  $k \rightarrow \infty$ . It is sufficient to verify the existence of an index set  $\bar{I} \subset I_t^{00}(\tilde{x})$  such that  $\{y_k\}_{k \in \mathbb{N}} \cap X(t, \tilde{x}, \bar{I})$  possesses infinite cardinality since this already gives us  $d \in \mathcal{T}_{X(t, \tilde{x}, \bar{I})}(\tilde{x})$ .  $\text{!DS}$

Fix  $k \in \mathbb{N}$  sufficiently large and  $l \in I_t^{0\pm}(\tilde{x})$ . Then, due to continuity of  $G_l$ ,  $H_l$ , as well as  $\varphi$  and feasibility of  $y_k$  to (P(t)), we either have  $l \in I_t^{0\pm}(y_k)$  and, thus,  $G_l(y_k) \in \{-t, t\}$ , or  $-t < G_l(y_k) < t$ . Similarly, we obtain  $-t \leq H_l(y_k) \leq t$  for all  $l \in I_t^{\pm 0}(\tilde{x})$ . Due to feasibility of  $y_k$  to (P(t)), we have  $-t \leq G_l(y_k) \leq t$  or  $-t \leq H_l(y_k) \leq t$ . Thus, setting

$$I_k := I_t^{00}(\tilde{x}) \cap \{l \in \mathcal{Q} \mid -t \leq G_l(y_k) \leq t\},$$

$y_k \in X(t, \tilde{x}, I_k)$  is valid. Since there are only finitely many subsets of  $I_t^{00}(\tilde{x})$  while  $\{y_k\}_{k \in \mathbb{N}}$  is infinite, there must exist  $\bar{I} \subset I_t^{00}(\tilde{x})$  such that  $\{y_k\}_{k \in \mathbb{N}} \cap X(t, \tilde{x}, \bar{I})$  is of infinite cardinality.

“ $\supset$ ” By definition of the tangent cone, we easily obtain  $\mathcal{T}_{X(t, \tilde{x}, I)}(\tilde{x}) \subset \mathcal{T}_{X(t)}(\tilde{x})$  for any  $I \subset I_t^{00}(\tilde{x})$ . Taking the union over all subsets of  $I_t^{00}(\tilde{x})$  yields the desired inclusion.  $\square$

**Theorem 3.8.** *Let  $\bar{x} \in X$  be a feasible point of (MPSC) where MPSC-LICQ is satisfied. Then there exist  $\bar{t} > 0$  and a neighborhood  $U \subset \mathbb{R}^n$  of  $\bar{x}$  such that GCQ holds for (P(t)) at all points from  $X(t) \cap U$  for all  $t \in (0, \bar{t}]$ .*

*Proof.* Due to the validity of MPSC-LICQ at  $\bar{x}$  and the continuous differentiability of  $g$ ,  $h$ ,  $G$ , and  $H$ , the gradients

$$\begin{aligned} & \{\nabla g_i(x) \mid i \in I^g(\bar{x})\} \cup \{\nabla h_j(x) \mid j \in \mathcal{P}\} \\ & \cup \{\nabla G_l(x) \mid l \in I^G(\bar{x}) \cup I^{GH}(\bar{x})\} \\ & \cup \{\nabla H_l(x) \mid l \in I^H(\bar{x}) \cup I^{GH}(\bar{x})\} \end{aligned}$$

are linearly independent for all  $x$  which are chosen from a sufficiently small neighborhood  $\mathcal{V}$  of  $\bar{x}$ , see Lemma 2.2. Invoking (2), we can choose a neighborhood  $U \subset \mathcal{V}$  of  $\bar{x}$  and  $\bar{t} > 0$  such that for any  $\tilde{x} \in X(t) \cap U$ , where  $t \in (0, \bar{t}]$  holds, we have

$$\begin{aligned} I^g(\tilde{x}) & \subset I^g(\bar{x}), \\ I_t^{00}(\tilde{x}) \cup I_t^{0\pm}(\tilde{x}) & \subset I^G(\bar{x}) \cup I^{GH}(\bar{x}), \\ I_t^{00}(\tilde{x}) \cup I_t^{\pm 0}(\tilde{x}) & \subset I^H(\bar{x}) \cup I^{GH}(\bar{x}). \end{aligned} \tag{4}$$

Particularly, for any such  $\tilde{x} \in \mathbb{R}^n$  and  $I \subset I_t^{00}(\tilde{x})$ , the gradients

$$\begin{aligned} & \{\nabla g_i(\tilde{x}) \mid i \in I^g(\tilde{x})\} \cup \{\nabla h_j(\tilde{x}) \mid j \in \mathcal{P}\} \\ & \cup \{\nabla G_l(\tilde{x}) \mid l \in I_t^{0\pm}(\tilde{x}) \cup I\} \\ & \cup \{\nabla H_l(\tilde{x}) \mid l \in I_t^{\pm 0}(\tilde{x}) \cup (I_t^{00}(\tilde{x}) \setminus I)\} \end{aligned}$$

are linearly independent, i.e., standard LICQ is valid for (P(t,  $\tilde{x}$ , I)) at  $\tilde{x}$  for any set  $I \subset I_t^{00}(\tilde{x})$ . This implies  $\mathcal{T}_{X(t, \tilde{x}, I)}(\tilde{x}) = \mathcal{L}_{X(t, \tilde{x}, I)}(\tilde{x})$  for any  $I \subset I_t^{00}(\tilde{x})$ . Exploiting Lemma 3.7, we obtain

$$\mathcal{T}_{X(t)}(\tilde{x}) = \bigcup_{I \subset I_t^{00}(\tilde{x})} \mathcal{L}_{X(t, \tilde{x}, I)}(\tilde{x}).$$

Computing the polar cone on both sides yields

$$\mathcal{T}_{X(t)}(\tilde{x})^\circ = \bigcap_{I \subset I_t^{00}(\tilde{x})} \mathcal{L}_{X(t, \tilde{x}, I)}(\tilde{x})^\circ. \tag{5}$$

Define

$$\begin{aligned} \mathcal{I}_G^+(t, \tilde{x}, I) & := I_{t,1}^{0+}(\tilde{x}) \cup I_{t,4}^{0-}(\tilde{x}) \cup [I \cap (I_{t,1}^{00}(\tilde{x}) \cup I_{t,4}^{00}(\tilde{x}))], \\ \mathcal{I}_G^-(t, \tilde{x}, I) & := I_{t,2}^{0+}(\tilde{x}) \cup I_{t,3}^{0-}(\tilde{x}) \cup [I \cap (I_{t,2}^{00}(\tilde{x}) \cup I_{t,3}^{00}(\tilde{x}))], \\ \mathcal{I}_H^+(t, \tilde{x}, I) & := I_{t,1}^{+0}(\tilde{x}) \cup I_{t,2}^{-0}(\tilde{x}) \cup [(I_t^{00}(\tilde{x}) \setminus I) \cap (I_{t,1}^{00}(\tilde{x}) \cup I_{t,2}^{00}(\tilde{x}))], \\ \mathcal{I}_H^-(t, \tilde{x}, I) & := I_{t,3}^{-0}(\tilde{x}) \cup I_{t,4}^{+0}(\tilde{x}) \cup [(I_t^{00}(\tilde{x}) \setminus I) \cap (I_{t,3}^{00}(\tilde{x}) \cup I_{t,4}^{00}(\tilde{x}))], \end{aligned}$$

and observe that these sets characterize the indices  $l \in \mathcal{Q}$  where the constraints  $G_l(\tilde{x}) \leq t$ ,  $G_l(\tilde{x}) \geq -t$ ,  $H_l(\tilde{x}) \leq t$ , and  $H_l(\tilde{x}) \geq -t$ , respectively, are active in  $(\mathbf{P}(t, \tilde{x}, I))$ . We therefore obtain

$$\mathcal{L}_{X(t, \tilde{x}, I)}(\tilde{x}) = \left\{ d \in \mathbb{R}^n \left| \begin{array}{l} \nabla g_i(\tilde{x}) \cdot d \leq 0 \quad i \in I^g(\tilde{x}) \\ \nabla h_j(\tilde{x}) \cdot d = 0 \quad j \in \mathcal{P} \\ \nabla G_l(\tilde{x}) \cdot d \leq 0 \quad l \in \mathcal{I}_G^+(t, \tilde{x}, I) \\ \nabla G_l(\tilde{x}) \cdot d \geq 0 \quad l \in \mathcal{I}_G^-(t, \tilde{x}, I) \\ \nabla H_l(\tilde{x}) \cdot d \leq 0 \quad l \in \mathcal{I}_H^+(t, \tilde{x}, I) \\ \nabla H_l(\tilde{x}) \cdot d \geq 0 \quad l \in \mathcal{I}_H^-(t, \tilde{x}, I) \end{array} \right. \right\}.$$

Exploiting Lemma 2.1, the polar of this cone is easily computed:

$$\mathcal{L}_{X(t, \tilde{x}, I)}(\tilde{x})^\circ = \left\{ \eta \in \mathbb{R}^n \left| \begin{array}{l} \eta = \sum_{i \in I^g(\tilde{x})} \lambda_i \nabla g_i(\tilde{x}) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(\tilde{x}) \\ \quad + \sum_{l \in \mathcal{Q}} [\mu_l \nabla G_l(\tilde{x}) + \nu_l \nabla H_l(\tilde{x})] \\ \forall i \in I^g(\tilde{x}): \lambda_i \geq 0 \\ \forall l \in \mathcal{Q}: \mu_l \begin{cases} \geq 0 & \text{if } l \in \mathcal{I}_G^+(t, \tilde{x}, I), \\ \leq 0 & \text{if } l \in \mathcal{I}_G^-(t, \tilde{x}, I), \\ = 0 & \text{otherwise,} \end{cases} \\ \forall l \in \mathcal{Q}: \nu_l \begin{cases} \geq 0 & \text{if } l \in \mathcal{I}_H^+(t, \tilde{x}, I), \\ \leq 0 & \text{if } l \in \mathcal{I}_H^-(t, \tilde{x}, I), \\ = 0 & \text{otherwise} \end{cases} \end{array} \right. \right\}. \quad (6)$$

Observing that  $(\mathbf{P}(t))$  is a standard nonlinear problem,  $\mathcal{T}_{X(t)}(\tilde{x}) \subset \mathcal{L}_{X(t)}(\tilde{x})$  and, thus,  $\mathcal{L}_{X(t)}(\tilde{x})^\circ \subset \mathcal{T}_{X(t)}(\tilde{x})^\circ$  are inherent. It remains to show  $\mathcal{T}_{X(t)}(\tilde{x})^\circ \subset \mathcal{L}_{X(t)}(\tilde{x})^\circ$ . Thus, choose  $\eta \in \mathcal{T}_{X(t)}(\tilde{x})^\circ$  arbitrarily. Then, in particular, (5) yields

$$\eta \in \mathcal{L}_{X(t, \tilde{x}, \emptyset)}(\tilde{x})^\circ \cap \mathcal{L}_{X(t, \tilde{x}, I_t^{00}(\tilde{x}))}(\tilde{x})^\circ.$$

Exploiting the representation (6), we find  $\lambda_i, \lambda'_i \geq 0$  ( $i \in I^g(\tilde{x})$ ),  $\rho, \rho' \in \mathbb{R}^p$ ,  $\mu, \mu' \in \mathbb{R}^q$ , as well as  $\nu, \nu' \in \mathbb{R}^q$  which satisfy

$$\begin{aligned} \eta &= \sum_{i \in I^g(\tilde{x})} \lambda_i \nabla g_i(\tilde{x}) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(\tilde{x}) + \sum_{l \in \mathcal{Q}} [\mu_l \nabla G_l(\tilde{x}) + \nu_l \nabla H_l(\tilde{x})] \\ &= \sum_{i \in I^g(\tilde{x})} \lambda'_i \nabla g_i(\tilde{x}) + \sum_{j \in \mathcal{P}} \rho'_j \nabla h_j(\tilde{x}) + \sum_{l \in \mathcal{Q}} [\mu'_l \nabla G_l(\tilde{x}) + \nu'_l \nabla H_l(\tilde{x})] \end{aligned}$$

and

$$\mu_l \begin{cases} \geq 0 & \text{if } l \in I_{t,1}^{0+}(\tilde{x}) \cup I_{t,4}^{0-}(\tilde{x}), \\ \leq 0 & \text{if } l \in I_{t,2}^{0+}(\tilde{x}) \cup I_{t,3}^{0-}(\tilde{x}), \\ = 0 & \text{otherwise,} \end{cases}$$



$$\begin{aligned}
\mu'_l & \begin{cases} \geq 0 & \text{if } l \in I_{t,1}^{0+}(\tilde{x}) \cup I_{t,4}^{0-}(\tilde{x}) \cup I_{t,1}^{00}(\tilde{x}) \cup I_{t,4}^{00}(\tilde{x}), \\ \leq 0 & \text{if } l \in I_{t,2}^{0+}(\tilde{x}) \cup I_{t,3}^{0-}(\tilde{x}) \cup I_{t,2}^{00}(\tilde{x}) \cup I_{t,3}^{00}(\tilde{x}), \\ = 0 & \text{otherwise,} \end{cases} \\
\nu_l & \begin{cases} \geq 0 & \text{if } l \in I_{t,1}^{+0}(\tilde{x}) \cup I_{t,2}^{-0}(\tilde{x}) \cup I_{t,1}^{00}(\tilde{x}) \cup I_{t,2}^{00}(\tilde{x}), \\ \leq 0 & \text{if } l \in I_{t,3}^{-0}(\tilde{x}) \cup I_{t,4}^{+0}(\tilde{x}) \cup I_{t,3}^{00}(\tilde{x}) \cup I_{t,4}^{00}(\tilde{x}), \\ = 0 & \text{otherwise,} \end{cases} \\
\nu'_l & \begin{cases} \geq 0 & \text{if } l \in I_{t,1}^{+0}(\tilde{x}) \cup I_{t,2}^{-0}(\tilde{x}), \\ \leq 0 & \text{if } l \in I_{t,3}^{-0}(\tilde{x}) \cup I_{t,4}^{+0}(\tilde{x}), \\ = 0 & \text{otherwise} \end{cases}
\end{aligned}$$

for all  $l \in \mathcal{Q}$ . Thus, we obtain

$$\begin{aligned}
0 &= \sum_{i \in I^g(\tilde{x})} (\lambda_i - \lambda'_i) \nabla g_i(\tilde{x}) + \sum_{j \in \mathcal{P}} (\rho_j - \rho'_j) \nabla h_j(\tilde{x}) \\
&\quad + \sum_{l \in \mathcal{Q}} [(\mu_l - \mu'_l) \nabla G_l(\tilde{x}) + (\nu_l - \nu'_l) \nabla H_l(\tilde{x})].
\end{aligned}$$

Observing  $\text{supp}(\mu - \mu') \subset I_t^{00}(\tilde{x}) \cup I_t^{0\pm}(\tilde{x})$  as well as  $\text{supp}(\nu - \nu') \subset I_t^{00}(\tilde{x}) \cup I_t^{\pm 0}(\tilde{x})$  and using (4), we obtain  $\lambda_i = \lambda'_i$  ( $i \in I^g(\tilde{x})$ ),  $\rho = \rho'$ ,  $\mu = \mu'$ , as well as  $\nu = \nu'$ . Particularly, !DS

$$\eta = \sum_{i \in I^g(\tilde{x})} \lambda_i \nabla g_i(\tilde{x}) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(\tilde{x}) + \sum_{l \in \mathcal{Q}} [\mu_l \nabla G_l(\tilde{x}) + \nu'_l \nabla H_l(\tilde{x})].$$

is obtained where  $\text{supp } \mu \subset I_t^{0\pm}(\tilde{x})$  and  $\text{supp } \nu' \subset I_t^{\pm 0}(\tilde{x})$  hold true. Finally, we exploit Corollary 3.3 in order to see  $\eta \in \mathcal{L}_{X(t)}(\tilde{x})^\circ$ . This shows the validity of the inclusion  $\mathcal{T}_{X(t)}(\tilde{x})^\circ \subset \mathcal{L}_{X(t)}(\tilde{x})^\circ$  and, thereby, GCQ holds true for (P(t)) at  $\tilde{x}$ . Since  $t \in (0, \bar{t}]$  and  $\tilde{x} \in X(t) \cap U$  were arbitrarily chosen, the proof is completed.  $\square$

## 4 Remarks on other possible relaxation schemes

In this section, we discuss three more relaxation approaches for the numerical treatment of (MPSC) which are inspired by the rich theory on MPCCs. Particularly, the relaxation schemes of Scholtes [2001], Steffensen and Ulbrich [2010], as well as Kadrani et al. [2009], are adapted to the setting of switching-constrained optimization. !DS

### 4.1 The relaxation scheme of Scholtes

For some parameter  $t \geq 0$ , let us consider the surrogate problem

$$\begin{aligned}
f(x) &\rightarrow \min \\
g_i(x) &\leq 0, & i \in \mathcal{M}, \\
h_j(x) &= 0, & j \in \mathcal{P}, \\
-t &\leq G_l(x)H_l(x) \leq t, & l \in \mathcal{Q}.
\end{aligned} \tag{Ps(t)}$$

This idea is inspired by Scholtes' global relaxation method which was designed for the computational treatment of MPCCs, see Scholtes [2001] and [Hoheisel et al., 2013, Section 3.1]. The feasible set of  $(\mathbf{P}_S(t))$  is denoted by  $X_S(t)$  and visualized in Figure 3. Note that the family  $\{X_S(t)\}_{t \geq 0}$  possesses the same properties as the family  $\{X(t)\}_{t \geq 0}$  described in Lemma 3.4. Thus, Scholtes' relaxation is reasonable for switching-constrained problems as well. In contrast to  $(\mathbf{P}(t))$ , where we need four inequality constraints in order to replace one original switching constraint, one only needs two inequality constraints in  $(\mathbf{P}_S(t))$  for the same purpose. This is a significant advantage of  $(\mathbf{P}_S(t))$  over the surrogate  $(\mathbf{P}(t))$ .

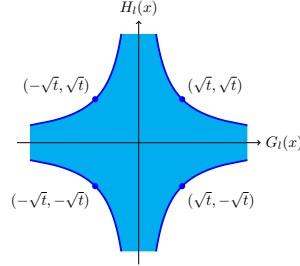


Figure 3: Geometric interpretation of the relaxed feasible set  $X_S(t)$ .

It is well known from [Hoheisel et al., 2013, Theorem 3.1] that Scholtes' relaxation approach finds Clarke-stationary points of MPCCs under an MPCC-tailored version of MFCQ. Note that, in the context of MPCCs, Clarke-stationarity is stronger than weak stationarity but weaker than Mordukhovich-stationarity. Below, we want to generalize the result from Hoheisel et al. [2013] to the problem  $(\mathbf{MPSC})$ .

For the fixed parameter  $t > 0$  and a feasible point  $x \in X_S(t)$  of  $(\mathbf{P}_S(t))$ , we introduce the index sets

$$\begin{aligned} I_t^+(x) &:= \{l \in \mathcal{Q} \mid G_l(x)H_l(x) = t\}, \\ I_t^-(x) &:= \{l \in \mathcal{Q} \mid G_l(x)H_l(x) = -t\}. \end{aligned}$$

In the upcoming theorem, we provide a convergence result of Scholtes' relaxation scheme for the problem  $(\mathbf{MPSC})$ .

**Theorem 4.1.** *Let  $\{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+$  be a sequence of positive relaxation parameters converging to zero. For each  $k \in \mathbb{N}$ , let  $x_k \in X_S(t_k)$  be a KKT point of  $(\mathbf{P}_S(t_k))$ . Assume that the sequence  $\{x_k\}_{k \in \mathbb{N}}$  converges to a point  $\bar{x} \in X$  where MPSC-MFCQ holds. Then  $\bar{x}$  is a W-stationary point of  $(\mathbf{MPSC})$ .*

*Proof.* Noting that  $x_k$  is a KKT point of  $(\mathbf{P}_S(t_k))$ , we find multipliers  $\lambda^k \in \mathbb{R}^m$ ,  $\rho^k \in \mathbb{R}^p$ , and  $\xi^k \in \mathbb{R}^q$  which satisfy the following conditions:

$$\begin{aligned} 0 = \nabla f(x_k) &+ \sum_{i \in I^g(x_k)} \lambda_i^k \nabla g_i(x_k) + \sum_{j \in \mathcal{P}} \rho_j^k \nabla h_j(x_k) \\ &+ \sum_{l \in \mathcal{Q}} \xi_l^k \left[ H_l(x_k) \nabla G_l(x_k) + G_l(x_k) \nabla H_l(x_k) \right], \end{aligned}$$

$$\begin{aligned}
& \forall i \in I^g(x_k): \lambda_i^k \geq 0, \forall i \in \mathcal{M} \setminus I^g(x_k): \lambda_i^k = 0, \\
& \forall l \in I_t^+(x_k): \xi_l^k \geq 0, \\
& \forall l \in I_t^-(x_k): \xi_l^k \leq 0, \\
& \forall l \in \mathcal{Q} \setminus (I_t^+(x_k) \cup I_t^-(x_k)): \xi_l^k = 0.
\end{aligned}$$

For any  $k \in \mathbb{N}$  and  $l \in \mathcal{Q}$ , let us define artificial multipliers  $\mu_l^k, \nu_l^k \in \mathbb{R}$  as stated below: !DS

$$\mu_l^k := \begin{cases} \xi_l^k H_l(x_k) & l \in I^G(\bar{x}) \cup I^{GH}(\bar{x}), \\ 0 & l \in I^H(\bar{x}), \end{cases} \quad \nu_l^k := \begin{cases} \xi_l^k G_l(x_k) & l \in I^H(\bar{x}) \cup I^{GH}(\bar{x}), \\ 0 & l \in I^G(\bar{x}). \end{cases}$$

Thus, we obtain

$$\begin{aligned}
0 = & \nabla f(x_k) + \sum_{i \in I^g(x_k)} \lambda_i^k \nabla g_i(x_k) + \sum_{j \in \mathcal{P}} \rho_j^k \nabla h_j(x_k) + \sum_{l \in \mathcal{Q}} \left[ \mu_l^k \nabla G_l(x_k) + \nu_l^k \nabla H_l(x_k) \right] \\
& + \sum_{l \in I^H(\bar{x})} \xi_l^k H_l(x_k) \nabla G_l(x_k) + \sum_{l \in I^G(\bar{x})} \xi_l^k G_l(x_k) \nabla H_l(x_k). \tag{7}
\end{aligned}$$

Next, we are going to show that the sequence  $\{(\lambda^k, \rho^k, \mu^k, \nu^k, \xi_I^k)\}_{k \in \mathbb{N}}$  is bounded where we used  $I := I^G(\bar{x}) \cup I^H(\bar{x})$  for brevity. We assume on the contrary that this is not the case and define

$$\forall k \in \mathbb{N}: \quad (\tilde{\lambda}^k, \tilde{\rho}^k, \tilde{\mu}^k, \tilde{\nu}^k, \tilde{\xi}_I^k) := \frac{(\lambda^k, \rho^k, \mu^k, \nu^k, \xi_I^k)}{\|(\lambda^k, \rho^k, \mu^k, \nu^k, \xi_I^k)\|_2}.$$

Clearly,  $\{(\tilde{\lambda}^k, \tilde{\rho}^k, \tilde{\mu}^k, \tilde{\nu}^k, \tilde{\xi}_I^k)\}_{k \in \mathbb{N}}$  is bounded and, thus, converges w.l.o.g. to a nonvanishing vector  $(\tilde{\lambda}, \tilde{\rho}, \tilde{\mu}, \tilde{\nu}, \tilde{\xi}_I)$  (otherwise, a suitable subsequence is chosen). The continuity of  $g$  ensures that  $I^g(x_k) \subset I^g(\bar{x})$  is valid for sufficiently large  $k \in \mathbb{N}$ . Dividing (7) by  $\|(\lambda^k, \rho^k, \mu^k, \nu^k, \xi_I^k)\|_2$  and taking the limit  $k \rightarrow \infty$  while respecting the continuous differentiability of all involved functions, we come up with

$$\begin{aligned}
0 = & \sum_{i \in I^g(\bar{x})} \tilde{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{P}} \tilde{\rho}_j \nabla h_j(\bar{x}) + \sum_{l \in \mathcal{Q}} \left[ \tilde{\mu}_l \nabla G_l(\bar{x}) + \tilde{\nu}_l \nabla H_l(\bar{x}) \right], \\
& \forall i \in I^g(\bar{x}): \tilde{\lambda}_i \geq 0, \forall i \in \mathcal{M} \setminus I^g(\bar{x}): \tilde{\lambda}_i = 0, \\
& \forall l \in I^H(\bar{x}): \tilde{\mu}_l = 0, \\
& \forall l \in I^G(\bar{x}): \tilde{\nu}_l = 0.
\end{aligned}$$

Now, the validity of MPSC-MFCQ yields  $\tilde{\lambda} = 0, \tilde{\rho} = 0, \tilde{\mu} = 0,$  and  $\tilde{\nu} = 0$ . Hence,  $\tilde{\xi}_{l_0} \neq 0$  !DS  
for at least one index  $l_0 \in I$ . Let us assume  $l_0 \in I^H(\bar{x})$ . Then we have  $\nu_{l_0}^k = \xi_{l_0}^k G_{l_0}(x_k)$ , which leads to

$$\tilde{\nu}_{l_0} = \lim_{k \rightarrow \infty} \frac{\nu_{l_0}^k}{\|(\lambda^k, \rho^k, \mu^k, \nu^k, \xi_I^k)\|_2} = \lim_{k \rightarrow \infty} \frac{\xi_{l_0}^k G_{l_0}(x_k)}{\|(\lambda^k, \rho^k, \mu^k, \nu^k, \xi_I^k)\|_2} = \tilde{\xi}_{l_0} G_{l_0}(\bar{x}) \neq 0.$$

This, however, is a contradiction since  $\tilde{\nu}$  vanishes due to the above arguments. Similarly, the case  $l_0 \in I^G(\bar{x})$  leads to a contradiction. As a consequence, the sequence  $\{(\lambda^k, \rho^k, \mu^k, \nu^k, \xi_I^k)\}_{k \in \mathbb{N}}$  is bounded.

Thus, we may assume w.l.o.g. that this sequence converges to  $(\lambda, \rho, \mu, \nu, \xi_I)$ . Again, we take the limit in (7) and obtain

$$\begin{aligned} 0 &= \nabla f(\bar{x}) + \sum_{i \in I^g(\bar{x})} \lambda_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(\bar{x}) + \sum_{l \in \mathcal{Q}} \left[ \mu_l \nabla G_l(\bar{x}) + \nu_l \nabla H_l(\bar{x}) \right], \\ \forall i \in I^g(\bar{x}): \lambda_i &\geq 0, \forall i \in \mathcal{M} \setminus I^g(\bar{x}): \lambda_i = 0, \\ \forall l \in I^H(\bar{x}): \mu_l &= 0, \\ \forall l \in I^G(\bar{x}): \nu_l &= 0 \end{aligned}$$

which shows that  $\bar{x}$  is a W-stationary point of (MPSC).  $\square$

Noting that no suitable definition of Clarke-stationarity (in particular, a *reasonable* stationarity concept which is stronger than W- but weaker than M-stationarity) seems to be available for (MPSC), Theorem 4.1 does not seem to be too surprising at all. The following example shows that we cannot expect any stronger results in general. Thus, the qualitative properties of Scholtes' relaxation method are substantially weaker than those of the relaxation scheme proposed in Section 3.1.

**Example 4.2.** *Let us consider the switching-constrained optimization problem*

$$\begin{aligned} \frac{1}{2}(x_1 - 1)^2 + \frac{1}{2}(x_2 - 1)^2 &\rightarrow \min \\ x_1 x_2 &= 0. \end{aligned}$$

*The globally optimal solutions of this program are given by  $(1, 0)$  as well as  $(0, 1)$ , and these points are S-stationary. Furthermore, there exists a W-stationary point at  $\bar{x} = (0, 0)$  which is no local minimizer.*

*One can easily check that the associated problem  $(P_S(t))$  possesses a KKT point at  $(\sqrt{t}, \sqrt{t})$  for any  $t \in (0, 1]$ . Taking the limit  $t \downarrow 0$ , this point tends to  $\bar{x}$  which is, as we already mentioned above, only W-stationary for the switching-constrained problem of interest. Clearly, MPSC-LICQ is valid at  $\bar{x}$ .*

IDS

Although the theoretical properties of Scholtes' relaxation approach do not seem to be promising in light of (MPSC), we check the applicability of the approach. More precisely, we analyze the restrictiveness of the assumption that a sequence of KKT points associated with  $(P_S(t))$  can be chosen.

In order to guarantee that locally optimal solutions of the nonlinear relaxed surrogate problems  $(P_S(t))$  which are located closely to the limit point from Theorem 4.1 are KKT points, a constraint qualification needs to be satisfied. Adapting [Hoheisel et al., 2013, Theorem 3.2] to the switching-constrained situation, it is possible to show that whenever MPSC-MFCQ is valid at a feasible point  $\bar{x} \in X$  of (MPSC), then standard MFCQ is valid for  $(P_S(t))$  in a neighborhood of  $\bar{x}$ .

**Theorem 4.3.** Let  $\bar{x} \in X$  be a feasible point of (MPSC) where MPSC-MFCQ is satisfied. Then there exists a neighborhood  $U \subset \mathbb{R}^n$  of  $\bar{x}$  such that MFCQ holds for  $(P_S(t))$  at all points from  $X_S(t) \cap U$  for all  $t > 0$ .

*Proof.* Due to the validity of MPSC-MFCQ at  $\bar{x}$ , the union

$$\begin{aligned} & \{\nabla g_i(\bar{x}) \mid i \in I^g(\bar{x})\} \cup \left[ \{\nabla h_j(\bar{x}) \mid j \in \mathcal{P}\} \cup \{\nabla G_l(\bar{x}) \mid l \in I^G(\bar{x}) \cup I^{GH}(\bar{x})\} \right. \\ & \quad \left. \cup \{\nabla H_l(\bar{x}) \mid l \in I^H(\bar{x}) \cup I^{GH}(\bar{x})\} \right] \end{aligned}$$

is positive-linearly independent. Invoking Lemma 2.2, there is a neighborhood  $U$  of  $\bar{x}$  such that the vectors

$$\begin{aligned} & \{\nabla g_i(x) \mid i \in I^g(\bar{x})\} \cup \left[ \{\nabla h_j(x) \mid j \in \mathcal{P}\} \cup \{\nabla G_l(x) \mid l \in I^G(\bar{x}) \cup I^{GH}(\bar{x})\} \right. \\ & \quad \left. \cup \{\nabla H_l(x) \mid l \in I^H(\bar{x}) \cup I^{GH}(\bar{x})\} \right] \end{aligned}$$

are positive-linearly independent for any choice of  $x \in U$ .

Now, fix  $t > 0$  as well as  $x \in X_S(t) \cap U$  and set  $I_t^a(x) := I_t^+(x) \cup I_t^-(x)$ . Note that  $t > 0$  guarantees  $I_t^+(x) \cap I_t^-(x) = \emptyset$ . Clearly, we have

$$\begin{aligned} \forall l \in I^H(\bar{x}): & \quad G_l(x) \neq 0 \quad H_l(x) \approx 0, \\ \forall l \in I^G(\bar{x}): & \quad G_l(x) \approx 0 \quad H_l(x) \neq 0 \end{aligned}$$

if  $U$  is sufficiently small. Exploiting Lemma 2.2 once more while recalling that  $G$  and  $H$  are continuously differentiable, we obtain that the vectors

$$\begin{aligned} & \{\nabla g_i(x) \mid i \in I^g(\bar{x})\} \cup \left[ \{\nabla h_j(x) \mid j \in \mathcal{P}\} \right. \\ & \quad \cup \{H_l(x) \nabla G_l(x) + G_l(x) \nabla H_l(x) \mid l \in I^G(\bar{x}) \cap I_t^a(x)\} \\ & \quad \cup \{H_l(x) \nabla G_l(x) + G_l(x) \nabla H_l(x) \mid l \in I^H(\bar{x}) \cap I_t^a(x)\} \quad (8) \\ & \quad \cup \{\nabla G_l(x) \mid l \in I^{GH}(\bar{x}) \cap I_t^a(x)\} \\ & \quad \left. \cup \{\nabla H_l(x) \mid l \in I^{GH}(\bar{x}) \cap I_t^a(x)\} \right] \end{aligned}$$

are positive-linearly independent if the neighborhood  $U$  is chosen small enough.

Suppose that there are vectors  $\lambda \in \mathbb{R}^m$ ,  $\rho \in \mathbb{R}^p$ , and  $\xi \in \mathbb{R}^q$  which satisfy

$$\begin{aligned} 0 &= \sum_{i \in I^g(x)} \lambda_i \nabla g_i(x) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(x) + \sum_{l \in \mathcal{Q}} \xi_l \left[ H_l(x) \nabla G_l(x) + G_l(x) \nabla H_l(x) \right], \\ \forall i \in I^g(x): & \quad \lambda_i \geq 0, \quad \forall i \in \mathcal{M} \setminus I^g(x): \quad \lambda_i = 0, \\ \forall l \in I_t^+(x): & \quad \xi_l \geq 0, \\ \forall l \in I_t^-(x): & \quad \xi_l \leq 0, \\ \forall l \in \mathcal{Q} \setminus I_t^a(x): & \quad \xi_l = 0. \end{aligned}$$

In order to show the validity of MFCQ for  $(\mathbf{P}_S(t))$  at  $x$ ,  $\lambda = 0$ ,  $\rho = 0$ , and  $\xi = 0$  have to be deduced. We get

$$\begin{aligned} 0 &= \sum_{i \in I^g(x)} \lambda_i \nabla g_i(x) + \sum_{j \in \mathcal{P}} \rho_j \nabla h_j(x) \\ &\quad + \sum_{l \in (I^G(\bar{x}) \cup I^H(\bar{x})) \cap I_t^a(x)} \xi_l \left[ H_l(x) \nabla G_l(x) + G_l(x) \nabla H_l(x) \right], \\ &\quad + \sum_{l \in I^{GH}(\bar{x}) \cap I_t^a(x)} \xi_l G_l(x) \nabla H_l(x) + \sum_{l \in I^{GH}(\bar{x}) \cap I_t^a(x)} \xi_l H_l(x) \nabla G_l(x). \end{aligned}$$

Noting that  $\lambda_i \geq 0$  holds for all  $i \in I^g(x)$  while  $I^g(x) \subset I^g(\bar{x})$  holds whenever  $U$  is chosen sufficiently small, we obtain  $\lambda = 0$ ,  $\rho = 0$ ,  $\xi_l = 0$  ( $l \in (I^G(\bar{x}) \cup I^H(\bar{x})) \cap I_t^a(x)$ ),  $\xi_l G_l(x) = 0$  ( $l \in I^{GH}(\bar{x}) \cap I_t^a(x)$ ), and  $\xi_l H_l(x) = 0$  ( $l \in I^{GH}(\bar{x}) \cap I_t^a(x)$ ) from the positive-linear independence of the vectors in (8). Since we have  $G_l(x) \neq 0$  and  $H_l(x) \neq 0$  for all  $l \in I_t^a(x)$  from  $t > 0$ ,  $\xi_l = 0$  follows for all  $l \in I_t^a(x)$  since  $I^G(\bar{x}) \cup I^H(\bar{x}) \cup I^{GH}(\bar{x}) = \mathcal{Q}$  is valid. This yields  $\xi = 0$ . Thus, MFCQ holds for  $(\mathbf{P}_S(t))$  at  $x$ .  $\square$

## 4.2 The relaxation scheme of Steffensen and Ulbrich

Here, we adapt the relaxation scheme from Steffensen and Ulbrich [2010] for the numerical treatment of  $(\mathbf{MPSC})$ . For any  $t > 0$ , let us introduce  $\phi(\cdot; t): \mathbb{R} \rightarrow \mathbb{R}$  by means of

$$\forall z \in \mathbb{R}: \quad \phi(z; t) := \begin{cases} |z| & \text{if } |z| \geq t, \\ t\theta(z/t) & \text{if } |z| < t, \end{cases}$$

where  $\theta: [-1, 1] \rightarrow \mathbb{R}$  is a twice continuously differentiable function with the following properties:

- (a)  $\theta(1) = \theta(-1) = 1$ ,                      (b)  $\theta'(-1) = -1$  and  $\theta'(1) = 1$ ,
- (c)  $\theta''(-1) = \theta''(1) = 0$ ,                    (d)  $\theta''(z) > 0$  for all  $z \in (-1, 1)$ .

A typical example for a function  $\theta$  with the above properties is given by

$$\forall z \in [-1, 1]: \quad \theta(z) := \frac{2}{\pi} \sin\left(\frac{\pi}{2}z + \frac{3\pi}{2}\right) + 1, \quad (9)$$

see [Steffensen and Ulbrich, 2010, Section 3]. Noting that the function  $\phi(\cdot; t)$  is smooth for any choice of  $t > 0$ , it can be used to regularize the feasible set of  $(\mathbf{MPSC})$ . A suitable surrogate problem is given by

$$\begin{aligned} f(x) &\rightarrow \min \\ g_i(x) &\leq 0, & i \in \mathcal{M}, \\ h_j(x) &= 0, & j \in \mathcal{P}, \\ G_l(x) + H_l(x) - \phi(G_l(x) - H_l(x); t) &\leq 0, & l \in \mathcal{Q}, \\ G_l(x) - H_l(x) - \phi(G_l(x) + H_l(x); t) &\leq 0, & l \in \mathcal{Q}, \\ -G_l(x) + H_l(x) - \phi(-G_l(x) - H_l(x); t) &\leq 0, & l \in \mathcal{Q}, \\ -G_l(x) - H_l(x) - \phi(-G_l(x) + H_l(x); t) &\leq 0, & l \in \mathcal{Q}. \end{aligned} \quad (\mathbf{P}_{\text{SU}}(t))$$

Its feasible set will be denoted by  $X_{\text{SU}}(t)$  and is visualized in Figure 4. Adapting the proof of [Steffensen and Ulbrich, 2010, Lemma 3.3], the family  $\{X_{\text{SU}}(t)\}_{t>0}$  possesses the properties (P2) and (P3) from Lemma 3.4. This justifies that  $(\text{P}_{\text{SU}}(t))$  is a relaxation of  $(\text{MPSC})$ . Note that we need to introduce four inequality constraints to replace one of the original switching constraints.

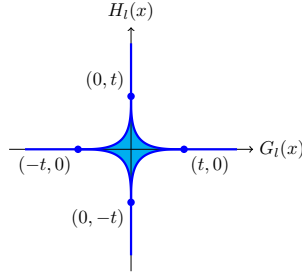


Figure 4: Geometric interpretation of the relaxed feasible set  $X_{\text{SU}}(t)$ .

It has been mentioned in Hoheisel et al. [2013] that the relaxation scheme of Steffensen and Ulbrich computes Clarke-stationary points of MPCCs under an MPCC-tailored version of CPLD, see [Hoheisel et al., 2013, Section 3.4] as well. Recalling some arguments from Section 4.1, the adapted method may only find W-stationary points of  $(\text{MPSC})$  in general. The upcoming example confirms this conjecture.

**Example 4.4.** *Let us consider the switching-constrained optimization problem*

$$\begin{aligned} x_1 x_2 - x_1 - x_2 &\rightarrow \min \\ x_1^2 + x_2^2 - 1 &\leq 0, \\ x_1 x_2 &= 0. \end{aligned} \tag{10}$$

*Obviously, the globally optimal solutions of this problem are given by  $(1, 0)$  as well as  $(0, 1)$ , and these points are S-stationary. Furthermore, there is a W-stationary point at  $\bar{x} = (0, 0)$  which is no local minimizer. The global maximizers  $(-1, 0)$  and  $(0, -1)$  do not satisfy any of the introduced stationarity concepts.*

*Let us consider the associated family of nonlinear problems  $(\text{P}_{\text{SU}}(t))$  for  $t \in (0, 1]$  where the function  $\theta$  is chosen as in (9). It can easily be checked that  $x(t) := (\frac{t}{2}(1 - \frac{2}{\pi}), \frac{t}{2}(1 - \frac{2}{\pi}))$  is a KKT point of  $(\text{P}_{\text{SU}}(t))$ . Note that  $x(t) \rightarrow \bar{x}$  as  $t \downarrow 0$ , and that MPSC-LICQ holds in  $\bar{x}$ . However,  $\bar{x}$  is only a W-stationary point of the switching-constrained problem (10).* !DS

### 4.3 The relaxation scheme of Kadrani, Dussault, and Benchakroun

Finally, we want to take a closer look at the relaxation approach which was suggested by Kadrani et al. [2009] for the treatment of MPCCs. For any  $t \geq 0$ , let us consider the

optimization problem

$$\begin{aligned}
& f(x) \rightarrow \min \\
& g_i(x) \leq 0, & i \in \mathcal{M}, \\
& h_j(x) = 0, & j \in \mathcal{P}, \\
& (G_l(x) - t)(H_l(x) - t) \leq 0, & l \in \mathcal{Q}, \\
& (-G_l(x) - t)(H_l(x) - t) \leq 0, & l \in \mathcal{Q}, \\
& (G_l(x) + t)(H_l(x) + t) \leq 0, & l \in \mathcal{Q}, \\
& (G_l(x) - t)(-H_l(x) - t) \leq 0, & l \in \mathcal{Q},
\end{aligned} \tag{P_{KDB}(t)}$$

whose feasible set will be denoted by  $X_{\text{KDB}}(t)$ . The family  $\{X_{\text{KDB}}(t)\}_{t \geq 0}$  only satisfies property (P1) from Lemma 3.4 while (P2) and (P3) are violated in general. Thus, the surrogate problem (P\_{KDB}(t)) does not induce a relaxation technique for (MPSC) in the narrower sense. Figure 5 depicts that  $X_{\text{KDB}}(t)$  is *almost* disconnected, i.e., it is close to crumbling into four disjoint sets for any  $t > 0$ . This may cause serious problems when standard techniques are used to solve the associated surrogate problem (P\_{KDB}(t)). Moreover, four inequality constraints are necessary to replace one switching constraint from (MPSC) in (P\_{KDB}(t)).

!DS

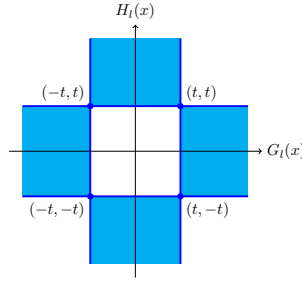


Figure 5: Geometric interpretation of the relaxed feasible set  $X_{\text{KDB}}(t)$ .

On the other hand, it is clear from [Hoheisel et al., 2013, Section 3.3] that the regularization approach of Kadrani, Dussault, and Benchakroun computes M-stationary points of MPCCs under an MPCC-tailored version of CPLD at the limit point. Furthermore, if an MPCC-tailored LICQ holds at the limit point, then standard GCQ holds for the surrogate problems in a neighborhood of the point for sufficiently small relaxation parameters. These results are closely related to those for the relaxation approach from Kanzow and Schwartz [2013] which we generalized to (MPSC) in Sections 3.1 and 3.2.

Although we abstain from a detailed analysis of the regularization method which is induced by the surrogate problem (P\_{KDB}(t)) due to the aforementioned shortcomings, the above arguments motivate the formulation of the following two conjectures.

**Conjecture 4.5.** *Let  $\{t_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+$  be a sequence of positive regularization parameters converging to zero. For each  $k \in \mathbb{N}$ , let  $x_k \in X_{\text{KDB}}(t_k)$  be a KKT point of (P\_{KDB}(t\_k)). Assume that the sequence  $\{x_k\}_{k \in \mathbb{N}}$  converges to a point  $\bar{x} \in X$  where MPSC-MFCQ holds. Then  $\bar{x}$  is an M-stationary point of (MPSC).*



**Conjecture 4.6.** *Let  $\bar{x} \in X$  be a feasible point of (MPSC) where MPSC-LICQ is satisfied. Then there exist  $\bar{t} > 0$  and a neighborhood  $U \subset \mathbb{R}^n$  of  $\bar{x}$  such that GCQ holds for  $(P_{\text{KDB}}(t))$  at all points from  $X_{\text{KDB}}(t) \cap U$  for all  $t \in (0, \bar{t}]$ .*

## 5 Numerical results

This section is dedicated to a detailed analysis and comparison of various numerical methods for (MPSC). To obtain a meaningful comparison, we apply the relaxation scheme from Section 3 as well as a collection of other algorithms (see below) to multiple classes of MPSCs which possess significant practical relevance. The particular examples we analyze are:

- an either-or constrained problem with known local and global solutions
- a switching-constrained optimal control problem involving the non-stationary heat equation in two dimensions
- optimization problems involving semi-continuous variables, in particular special instances of portfolio optimization

For each of the examples, we first provide an overview of the corresponding problem structure, and then give some numerical results. To facilitate a quantitative comparison of the used algorithms, we use performance profiles (see Dolan and Moré [2002]) based on the computed function values.

!DS

### 5.1 Implementation

The numerical experiments in this section were all done in MATLAB R2018a. The particular algorithms we use for our computations are the following:

!DS

**KS:** the adapted Kanzow–Schwartz relaxation scheme from this paper

**FMC:** the `fmincon` function from the MATLAB Optimization Toolbox

**SNOPT:** the SNOPT nonlinear programming solver from Gill et al. [2002], called through the TOMLAB programming environment

**IPOPT:** the IPOPT interior-point algorithm from Wächter and Biegler [2006]

The overall implementation is done in MATLAB, and each algorithm is called with user-supplied gradients of the objective functions and constraints. The stopping tolerance for all algorithms is set to  $10^{-4}$  (although it should be noted that the methods use different stopping criteria, i.e., they impose the accuracy in different ways). For the KS algorithm, the relaxation parameters are chosen as  $t_k := 0.01^k$ , and the method is also terminated as soon as  $t_k$  drops below  $10^{-8}$ . Finally, to solve the relaxed subproblems in the relaxation method, we employ the SNOPT algorithm with an accuracy of  $10^{-6}$ .

Judging by past experience with MPCCs, the SNOPT algorithm can be expected to rival the relaxation scheme in terms of robustness. To accurately measure the performance of the solvers, it is important to note that MPSCs can, in general, admit a substantial amount of local minimizers. Therefore, the robustness is best measured by comparing the obtained function values (using different methods and starting points) with the globally optimal function value—if the latter is known; otherwise, a suitable approximate is used. To avoid placing too much emphasis on the accuracy of the final output (which does not make sense since the algorithms use completely different stopping criteria), we use the quantity

$$Q_\delta(x) := \begin{cases} f(x) - f_{\min} + \delta, & \text{if } x \text{ is feasible within tolerance,} \\ +\infty, & \text{otherwise} \end{cases} \quad (11)$$

as the base metric for the performance profiles, where  $x$  is the final iterate of the given algorithmic run,  $f_{\min}$  the (approximate) global minimal value of the underlying problem, and  $\delta \geq 0$  is an additional parameter which reduces the sensitivity of the values to numerical accuracy. We have found that an appropriate choice of  $\delta$  can significantly improve the meaningfulness of the results.

## 5.2 Numerical examples

The following pages contain three examples of MPSCs. In Section 5.2.1, we deal with an *either-or constrained* problem, which can be reformulated as an MPSC, see Mehlitz [2018]. Section 5.2.2 is dedicated to a switching-constrained optimal control problem based on the framework from Clason et al. [2017]. Finally, in Section 5.2.3, we deal with a class of optimization problems with *semi-continuous variables*, which can again be reformulated as MPSCs. This section contains a particular example from portfolio optimization which originates from Frangioni and Gentile [2007].

### 5.2.1 An either-or constrained example

Let us consider the optimization problem

$$\begin{aligned} (x_1 - 8)^2 + (x_2 + 3)^2 &\rightarrow \min \\ x_1 - 2x_2 + 4 &\leq 0 \vee x_1 - 2 \leq 0, \\ x_1^2 - 4x_2 &\leq 0 \vee (x_1 - 3)^2 + (x_2 - 1)^2 - 10 \leq 0. \end{aligned} \quad (\text{E2})$$

Here,  $\vee$  denotes the logical “or”. The feasible set of this program is visualized in Figure 6. It is easily seen that (E2) possesses the unique global minimizer  $\bar{x} = (2, -2)$  and another local minimizer  $\tilde{x} = (4, 4)$ . Arguing as in [Mehlitz, 2018, Section 7], we can transform (E2) into a switching-constrained optimization problem by introducing additional variables:

$$\begin{aligned} (x_1 - 8)^2 + (x_2 + 3)^2 &\rightarrow \min_{x,z} \\ z_1, z_2, z_3, z_4 &\leq 0, \\ (x_1 - 2x_2 + 4 - z_1)(x_1 - 2 - z_2) &= 0, \\ (x_1^2 - 4x_2 - z_3)((x_1 - 3)^2 + (x_2 - 1)^2 - 10 - z_4) &= 0. \end{aligned} \quad (12)$$

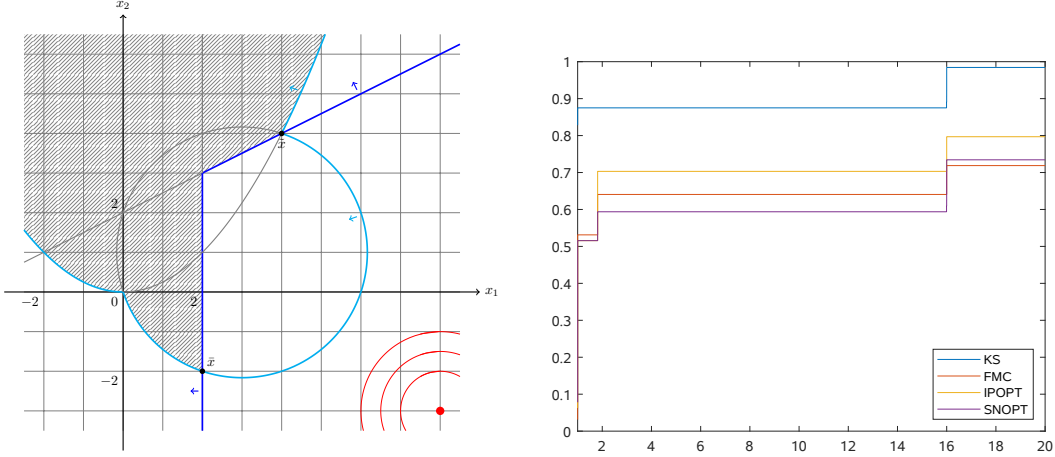


Figure 6: The either-or constrained problem from Section 5.2.1: feasible set, some level sets, and local minimizers (left), performance profile (right).

Note that the local minimizers of (E2) can be found among the local minimizers of (12) choosing suitable values for the variable  $z$ , see [Mehlitz, 2018, Section 7.1].

The algorithms in question are each tested with the starting points in the set  $\{0, 1\}^6$ , which makes for a total of 64 starting points. The resulting performance profile can also be found in Figure 6; it is based on the metric (11) with  $\delta := 1$ . Clearly, the KS relaxation method is the most robust of the four algorithms, finding the best function values (among the tested algorithms) in more than 80% of the test runs.

PM:  $z_1, \dots, z_4 \leq 0$  does to hold, so are we really using these starting points?

### 5.2.2 Switching-constrained optimal control

Here, we intend to solve a switching-constrained optimal control problem with the proposed relaxation method. The underlying example is taken from [Clason et al., 2017, Section 5.2].

Let  $I := (0, T)$ , with  $T := 10$  the final time,  $\Omega := (-1, 1)^2$ , and let  $\Gamma$  be the boundary of  $\Omega$ . Furthermore, we define  $\Omega_u := (-1, 0] \times (-1, 1)$  as well as  $\Omega_v := (0, 1) \times (-1, 1)$ . Let us consider the optimal control of the non-stationary heat equation with zero initial and Neumann boundary conditions given below:

$$\begin{aligned} \partial_t y(t, \omega) - \Delta_\omega y(t, \omega) - \frac{1}{10} \chi_{\Omega_u}(\omega) u(t) - \frac{1}{10} \chi_{\Omega_v}(\omega) v(t) &= 0 & \text{a.e. on } I \times \Omega, \\ \vec{\mathbf{n}}(\omega) \cdot \nabla_\omega y(t, \omega) &= 0 & \text{a.e. on } I \times \Gamma, \\ y(0, \omega) &= 0 & \text{a.e. on } \Omega. \end{aligned} \quad (13)$$

Here,  $\chi_A: \Omega \rightarrow \mathbb{R}$  denotes the characteristic function of a measurable set  $A \subset \Omega$  which equals 1 on  $A$  and vanishes otherwise. Let  $y_d \in L^2(I; H^1(\Omega))$  be the solution of the state equation associated with the desirable control functions  $u_d, v_d \in L^2(I)$  given by

$$\forall t \in I: \quad u_d(t) := 20 \sin^4(2\pi t/T), \quad v_d(t) := 10 \cos^4(1.4\pi t/T).$$

All feasible controls  $u, v \in H^1(I)$  shall satisfy the switching requirement

$$u(t)v(t) = 0 \quad \text{a.e. on } I. \quad (14)$$

Note that  $u_d$  and  $v_d$  violate this switching condition. We aim to find the minimum of the objective function defined by

$$\begin{aligned} \mathcal{J}(y, u, v) := & \frac{1}{2} \|y - y_d\|_{L^2(I; L^2(\Omega))}^2 + \frac{\alpha}{2} \left( \|u\|_{L^2(I)}^2 + \|v\|_{L^2(I)}^2 \right) \\ & + \frac{\beta}{2} \left( \|\partial_t u\|_{L^2(I)}^2 + \|\partial_t v\|_{L^2(I)}^2 \right) \end{aligned} \quad (15)$$

with respect to  $(y, u, v) \in L^2(I; H^1(\Omega)) \times H^1(I) \times H^1(I)$  such that  $(u, v)$  satisfy the switching requirement (14) while  $y$  solves the associated state equation (13). We chose  $\alpha := 10^{-6}$  and  $\beta := 10^{-5}$  for our experiments.

For the numerical solution of the problem, the domain  $\Omega$  is tessellated using the function `generateMesh` from the MATLAB PDE toolbox and a discretization tolerance of  $h := 10^{-1}$ . The time interval  $I$  is subdivided into equidistant intervals of size  $\tau := 10^{-1}$ . Both the spatial and temporal discretizations use standard piecewise linear (continuous) finite elements, which leads to a conforming approximation of the  $H^1$ -norm in (15).

After discretization, the problem turns into a finite-dimensional MPSC comprising the variables  $\vec{u}, \vec{v} \in \mathbb{R}^{101}$ , a quadratic objective function, and the switching constraints  $\vec{u}_i \vec{v}_i = 0$  for all  $i = 1, \dots, 101$ . These correspond to the simple constraint mappings  $G(\vec{u}, \vec{v}) := \vec{u}$  and  $H(\vec{u}, \vec{v}) := \vec{v}$ . Note that the feasible set can be seen as the union of  $2^{101}$  convex “branches” (obtained by setting either  $\vec{u}_i = 0$  or  $\vec{v}_i = 0$  for each  $i = 1, \dots, 101$ ). Hence, the problem can be expected to admit a substantial amount of local minimizers, and it is unrealistic to expect algorithmic implementations to find the global minimizer, even when tested with a large number of initial points. To facilitate a quantitative comparison of our numerical algorithms (as in the previous section), we use the following heuristic to obtain an upper estimate of the optimal value: using a coarser time discretization (with  $\tau = 0.4$ ), we compute the *exact* global minimizer of the resulting problem by minimizing the objective over each of the branches induced by the switching constraints. The corresponding global minimizer is then lifted to the finer time grid (with  $\tau = 0.1$ ) by linear interpolation, and the resulting point is used as an initial guess for all the used algorithms. The resulting estimate of the optimal value is 0.2997, and the associated controls are depicted in Figure 7.

For the numerical tests, we generated 1000 starting points with coordinates chosen randomly in the interval  $[0, 10]$ . The performance profile was constructed by using the metric (11) with  $\delta := 0$ , and it too can be found in Figure 7.

!DS  
!DS

As in the previous example, the relaxation method turns out to be the most robust of the tested algorithms, finding lower function values in around 70% of the test runs. When analyzing the results in more detail, it turns out that, as expected, the algorithms found an exorbitant amount of distinct points (possibly local minimizers). Interestingly, however, the associated function values actually lie quite close to each other. This explains the  $x$ -axis scaling in the performance profile plot.

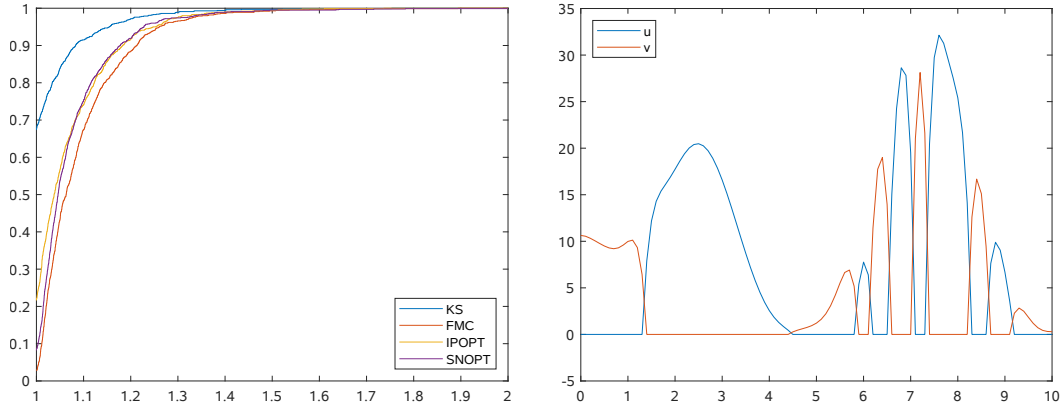


Figure 7: Performance profile and (possible) global minimizer for the switching-constrained optimal control problem from Section 5.2.2.

### 5.2.3 Semi-continuous variables

In many optimization scenarios, it is desirable that a nonnegative decision variable is either exactly zero or contained in some interval whose lower bound is strictly positive. This is the case, for instance, in production planning, portfolio optimization, compressed sensing in signal processing, and subset selection in regression. More details can be found in Burdakov et al. [2016], Sun et al. [2013], and the references therein.

Given a decision variable  $x \in \mathbb{R}^n$  and bounds  $\ell, u \in \mathbb{R}^n$ ,  $\ell \leq u$ , a requirement of the above form can be reformulated as the either-or type constraints

$$x_i = 0 \vee x_i \in [\ell_i, u_i], \quad i = 1, \dots, n. \quad (16)$$

In this context, the variable  $x$  is often called *semi-continuous* since it is required to lie in some continuous interval, except for the outlier case when it is equal to zero. (One might also be inclined to call  $x$  a *semi-discrete* variable, but we have not seen this terminology elsewhere in the literature.)

Constraint systems of the form (16) can be reformulated as switching constraints by using slack variables. Indeed, there are two ways of doing so: On the one hand, we could introduce two nonnegative slack variables to transform the box constraints in (16) into equality constraints; this procedure eventually yields an MPSC with  $2n$  switching constraints. On the other hand, assuming that  $u_i \geq 0$  holds for all  $i = 1, \dots, n$  (which is the case in nearly all relevant applications), we can simply treat the requirement  $x_i \leq u_i$  as a standard inequality constraint which should be fulfilled at all times. Clearly, if  $x_i = 0$  is valid, then the inequality  $x_i \leq u_i$  holds automatically, so that we can rewrite (16) as

$$\begin{aligned} x_i &\leq u_i, & i &= 1, \dots, n, \\ x_i &= 0 \vee x_i \geq \ell_i, & i &= 1, \dots, n. \end{aligned}$$

Using a single slack vector  $y \in \mathbb{R}^n$ , we can now rewrite this system as

$$\begin{aligned} x &\leq u, \\ y &\geq 0, \\ x_i(x_i - \ell_i - y_i) &= 0, \quad i = 1, \dots, n. \end{aligned}$$

In the notation of our general framework (MPSC), this corresponds to the switching mappings

$$G(x, y) := x \quad \text{and} \quad H(x, y) := x - \ell - y.$$

The inequality constraints  $x \leq u$  and  $y \geq 0$  can be implemented as components of the mapping  $g$  (which may contain other constraints depending on the particular problem). Note that the above reformulation approach only results in  $n$  switching constraints.

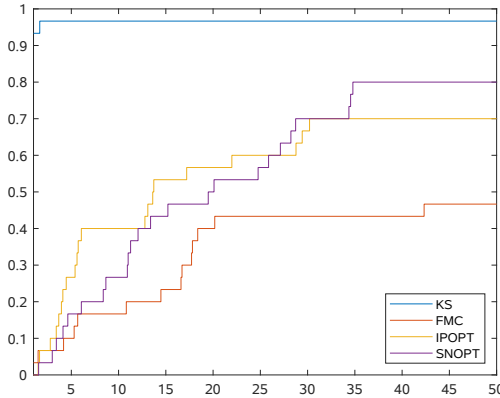


Figure 8: Performance profile for the portfolio optimization problem from Section 5.2.3.

We now present a concrete example of portfolio optimization based on the test examples in Frangioni and Gentile [2007]. The problems in this reference have the form

$$\begin{aligned} x^\top Qx &\rightarrow \min \\ e^\top x &= 1, \\ \mu^\top x &\geq \rho, \\ x_i = 0 \vee x_i &\in [\ell_i, u_i], \quad i = 1, \dots, n, \end{aligned} \tag{17}$$

with randomly generated  $Q \in \mathbb{R}^{n \times n}$ ,  $\mu, \ell, u \in \mathbb{R}^n$ , and  $\rho \in \mathbb{R}$ . Here,  $e \in \mathbb{R}^n$  represents the all-ones vector. More details can be found in Frangioni and Gentile [2007] and on their webpage <http://www.di.unipi.it/optimize/Data/MV.html>. The particular examples we chose are the 30 instances with size 200. The corresponding problems (17) are reformulated as MPSCs by means of the aforementioned procedure, and the resulting problems are then attacked by the four test algorithms in question. Figure 8 depicts the resulting performance profile based on the metric (11) with  $\delta := 0$ .

!DS

For this particular problem class, it turns out that the performance advantage of the relaxation method is particularly large when compared to its non-relaxed counterparts.

In 28 out of 30 examples, the KS algorithm finds the best function value among the tested methods; in addition, it also seems to find feasible points much more reliably than the other algorithms.

## 6 Final remarks

In this paper, we have presented a relaxation method for the solution of mathematical programs with switching constraints (MPSCs). Our theoretical analysis yields strong convergence properties for the method; in particular, the algorithm computes M-stationary points of MPSCs under a problem-tailored constraint qualification (MPSC-NNAMCQ) which is weaker than MPSC-MFCQ. The numerical experiments include a wide array of practically relevant problems and demonstrate the computational efficiency of the proposed algorithm.

In addition, we have conducted a dedicated analysis for other relaxation schemes which are known in the MPCC literature and can be carried over to the setting of switching constraints. In particular, adapted versions of the relaxation schemes of Scholtes as well as Steffensen and Ulbrich are shown to converge to weakly stationary points only, even if fairly strong regularity properties such as MPSC-LICQ are satisfied.

## References

- W. Achtziger and C. Kanzow. Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications. *Mathematical Programming, Series A*, 114(1):69–99, 2008. doi: [10.1007/s10107-006-0083-3](https://doi.org/10.1007/s10107-006-0083-3).
- M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. Wiley and Sons, New York, 1993.
- O. P. Burdakov, C. Kanzow, and A. Schwartz. Mathematical programs with cardinality constraints: reformulation by complementarity-type conditions and a regularization method. *SIAM Journal on Optimization*, 26(1):397–425, 2016. doi: [10.1137/140978077](https://doi.org/10.1137/140978077).
- C. Clason, A. Rund, and K. Kunisch. Nonconvex penalization of switching control of partial differential equations. *Systems & Control Letters*, 106:1–8, 2017. doi: [10.1016/j.sysconle.2017.05.006](https://doi.org/10.1016/j.sysconle.2017.05.006).
- Elizabeth D. Dolan and Jorge J. Moré. Benchmarking optimization software with performance profiles. *Math. Program.*, 91(2, Ser. A):201–213, 2002. ISSN 0025-5610. doi: [10.1007/s101070100263](https://doi.org/10.1007/s101070100263). URL <https://doi.org/10.1007/s101070100263>.
- A. Frangioni and C. Gentile. SDP diagonalizations and perspective cuts for a class of non-separable MIQP. *Oper. Res. Lett.*, 35(2):181–185, 2007. doi: [10.1016/j.orl.2006.03.008](https://doi.org/10.1016/j.orl.2006.03.008).

- P. E. Gill, W. Murray, and M. A. Saunders. SNOPT: an SQP algorithm for large-scale constrained optimization. *SIAM Journal on Optimization*, 12(4):979–1006, 2002. doi: [10.1137/S1052623499350013](https://doi.org/10.1137/S1052623499350013).
- M. Gugat. Optimal switching boundary control of a string to rest in finite time. *ZAMM - Journal of Applied Mathematics and Mechanics*, 88(4):283–305, 2008. doi: [10.1002/zamm.200700154](https://doi.org/10.1002/zamm.200700154).
- F. M. Hante and S. Sager. Relaxation methods for mixed-integer optimal control of partial differential equations. *Computational Optimization and Applications*, 55(1):197–225, 2013. doi: [10.1007/s10589-012-9518-3](https://doi.org/10.1007/s10589-012-9518-3).
- T. Hoheisel and C. Kanzow. Stationary conditions for mathematical programs with vanishing constraints using weak constraint qualifications. *Journal of Mathematical Analysis and Applications*, 337(1):292 – 310, 2008. doi: [10.1016/j.jmaa.2007.03.087](https://doi.org/10.1016/j.jmaa.2007.03.087).
- T. Hoheisel, C. Kanzow, and A. Schwartz. Convergence of a local regularization approach for mathematical programmes with complementarity or vanishing constraints. *Optimization Methods and Software*, 27(3):483–512, 2012. doi: [10.1080/10556788.2010.535170](https://doi.org/10.1080/10556788.2010.535170).
- T. Hoheisel, C. Kanzow, and A. Schwartz. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. *Mathematical Programming*, 137(1):257–288, 2013. doi: [10.1007/s10107-011-0488-5](https://doi.org/10.1007/s10107-011-0488-5).
- A. Kadrani, J.-P. Dussault, and A. Benchakroun. A new regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 20(1):78–103, 2009. doi: [10.1137/070705490](https://doi.org/10.1137/070705490).
- C. Kanzow and A. Schwartz. A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. *SIAM Journal on Optimization*, 23(2):770–798, 2013. doi: [10.1137/100802487](https://doi.org/10.1137/100802487).
- D. Liberzon. *Switching in Systems and Control*. Birkhäuser, Boston, 2003.
- Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.
- P. Mehlitz. Stationarity conditions and constraint qualifications for mathematical programs with switching constraints. *Preprint TU Freiberg*, 2018. URL <http://www.mathe.tu-freiberg.de/files/personal/254/2018-03-mehlitz.pdf>.
- J. V. Outrata, M. Kočvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic, Dordrecht, 1998.
- L. Qi and Z. Wei. On the constant positive linear dependence condition and its application to SQP methods. *SIAM Journal on Optimization*, 10(4):963–981, 2000. doi: [10.1137/S1052623497326629](https://doi.org/10.1137/S1052623497326629).



- R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 1998.
- S. Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918–936, 2001. doi: [10.1137/S1052623499361233](https://doi.org/10.1137/S1052623499361233).
- T. I. Seidman. Optimal control of a diffusion/reaction/switching system. *Evolution Equations & Control Theory*, 2(4):723–731, 2013. doi: [10.3934/eect.2013.2.723](https://doi.org/10.3934/eect.2013.2.723).
- S. Steffensen and M. Ulbrich. A new relaxation scheme for mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 20(5):2504–2539, 2010. doi: [10.1137/090748883](https://doi.org/10.1137/090748883).
- X. Sun, X. Zheng, and D. Li. Recent advances in mathematical programming with semi-continuous variables and cardinality constraint. *Journal of the Operations Research Society of China*, 1(1):55–77, 2013. doi: [10.1007/s40305-013-0004-0](https://doi.org/10.1007/s40305-013-0004-0).
- A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program. Ser. A*, 106(1):25–57, 2006. doi: [10.1007/s10107-004-0559-y](https://doi.org/10.1007/s10107-004-0559-y).
- L. Wang and Q. Yan. Time optimal controls of semilinear heat equation with switching control. *Journal of Optimization Theory and Applications*, 165(1):263–278, 2015. doi: [10.1007/s10957-014-0606-7](https://doi.org/10.1007/s10957-014-0606-7).
- J. J. Ye and X. Y. Ye. Necessary optimality conditions for optimization problems with variational inequality constraints. *Math. Oper. Res.*, 22(4):977–997, 1997. ISSN 0364-765X. doi: [10.1287/moor.22.4.977](https://doi.org/10.1287/moor.22.4.977). URL <https://doi.org/10.1287/moor.22.4.977>.
- E. Zuazua. Switching control. *Journal of the European Mathematical Society*, 13(1):85–117, 2011. doi: [10.4171/JEMS/245](https://doi.org/10.4171/JEMS/245).