Preprint 2016-08

Francisco Benita, Patrick Mehlitz

Optimal Control Problems with Terminal Complementarity Constraints
Francisco Benita, Patrick Mehlitz

Optimal Control Problems with Terminal Complementarity Constraints

TU Bergakademie Freiberg
Fakultät für Mathematik und Informatik
Prüferstraße 9
09599 FREIBERG
http://tu-freiberg.de/fakult1
Optimal Control Problems with Terminal Complementarity Constraints

Francisco Benita, Patrick Mehlitz

Abstract In this paper, we study an optimal control problem of ordinary differential equations with linear dynamics, affine mixed control-state constraints, and terminal complementarity constraints on the state function. We derive its weak, Clarke, Mordukhovich, and strong stationarity conditions, and we present constraint qualifications which ensure that these conditions are satisfied at a locally optimal solution of the optimal control problem. Finally, we show that Scholtes’s global relaxation technique is applicable to the problem and yields Clarke stationary points under appropriate assumptions.

Keywords Mathematical Program with Complementarity Constraints · Optimal Control · Optimality Conditions · Programming in Banach Spaces · W-, C-, M-, S-Stationarity

Mathematics Subject Classification (2000) 46N10 · 49K15 · 49K27 · 90C30 · 90C33

1 Introduction

We are going to consider the following optimal control problem of ordinary differential equations (ODEs for short) with terminal state complementarity constraints (OCTCC):

\[
f(x(T)) + \frac{1}{2} \| x - x_d \|^2_{L^2(0,T)^n} + \frac{\sigma}{2} \| u - u_d \|^2_{L^2(0,T)^m} \to \min_{x,u}
\]

\[
\dot{x}(t) - Ax(t) - Bu(t) = 0 \quad \text{a.e. on } (0,T)
\]

\[
x(0) = 0
\]

\[
Cx(t) + Du(t) - f(t) \leq 0 \quad \text{a.e. on } (0,T)
\]

\[
g(x(T)) \in K
\]

\[
G(x(T)) \geq 0
\]

\[
H(x(T)) \geq 0
\]

\[
G(x(T)) \cdot H(x(T)) = 0,
\]

Let us first summarize our standing assumptions on (OCTCC) before we comment on the question where problems of type (OCTCC) appear in practice.

Assumption 1.1 The matrices \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \), \( C \in \mathbb{R}^{q \times n} \), \( D \in \mathbb{R}^{q \times m} \), and \( f \in L^2(0,T)^q \) are fixed, \( \sigma > 0 \) is a scalar, the functions \( f: \mathbb{R}^n \to \mathbb{R} \), \( g: \mathbb{R}^n \to \mathbb{R}^k \), as well as \( G, H: \mathbb{R}^n \to \mathbb{R}^k \) are continuously differentiable, and \( K \subseteq \mathbb{R}^k \) is a nonempty, closed, convex set. Furthermore, \( x_d \in L^2(0,T)^n \) and \( u_d \in L^2(0,T)^m \) are the given desired state and control, respectively. The decision space for the state variable will be \( H^1(0,T)^n \) while the control \( u \) is chosen from \( L^2(0,T)^m \).

F. Benita
Singapore University of Technology and Design, Department of Engineering Systems and Design, 487372 Singapore
E-mail: francisco.benita@sutd.edu.sg

P. Mehlitz (Corresponding author)
Technische Universität Bergakademie Freiberg, Faculty of Mathematics and Computer Science, 09596 Freiberg, Germany
E-mail: mehlitz@math.tu-freiberg.de
Problems of type \((\text{OCTCC})\) arise in practical applications where connected multi-agent networks have to be controlled. In such systems, multiple agents perform cooperative teamwork in order to accomplish predefined goals and requirements, i.e. spacecraft formation [4], satellite clustering [8], and flocking [31]. For example, in [10], the authors consider the problem of generating optimal trajectories that bring agents (satellites, robots,...) into proximity to form a connected network at terminal time by minimizing the total control effort necessary to do so. The formulation via optimal control includes specified initial states, linear dynamics, and a connectivity constraint on the final induced topology. The focus of the problem is not only optimizing the agents’s trajectories but also evaluating and improving the topology of the final connected network using nonlinear terminal constraints. Thinking about frictional constraints the agents need to fulfill at the final time (in order to transfer data, goods,...), see [33], the resulting optimal control problem can be stated in the form \((\text{OCTCC})\). The problem \((\text{OCTCC})\) also arises from a certain reformulation of the natural gas cash-out problem, see [24]. It is a multi-agent cooperative control problem possessing a hierarchical structure. In the system which is governed by ODEs (upper level problem), the terminal values of the state variables influence a common finite-dimensional optimization problem (lower level problem) whose optimal solution affects the objective functional of the upper level problem. It is clear that optimization problems of this type are bilevel programming problems, see [11], and particularly, since one of the involved decision levels comprises an optimal control problem, they belong to the class of bilevel optimal control problems. In [5,6,24], the authors use the implicitly known optimal value function of the lower level problem to derive optimality conditions for the natural gas cash-out problem. However, the possibility to reformulate such a problem using the lower level Karush Kuhn Tucker (KKT) conditions, a common approach in the bilevel programming literature, will lead to a problem of type \((\text{OCTCC})\).

Mathematical programs with complementarity constraints (MPCCs) arise frequently from different applications in finite- and infinite-dimensional decision spaces. It is well-known that these problems are difficult to handle with the standard tools of mathematical programming since they are highly degenerated, i.e. all constraint qualifications of reasonable strength fail to hold at all feasible points. That is why a huge effort was put in constructing suitable stationarity concepts and constraint qualifications. This is generally achieved by considering several different reformulations of the complementarity program which can be handled using standard arguments from the theory of smooth and nonsmooth constrained optimization. Standard MPCCs, see Section 3 for a precise explanation, have been investigated extensively from a theoretical and numerical point of view, see [16, 21, 27, 32, 37, 38, 41] and the references therein. Recently, some theoretical investigations on MPCCs where the complementarity is induced by the cone of all symmetric, semidefinite matrices or the second-order cone have been done, see [12, 26]. Optimal control problems with mixed control-state complementarity constraints are the topic of interest in [18, 28]. On the other hand, optimal control problems with variational inequality constraints (which cover complementarity constraints) are studied in a special setting in [19, 20] and other contributions by these authors. In [29, 40], the authors investigate a very general MPCC in Banach spaces which covers all the above classes of complementarity problems.

In this paper, we consider the case of an optimal control problem of ODEs with terminal complementarity constraints, and to the best of our knowledge, there do not exist studies dealing with this special class of programs. Although the decision variables come from an infinite-dimensional space, the complementarity constraint only appears in a finite-dimensional context. Consequently, the special structure of the studied problem allows us to apply standard techniques and results from the rich theory on finite-dimensional MPCCs. Here, we clarify how the so-called weak, Clarke, Mordukhovich, and strong stationarity conditions for \((\text{OCTCC})\) look like and under which constraint qualifications they yield necessary optimality conditions for \((\text{OCTCC})\).

The paper is organized as follows: In Section 2, we present the notation we are going to use throughout the paper as well as some preliminary results. Afterwards, we comment on stationarity conditions and constraint qualifications for MPCCs with infinite-dimensional decision space in Section 3. In Section 4, we study the properties of the solution operator to the given linear ODE system. We investigate the existence of globally optimal solutions of \((\text{OCTCC})\) in Section 5. Section 6 is dedicated to the derivation of the weak, Clarke, Mordukhovich, and strong stationarity conditions for \((\text{OCTCC})\). In Section 7, we present constraint qualifications which ensure that the different stationarity concepts are satisfied at a given locally optimal solution of \((\text{OCTCC})\). Finally, in Section 8, we apply Scholtes’s relaxation scheme
for the numerical handling of common MPCCs to (OCTCC) and show that the corresponding limit point is a C-stationary point for (OCTCC) under suitable assumptions.

2 Notation and preliminaries

2.1 Basic notation

In this paper, we use \( \mathbb{N}, \mathbb{R}, \mathbb{R}^+, \mathbb{R}^n, \mathbb{R}^{n,+}, \) and \( \mathbb{R}^{m \times n} \) to denote the natural numbers, the real numbers, the positive real numbers, the set of all real vectors with \( n \) components, the set of all real vectors with \( n \) nonnegative components, and the set of all real matrices with \( m \) rows and \( n \) columns, respectively. For brevity, we set \( \mathbb{R}^+_0 := \mathbb{R}_0^{1,+} \). Naturally, we equip \( \mathbb{R}^n \) with the Euclidean norm \( \| \cdot \|_2 \). However, we also deal with the norm \( \| \cdot \|_1 \). Recall that we have

\[
\forall x \in \mathbb{R}^n : \quad \| x \|_2 := \left( \sum_{j=1}^n x_j^2 \right)^{\frac{1}{2}}, \quad \| x \|_1 := \sum_{j=1}^n | x_j |.
\]

For given vectors \( x, y \in \mathbb{R}^n \), \( x \cdot y \) denotes their Euclidean inner product. For a sequence \( \{ t_k \} \subseteq \mathbb{R} \), we write \( t_k \searrow 0 \) in order to express that \( \{ t_k \} \subseteq \mathbb{R}^+ \) converges to zero. Especially, \( t_k \neq 0 \) holds for all \( k \in \mathbb{N} \). If \( Q \in \mathbb{R}^{m \times n} \) is an arbitrary matrix, then \( Q^T \in \mathbb{R}^{n \times m} \) represents its transpose, whereas \( Q^\star \in \mathbb{R}^{m \times n} \) is the pseudo inverse matrix of \( Q \). Note that whenever \( Q \) possesses full row rank \( m \), then \( Q^\star = Q^T (QQ^T)^{-1} \) is satisfied. We use \( O \in \mathbb{R}^{m \times n} \), \( E \in \mathbb{R}^{m \times n} \), and \( I_n \in \mathbb{R}^{n \times n} \) to represent the zero matrix, the all-ones matrix, and the identity matrix of appropriate dimensions, respectively.

Let \( \mathcal{X} \) be an arbitrary Banach space with norm \( \| \cdot \|_\mathcal{X} \) and zero vector \( 0 \). Then its topological dual is denoted by \( \mathcal{X}^* \), whereas \( \langle \cdot, \cdot \rangle_\mathcal{X} : \mathcal{X} \times \mathcal{X}^* \to \mathbb{R} \) expresses the corresponding dual pairing. Recall that \( \mathcal{X} \) is called reflexive if the canonical embedding \( \mathcal{X} \ni x \mapsto (x, \cdot)_\mathcal{X} \in \mathcal{X}^{**} \) is surjective. In this case, we have \( \mathcal{X} \cong \mathcal{X}^{**} \). Obviously, any Hilbert space is reflexive due to Riesz’s representation theorem. Let \( x \in \mathcal{X} \) and \( \varepsilon > 0 \) be arbitrarily chosen. Then \( U^\varepsilon(x) \) denotes the open \( \varepsilon \)-ball around \( x \) in \( \mathcal{X} \). For a nonempty set \( A \subseteq \mathcal{X} \), we denote by \( \text{lin} A \), \( \text{conv} A \), \( \text{cone} A \), and \( \text{cl} A \) the smallest convex cone which contains \( A \), the convex hull of \( A \), the smallest convex cone which contains \( A \), and the topological closure of \( A \), respectively.

Let \( \mathcal{Y} \) be another arbitrary Banach space. The product space \( \mathcal{X} \times \mathcal{Y} \) is a Banach space as well, equipped, e.g., with the sum norm induced by \( \| \cdot \|_\mathcal{X} \) and \( \| \cdot \|_\mathcal{Y} \). In the case where \( \mathcal{Y} = \mathcal{X} \) holds, we use \( \mathcal{X}^2 := \mathcal{X} \times \mathcal{X} \), and the components of any \( x \in \mathcal{X}^2 \) are addressed by \( x_1, x_2 \in \mathcal{X} \). We exploit a similar notion for all \( n \in \mathbb{N} \) satisfying \( n \geq 3 \) as well as arbitrary sets \( A \subseteq \mathcal{X} \), i.e. \( A^n \) represents the Cartesian product of order \( n \) of \( A \).

We denote the space of bounded (or, equivalently, continuous) linear operators mapping from \( \mathcal{X} \) to \( \mathcal{Y} \) by \( \mathbb{L}[\mathcal{X}, \mathcal{Y}] \). For any operator \( F \in \mathbb{L}[\mathcal{X}, \mathcal{Y}] \), the operator \( F^* \in \mathbb{L}[\mathcal{Y}, \mathcal{X}^*] \) denotes the adjoint of \( F \). Let \( Z \) be a Banach space as well and fix an operator \( G \in \mathbb{L}[\mathcal{X}, Z] \). The linear operator \( (F, G) \in \mathbb{L}[\mathcal{X}, \mathcal{Y} \times Z] \) is defined by \( (F, G)[x] := (F[x], G[x]) \) for any \( x \in \mathcal{X} \).

A Banach space \( \mathcal{X} \) is continuously embedded into a Banach space \( \mathcal{Y} \) (\( \mathcal{X} \hookrightarrow \mathcal{Y} \) for short) if \( \mathcal{X} \subseteq \mathcal{Y} \) holds true and if the linear mapping \( \mathcal{X} \ni x \mapsto x \in \mathcal{Y} \) belongs to \( \mathbb{L}[\mathcal{X}, \mathcal{Y}] \). Furthermore, \( \mathcal{X} \) is compactly embedded into \( \mathcal{Y} \) if we have \( \mathcal{X} \hookrightarrow \mathcal{Y} \) and if the closed unit ball in \( \mathcal{X} \) is compact in \( \mathcal{Y} \).

2.2 Variational analysis

For a reflexive Banach space \( \mathcal{X} \) and a nonempty set \( A \subseteq \mathcal{X} \), we introduce the polar cone and the annihilator of \( A \) as stated below:

\[
A^\circ := \{ x^* \in \mathcal{X}^* | \forall x \in A : \langle x, x^* \rangle_{\mathcal{X}} \leq 0 \}, \quad A^\perp := \{ x^* \in \mathcal{X}^* | \forall x \in A : \langle x, x^* \rangle_{\mathcal{X}} = 0 \}.
\]

In the following lemmas, we present some helpful calculus rules for the above set operations. The first result is taken from [29, Lemma 2.1].
Lemma 2.1 Let $K, K_1, K_2 \subseteq \mathcal{X}$ be nonempty, closed, convex cones in a reflexive Banach space $\mathcal{X}$. Then we have

\begin{align*}
K^\circ &= K, \\
(K_1 + K_2)^\circ &= K_1^\circ \cap K_2^\circ, \\
(K_1 \cap K_2)^\circ &= \text{cl}(K_1^\circ + K_2^\circ), \\
K^\perp &= K \cap (-K), \\
K^\perp &= K^\circ \cap (-K)^\circ.
\end{align*}

Lemma 2.2 Let $S \subseteq \mathbb{R}^n$ be a linear subspace of $\mathbb{R}^n$ and let $Q \in \mathbb{R}^{m \times n}$ be an arbitrary matrix. Then we have

\[ \{ z \in S \mid Qz = 0 \}^\circ = Q^\top \mathbb{R}^m - S^\perp. \]

Proof The proof directly follows from the generalized Farkas lemma, see [17, Theorem 1], which yields

\[ \{ z \in S \mid Qz = 0 \}^\circ = \text{cl}(Q^\top \mathbb{R}^m + S^\circ). \]

Noting that $S$ is a subspace of $\mathbb{R}^n$, we have $S^\circ = S^\perp = -S^\perp$. Furthermore, since $Q^\top \mathbb{R}^m$ and $S^\perp$ are linear subspaces of $\mathbb{R}^n$, the same holds true for $Q^\top \mathbb{R}^m - S^\perp$ and, thus, this set is closed. This yields the claim. □

Choose $\bar{x} \in A$ arbitrarily. Then the radial cone, the tangent (or Bouligand) cone, the weak tangent cone, and the Clarke tangent cone to $A$ at $\bar{x}$ are defined as stated below:

\[
\begin{align*}
\mathcal{R}_A(\bar{x}) &= \{ d \in \mathcal{X} \mid \exists \alpha > 0 : \bar{x} + t d \in A \forall t \in (0, \alpha) \}, \\
\mathcal{T}_A(\bar{x}) &= \{ d \in \mathcal{X} \mid \exists \{d_k\} \subseteq \mathcal{X} : \exists \{t_k\} \subseteq \mathbb{R} : d_k \rightarrow d, t_k \downarrow 0, \bar{x} + t_k d_k \in A \forall k \in \mathbb{N} \}, \\
\mathcal{T}_A^w(\bar{x}) &= \{ d \in \mathcal{X} \mid \exists \{d_k\} \subseteq \mathcal{X} : \exists \{t_k\} \subseteq \mathbb{R} : d_k \rightarrow d, t_k \downarrow 0, \bar{x} + t_k d_k \in A \forall k \in \mathbb{N} \}, \\
\mathcal{T}_A^w(\bar{x}) &= \left\{ d \in \mathcal{X} \mid \forall \{x_k\} \subseteq A : x_k \rightarrow \bar{x} \text{ and } \exists \{t_k\} \subseteq \mathbb{R}, t_k \downarrow 0 \right\}.
\end{align*}
\]

Clearly, $\mathcal{R}_A(\bar{x}) \subseteq \mathcal{T}_A(\bar{x}) \subseteq \mathcal{T}_A^w(\bar{x})$ and $\mathcal{T}_A^w(\bar{x}) \subseteq \mathcal{T}_A(\bar{x})$ hold. Whenever $A$ is convex, we have

\[
\mathcal{R}_A(\bar{x}) = \text{cone}(A - \{\bar{x}\}), \quad \mathcal{T}_A^w(\bar{x}) = \mathcal{T}_A(\bar{x}) = \mathcal{T}_A^w(\bar{x}) = \text{cl}(\mathcal{R}_A(\bar{x})),
\]

and if $A$ is a convex cone, we obtain $\mathcal{R}_A(\bar{x}) = A + \text{lin}\{\bar{x}\}$ as well as the following polar relations, see [7, Example 2.62]:

\[
\mathcal{R}_A(\bar{x})^\circ = \mathcal{T}_A^w(\bar{x}) = \mathcal{T}_A(\bar{x})^\circ = \mathcal{T}_A^w(\bar{x})^\circ = A^\circ \cap \{\bar{x}\}^\perp.
\]

We define the Fréchet (or regular), the limiting (also referred to as basic or Mordukhovich), and the Clarke (or convexified) normal cone to $A$ at $\bar{x}$ as stated below:

\[
\begin{align*}
\mathcal{N}_A(\bar{x}) &= \left\{ x^* \in \mathcal{X}^* \mid \limsup_{x \rightarrow \bar{x}, x \in A} \frac{(x - \bar{x}, x^*)_{\mathcal{X}}}{\| x - \bar{x} \|_{\mathcal{X}}} \leq 0 \right\}, \\
\mathcal{N}_A^\circ(\bar{x}) &= \left\{ x^* \in \mathcal{X}^* : \exists \{x_k\} \subseteq A : \exists \{x_k^*\} \subseteq \mathcal{X}^* : x_k \rightarrow \bar{x}, x_k^* \rightarrow x^* \right\}, \\
\mathcal{N}_A^\perp(\bar{x}) &= \mathcal{T}_A^w(\bar{x})^\circ.
\end{align*}
\]

The relation $\mathcal{N}_A(\bar{x}) = \mathcal{T}_A^w(\bar{x})^\circ$ is obtained from [30, Corollary 1.11] which shows $\mathcal{N}_A(\bar{x}) = \mathcal{T}_A(\bar{x})^\circ$ whenever $\mathcal{X}$ is finite dimensional. Moreover, we have the relation $\mathcal{N}_A^\perp(\bar{x}) = \text{cl conv} \mathcal{N}_A(\bar{x})$, see [30, Theorem 3.57]. On the other hand, if $A$ is a closed, convex set, then $\mathcal{N}_A(\bar{x}) = \mathcal{N}_A^\circ(\bar{x}) = \mathcal{N}_A^\perp(\bar{x}) = \mathcal{T}_A(\bar{x})^\circ$ is always satisfied. Moreover, we obtain the following stability property in this situation.

Lemma 2.3 Let $A \subseteq \mathcal{X}$ be a nonempty, closed, convex set in a reflexive Banach space $\mathcal{X}$. Let $\{x_k\} \subseteq A$ and $\{x_k^*\} \subseteq \mathcal{X}^*$ be sequences which satisfy $x_k \rightarrow \bar{x}$, $x_k^* \rightarrow \bar{x}^*$ for some points $\bar{x} \in \mathcal{X}$ and $\bar{x}^* \in \mathcal{X}^*$. Finally, let $x_k^* \in \mathcal{N}_A(x_k)$ be valid for all $k \in \mathbb{N}$. Then $\bar{x}^* \in \mathcal{N}_A(\bar{x})$ holds true.
Proof The fact that $A$ is closed readily implies $\bar{x} \in A$. On the other hand, the convexity of $A$ leads to

$$\mathcal{N}_A(\bar{x}) = \{x^* \in X^* \mid \forall x \in A : \langle x - \bar{x}, x^* \rangle_X \leq 0\}$$

for all $\bar{x} \in A$. Thus, we have $\langle x - x_k, x_k^* \rangle_X \leq 0$ for all $x \in A$ and $k \in \mathbb{N}$. Consequently, for fixed $x \in A$,

$$\langle x - \bar{x}, x^* \rangle_X = \lim_{k \to \infty} \langle x, x^*_k \rangle_X - \lim_{k \to \infty} \langle x_k, x^*_k \rangle_X = \lim_{k \to \infty} \langle x - x_k, x^*_k \rangle_X \leq 0$$

is obtained from the continuity of the dual pairing. Since $x \in A$ was chosen arbitrarily, we finally obtain $\bar{x}^* \in \mathcal{N}_A(\bar{x})$.

2.3 Differentiation and set-valued mappings

Let $X$ and $Y$ be arbitrary Banach spaces, let $\theta : X \to Y$ be a given mapping, and fix some point $\bar{x} \in X$. If there exists a linear operator $\theta'(\bar{x}) \in L[X, Y]$ which satisfies

$$\lim_{h \to 0} \frac{\theta(\bar{x} + h) - \theta(\bar{x}) - \theta'(\bar{x})[h]}{\|h\|_X} = 0,$$

then $\theta$ is called Fréchet differentiable at $\bar{x}$ and $\theta'(\bar{x})$ is referred to as the Fréchet derivative of $\theta$ at $\bar{x}$. If the mapping $X \ni x \mapsto \theta'(x) \in L[X, Y]$ is well-defined in a neighborhood of $\bar{x}$ and continuous at this point, then $\theta$ is called continuously Fréchet differentiable at $\bar{x}$. The mapping $\theta$ is called (continuously) Fréchet differentiable if it is (continuously) Fréchet differentiable everywhere in $X$. In the special case $X = \mathbb{R}^m$ and $Y = \mathbb{R}^n$, Fréchet differentiability coincides with the common concept of differentiability. In this case, we use $\nabla \theta(\bar{x}) \in \mathbb{R}^{m \times n}$ in order to represent the Jacobian of $\theta$ at $\bar{x}$ which equals $\theta'(\bar{x})$. If $m = 1$ holds, the gradient $\nabla \theta(\bar{x}) \in \mathbb{R}^n$ is interpreted as a column vector.

Let $\varphi : X \to \mathbb{R}$ be a locally Lipschitz continuous functional and choose $\bar{x} \in X$ and $d \in X$ arbitrarily. Then the limit

$$\varphi^d(\bar{x}; d) := \lim_{x \to \bar{x}} \sup_{t \geq 0} \frac{\varphi(x + td) - \varphi(x)}{t}$$

exists and is called Clarke directional derivative of $\varphi$ at $\bar{x}$ in direction $d$. Furthermore, the nonempty, bounded, closed, and convex set

$$\partial^c \varphi(\bar{x}) := \{x^* \in X^* \mid \forall d \in X : \langle d, x^* \rangle_X \leq \varphi^d(\bar{x}; d)\}$$

denotes the Clarke subdifferential of $\varphi$ at $\bar{x}$. Calculus rules comprising a sum, scalarization, chain, and maximum rule for this tool are stated in [9].

Let $\Gamma : X \rightrightarrows Y$ be a set-valued mapping, i.e. a map which assigns to any $x \in X$ a (possibly empty) set $\Gamma(x) \subseteq Y$. The set $gph \Gamma := \{(x, y) \in X \times Y \mid y \in \Gamma(x)\}$ is called graph of $\Gamma$. Let $(\bar{x}, \bar{y}) \in gph \Gamma$ be arbitrarily chosen. We call $\Gamma$ calm at this point if there exist constants $\varepsilon > 0$, $\delta > 0$, as well as $L > 0$, such that

$$\forall x \in U_{\varepsilon}(\bar{x}) : \ \Gamma(x) \cap U_{\delta}(\bar{y}) \subseteq \Gamma(\bar{x}) + L\|x - \bar{x}\|_X B_Y$$

is satisfied. Here, $B_Y$ denotes the closed unit ball in $Y$.

2.4 Function spaces

Let $T > 0$ and $n$ be a fixed real number and a fixed positive natural number, respectively. In $C[0, T]^n$, we collect all continuous functions mapping from $[0, T]$ to $\mathbb{R}^n$. If we equip this vector space with the norm

$$\forall x \in C[0, T]^n : \ \|x\|_{C[0, T]^n} := \sup_{t \in [0, T]} \|x(t)\|_2,$$

it becomes a nonreflexive Banach space. We use $L^2(0, T)^n$ to express the Banach space of all (equivalence classes of) quadratically Lebesgue integrable functions mapping from $(0, T)$ to $\mathbb{R}^n$ which is equipped with the norm

$$\forall x \in L^2(0, T)^n : \ \|x\|_{L^2(0, T)^n} := \left(\int_0^T \|x(t)\|^2_2 dt\right)^{1/2}.$$
Exploiting the fact that $L^2(0, T)^n$ is a Hilbert space, we can use Riesz’s representation theorem to identify $L^2(0, T)^n$ and $(L^2(0, T)^n)^*$ with each other. The dual pairing in $L^2(0, T)^n$ is given by

$$
\forall x, y \in L^2(0, T)^n: \langle x, y \rangle_{L^2(0, T)^n} = \int_0^T x(t) \cdot y(t) dt.
$$

Finally, $H^1(0, T)^n$ shall comprise all functions $x \in L^2(0, T)^n$ whose components $x_1, \ldots, x_n: (0, T) \to \mathbb{R}$ possess weak derivatives such that the function $\dot{x}: (0, T) \to \mathbb{R}^n$ which is composed of the weak derivatives $\dot{x}_1, \ldots, \dot{x}_n: (0, T) \to \mathbb{R}$ belongs to $L^2(0, T)^n$. A suitable norm in $H^1(0, T)^n$ is given by

$$
\forall x \in H^1(0, T)^n: \|x\|_{H^1(0, T)^n} := \left(\|\dot{x}\|_{L^2(0, T)^n}^2 + \|x\|_{L^2(0, T)^n}^2\right)^{\frac{1}{2}}.
$$

For brevity, let us stipulate $C[0, T] := C[0, T]^1$, $L^2(0, T) := L^2(0, T)^1$, as well as $H^1(0, T) := H^1(0, T)^1$. More information on the Banach spaces introduced above can be found in Appendix A. Especially, Theorem A.1 shows that $H^1(0, T)^n \hookrightarrow L^2(0, T)^n$ and $H^1(0, T)^n \hookrightarrow C[0, T]^n$ are compact embeddings.

2.5 Constraint qualifications in Banach space programming

Let $W$ and $\mathcal{Y}$ be reflexive Banach spaces, and let $\tilde{f}: W \to \mathbb{R}$ and $\tilde{F}: W \to \mathcal{Y}$ be continuously Fréchet differentiable mappings. Moreover, we fix a nonempty, closed, convex set $\Lambda \subseteq \mathcal{Y}$. Let us consider

$$
\begin{align*}
\tilde{f}(w) &\to \min \\
\tilde{F}(w) &\in \Lambda
\end{align*}
$$

(1)

at one of its feasible points $\bar{w} \in W$.

Here, we want to briefly present some available optimality conditions and constraint qualifications which can be used to characterize locally optimal solutions of (1). The so-called (generalized) KKT conditions are satisfied at $\bar{w}$ if the following is valid:

$$
\exists \lambda \in T_\Lambda(\tilde{F}(\bar{w}))^\circ: \quad 0 = \tilde{f}'(\bar{w}) + \tilde{F}'(\bar{w})^* [\lambda],
$$

see [25]. Following the seminal papers [25, 35], the constraint qualification of Kurcyusz, Robinson, and Zowe (KRZCQ for short) is said to be satisfied at $\bar{w}$ if

$$
\tilde{F}'(\bar{w})|W - R_\Lambda(\tilde{F}(\bar{w})) = \mathcal{Y}
$$

(2)

holds. Suppose that $\bar{w}$ is a locally optimal solution of (1) where KRZCQ holds. Then the KKT conditions are valid at $\bar{w}$, see [7, Theorem 3.9].

Assume that (2) holds at $\bar{w}$. Polarizing this equation (i.e. computing the polar cone on both sides of the equality sign), we obtain that the condition

$$
\begin{align*}
0 = \tilde{F}(\bar{w})^* [\lambda], \\
\lambda \in T_\Lambda(\tilde{F}(\bar{w}))^\circ
\end{align*}
$$

(3)

is satisfied as well. It is well-known that (3) does not necessarily need to imply the validity of (2). However, if $\Lambda$ possesses a nonempty interior or $\mathcal{Y}$ is finite dimensional, then (3) is equivalent to (2), see [7, Proposition 2.97].

**Lemma 2.4** We consider $S := \{w \in W| \tilde{F}(w) \in \Lambda\}$ and some point $\bar{w} \in S$. Under any of the following three conditions, we have

$$
N_S(\bar{w}) = \tilde{F}(\bar{w})^* [T_\Lambda(\tilde{F}(\bar{w}))^\circ].
$$

1. The constraint qualification (2) is valid.
2. The constraint qualification (3) is valid and $\Lambda$ possesses a nonempty interior.
3. The constraint qualification (3) is valid and $\mathcal{Y}$ is finite dimensional.
3 Stationarity concepts and constraint qualifications for MPCCs

Proof Let us start with the proof of the lemma’s assertion under validity of (2). First, we show
\[
T^w_S(\bar{w}) = \{ d \in W \mid \bar{F}'(\bar{w})[d] \in T_A(\bar{F}(\bar{w})) \}. \tag{4}
\]
Noting that (2) holds, we obtain \( T^w_S(\bar{w}) = T_S(\bar{w}) = \{ d \in W \mid \bar{F}'(\bar{w})[d] \in T_A(\bar{F}(\bar{w})) \} \) from [7, Corollary 2.91]. Since we have \( T^w_S(\bar{w}) \subseteq T^w_S(\bar{w}) \) by definition of these cones, the inclusion \( \subseteq \) in (4) is already proven. For the proof of the converse inclusion, choose \( d \in T^w_S(\bar{w}) \setminus \{ 0 \} \) arbitrarily (for \( d = 0 \), the statement is trivial). Then we find sequences \( \{ d_k \} \subseteq W \) and \( \{ t_k \} \subseteq \mathbb{R} \) such that \( d_k \to d \) and \( t_k \downarrow 0 \) as well as \( \bar{F}(\bar{w} + t_k d_k) \in A \) hold true for all \( k \in \mathbb{N} \). The convexity of \( A \) implies that the relation
\[
\xi_k := t_k^{-1}(\bar{F}(\bar{w} + t_k d_k) - \bar{F}(\bar{w})) \in R_A(\bar{F}(\bar{w}))
\]
is satisfied for all \( k \in \mathbb{N} \). Choose \( y^* \in \mathcal{Y}^* \) arbitrarily. Then we obtain
\[
\langle \bar{F}'(\bar{w})[d], y^* \rangle_Y = \lim_{k \to \infty} \langle \bar{F}'(\bar{w})[d_k], y^* \rangle_Y
\]
\[
= \lim_{k \to \infty} \left\langle \frac{\bar{F}(\bar{w} + t_k d_k) - \bar{F}(\bar{w}) - \bar{F}'(\bar{w})[t_k d_k]}{t_k}, y^* \right\rangle_Y
\]
\[
= \lim_{k \to \infty} \left( \langle \xi_k, y^* \rangle_Y - \left\langle \frac{\|d_k\|_W}{\|t_k d_k\|_W} \bar{F}(\bar{w} + t_k d_k) - \bar{F}(\bar{w}) - \bar{F}'(\bar{w})[t_k d_k], y^* \right\rangle_Y \right).
\]
Due to the weak convergence of \( \{ d_k \} \), this sequence is bounded and, thus, \( \{ t_k d_k \} \) converges to the zero vector. Exploiting the definition of Fréchet differentiability and the inequality of Cauchy and Schwarz, we easily see
\[
\lim_{k \to \infty} \left( \frac{\|d_k\|_W}{\|t_k d_k\|_W} \langle \xi_k, y^* \rangle_Y - \langle \frac{\|d_k\|_W}{\|t_k d_k\|_W} \bar{F}(\bar{w} + t_k d_k) - \bar{F}(\bar{w}) - \bar{F}'(\bar{w})[t_k d_k], y^* \rangle_Y \right) = 0.
\]
Thus, we have \( \lim_{k \to \infty} \langle \xi_k, y^* \rangle_Y = \langle \bar{F}'(\bar{w})[d], y^* \rangle_Y \) for all \( y^* \in \mathcal{Y}^* \), i.e. \( \xi_k \) converges weakly to \( \bar{F}'(\bar{w})[d] \).
Mazur’s lemma yields
\[
\bar{F}'(\bar{w})[d] \in \text{cl conv}(\xi_k \mid k \in \mathbb{N}) \subseteq \text{cl } R_A(\bar{F}(\bar{w})) = T_A(\bar{F}(\bar{w})).
\]
This shows the other inclusion \( \subseteq \) in (4).

From above we conclude
\[
\hat{N}_S(\bar{w}) \subseteq N_S(\bar{w}) \subseteq N^w_S(\bar{w}) = T^w_S(\bar{w})^0 = T^w_S(\bar{w})^0 = \hat{N}_S(\bar{w}),
\]
i.e. \( N_S(\bar{w}) = \{ d \in W \mid \bar{F}'(\bar{w})[d] \in T_A(\bar{F}(\bar{w})) \}^0 \) is valid. Finally, we exploit the validity of (2) and the generalized Farkas lemma, see [17, Theorem 1, Lemma 3], in order to obtain
\[
N_S(\bar{w}) = \{ d \in W \mid \bar{F}'(\bar{w})[d] \in T_A(\bar{F}(\bar{w})) \}^0 = \bar{F}'(\bar{w})^* \{ T_A(\bar{F}(\bar{w})) \}^0
\]
which completes the proof of the lemma’s assertion under validity of (2).

Supposing that the constraint qualification (3) holds while \( A \) possesses a non-empty interior or \( \mathcal{Y} \) is finite dimensional, the constraint qualification (2) is valid as well, see [7, Proposition 2.97]. Thus, under the second or third set of assumptions, the lemma’s assertion holds as well. \( \square \)

3 Stationarity concepts and constraint qualifications for MPCCs

In this section, we recall some stationarity concepts and constraint qualifications for MPCCs of special structure. In terms of (OCTCC), we need to consider
\[
\tilde{f}(w) = \min \begin{cases} 
\tilde{F}(w) \in A \\
\tilde{G}(w) \geq 0 \\
\tilde{H}(w) \geq 0 \\
\tilde{G}(w) \cdot \tilde{H}(w) = 0 \end{cases} \tag{MPCC}
\]
in more detail. Here, \( \tilde{f} : W \to \mathbb{R} \), \( \tilde{F} : W \to \mathcal{Y} \), and \( \tilde{G}, \tilde{H} : W \to \mathbb{R}^k \) are assumed to be continuously Fréchet differentiable mappings between reflexive Banach spaces \( W \) and \( \mathcal{Y} \) while \( A \subseteq \mathcal{Y} \) is a nonempty, closed,
and convex set. By \( \tilde{G}_1, \ldots, \tilde{G}_k : \mathcal{W} \to \mathbb{R} \) and \( \tilde{H}_1, \ldots, \tilde{H}_k : \mathcal{W} \to \mathbb{R} \) we denote the component mappings of \( \tilde{G} \) and \( \tilde{H} \), respectively. The last three constraints in (MPCC) form the so-called complementarity constraints.

The standard MPCC, i.e. the program (MPCC) where \( \mathcal{W} = \mathbb{R}^n \), \( \mathcal{Y} = \mathbb{R}^p \), and \( \lambda = -\mathbb{R}^n_+ \) hold, has been extensively studied in the literature, see [13,14,16,21,27,32,37,41] and the references therein.

In [40], the author introduces a far more general complementarity constrained optimization problem. He considers the model (MPCC) where the complementarity condition is replaced by

\[
\tilde{G}(w) \in \tilde{K}, \quad \tilde{H}(w) \in \tilde{K}^*, \quad \langle \tilde{G}(w), \tilde{H}(w) \rangle = 0.
\]

Here, \( \tilde{K} \subseteq \mathcal{V} \) is a nonempty, closed, convex cone in a reflexive Banach space \( \mathcal{V} \) and \( \tilde{G} : \mathcal{W} \to \mathcal{V} \) as well as \( \tilde{H} : \mathcal{W} \to \mathcal{V}^* \) are continuously Fréchet differentiable mappings. The author introduces a reasonable concept of strong stationarity and states constraint qualifications implying locally optimal solutions of the underlying complementarity problem to satisfy these strong stationarity conditions. In [29], the authors continue these considerations by defining a generalized concept of weak stationarity for MPCCs in Banach spaces. The program (MPCC) may be seen somewhere in between a standard MPCC and an MPCC in Banach spaces.

It is well-known from [29, Lemma 3.1] that KRZCQ fails to hold at any feasible point of (MPCC). Hence, for some locally optimal solution \( \bar{w} \in \mathcal{W} \) of (MPCC), the corresponding KKT conditions, which take the form

\[
0 = f'(\bar{w}) + F'(\bar{w})^* \lambda \geq \sum_{i=1}^{k} [\mu_i \tilde{G}_i'(\bar{w}) + \nu_i \tilde{H}_i'(\bar{w})],
(5a)
\]

\[
\lambda \in T_\mathcal{W}(F(\bar{w}))^*,
(5b)
\]

\[
\forall i \in I^{+0}(\bar{w}) : \mu_i = 0,
(5c)
\]

\[
\forall i \in I^{0+}(\bar{w}) : \nu_i = 0,
(5d)
\]

\[
\forall i \in I^{00}(\bar{w}) : \mu_i \geq 0 \land \nu_i \geq 0,
(5e)
\]

see [40, Lemma 5.1], might turn out to be too strong necessary optimality conditions. Here, the index sets \( I^{+0}(\bar{w}) \), \( I^{0+}(\bar{w}) \), and \( I^{00}(\bar{w}) \) are defined as stated below:

\[
I^{+0}(\bar{w}) := \{ i \in \{1, \ldots, k\} \mid \tilde{G}_i(\bar{w}) > 0 \land \tilde{H}_i(\bar{w}) = 0 \},
\]

\[
I^{0+}(\bar{w}) := \{ i \in \{1, \ldots, k\} \mid \tilde{G}_i(\bar{w}) = 0 \land \tilde{H}_i(\bar{w}) > 0 \},
\]

\[
I^{00}(\bar{w}) := \{ i \in \{1, \ldots, k\} \mid \tilde{G}_i(\bar{w}) = 0 \land \tilde{H}_i(\bar{w}) = 0 \}.
\]

The set \( I^{00}(\bar{w}) \) is known as the set of biactive or degenerated indices.

In order to overcome this shortcoming, stationarity conditions which are weaker than the KKT conditions can be introduced. The following definition transfers the concepts of weak, Clarke, Mordukhovich, and strong stationarity for standard MPCCs, see e.g. [41, Definitions 2.3, 2.4, 2.6, 2.7], to our generalized model (MPCC).

**Definition 3.1** Let \( \bar{w} \in \mathcal{W} \) be a feasible point of (MPCC).

1. The point \( \bar{w} \) is called weakly stationary (W-stationary for short) for (MPCC) provided there exist multipliers \( \lambda \in \mathcal{Y}^* \) and \( \mu, \nu \in \mathbb{R}^k \) which satisfy the conditions (5a)-(5d).
2. The point \( \bar{w} \) is called Clarke stationary (C-stationary for short) for (MPCC) provided there exist multipliers \( \lambda \in \mathcal{Y}^* \) and \( \mu, \nu \in \mathbb{R}^k \) which satisfy the conditions (5a)-(5d) and

\[
\forall i \in I^{00}(\bar{w}) : \mu_i \nu_i \geq 0.
\]
3. The point \( \bar{w} \) is called Mordukhovich stationary (M-stationary for short) for (MPCC) provided there exist multipliers \( \lambda \in \mathcal{Y}^* \) and \( \mu, \nu \in \mathbb{R}^k \) which satisfy the conditions (5a)-(5d) and

\[
\forall i \in I^{00}(\bar{w}) : \mu_i \nu_i = 0 \lor (\mu_i > 0 \land \nu_i > 0).
\]
4. The point $\bar{w}$ is called strongly stationary (S-stationary for short) for (MPCC) provided there exist multipliers $\lambda \in Y^*$ and $\mu, \nu \in \mathbb{R}^k$ which satisfy the conditions (5a)-(5e).

By definition, the S-stationarity conditions defined above equal the KKT conditions of (MPCC). Furthermore, we have the relations

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

between the introduced stationarity notions.

Next, we present constraint qualifications which imply that a locally optimal solution of (MPCC) is C-, M-, or even S-stationary.

**Proposition 3.1** Let $\bar{w} \in W$ be a locally optimal solution of (MPCC) where the constraint qualification

$$0 = F'(\bar{w})^*[\lambda] - \sum_{i=1}^k [\mu_i G'_i(\bar{w}) + \nu_i H'_i(\bar{w})],$$

$$\lambda \in T_{\lambda}(F(\bar{w}))^\circ,$$

$$\forall i \in I^{+0}(\bar{w}): \mu_i = 0,$$

$$\forall i \in I^{0+}(\bar{w}): \nu_i = 0,$$

$$\forall i \in I^{00}(\bar{w}): \mu_i \nu_i \geq 0$$

(6)

is satisfied and let one of the following conditions be valid:

1. the constraint qualification (2) is valid,
2. $\Lambda$ possesses a nonempty interior,
3. $Y$ is finite dimensional.

Then $\bar{w}$ is a C-stationary point of (MPCC).

**Proof** For all $i = 1, \ldots, k$, we define a function $\varphi_i: W \to \mathbb{R}$ by means of

$$\forall w \in W: \quad \varphi_i(w) := \min\{G_i(w); H_i(w)\}.$$ 

Since the mappings $G$ and $H$ are continuously Fréchet differentiable, they are locally Lipschitz continuous. Consequently, the functions $\varphi_1, \ldots, \varphi_k$ are locally Lipschitz continuous.

Let us consider the program

$$\bar{f}(w) \to \min$$

$$\varphi_i(w) = 0 \quad i = 1, \ldots, k$$

$$w \in S$$

(7)

where we set $S := \{w \in W | F(w) \in A\}$. Obviously, (MPCC) and (7) are equivalent optimization problems. Thus, $\bar{w}$ is a locally optimal solution of (7), and we can invoke [30, Theorems 3.57 and 5.21(iii)] to find multipliers $\gamma_0 \geq 0$ and $\gamma_1, \ldots, \gamma_k \in \mathbb{R}$ which do not vanish at the same time and satisfy

$$-\gamma_0 \bar{f}(\bar{w}) \in \sum_{i=1}^k \gamma_i \partial^c \varphi_i(\bar{w}) + N_S(\bar{w}).$$

Applying [9, Propositions 2.3.3 and 2.3.12], we find

$$\forall i \in \{1, \ldots, k\}: \quad \partial^c \varphi_i(\bar{w}) = \begin{cases} \{\bar{H}'_i(\bar{w})\} & \text{if } i \in I^{+0}(\bar{w}), \\ \{\bar{G}'_i(\bar{w})\} & \text{if } i \in I^{0+}(\bar{w}), \\ \text{conv}\{\bar{G}'_i(\bar{w}), \bar{H}'_i(\bar{w})\} & \text{if } i \in I^{00}(\bar{w}). \end{cases}$$

Since the constraint qualification (6) implies the qualification condition (3), we obtain the formula $N_S(\bar{w}) = F'(\bar{w})^*[T_{\lambda}(F(\bar{w}))^\circ]$ from Lemma 2.4 under any of the additional constraint qualifications postulated above.
Consequently, there exist numbers κ₁, . . . , κₖ ∈ R and some λ′ ∈ Y ∗ such that
\[
0 = γ₀ \tilde{F}'(\bar{w}) + \tilde{F}'(\bar{w})^*[\lambda'] + \sum_{i=1}^{k} γ_i [κ_i \tilde{G}'_i(\bar{w}) + (1 - κ) \tilde{H}'_i(\bar{w})],
\]
\[
λ' ∈ T_{A}(\tilde{F}(\bar{w}))°,
\]
\[
∀i ∈ I^{+0}(\bar{w}): \ k_i = 0,
\]
\[
∀i ∈ I^{0+}(\bar{w}): \ ν_i = 1,
\]
\[
∀i ∈ I^{00}(\bar{w}): \ k_i ∈ [0,1]
\]
holds true. Let us set µ′ᵢ := −γᵢκᵢ and ν′ᵢ := γᵢ(1 − κᵢ) for all i = 1, . . . , k. Then we have
\[
0 = γ₀ \tilde{F}'(\bar{w}) + \tilde{F}'(\bar{w})^*[\lambda'] − \sum_{i=1}^{k} [µ′ᵢ \tilde{G}'_i(\bar{w}) + ν′ᵢ \tilde{H}'_i(\bar{w})],
\]
\[
λ' ∈ T_{A}(\tilde{F}(\bar{w}))°,
\]
\[
∀i ∈ I^{+0}(\bar{w}): \ µ′ᵢ = 0,
\]
\[
∀i ∈ I^{0+}(\bar{w}): \ ν′ᵢ = 0,
\]
\[
∀i ∈ I^{00}(\bar{w}): \ µ′ᵢν′ᵢ ≥ 0.
\]

Now, assume γ₀ = 0. Then the constraint qualification (6) implies λ′ = 0, µ′ = 0, and ν′ = 0. For i ∈ I^{+0}(\bar{w}), we obtain γᵢ = 0 from the definition of ν′ᵢ. Analogously, for i ∈ I^{0+}(\bar{w}), we have γᵢ = 0 from the definition of µ′ᵢ. Choose i ∈ I^{00}(\bar{w}). In the case κᵢ = 0, we have γᵢ = 0 from the definition of ν′ᵢ. On the other hand, κᵢ ∈ (0,1] leads to γᵢ = 0 by definition of µ′ᵢ. Summarizing the above considerations, we end up with γ₀ = γ₁ = · · · = γₖ = 0 which is a contradiction.

Thus, we can assume γ₀ > 0. We set λ := γ₀⁻¹λ′, µ := γ₀⁻¹µ′, and ν := γ₀⁻¹ν′ in order to see from (8) that w is a C-stationary point for (MPCC).

**Proposition 3.2** Let w ∈ W be a locally optimal solution of (MPCC) where the constraint qualification
\[
0 = \tilde{F}'(\bar{w})^*[\lambda] − \sum_{i=1}^{k} \left[ µ_i \tilde{G}'_i(\bar{w}) + ν_i \tilde{H}'_i(\bar{w}) \right],
\]
\[
λ ∈ T_{A}(\tilde{F}(\bar{w}))°,
\]
\[
∀i ∈ I^{+0}(\bar{w}): \ µ_i = 0,
\]
\[
∀i ∈ I^{0+}(\bar{w}): \ ν_i = 0,
\]
\[
∀i ∈ I^{00}(\bar{w}): \ µ_iν_i = 0 ∨ (µ_i > 0 ∧ ν_i > 0)
\]
is satisfied and let one of the following conditions be valid:

1. the constraint qualification (2) holds,
2. A possesses a nonempty interior,
3. Y is finite dimensional.

Then w is an M-stationary point of (MPCC).

**Proof** Let us define closed sets S, R ⊆ W by
\[
S := \{ w ∈ W | \tilde{F}(w) ∈ A \}, \quad R := \{ w ∈ W | (\tilde{G}(w), \tilde{H}(w)) ∈ C \}
\]
where
\[
C := \{ (a,b) ∈ R^k × R^k | a,b ≥ 0, a · b = 0 \}
\]
denotes the complementarity angle induced in R^k. Note that the feasible set of (MPCC) equals S ∩ R. Since w is a locally optimal solution of (MPCC), the necessary optimality condition −f(\bar{w}) ∈ N_{S∩R}(\bar{w}) follows from [30, Proposition 5.1].

Similar as in the proof of Proposition 3.1, we obtain N_{S}(\bar{w}) = \tilde{F}'(\bar{w})^*[T_{A}(\tilde{F}(\bar{w}))°] from the postulated assumptions. Note that C as a subset of a finite-dimensional Banach space is so-called sequentially normally compact, see [30, Definition 1.20], everywhere. Moreover, [14, Proposition 2.4] leads to
\[
N_{C}(\tilde{G}(\bar{w}), \tilde{H}(\bar{w})) = \begin{cases}
(\bar{\mu}, \bar{\nu}) ∈ R^k × R^k & \begin{cases}
\bar{\mu}_i = 0 & \text{if } i ∈ I^{+0}(\bar{w})
\bar{\nu}_i = 0 & \text{if } i ∈ I^{0+}(\bar{w})
\bar{\mu}_i\bar{\nu}_i = 0 ∨ (\bar{\mu}_i < 0 ∧ \bar{\nu}_i < 0) & \text{if } i ∈ I^{00}(\bar{w})
\end{cases}
\end{cases}
\]
and thus, the validity of the constraint qualification (9) implies that the condition

\[ 0 = \tilde{G}'(\bar{w})^*[\tilde{\mu}] + \tilde{H}'(\bar{w})^*[\tilde{\nu}], \]

holds as well. Consequently, we can invoke [30, Theorem 3.8] in order to obtain

\[ \mathcal{N}_R(\bar{w}) \subseteq \{ \tilde{G}'(\bar{w})^*[\tilde{\mu}] + \tilde{H}'(\bar{w})^*[\tilde{\nu}] \in \mathcal{W}^* | (\tilde{\mu}, \tilde{\nu}) \in \mathcal{N}_E(\tilde{G}(\bar{w}), \tilde{H}(\bar{w})) \} . \]

Moreover, [30, Theorem 3.84] and the continuous Fréchet differentiability of \( \tilde{G} \) and \( \tilde{H} \) imply that \( R \) is sequentially normally compact at \( \bar{w} \).

Due to the upper estimate of \( \mathcal{N}_R(\bar{w}) \) and the precise representation of \( \mathcal{N}_S(\bar{w}) \) derived above, the constraint qualification (9) implies \( \mathcal{N}_S(\bar{w}) \cap (-\mathcal{N}_R(\bar{w})) = \{0\} \). Recalling that \( R \) is sequentially normally compact at \( \bar{w} \), we obtain the inclusion \( \mathcal{N}_{S \cap R}(\bar{w}) \subseteq \mathcal{N}_S(\bar{w}) + \mathcal{N}_R(\bar{w}) \) from [30, Corollary 3.5]. Now, the proposition’s statement simply follows from the above considerations. This completes the proof. \( \square \)

The following proposition, which results from applying [40, Theorem 5.2] to (MPCC), presents constraint qualifications which imply locally optimal solutions of (MPCC) to be S-stationary.

**Proposition 3.3** Let \( \bar{w} \in \mathcal{W} \) be a locally optimal solution of (MPCC) and set

\[ L := \left\{ (a, b) \in \mathbb{R}^k \times \mathbb{R}^k \left| \begin{array}{ll}
  a_i = 0 & \text{if } i \in I^{0+}(\bar{w}) \cup I^{00}(\bar{w}) \\
  b_i = 0 & \text{if } i \in I^{+0}(\bar{w}) \cup I^{00}(\bar{w})
\end{array} \right. \right\}. \]

Suppose that the constraint qualifications

\[
(F'(\bar{w}), G'(\bar{w}), H'(\bar{w}))[\mathcal{W}] - \mathcal{R}_A(F(\bar{w})) \times L = \mathcal{Y} \times \mathbb{R}^k \times \mathbb{R}^k \tag{11}
\]

and

\[
\text{cl}\left( (F'(\bar{w}), G'(\bar{w}), H'(\bar{w}))[\mathcal{W}] - \mathcal{T}_A(F(\bar{w}))^\perp \right) \times L = \mathcal{Y} \times \mathbb{R}^k \times \mathbb{R}^k \tag{12}
\]

hold true. Then \( \bar{w} \) is an S-stationary point of (MPCC).

In the subsequent remark, we comment on the constraint qualifications for (MPCC) we used in Propositions 3.1, 3.2, and 3.3.

**Remark 3.1** It is easily seen that the constraint qualification (6) implies (9).

In general, the constraint qualification (12) does not imply (11) or vice versa. However, if \( \mathcal{Y} \) is finite dimensional or \( \Lambda \) possesses a nonempty interior, then (12) implies (11), see [7, Corollary 2.98]. One can easily check that in case of a standard MPCC (\( \mathcal{W} = \mathbb{R}^n \), \( \mathcal{Y} = \mathbb{R}^p \), and \( \Lambda = -\mathbb{R}_0^{p,+} \)), the constraint qualifications (11) and (12) equal MPCC-MFCQ and MPCC-LICQ (see [13, Definition 2.1] for the definitions), respectively.

Clearly, both conditions (11) and (12) are implied by

\[
(F'(\bar{w}), G'(\bar{w}), H'(\bar{w}))[\mathcal{W}] - \mathcal{T}_A(F(\bar{w}))^\perp \times L = \mathcal{Y} \times \mathbb{R}^k \times \mathbb{R}^k,
\]

and this condition holds whenever the operator \((F'(\bar{w}), G'(\bar{w}), H'(\bar{w}))\) is surjective.

Suppose that (11) is valid. Then (2) is valid as well. Moreover, computing the polar cone on both sides of the equality sign in (11) shows that the constraint qualification

\[
0 = F'(\bar{w})^*[\lambda] - \sum_{i=1}^k [\mu_i G'_i(\bar{w}) + \nu_i H'_i(\bar{w})], \\
\lambda \in \mathcal{T}_A(F(\bar{w}))^0, \\
\forall i \in I^{0+}(\bar{w}): \mu_i = 0, \\
\forall i \in I^{+0}(\bar{w}): \nu_i = 0
\]

holds true. Clearly, this condition implies the validity of (6) and (9). Thus, if \( \bar{w} \) is a locally optimal solution of (MPCC) where (11) is valid, then \( \bar{w} \) is an M-stationary point of (MPCC) by means of Proposition 3.2. This generalizes the well-known result that MPCC-MFCQ for standard MPCCs implies M-stationarity of locally optimal solutions, see [15, comments after Remark 3.1].
4 On the solution operator of linear ODEs

Let $S: L^2(0, T)^m \to H^1(0, T)^n$ be the operator which assigns to any control function $u \in L^2(0, T)^m$ the unique solution of the ODE

$$\begin{align*}
\dot{x}(t) - Ax(t) - Bu(t) &= 0 \quad \text{a.e. on } (0, T) \\
x(0) &= 0
\end{align*}$$

(13)

which appears in the constraints of (OCTCC). Due to the linearity of the system (13), $S$ is a linear mapping.

The matrix function $\Phi: [0, T] \to \mathbb{R}^{n \times n}$ shall denote the uniquely determined solution of the matrix differential equation

$$\begin{align*}
\dot{\Phi}(t) - A\Phi(t) &= 0 \quad \text{a.e. on } (0, T) \\
\Phi(0) &= I_n.
\end{align*}$$

(14)

Note that $\Phi$ is continuously differentiable on $[0, T]$. Furthermore, the images of $\Phi$ are invertible almost everywhere on $[0, T]$. Thus, it is reasonable to introduce $\Phi^{-1}(t) := \Phi(t)^{-1}$ for all $t \in [0, T]$. Exploiting $\Phi(t)\Phi^{-1}(t) = I_n$, we obtain $\dot{\Phi}(t)\Phi^{-1}(t) + \Phi(t)\dot{\Phi}^{-1}(t) = 0$ from the product rule of differentiation and, thus,

$$
\dot{\Phi}^{-1}(t) = -\Phi^{-1}(t)\dot{\Phi}(t)\Phi^{-1}(t) = -\Phi^{-1}(t)A\Phi(t)\Phi^{-1}(t) = -\Phi^{-1}(t)A
$$

(15)

for almost every $t \in [0, T]$. For detailed information on the matrix function $\Phi$ and its properties, we refer the interested reader to [2, Sections 17, 18].

Invoking [2, equation (18.14)], we can use $\Phi$ to obtain an explicit characterization of $S$:

$$
\forall u \in L^2(0, T)^m \forall t \in (0, T): \quad S[u](t) = \Phi(t) \int_0^t \Phi^{-1}(s)Bu(s)ds.
$$

(16)

In Appendix B, we show $S \in \mathbb{L}[L^2(0, T)^m, H^1(0, T)^n]$, i.e. that $S$ is continuous.

Remark 4.1 Note that the homogeneous initial condition $x(0) = 0$ is not restrictive in (OCTCC). If $x(0) = x_0$ for some $x_0 \in \mathbb{R}^n$ holds, then for any $u \in L^2(0, T)^m$, the solution $x \in H^1(0, T)^n$ of the corresponding dynamics is given by

$$
x(t) := \Phi(t) \left[ x_0 + \int_0^t \Phi^{-1}(s)Bu(s)ds \right]
$$

for almost every $t \in (0, T)$, see [2, Section 18]. More precisely, the corresponding affine solution operator results from $S$ by adding the constant shift $\Phi(\cdot)x_0$. Thus, shifting the arguments of the functions $f$, $g$, $G$, and $H$ in (OCTCC) from $x(T)$ to $x(T) + \Phi(T)x_0$, defining a new desired state by $\tilde{x}_d := \tilde{x}_d - \Phi(\cdot)x_0$, and considering the dynamics with homogeneous initial conditions, we can transfer any inhomogeneous initial condition on $x$ into a homogeneous one in (OCTCC).

We will exploit the linear operators $E: H^1(0, T)^n \to L^2(0, T)^n$ and $E_T: H^1(0, T)^n \to \mathbb{R}^n$ defined below:

$$
\forall x \in H^1(0, T)^n: \quad E[x] := x, \quad E_T[x] := x(T).
$$

(17)

Lemma 4.1 We have $E \in \mathbb{L}[H^1(0, T)^n, L^2(0, T)^n]$ and $E_T \in \mathbb{L}[H^1(0, T)^n, \mathbb{R}^n]$. Furthermore, the operators $E$ and $E_T$ are compact.

Proof Obviously, $E$ equals the natural embedding from $H^1(0, T)^n$ into $L^2(0, T)^n$ which is compact, see Theorem A.1.

We exploit Theorem A.1 once more in order to see that $H^1(0, T)^n$ is compactly embedded into $C[0, T]^n$. Particularly, we find a constant $\gamma > 0$ which satisfies

$$
\forall x \in H^1(0, T)^n: \quad \|x\|_{C[0, T]^n} \leq \gamma \|x\|_{H^1(0, T)^n}.
$$
By definition of the norm in $C[0,T]^n$, we obtain
\[ \forall x \in H^1(0,T)^n : \|E_T[x]\|_2 = \|x(T)\|_2 \leq \|x\|_{C[0,T]^n} \leq \gamma \|x\|_{H^1(0,T)^n}. \]

Thus, $E_T \in \mathbb{L}[H^1(0,T)^n, \mathbb{R}^n]$ holds true. The compactness of $H^1(0,T)^n \hookrightarrow C[0,T]^n$ implies the compactness of $E_T$. □

In this paper, we will work with the operators
\[ \mathcal{S} := E \circ S, \quad \mathcal{S}_T := E_T \circ S. \] 

Due to Lemma 4.1, we have $\mathcal{S} \in \mathbb{L}[L^2(0,T)^m, L^2(0,T)^n]$ and $\mathcal{S}_T \in \mathbb{L}[L^2(0,T)^m, \mathbb{R}^n]$. The following technical lemma will be helpful for the derivation of the stationarity concepts for (OCTCC).

**Lemma 4.2** For a real vector $b \in \mathbb{R}^n$, we have
\[ \forall u \in L^2(0,T)^m : b \cdot \mathcal{S}_T[u] = \int_0^T (B^\top \zeta(t)) \cdot u(t)dt. \]

Here, $\zeta \in H^1(0,T)^n$ denotes the unique solution of the linear ODE
\[ \begin{align*}
\zeta(t) + A^\top \zeta(t) &= 0 \quad a.e. \text{ on } (0,T) \\
\zeta(T) - b &= 0.
\end{align*} \]

Especially, we have $\mathcal{S}_T^*[b] = B^\top \zeta$.

**Proof** We exploit (16) to obtain
\[ b \cdot \mathcal{S}_T[u] = b \cdot \Phi(T) \int_0^T \Phi^{-1}(s)Bu(s)ds = \int_0^T (B^\top \Phi^{-1}(s)^\top \Phi(T)^\top b) \cdot u(s)ds. \]

Let us set $\zeta(t) := \Phi^{-1}(t)^\top \Phi(T)^\top b$ for all $t \in [0,T]$. Then we have $\zeta(T) = b$, and
\[ \zeta(t) = \Phi^{-1}(t)^\top \Phi(T)^\top b = -A^\top \Phi^{-1}(t)^\top \Phi(T)^\top b = -A^\top \zeta(t) \]

follows from (15) for any $t \in (0,T)$. The last statement of the lemma follows from
\[ \langle u, \mathcal{S}_T^*[b] \rangle_{L^2(0,T)^m} = b \cdot \mathcal{S}_T[u] = \int_0^T (B^\top \zeta(t)) \cdot u(t)dt = \langle u, B^\top \zeta \rangle_{L^2(0,T)^m} \]

which holds for all $u \in L^2(0,T)^m$. This completes the proof. □

**Lemma 4.3** For given $v \in L^2(0,T)^n$, we have $\mathcal{S}^*[v] = B^\top \eta$ where $\eta \in H^1(0,T)^n$ denotes the unique solution of the linear ODE
\[ \begin{align*}
\dot{\eta}(t) + A^\top \eta(t) + v(t) &= 0 \quad a.e. \text{ on } (0,T) \\
\eta(T) &= 0.
\end{align*} \]

**Proof** For $u \in L^2(0,T)^m$ and $v \in L^2(0,T)^n$, we exploit integration by parts and (16) in order to obtain
\[ \langle u, \mathcal{S}^*[v] \rangle_{L^2(0,T)^m} = \langle v, \mathcal{S}[u] \rangle_{L^2(0,T)^n} = \int_0^T v(t) \cdot \left( \Phi(t) \int_0^t \Phi^{-1}(s)Bu(s)ds \right)dt \]
\[ = \left( \int_0^T \Phi(t)^\top v(t)dt \right) \cdot \left( \int_0^T \Phi^{-1}(s)Bu(s)ds \right) \\
- \int_0^T \left( \int_0^t \Phi(s)^\top v(s)ds \right) \cdot (\Phi^{-1}(t)Bu(t))dt \\
= \int_0^T \left( \int_t^T \Phi(s)^\top v(s)ds \right) \cdot (\Phi^{-1}(t)Bu(t))dt \\
= \int_0^T \left[ B^\top \Phi^{-1}(t)^\top \left( \int_t^T \Phi(s)^\top v(s)ds \right) \right] \cdot u(t)dt. \]
For any $t \in [0, T]$, we set
\[ \eta(t) := \Phi^{-1}(t)^\top \left( \int_t^T \phi(s)^\top v(s) \, ds \right). \]
Obviously, $\eta(T) = 0$ is valid. Moreover, for any $t \in (0, T)$, we obtain
\[
\begin{align*}
\dot{\eta}(t) &= \dot{\Phi}^{-1}(t)^\top \left( \int_t^T \phi(s)^\top v(s) \, ds \right) + \Phi^{-1}(t)^\top (-\phi(t)^\top v(t)) \\
&= -A^\top \Phi^{-1}(t)^\top \left( \int_t^T \phi(s)^\top v(s) \, ds \right) - v(t) = -A^\top \eta(t) - v(t)
\end{align*}
\]
from the product rule of differentiation and (15). This shows the claim. □

Lemma 4.2 allows us to characterize the (generalized) differentiability properties of a certain class of functions which will appear later on.

**Lemma 4.4** Let $\theta : \mathbb{R}^m \to \mathbb{R}^n$ be a given function, define $\Theta : L^2(0, T)^m \to \mathbb{R}^n$ by $\Theta := \theta \circ S_T$, fix some point $\bar{u} \in L^2(0, T)^m$, and set $\bar{x} := S[\bar{u}]$. Then the following assertions hold true.

1. Assume that $\theta$ is continuously differentiable at $\bar{x}(T)$. Then $\Theta$ is continuously Fréchet differentiable at $\bar{u}$ and the formula
\[
\forall u \in L^2(0, T)^m : \quad \Theta'(\bar{u})[u] = \nabla \theta(\bar{x}(T)) S_T[u]
\]
is valid. Furthermore, for any $a \in \mathbb{R}^n$, we have $\Theta'(\bar{u})^*[a] = B^\top \zeta$ where $\zeta \in H^1(0, T)^n$ is the unique solution of the linear ODE
\[
\dot{\zeta}(t) + A^\top \zeta(t) = 0 \quad \text{a.e. on } (0, T) \\
\zeta(T) - \nabla \theta(\bar{x}(T))^\top a = 0.
\]

2. Let $s = 1$ hold true. Assume that $\theta$ is locally Lipschitz continuous at $\bar{x}(T)$. Then $\Theta$ is locally Lipschitz continuous at $\bar{u}$. Furthermore, for any subgradient $\Xi \in \partial \Theta(\bar{u})$, there exists $\xi \in \partial \theta(\bar{u})$ such that the formula
\[
\forall u \in L^2(0, T)^m : \quad \langle u, \Xi \rangle_{L^2(0, T)^m} = \xi \cdot S_T[u]
\]
is valid.

**Proof** Since $S_T$ is a continuous linear operator, it is a continuously Fréchet differentiable mapping.

Let us start with the proof of the lemma’s first assertion. Its first part is a straightforward consequence of the chain rule for Fréchet differentiable functions, see [39, Satz 2.20]. The desired representation of $\Theta'(\bar{u})^*[a] = S_T[\nabla \theta(\bar{x}(T))^\top a]$ follows from Lemma 4.2.

For the proof of the lemma’s second assertion, we only need to apply Clarke’s chain rule, see [9, Theorem 2.3.10]. □

**5 Existence of optimal solutions for (OCTCC)**

Before we can start with the derivation of necessary optimality conditions for (OCTCC), it has to be considered under which conditions this problem actually possesses a globally optimal solution. Note that standard arguments from optimal control, see [39], do not seem to be suited for the discussion of (OCTCC) since the feasible set of this problem is likely to be nonconvex.

**Theorem 5.1** In addition to Assumption 1.1, let the function $f : \mathbb{R}^n \to \mathbb{R}$ be bounded from below. Furthermore, assume that (OCTCC) possesses at least one feasible point. Then (OCTCC) possesses a globally optimal solution.
Proof Due to the fact that the feasible set of \((\text{OCTCC})\) is nonempty, we find a so-called minimizing sequence \(\{(x_k, u_k)\} \subseteq H^1(0,T)^n \times L^2(0,T)^m\) of \((\text{OCTCC})\), i.e. all the points \((x_k, u_k), k \in \mathbb{N}\), are feasible to \((\text{OCTCC})\) and

\[
\lim_{k \to \infty} \left( f(x_k(T)) + \frac{1}{2} \|x_k - x_d\|^2_{L^2(0,T)^n} + \frac{\sigma}{2} \|u_k - u_d\|^2_{L^2(0,T)^m} \right) = \alpha
\]

is satisfied where \(\alpha\) denotes the infimal value of \((\text{OCTCC})\). Since \(f\) is bounded from below, the same holds true for the objective of \((\text{OCTCC})\) and, consequently, we have \(\alpha \in \mathbb{R}\). This implies that \(\{u_k\} \subseteq L^2(0,T)^m\) is bounded. Due to the reflexivity of \(L^2(0,T)^n\), there is a weakly convergent subsequence of \(\{u_k\}\) with weak limit point \(\bar{u} \in L^2(0,T)^m\), and w.l.o.g. we can assume \(u_k \to \bar{u}\).

Let \(\mathcal{S}\) denote the solution operator of the linear ODE \((13)\). From Section 4 we have that \(\mathcal{S}\) belongs to \(\mathbb{L}[L^2(0,T)^m, H^1(0,T)^n]\). We set \(\bar{x} := \mathcal{S}[\bar{u}]\). The linearity and continuity of \(\mathcal{S}\) imply that the set

\[
U_{ad} := \big\{ u \in L^2(0,T)^m \mid \mathcal{C}[u](t) + \mathcal{D}u(t) \leq 0 \text{ a.e. on } (0,T) \big\}
\]

is closed as well as convex. Particularly, \(U_{ad}\) is weakly sequentially closed. Due to \(x_k = \mathcal{S}[u_k]\) for all \(k \in \mathbb{N}\), we obtain \(\{u_k\} \subseteq U_{ad}\) and, consequently, \(\bar{u} \in U_{ad}\).

From \(u_k \to \bar{u}\) in \(L^2(0,T)^m\) we deduce \(x_k \to \bar{x}\) in \(H^1(0,T)^n\). Invoking the compactness of the evaluation operator \(\mathcal{E}_T\) defined in \((17)\), see Lemma 4.1, \(x_k(T) \to \bar{x}(T)\) in \(\mathbb{R}^n\) is obtained. The continuity of the functions \(g, H\), and \(H\) as well as the closedness of \(K\) imply that \((\bar{x}, \bar{u})\) is a feasible point of \((\text{OCTCC})\).

Clearly, the mapping \(L^2(0,T)^m \ni u \mapsto \frac{\sigma}{2} \|u - u_d\|^2_{L^2(0,T)^m} \in \mathbb{R}\) is continuous as well as convex and, thus, weakly lower semicontinuous. On the other hand, from \(x_k \to \bar{x}\) in \(H^1(0,T)^n\) and the compactness of the embedding \(H^1(0,T)^n \hookrightarrow L^2(0,T)^n\), see Theorem A.1, we obtain \(x_k \to \bar{x}\) in \(L^2(0,T)^m\). Finally, we have \(x_k(T) \to \bar{x}(T)\) in \(\mathbb{R}^n\) from above. Thus, the continuity of \(f\) and \(L^2(0,T)^m \ni x \mapsto \frac{1}{2} \|x - x_d\|^2_{L^2(0,T)^m} \in \mathbb{R}\) yields

\[
f(\bar{x}(T)) + \frac{1}{2} \|\bar{x} - d\|^2_{L^2(0,T)^n} + \frac{\sigma}{2} \|\bar{u} - u_d\|^2_{L^2(0,T)^m}
\]

\[
= \lim_{k \to \infty} f(x_k(T)) + \lim_{k \to \infty} \frac{1}{2} \|x_k - x_d\|^2_{L^2(0,T)^n} + \frac{\sigma}{2} \|u_k - u_d\|^2_{L^2(0,T)^m}
\]

\[
\leq \liminf_{k \to \infty} f(x_k(T)) + \lim_{k \to \infty} \frac{1}{2} \|x_k - x_d\|^2_{L^2(0,T)^n} + \liminf_{k \to \infty} \frac{\sigma}{2} \|u_k - u_d\|^2_{L^2(0,T)^m}
\]

\[
= \liminf_{k \to \infty} \left[ f(x_k(T)) + \frac{1}{2} \|x_k - x_d\|^2_{L^2(0,T)^n} + \frac{\sigma}{2} \|u_k - u_d\|^2_{L^2(0,T)^m} \right] = \alpha
\]

Now, the feasibility of \((\bar{x}, \bar{u})\) to \((\text{OCTCC})\) implies that this point is a globally optimal solution of \((\text{OCTCC})\). This completes the proof. \(\square\)

Note that the above proof is possible without exploiting the compactness of \(H^1(0,T)^n \hookrightarrow L^2(0,T)^n\); one can use the weak lower semicontinuity of the mapping \(L^2(0,T)^m \ni x \mapsto \frac{1}{2} \|x - x_d\|^2_{L^2(0,T)^m} \in \mathbb{R}\) and the fact that for real sequences \(\{a_k\}\) and \(\{b_k\}\), \(\liminf_{k \to \infty} a_k + \liminf_{k \to \infty} b_k \leq \liminf_{k \to \infty} [a_k + b_k]\) is valid.

6 Stationarity concepts for \((\text{OCTCC})\)

In this section, we state reasonable concepts of stationarity for \((\text{OCTCC})\). Therefore, we eliminate the state variable \(x\) from \((\text{OCTCC})\) first.

Obviously, by means of

\[
\forall v \in L^2(0,T)^n \forall t \in (0,T) : \quad \mathcal{C}[v](t) := C v(t),
\]

\[
\forall u \in L^2(0,T)^m \forall t \in (0,T) : \quad \mathcal{D}[u](t) := D u(t),
\]

we define operators \(C \in \mathbb{L}[L^2(0,T)^n, L^2(0,T)^n]\) and \(D \in \mathbb{L}[L^2(0,T)^m, L^2(0,T)^n]\).

Using the operator \(\mathcal{E}_T\) defined in \((18)\), we introduce mappings \(\mathcal{F} : L^2(0,T)^m \to \mathbb{R}, \mathcal{G} : L^2(0,T)^m \to \mathbb{R}^k\), and \(\mathcal{H} : L^2(0,T)^m \to \mathbb{R}^k\) as given below:

\[
\mathcal{F} := f \circ \mathcal{S}_T, \quad \mathcal{G} := g \circ \mathcal{S}_T, \quad \mathcal{H} := H \circ \mathcal{S}_T.
\]
By $\tilde{G}_1, \ldots, \tilde{G}_k: L^2(0,T)^m \rightarrow \mathbb{R}$ and $\tilde{H}_1, \ldots, \tilde{H}_k: L^2(0,T)^m \rightarrow \mathbb{R}$, we denote the component functions of $\tilde{G}$ and $\tilde{H}$, respectively.

Finally, we define a closed, convex cone $C \subseteq L^2(0,T)^q$ as stated below:

$$C := \left\{ w \in L^2(0,T)^q \mid w(t) \leq 0 \ a.e. \ on \ (0,T) \right\}.$$  

(20)

Then (OCTCC) is equivalent to the MPCC

$$\begin{align*}
\hat{f}(u) + \frac{1}{2} \left[ d^2 \mathbb{S}(u) - x_d \right]_{L^2(0,T)^n} + \frac{1}{2} \left[ u - u_d \right]_{L^2(0,T)^m} & \rightarrow \min_u \\
(C \circ \mathbb{S} + \mathbb{D})[u] - f & \in C \\
\tilde{g}(u) & \in K \\
\tilde{G}(u) & \geq 0 \\
\tilde{H}(u) & \geq 0 \\
\tilde{G}(u) \cdot \tilde{H}(u) & = 0
\end{align*}$$

(21)

where only the control function $u$ is left as a variable. Note that the operator $\mathbb{S}$ is defined in (18).

Let us check how the concepts of W-, C-, M-, and S-stationarity from Definition 3.1 look like for the program (21). Due to the equivalence of (21) and (OCTCC), we can identify the W-, C-, M-, and S-stationarity conditions of (21) with the W-, C-, M-, and S-stationarity conditions of (OCTCC), respectively.

**Proposition 6.1** Let $\bar{u} \in L^2(0,T)^m$ be a feasible point of surrogate problem (21) and define $\bar{x} := \mathbb{S}[\bar{u}]$. Then $\bar{u}$ is $W$-stationary for problem (21) if and only if there exist $p \in H^1(0,T)^n$, $q \in L^2(0,T)^q$, $\lambda \in \mathbb{R}^n$, and $\mu, \nu \in \mathbb{R}^k$ which satisfy the following conditions:

$$\begin{align*}
0 = \dot{p}(t) + A^T p(t) + C^T q(t) + \tilde{x}(t) - \tilde{x}_d(t) & \quad a.e. \ on \ (0,T), \\
p(T) = \nabla \hat{f}(\tilde{x}(T)) + \tilde{g}(\tilde{x}(T))^\top \lambda - \sum_{i=1}^k [\mu_i \nabla G_i(\tilde{x}(T)) + \nu_i \nabla H_i(\tilde{x}(T))], \\
0 = \sigma(\bar{u}(t) - u_d(t)) + B^T p(t) + D^T \tilde{g}(t) & \quad a.e. \ on \ (0,T), \\
\bar{q}(t) & \geq 0 \quad a.e. \ on \ (0,T), \\
0 = \bar{q}(t) \cdot (C \tilde{x}(T) + D \bar{u}(t) - f(t)) & \quad a.e. \ on \ (0,T), \\
\lambda & \in T_K(g(\tilde{x}(T)))^0, \\
\forall i \in I^{0+}(\bar{u}) : \mu_i & = 0, \\
\forall i \in I^{0+}(\bar{u}) : \nu_i & = 0.
\end{align*}$$

(22)

(22a)

(22b)

(22c)

(22d)

(22e)

(22f)

(22g)

(22h)

Here, the index sets $I^{0+}(\bar{u})$, $I^{0+}(\bar{u})$, and $I^{00}(\bar{u})$ are defined as stated below:

$$I^{0+}(\bar{u}) := \left\{ i \in \{1, \ldots, k \} \mid G_i(\tilde{x}(T)) > 0 \land H_i(\tilde{x}(T)) = 0 \right\},$$

$$I^{0+}(\bar{u}) := \left\{ i \in \{1, \ldots, k \} \mid G_i(\tilde{x}(T)) = 0 \land H_i(\tilde{x}(T)) > 0 \right\},$$

$$I^{00}(\bar{u}) := \left\{ i \in \{1, \ldots, k \} \mid G_i(\tilde{x}(T)) = 0 \land H_i(\tilde{x}(T)) = 0 \right\}.$$  

(23)

Proof Due to Definition 3.1, $\bar{u}$ is $W$-stationary for (21) if and only if there are Lagrange multipliers $\bar{q} \in T_C((C \circ \mathbb{E})[\bar{x}] + \mathbb{D}[\bar{u}] - f)^\circ$, $\lambda \in T_K(\tilde{g}(\bar{u}))^\circ$, and $\mu, \nu \in \mathbb{R}^k$ which satisfy (22g), (22h), and

$$\begin{align*}
0 = \tilde{f}'(\bar{u}) + \mathbb{S}^* (\mathbb{E}[x] - x_d) & \quad + \sigma(\bar{u} - u_d) + (\mathbb{S}^* \circ C^* + D^*)[\bar{q}] \\
& \quad + \tilde{g}'(\bar{u})^* [\lambda] - \sum_{i=1}^k [\mu_i \tilde{G}_i'(\bar{u}) + \nu_i \tilde{H}_i'(\bar{u})] \\
& \quad + \tilde{s}^* (\mathbb{E}[x] - x_d) + C^* [\bar{q}] + \sigma(\bar{u} - u_d) + D^* [\bar{q}] \\
& \quad + \tilde{f}'(\bar{u}) + \tilde{g}'(\bar{u})^* [\lambda] - \sum_{i=1}^k [\mu_i \tilde{G}_i'(\bar{u}) + \nu_i \tilde{H}_i'(\bar{u})] \\
& \quad + \tilde{g}'(\bar{u})^* [\lambda] - \sum_{i=1}^k [\mu_i \tilde{G}_i'(\bar{u}) + \nu_i \tilde{H}_i'(\bar{u})].
\end{align*}$$

(24)

Here, we used [39, Satz 2.20] for the computation of the objective's Fréchet derivative. Applying [7, Example 2.64], we easily see that for a function $g \in L^2(0,T)^q$, the relation $g \in T_C((C \circ \mathbb{E})[\bar{x}] + \mathbb{D}[\bar{u}] - f)^\circ$ is equivalent to the validity of (22d) and (22e). By definition of $\tilde{g}$, we have $\lambda \in T_K(g(\tilde{x}(T)))^\circ$, i.e. (22f) holds.
Using the definition of the adjoint operator, the representations
\[ \forall w \in L^2(0, T)^n \forall t \in (0, T): \quad C^* [w](t) = C^T w(t), \quad D^*[w](t) = D^T w(t) \]
follow easily.

Now, we can apply Lemma 4.3 in order to see that (24) is equivalent to
\[ 0 = B^T \eta + \sigma(\bar{u} - u_d) + D^T g + \tilde{f}(\bar{u}) + \tilde{g}(\bar{u})^*[\lambda] - \sum_{i=1}^{k} [\mu_i \tilde{G}_i(\bar{u}) + \nu_i \tilde{H}_i(\bar{u})] \]
where \( \eta \in H^1(0, T)^n \) is the unique solution of the linear ODE
\[ \dot{\eta}(t) + A^T \eta(t) + \bar{x}(t) - x_d(t) + C^T g(t) = 0 \quad \text{a.e. on (0, T)} \]
\[ \eta(T) = 0. \]

By means of Lemma 4.4, (25) is equivalent to
\[ 0 = B^T \eta + \sigma(\bar{u} - u_d) + D^T g + B^T \zeta, \]
where \( \zeta \in H^1(0, T)^n \) is the unique solution of the linear ODE
\[ \zeta(t) + A^T \zeta(t) = 0 \quad \text{a.e. on (0, T)} \]
\[ \zeta(T) - \left( \nabla f(\bar{x}(T)) + \nabla g(\bar{x}(T))^T \lambda \right. \\
\left. - \sum_{i=1}^{k} [\mu_i \nabla G_i(\bar{x}(T)) + \nu_i \nabla H_i(\bar{x}(T))] \right) = 0. \]

Let us define a function \( p \in H^1(0, T)^n \) by means of \( p := \eta + \zeta \). The above considerations show that the conditions (22a)-(22c) are satisfied as well. This completes the proof. \( \square \)

The system (22) comprises the classical elements appearing in the necessary optimality systems of ODE control: the adjoint equation (22a), transversality conditions (22b), and Pontryagin’s (linearized) Maximum Principle (22c), see [34]. Since the complementarity condition in (OCTCC) only affects the terminal condition on the state function, the only difference to optimality conditions for ODE control problems with terminal constraints show up in the transversality conditions (22b) and the presence of additional multipliers \( \mu, \nu \in \mathbb{R}^k \) which are characterized in (22g) and (22h).

Note that the adjoint state \( p \in H^1(0, T)^n \) which appears in the system (22) can be interpreted as the Lagrange multiplier which is associated with the ODE constraint in (OCTCC).

Since the problems (OCTCC) and (21) are equivalent, we obtain the following corollary from Proposition 6.1 and Definition 3.1.

**Corollary 6.1** Let \( (\bar{x}, \bar{u}) \in H^1(0, T)^n \times L^2(0, T)^m \) be a feasible point of (21). Then the following assertions hold.

1. The point \( (\bar{x}, \bar{u}) \) is W-stationary for (OCTCC) if and only if there exist multipliers \( p \in H^1(0, T)^n \), \( g \in L^2(0, T)^q \), \( \lambda \in \mathbb{R}^s \), and \( \mu, \nu \in \mathbb{R}^k \) which satisfy (22).
2. The point \( (\bar{x}, \bar{u}) \) is C-stationary for (OCTCC) if and only if there exist multipliers \( p \in H^1(0, T)^n \), \( g \in L^2(0, T)^q \), \( \lambda \in \mathbb{R}^s \), and \( \mu, \nu \in \mathbb{R}^k \) which satisfy (22) and
\[ \forall i \in I^0(\bar{u}): \quad \mu_i \geq 0. \]
3. The point \( (\bar{x}, \bar{u}) \) is M-stationary for (OCTCC) if and only if there exist multipliers \( p \in H^1(0, T)^n \), \( g \in L^2(0, T)^q \), \( \lambda \in \mathbb{R}^s \), and \( \mu, \nu \in \mathbb{R}^k \) which satisfy (22) and
\[ \forall i \in I^0(\bar{u}): \quad \mu_i \nu_i = 0 \vee (\mu_i > 0 \wedge \nu_i > 0). \]
4. The point \( (\bar{x}, \bar{u}) \) is S-stationary for (OCTCC) if and only if there exist multipliers \( p \in H^1(0, T)^n \), \( g \in L^2(0, T)^q \), \( \lambda \in \mathbb{R}^s \), and \( \mu, \nu \in \mathbb{R}^k \) which satisfy (22) and
\[ \forall i \in I^0(\bar{u}): \quad \mu_i \geq 0 \wedge \nu_i \geq 0. \]
7 Necessary optimality conditions for (OCTCC)

In this section, we want to present constraint qualifications ensuring that a given locally optimal solution of (OCTCC) satisfies the C-, M-, or S-stationarity conditions we derived in Corollary 6.1. Furthermore, we illustrate the obtained results using an intuitive numerical example.

In addition to our standing Assumption 1.1, we need another assumption in order to proceed.

**Assumption 7.1** The matrix $D \in \mathbb{R}^{q \times m}$ is assumed to possess full row rank $q < m$. Furthermore, $Y \in \mathbb{R}^{m \times (m-q)}$ is a matrix with full column rank $m-q$ such that $DY = O$ holds true, i.e. the columns of $Y$ generate the null space of $D$. If no mixed control-state constraints are formulated, we set $Y := I_m$.

7.1 Some preparations

Here, we present some results which will be helpful later on when we are going to derive applicable constraint qualifications for (OCTCC).

**Lemma 7.1** The operator $C \circ S + D \in \mathbb{L}[L^2(0,T)^m, L^2(0,T)^q]$, where $C$ and $D$ are defined in (19) while $S$ is defined in (18), is surjective.

**Proof** Choose $w \in L^2(0,T)^q$ arbitrarily and consider $(C \circ S + D)[u] = w$. Exploiting the definition of the appearing operators, this linear system is equivalent to

$$\begin{align*}
\dot{x}(t) - Ax(t) - Bu(t) &= 0 \quad \text{a.e. on } (0,T) \\
x(0) &= 0 \\
C x(t) + Du(t) &= w(t) \quad \text{a.e. on } (0,T).
\end{align*}$$

Let us consider $\bar{u}_x(t) := D^1(w(t) - C x(t))$ for arbitrary functions $x \in H^1(0,T)^n$ and all $t \in (0,T)$. Putting $\bar{u}_x$ into the linear ODE leads to

$$\begin{align*}
\dot{x}(t) - (A - BD^1 C)x(t) &= BD^1 w(t) \quad \text{a.e. on } (0,T) \\
x(0) &= 0.
\end{align*}$$

Using the theory of linear ODEs, see [2], this problem possesses a unique solution $\bar{x} \in H^1(0,T)^n$. Now it is easy to see that $(\bar{x}, \bar{u}_x)$ provides a solution of (26), i.e. $\bar{u}_x$ satisfies $(C \circ S + D)[\bar{u}_x] = w$. \hfill \Box

**Lemma 7.2** Suppose that the matrix

$$[BY \ (A - BD^1 C)BY \ \ldots \ (A - BD^1 C)^{n-1}BY] \in \mathbb{R}^{m \times (m-q)}$$

possesses full row rank $n$. Then $(C \circ S + D, S_T) \in \mathbb{L}[L^2(0,T)^m, L^2(0,T)^q \times \mathbb{R}^n]$ is a surjective operator. Here, $S_T$ is defined in (18).

**Proof** For arbitrary $w \in L^2(0,T)^n$ and $b \in \mathbb{R}^n$, we consider the linear equation $(C \circ S + D, S_T)[u] = (w, b)$. Again, we plug in the definitions of the appearing operators to transfer this equation equivalently into the linear system

$$\begin{align*}
\dot{x}(t) - Ax(t) - Bu(t) &= 0 \quad \text{a.e. on } (0,T) \\
x(0) &= 0 \\
x(T) - b &= 0 \\
C x(t) + Du(t) &= w(t) \quad \text{a.e. on } (0,T).
\end{align*}$$

Now, we introduce $\bar{u}_{x,v}(t) := Y v(t) + D^1(w(t) - C x(t))$ for $x \in H^1(0,T)^n$ as well as $v \in L^2(0,T)^{m-q}$ and all $t \in (0,T)$. By definition of $Y$, see Assumption 7.1, we have $D \bar{u}_{x,v}(t) = w(t) - C x(t)$ for all $t \in (0,T)$. We put $\bar{u}_{x,v}$ into the linear ODE which appears in (29) in order to obtain

$$\begin{align*}
\dot{x}(t) - (A - BD^1 C)x(t) - BY v(t) &= BD^1 w(t) \quad \text{a.e. on } (0,T) \\
x(0) &= 0 \\
x(T) - b &= 0.
\end{align*}$$

Let $\bar{x} \in H^1(0,T)^n$ be the unique solution of the linear ODE (27) and consider

$$\dot{x}(t) - (A - BD^1C)x(t) - BYv(t) = 0 \quad \text{a.e. on } (0,T)$$

$$x(0) = 0$$

$$x(T) - (b - \bar{x}(T)) = 0.$$  \hfill (31)

Note that the matrix defined in (28) is the so-called controllability matrix of (31). Since it is assumed to possess full row rank $n$, the system (31) is controllable due to the famous Kalman Theorem, see [3, Theorem 4.1]. Consequently, there exists a pair $(x', v') \in H^1(0,T)^n \times L^2(0,T)^{m-q}$ which solves (31). For a detailed introduction to the theory of controllability of linear systems, we refer the interested reader to [3].

Finally, it is easy to check that $(\bar{x} + x', v')$ provides a solution of (30). Thus, the pair $(\bar{x} + x', \bar{u}_{x+x',v'})$ solves (29), i.e. $\bar{u}_{x+x',v'}$ satisfies $(C \circ \bar{S} + D, S_T)[\bar{u}_{x+x',v'}] = (w, b)$. This completes the proof. $\square$

**Lemma 7.3** Let $(\bar{x}, \bar{u}) \in H^1(0,T)^n \times L^2(0,T)^m$ be a feasible point of (OCTCC) where the constraint qualification

$$0 = \nabla g(\bar{x}(T))^\top \lambda,$$

$$\lambda \in T_K(g(\bar{x}(T))^q)$$

is valid. Furthermore, suppose that the matrix (28) possesses full row rank $n$. Then the constraint qualification

$$(C \circ \bar{S} + D, \bar{g}'(\bar{u}))[L^2(0,T)^m] - \{0\} \times R_K(\bar{g}(\bar{u})) = L^2(0,T)^{q} \times \mathbb{R}^c$$

for (21) holds.

**Proof** First, we note that (32) is equivalent to

$$\nabla g(\bar{x}(T))[\mathbb{R}^n] - R_K(g(\bar{x}(T))) = \mathbb{R}^c,$$

see Section 2.5. Now, choose $(w, c) \in L^2(0,T)^{q} \times \mathbb{R}^c$ arbitrarily. Then we find vectors $b \in \mathbb{R}^n$ and $\xi \in R_K(g(\bar{x}(T)))$ such that $c = \nabla g(\bar{x}(T))b - \xi = c$ holds. Next, let us consider the operator equation

$$(C \circ \bar{S} + D, S_T)[\bar{u}] = (w, b).$$

Due to Lemma 7.2, it possesses a solution $\bar{u} \in L^2(0,T)^m$. We note

$$\bar{g}'(\bar{u})[\bar{u}] - \xi = \nabla g(\bar{x}(T))S_T[\bar{u}] - \xi = \nabla g(\bar{x}(T))b - \xi = c,$$

see Lemma 4.4. Consequently, $(C \circ \bar{S} + D, \bar{g}'(\bar{u}))[\bar{u}] - (0, \xi) = (w, c)$ holds true. Noting $\xi \in R_K(\bar{g}(\bar{u}))$ by definition of $\bar{g}$, we see that the constraint qualification (33) holds true. $\square$

**Lemma 7.4** Let $\Theta: \mathbb{R}^n \rightarrow \mathbb{R}^q$ be a given function and $S \subseteq \mathbb{R}^q$ be a nonempty, closed set. Define $\Theta: L^2(0,T)^m \rightarrow \mathbb{R}^q$ by $\Theta := \Theta \circ S_T$. Let $\bar{u} \in L^2(0,T)^m$ be fixed and set $\bar{x} := S[\bar{u}]$. Assume that the constraint qualification

$$0 = \nabla \Theta(\bar{x}(T))^\top \lambda,$$

$$\lambda \in S$$

is valid and that the matrix (28) possesses full row rank $n$. Then the constraint qualification

$$0 = (S^* \circ C^* + D^*)[g] + \Theta'(\bar{u})^*[\lambda],$$

$$(g, \lambda) \in L^2(0,T)^{q} \times S$$

holds as well.

**Proof** Due to the full row rank assumption on the matrix (28), the operator $(C \circ \bar{S} + D, S_T)$ is surjective, i.e. the system

$$0 = (S^* \circ C^* + D^*)[g] + S_T^*[b],$$

$$(g, b) \in L^2(0,T)^{q} \times \mathbb{R}^n$$

possesses only the trivial solution $(g, b) = (0, 0)$.
Now, choose $g \in L^2(0,T)^q$ and $\lambda \in \mathbb{R}^n$ which solve the system

$$
0 = (\mathcal{S}^* \circ \mathcal{C}^* + \mathcal{D}^*)[g] + \Theta'(\bar{u})^* [\lambda],
$$

$$(g, \lambda) \in L^2(0,T)^q \times S.$$  

Due to Lemma 4.4, this is equivalent to

$$
0 = (\mathcal{S}^* \circ \mathcal{C}^* + \mathcal{D}^*)[g] + \mathbb{S}_T^T \left[ \nabla \theta(\bar{x}(T))^\top \lambda \right],
$$

$$(g, \lambda) \in L^2(0,T)^q \times S.$$  

Invoking the above considerations, we obtain $\bar{g} = 0$ and $\nabla \theta(\bar{x}(T))^\top \lambda = 0$. Since we have $\lambda \in S$, $\lambda = 0$ follows from (34). Thus, we have $(g, \lambda) = (0,0)$, i.e. the constraint qualification (35) is valid as well. \(\square\)

7.2 Necessary optimality conditions

Now, we are well-prepared to state necessary optimality conditions for (OCTCC). In our subsequent proofs we exploit the following strategy: The local optimality of $(\bar{x}, \bar{u}) \in H^1(0,T)^n \times L^2(0,T)^m$ for (OCTCC) implies that $\bar{u} \in L^2(0,T)^m$ is a locally optimal solution of (21). We exploit the MPCC constraint qualifications introduced in Section 3 in order to ensure that the C-, M-, and S-stationarity conditions are necessary optimality conditions for (21). As we mentioned in Section 6, these stationarity notions can be interpreted as the C-, M-, and S-stationarity conditions of (OCTCC).

For a feasible point $(\bar{x}, \bar{u}) \in H^1(0,T)^n \times L^2(0,T)^m$ of (OCTCC), we are going to exploit the index sets $I^{00}(\bar{u})$, $I^{0+}(\bar{u})$, and $I^{00}(\bar{u})$ defined in (23).

First, we examine under which constraint qualification the C-stationarity conditions provide a necessary optimality criterion for (OCTCC).

**Theorem 7.2** Let $(\bar{x}, \bar{u}) \in H^1(0,T)^n \times L^2(0,T)^m$ be a locally optimal solution of (OCTCC) where the following constraint qualification holds:

$$
0 = \nabla g(\bar{x}(T))^\top \lambda - \sum_{i=1}^k [\mu_i \nabla G_i(\bar{x}(T)) + \nu_i \nabla H_i(\bar{x}(T))], \\
\lambda \in T_K(g(\bar{x}(T))^\top), \\
\forall i \in I^{00}(\bar{u}) : \mu_i = 0, \\
\forall i \in I^{0+}(\bar{u}) : \nu_i = 0, \\
\forall i \in I^{00}(\bar{u}) : \mu_i \nu_i \geq 0
$$

Furthermore, assume that the matrix (28) possesses full row rank $n$. Then $(\bar{x}, \bar{u})$ is a C-stationary point of (OCTCC).

**Proof** First, note that $\bar{u}$ is a locally optimal solution of (21). Let us invoke Proposition 3.1.

Clearly, the constraint qualification (36) implies the validity of (32). Due to Lemma 7.3, the constraint qualification (33) holds as well. This qualification condition implies (2) for the MPCC (21).

Moreover, Lemma 7.4 shows that the constraint qualification

$$
0 = (\mathcal{S}^* \circ \mathcal{C}^* + \mathcal{D}^*)[g] + \tilde{g}'(\bar{u})^* [\lambda] - \sum_{i=1}^k [\mu_i \tilde{G}_i(\bar{u}) + \nu_i \tilde{H}_i(\bar{u})],
$$

$$(g, \lambda) = 0 = \mathcal{S}^* \circ \mathcal{C}^* + \mathcal{D}^*, \\
\lambda \in T_K(\tilde{g}(\bar{u}))^\top, \\
\forall i \in I^{00}(\bar{u}) : \mu_i = 0, \\
\forall i \in I^{0+}(\bar{u}) : \nu_i = 0, \\
\forall i \in I^{00}(\bar{u}) : \mu_i \nu_i \geq 0
$$

is satisfied. Note that this qualification condition implies (6) for the MPCC (21).

Due to Proposition 3.1, $\bar{u}$ is a C-stationary point of (21) and, thus, $(\bar{x}, \bar{u})$ is a C-stationary point of (OCTCC). \(\square\)
Following the same proof strategy as used for the validation of Theorem 7.2 (we only need to exploit Proposition 3.2 instead of Proposition 3.1), we can show that locally optimal solutions of (OCTCC) are M-stationary points under suitable constraint qualifications.

**Theorem 7.3** Let \( (\bar{x}, \bar{u}) \in H^1(0,T)^n \times L^2(0,T)^m \) be a locally optimal solution of (OCTCC) where the following constraint qualification holds:

\[
0 = \nabla g(\bar{x}(T))^\top \lambda - \sum_{i=1}^k [\mu_i \nabla G_i(\bar{x}(T)) + \nu_i \nabla H_i(\bar{x}(T))],
\]

\( \lambda \in T_K(g(\bar{x}(T))^\circ), \)

\( \forall i \in I^0(\bar{u}): \mu_i = 0, \)

\( \forall i \in I^0(\bar{u}): \nu_i = 0, \)

\( \forall i \in I^0(\bar{u}): \mu_i \nu_i = 0 \lor (\mu_i > 0 \land \nu_i > 0) \)

Furthermore, assume that the matrix (28) possesses full row rank \( n \). Then \( (\bar{x}, \bar{u}) \) is an M-stationary point of (OCTCC).

Observe that condition (37) is weaker than (36) but implies the M-stationarity conditions which are stronger than the C-stationarity conditions.

From the theory of finite-dimensional MPCCs we expect any locally optimal solution of (OCTCC) where the boundary constraints form the union of a finite number of polyhedral sets to be M-stationary, see [14, Theorem 3.5 and Propositions 3.7, 3.8]. The corresponding result is stated in the following theorem.

**Theorem 7.4** Let \( (\bar{x}, \bar{u}) \in H^1(0,T)^n \times L^2(0,T)^m \) be a locally optimal solution of (OCTCC). Suppose that the functions \( g, G, \) and \( H \) are affine while the set \( K \) is polyhedral. Finally, assume that the matrix (28) possesses full row rank \( n \). Then \( (\bar{x}, \bar{u}) \) is an M-stationary point of (OCTCC).

**Proof** Clearly, \( \bar{u} \) is a locally optimal solution of (21). Let us define

\[ C := \{ z \in \mathbb{R}^n | g(z) \in K, \ (G(z), H(z)) \in \mathcal{C} \}, \]

see (10) for the definition of the set \( \mathcal{C} \). Note that \( \mathcal{C} \) is the union of finitely many polyhedral sets. The feasible set \( M \subseteq L^2(0,T)^m \) of (21) can be represented in the form

\[ M = \{ u \in L^2(0,T)^m | (C \circ S + D, S_T)[u] - (f, 0) \in C \times C \} \]

where \( C \) is given in (20).

Applying [30, Proposition 5.1], we obtain

\[ -\bar{p}'(\bar{u}) - \mathcal{S}(E[\bar{x}] - x_d) - \mathcal{S}(\bar{u} - \bar{u}_d) \in \mathcal{N}_M(\bar{u}). \]

Due to the theorem's assumptions, \( (C \circ S + D, S_T) \) is surjective, see Lemma 7.2. Hence, [30, Proposition 1.2, Theorem 1.17] yield

\[ \mathcal{N}_M(\bar{u}) = (C \circ S + D, S_T)\mathcal{N}_M((C \circ S + D, S_T)[\bar{u}]\).
\]

Let us introduce the multifunction \( T: \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^n \) given as stated below for any \( \alpha \in \mathbb{R}^n \) and \( \beta, \gamma \in \mathbb{R}^k; \)

\( T(\alpha, \beta, \gamma) := \{ z \in \mathbb{R}^n | g(z) + \alpha \in K, \ (G(z) + \beta, H(z) + \gamma) \in \mathcal{C} \}. \)

Due to the postulated assumptions, \( T \) is a polyhedral multifunction (i.e. its graph is the finite union of polyhedral sets) and, hence, calm at \( (0,0,0, \bar{x}(T)) \), see [32, Theorem 2.4] or [36, Proposition 1]. That is why [22, Proposition 3.4] leads to

\[ \mathcal{N}_C(\bar{x}(T)) \subseteq \{ \nabla g(\bar{x}(T))^\top \lambda + \sum_{i=1}^k [\mu_i \nabla G_i(\bar{x}(T)) + \nu_i \nabla H_i(\bar{x}(T))] \} \in \mathbb{R}^n | \lambda \in T_K(g(\bar{x}(T))^\circ), \ (\bar{\mu}, \bar{\nu}) \in \mathcal{N}_C(G(\bar{x}(T)), H(\bar{x}(T))) \}. \]
From the proof of Proposition 3.2 we obtain a formula for \( \mathcal{N}_E(G(\bar{x}(T)), H(\bar{x}(T))) \). Moreover, we easily see
\[
\mathcal{N}_{C+\{f\}}((C \circ E)[\bar{x}] + D[\bar{u}]) = T_{C+\{f\}}((C \circ E)[\bar{x}] + D[\bar{u}] - f^0)
\]
due to the convexity of \( C + \{f\} \). We already mentioned
\[
T_C((C \circ E)[\bar{x}] + D[\bar{u}] - f^0) = \{ q \in L^2(0,T)^q \mid q \text{ satisfies (22d), (22e)} \}
\]
in the proof of Proposition 6.1.

Putting the above observations together, we find \( q \in L^2(0,T)^q, \lambda \in \mathbb{R}^\kappa, \) and \( \mu, \nu \in \mathbb{R}^k \) which satisfy (22d) - (22h) and
\[
\forall i \in I^{00}(\bar{u}) : \quad \mu_i > 0 \lor (\mu_i > 0 \land \nu_i > 0)
\]
as well as
\[
0 = \bar{f}'(\bar{u}) + \mathbf{S}'[E[\bar{x}] - x_4] + \sigma(\bar{u} - u_4) + (\mathbf{S}^* \circ C^* + D^*)[\mathcal{g}]
\[
+ \mathbf{S}^*_H \left[ \nabla g(\bar{x}(T))^\top \lambda - \sum_{i=1}^k [\mu_i \nabla G_i(\bar{x}(T)) + \nu_i \nabla H_i(\bar{x}(T))] \right]
\]
\[
= \mathbf{S}^*_H [E[\bar{x}] - x_4 + C^*[\mathcal{g}] + \sigma(\bar{u} - u_4) + D^*[\mathcal{g}]
\[
+ \bar{f}'(\bar{u}) + \bar{g}'(\bar{u})^* [\lambda] - \sum_{i=1}^k [\mu_i \bar{G}'_i(\bar{u}) + \nu_i \bar{H}'_i(\bar{u})],
\]
see Lemma 4.4. The remaining part of the proof, i.e. the validation of (22a) - (22c), parallels the argumentation in the proof of Proposition 6.1. This shows that \((\bar{x}, \bar{u})\) is an M-stationary point of (OCTCC).

\[
\text{Next, we present conditions implying S-stationarity of locally optimal solutions of (OCTCC).}
\]

**Theorem 7.5** Let \((\bar{x}, \bar{u}) \in H^1(0,T)^n \times L^2(0,T)^m\) be a locally optimal solution of (OCTCC) where the following constraint qualification holds:
\[
\begin{align*}
0 &= \nabla g(\bar{x}(T))^\top \lambda - \sum_{i=1}^k [\mu_i \nabla G_i(\bar{x}(T)) + \nu_i \nabla H_i(\bar{x}(T))], \\
\lambda &\in T_K(g(\bar{x}(T)))^o - T_K(g(\bar{x}(T)))^o, \\
\forall i \in I^{00}(\bar{u}) : \mu_i = 0, \\
\forall i \in I^{+0}(\bar{u}) : \nu_i = 0
\end{align*}
\]
\[
\Rightarrow \left\{ \begin{array}{c}
\lambda = 0, \\
\mu = 0, \\
\nu = 0
\end{array} \right. \quad (38)
\]
Furthermore, assume that the matrix (28) possesses full row rank \( n \). Then \((\bar{x}, \bar{u})\) is an S-stationary point of (OCTCC).

**Proof** For the proof, we want to use Proposition 3.3. Again, we start our argumentation with the observation that \( \bar{u} \) is a locally optimal solution of (21). Clearly, the constraint qualification (38) is equivalent to
\[
\begin{align*}
(\lambda, \mu, \nu) &= (0, 0, 0), \\
0 &= \nabla g(\bar{x}(T))^\top \lambda - \nabla g(\bar{x}(T))^\top \mu - \nabla H(\bar{x}(T))^\top \nu, \\
\lambda &= T_K(g(\bar{x}(T)))^o - T_K(g(\bar{x}(T)))^o, \\
\forall i \in I^{+0}(\bar{u}) : \mu_i = 0, \\
\forall i \in I^{0+}(\bar{u}) : \nu_i = 0
\end{align*}
\]
We compute the polar cone on both sides of the equation, see Lemmas 2.1 and 2.2, in order to obtain
\[
(\nabla g(\bar{x}(T)), \nabla G(\bar{x}(T)), \nabla H(\bar{x}(T)))^\top [\mathbb{R}^\kappa] - T_K(g(\bar{x}(T)))^o \times L = \mathbb{R}^\kappa \times \mathbb{R}^k \times \mathbb{R}^k
\]
where \( L \subseteq \mathbb{R}^k \times \mathbb{R}^k \) is the subspace defined below:
\[
L := \left\{ (a, b) \in \mathbb{R}^k \times \mathbb{R}^k \mid \begin{array}{l}
a_i = 0 \quad \text{if } i \in I^{0+}(\bar{u}) \cup I^{00}(\bar{u}) \\
b_i = 0 \quad \text{if } i \in I^{+0}(\bar{u}) \cup I^{00}(\bar{u}) \end{array} \right\}.
\]
In the following, we will show that the constraint qualification
\[
(C \circ \mathcal{S} + D, \tilde{g}'(\tilde{u}), \tilde{H}'(\tilde{u}))[L^2(0,T)^n]
- \{0\} \times T_K(\tilde{g}(\tilde{u}))^\perp \times L = L^2(0,T)^q \times \mathbb{R}^\kappa \times \mathbb{R}^k \times \mathbb{R}^k
\]
is valid since, by means of Proposition 3.3 and Remark 3.1, it is sufficient for \(\tilde{u}\) to be an S-stationary point of (OC). i.e. for \((\tilde{x}, \tilde{u})\) to be an S-stationary point of (OCTCC).

Let us choose \(w \in L^2(0,T)^q\), \(c_g \in \mathbb{R}^\kappa\), and \(c_g, c_H \in \mathbb{R}^k\) arbitrarily. Since (39) is valid, we find some \(b \in \mathbb{R}^n\), \(\xi \in T_K(\tilde{g}(\tilde{x}(T)))^\perp \) and \((l_G, l_H) \in L\) which satisfy
\[
\nabla g(\tilde{x}(T))b - \xi = c_g, \quad \nabla G(\tilde{x}(T))b - l_G = c_G, \quad \nabla H(\tilde{x}(T))b - l_H = c_H.
\]
Now, we can invoke Lemma 7.2 in order to find a function \(\tilde{u}\) which is a solution of the linear equation \((C \circ \mathcal{S} + D, S_T)[\tilde{u}] = (w, b)\). By construction, we have
\[
(C \circ \mathcal{S} + D)[\tilde{u}] = w,
\]
\[
\tilde{g}'(\tilde{u})[\tilde{u}] - \xi = \nabla g(\tilde{x}(T))S_T[\tilde{u}] - \xi = \nabla g(\tilde{x}(T))b - \xi = c_g,
\]
\[
\tilde{H}'(\tilde{u})[\tilde{u}] - l_H = \nabla H(\tilde{x}(T))S_T[\tilde{u}] - l_H = \nabla H(\tilde{x}(T))b - l_H = c_H.
\]
Since \(w, c_g, c_G, c_H\) were arbitrarily chosen, (40) is valid. As mentioned earlier, this shows that \((\tilde{x}, \tilde{u})\) is an S-stationary point of (OCTCC). \(

The following corollary is a direct consequence of Theorem 7.5.

**Corollary 7.1** Let \((\tilde{x}, \tilde{u}) \in H^1(0,T)^n \times L^2(0,T)^m\) be a locally optimal solution of (OCTCC). Assume that the gradients
\[
\nabla g_i(\tilde{x}(T)) \quad i = 1, \ldots, \kappa,
\]
\[
\nabla G_i(\tilde{x}(T)) \quad i \in I^{0+}(\tilde{u}) \cup I^{00}(\tilde{u}),
\]
\[
\nabla H_i(\tilde{x}(T)) \quad i \in I^{+}(\tilde{u}) \cup I^{00}(\tilde{u})
\]
are linearly independent where \(g_1, \ldots, g_\kappa : \mathbb{R}^n \to \mathbb{R}\) denote the component functions of \(g\). Furthermore, assume that the matrix (28) possesses full row rank \(n\). Then \((\tilde{x}, \tilde{u})\) is an S-stationary point of (OCTCC).

The following example is included to visualize the theory obtained above.

**Example 7.1** For \(n = m = 2, q = 0,\) and \(T = \ln 2\), we consider the problem
\[
\frac{1}{2}(\|x_1 - 1\|_{L^2(0,T)}^2 + \|x_2\|_{L^2(0,T)}^2)
+ \frac{1}{2}(\|u_1\|_{L^2(0,T)}^2 + \|u_2\|_{L^2(0,T)}^2) \to \min_{x,u}
\]
\[
\dot{x}_1(t) - u_1(t) = 0 \quad \text{a.e. on } (0,T)
\]
\[
\dot{x}_2(t) - u_2(t) = 0 \quad \text{a.e. on } (0,T)
\]
\[
x_1(0) = 0
\]
\[
x_2(0) = 0
\]
\[
x_1(T) - x_2(T) \leq 0
\]
\[
x_1(T) \geq 0
\]
\[
x_2(T) \geq 0
\]
\[
x_1(T)x_2(T) = 0.
\]
One can easily see that the terminal constraints are equivalent to \(x_1(T) = 0\) and \(x_2(T) \geq 0\). Hence, the optimal solution \((\bar{x}, \bar{u})\) of (41) satisfies \(\bar{x}_2 = \bar{u}_2 = 0\), while \((\bar{x}_1, \bar{u}_1)\) solves the convex program
\[
\frac{1}{2}\|x_1 - 1\|_{L^2(0,T)}^2 + \frac{1}{2}\|u_1\|_{L^2(0,T)}^2 \to \min_{x_1,u_1}
\]
\[
\dot{x}_1(t) - u_1(t) = 0 \quad \text{a.e. on } (0,T)
\]
\[
x_1(0) = x_1(T) = 0.
\]
Standard arguments from optimal control (apply \cite[Theorem 5.22]{23}) lead to
\[ \forall t \in (0, T): \quad \dot{x}_1(t) = \frac{1}{3} \sinh(t) - \cosh(t) + 1 \quad \dot{u}_1(t) = \frac{1}{3} \cosh(t) - \sinh(t). \]

We have the relations \( I^{0+}(\bar{u}) = I^{0+}(\bar{u}) = \emptyset \) and \( I^{00}(\bar{u}) = \{1\} \). Clearly, the constraint qualification (37) is satisfied, whereas (38) fails to hold at \((\bar{x}, \bar{u})\). Additionally, the matrix (28) corresponding to (41) is given by
\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix}
\]
and, thus, possesses full row rank 2.

Due to Theorem 7.3, \((\bar{x}, \bar{u})\) is an M-stationary point of (41). Hence, there need to exist multipliers \(p \in H^1(0, T)^2 \) and \(\lambda, \mu, \nu \in \mathbb{R}^n\) which satisfy
\[
0 = p_1(t) + \frac{1}{3} \sinh(t) - \cosh(t) \quad \text{a.e. on } (0, T), \quad \text{(42a)}
\]
\[
0 = p_2(t) \quad \text{a.e. on } (0, T), \quad \text{(42b)}
\]
\[
p_1(T) = \lambda - \mu, \quad \text{(42c)}
\]
\[
p_2(T) = -\lambda - \nu, \quad \text{(42d)}
\]
\[
0 = \frac{1}{3} \cosh(t) - \sinh(t) + p_1(t) \quad \text{a.e. on } (0, T), \quad \text{(42e)}
\]
\[
0 = p_2(t) \quad \text{a.e. on } (0, T), \quad \text{(42f)}
\]
\[
\lambda \geq 0, \quad \text{(42g)}
\]
\[
\mu \nu = 0 \lor (\mu > 0 \land \nu > 0) \quad \text{(42h)}
\]

almost everywhere on \((0, T), \) see Corollary 6.1. From (42e) and (42f) we obtain the relations \(p_1 = -\bar{u}_1\) and \(p_2 = 0\) which do not contradict (42a) and (42b), respectively. Hence, (42c) and (42d) reduce to \(\frac{1}{3} = \lambda - \mu\) and \(0 = \lambda + \nu\). Together with (42g) and (42h) we obtain \((\lambda, \mu, \nu) \in \{(0, -\frac{1}{3}, 0), (\frac{1}{3}, 0, -\frac{1}{3})\}\), i.e. the M-stationarity conditions are satisfied. Note that these multipliers do not solve the system of S-stationarity since either \(\mu\) or \(\nu\) is strictly negative. The fact that \((\bar{x}, \bar{u})\) is an M-stationary point of (41) is a direct consequence of Theorem 7.4 as well.

8 Scholtes’s global relaxation scheme for (OCTCC)

In \cite{21}, the authors present an overview of relaxation techniques for common finite-dimensional MPCCs and discuss their theoretical and numerical properties. Since the complementarity constraint arising in (OCTCC) is of the same type, it should be possible to adapt these results here. Clearly, a discussion of all these procedures goes beyond the scope of the paper. Hence, we only consider the relaxation scheme of Scholtes, introduced in \cite{38} originally. However, it seems to be reasonable from the subsequent proofs that all the results obtained in \cite{21} can be adapted to (OCTCC).

For some sequence \(\{\alpha^r\} \subseteq \mathbb{R}^+\) of regularization parameters monotonically decreasing to zero, we consider the relaxed problem
\[
f(x(T)) + \frac{1}{2} \|x - x_d\|_{L^2(0, T)}^2 + \frac{1}{2} \|u - u_d\|_{L^2(0, T)}^2 \rightarrow \min_{x, u}
\]
\[
\dot{x}(t) - Ax(t) - Bu(t) = 0 \quad \text{a.e. on } (0, T),
\]
\[
x(0) = 0
\]
\[
Cx(t) + Du(t) - f \leq 0 \quad \text{a.e. on } (0, T)
\]
\[
g(x(T)) \leq 0
\]
\[
G(x(T)) \geq 0
\]
\[
H(x(T)) \geq 0
\]
\[
G_i(x(T))H_i(x(T)) \leq \alpha^r \quad i = 1, \ldots, k
\]

(OCTCC(\(\alpha^r\)))
whose feasible set will be denoted by $M(\alpha')$ for any $r \in \mathbb{N}$. Note that for simplicity, we fix $K := -\mathbb{R}^n_{0}^{+}$ in this section. The possible appearance of additional terminal equality constraints can be handled without any difficulties. Proceeding as in [21, Section 3.1], we introduce some index sets for a point $(x, u) \in M(\alpha')$:

\[
I^g(u) := \{i \in \{1, \ldots, k\} \mid g_i(x(T)) = 0\},
\]

\[
I^G(u) := \{i \in \{1, \ldots, k\} \mid G_i(x(T)) = 0\},
\]

\[
I^H(u) := \{i \in \{1, \ldots, k\} \mid H_i(x(T)) = 0\},
\]

\[
I^{GH}(u, \alpha') := \{i \in \{1, \ldots, k\} \mid G_i(x(T))H_i(x(T)) = \alpha'\}.
\]

Observe that due to $x = S[u]$, these index sets only depend on $u$ and (possibly) $\alpha'$, i.e. the above notation is consistent.

In order to get our convergence result as $\alpha' \searrow 0$, it will be necessary to compute a sequence of KKT points of the programs (OCTCC($\alpha'$)) as $r \to \infty$. Therefore, we present the KKT conditions of (OCTCC($\alpha'$)) in the subsequent lemma. Its proof mainly parallels the one of Proposition 6.1 and, thus, is omitted.

**Lemma 8.1** For some $\alpha' > 0$, a feasible point $(\bar{x}, \bar{u}) \in H^1(0, T)^n \times L^2(0, T)^m$ of (OCTCC($\alpha'$)) is a KKT point of the latter problem if and only if there exist $p \in H^1(0, T)^n$, $\varphi \in L^2(0, T)^q$, $\lambda \in \mathbb{R}^n$, and $\mu, \nu, \xi \in \mathbb{R}^k$ which satisfy the conditions (22a), (22c)-(22e), and

\[
p(T) = \nabla f(\bar{x}(T)) + \nabla g(\bar{x}(T))\top\lambda
\]

\[
- \sum_{i=1}^{k} (\mu_i - \xi_i H_i(\bar{x}(T)))\nabla G_i(\bar{x}(T)) + (\nu_i - \xi_i G_i(\bar{x}(T)))\nabla H_i(\bar{x}(T)) \bigg],
\]

\[
\lambda \geq 0, \forall i \notin I^g(\bar{u}) : \lambda_i = 0,
\]

\[
\mu \geq 0, \forall i \notin I^G(\bar{u}) : \mu_i = 0,
\]

\[
\nu \geq 0, \forall i \notin I^H(\bar{u}) : \nu_i = 0,
\]

\[
\xi \geq 0, \forall i \notin I^{GH}(\bar{u}, \alpha') : \xi_i = 0.
\]

Now, we can formulate the main result of this section.

**Theorem 8.1** Let Assumption 7.1 be valid and let $\{\alpha'\} \subseteq \mathbb{R}^+$ be a scalar sequence monotonically decreasing to zero. For any $r \in \mathbb{N}$, let $(x^r, u^r) \in M(\alpha')$ be a KKT point of (OCTCC($\alpha'$)). Additionally, assume $u^r \to \bar{u}$ in $L^2(0, T)^m$, set $\bar{x} := S[\bar{u}]$, and let the constraint qualification

\[
0 = \nabla g(\bar{x}(T))\top\lambda - \sum_{i=1}^{k} [\mu_i \nabla G_i(\bar{x}(T)) + \nu_i \nabla H_i(\bar{x}(T))],
\]

\[
\lambda \geq 0, \forall i \notin I^g(\bar{u}) : \lambda_i = 0,
\]

\[
\forall i \in I^g(\bar{u}) : \mu_i = 0,
\]

\[
\forall i \in I^H(\bar{u}) : \nu_i = 0
\]

be satisfied. Furthermore, suppose that the matrix (28) possesses full row rank $n$. Then $(\bar{x}, \bar{u})$ is a $C$-stationary point of (OCTCC) with $K = -\mathbb{R}^n_{0}^{+}$.

**Proof** First, we note that $u^r \to \bar{u}$ in $L^2(0, T)^m$ and $x^r = S[u^r]$ for all $r \in \mathbb{N}$ imply $x^r \to \bar{x}$ in $H^1(0, T)^n$. Furthermore, we have $H^1(0, T)^n \to C[0, T]^n$ from Theorem A.1 which yields $x^r(T) \to \bar{x}(T)$ in $\mathbb{R}^n$. Thus, the limit point $(\bar{x}, \bar{u})$ is feasible to (OCTCC).

Fix some $r \in \mathbb{N}$. Since $(x^r, u^r)$ is a KKT point of (OCTCC($\alpha'$)), $u^r$ is a KKT point of

\[
\tilde{f}(u) + \frac{1}{2}\|S[u] - x_4\|^2_{L^2(0, T)^n} + \frac{1}{2}\|u - u_4\|^2_{L^2(0, T)^m} \to \min_u
\]

\[
(C \circ S + D)[u] - f \in C
\]

\[
\tilde{g}(u) \leq 0
\]

\[
\tilde{G}(u) \geq 0
\]

\[
\tilde{H}(u) \geq 0
\]

\[
\tilde{G}_i(u)\tilde{H}_i(u) \leq \alpha'^r \quad i = 1, \ldots, k,
\]

\[
\frac{1}{2}\|S[u] - x_4\|^2_{L^2(0, T)^n} + \frac{1}{2}\|u - u_4\|^2_{L^2(0, T)^m} \to \min_u
\]

\[
(C \circ S + D)[u] - f \in C
\]

\[
\tilde{g}(u) \leq 0
\]

\[
\tilde{G}(u) \geq 0
\]

\[
\tilde{H}(u) \geq 0
\]

\[
\tilde{G}_i(u)\tilde{H}_i(u) \leq \alpha'^r \quad i = 1, \ldots, k,
\]
see Section 6 for the definition of all appearing mappings, operators, and sets. Let us set
\[ S := \{ u \in L^2(0, T)^m \mid (C \circ \mathcal{S} + D)[u] - f \in C \}. \]
Invoking Lemmas 2.4, 4.4, 7.1, and 8.1, we obtain the existence of multipliers \( v^r \in \mathcal{N}_S(u^r) \), \( \lambda^r \in \mathbb{R}^k \), and \( \mu^r, \nu^r, \xi^r \in \mathbb{R}^k \) which satisfy
\[
0 = S^* [E[x^r] - x_d] + \sigma(u^r - u_d) + v^r \\
+ S^*_f \left[ \nabla f(x^r(T)) + \nabla g(x^r(T))^\top \lambda^r \\
- \sum_{i=1}^k (\mu^r_i - \xi^r_i H_i(x^r(T))) \nabla G_i(x^r(T)) \\
- \sum_{i=1}^k (\nu^r_i - \xi^r_i G_i(x^r(T))) \nabla H_i(x^r(T)) \right] \\
\text{as well as}
\]
\[
\lambda^r \geq 0 \land \forall i \notin I^G(u^r): \lambda^r_i = 0, \\
\mu^r \geq 0 \land \forall i \notin I^G(u^r): \mu^r_i = 0, \\
\nu^r \geq 0 \land \forall i \notin I^H(u^r): \nu^r_i = 0, \\
\xi^r \geq 0 \land \forall i \notin I^{GH}(u^r, \alpha^r): \xi^r_i = 0.
\]
Clarity from \( \alpha^r > 0 \) we obtain \( I^G(u^r) \cap I^{GH}(u^r, \alpha^r) = \emptyset \) as well as \( I^H(u^r) \cap I^{GH}(u^r, \alpha^r) = \emptyset \), i.e. the multipliers \( \mu^r_i \) and \( \xi^r_i \) (\( \nu^r_i \) and \( \xi^r_i \), respectively) cannot be positive at the same time for all indices \( i = 1, \ldots, k \). Moreover, for sufficiently large \( r \in \mathbb{N} \), we obtain \( I^G(u^r) \subseteq I^G(\bar{u}) \), \( I^H(u^r) \subseteq I^H(\bar{u}) \), and \( I^{GH}(u^r) \subseteq I^{GH}(\bar{u}) \cup I^{GH}(\bar{\bar{u}}) \) from \( x^r(T) \to \bar{x}(T) \).

Let us introduce
\[
\bar{\mu}^r_i := \begin{cases} 
\mu^r_i & \text{if } \mu^r_i > 0, \\
-\xi^r_i H_i(x^r(T)) & \text{if } \xi^r_i > 0 \land i \notin I^0(\bar{u}), \\
0 & \text{otherwise}
\end{cases}
\]
and
\[
\bar{\nu}^r_i := \begin{cases} 
\nu^r_i & \text{if } \nu^r_i > 0, \\
-\xi^r_i G_i(x^r(T)) & \text{if } \xi^r_i > 0 \land i \notin I^{0+}(\bar{u}), \\
0 & \text{otherwise}.
\end{cases}
\]
Then we can restate (45) as
\[
0 = S^*[E[x^r] - x_d] + \sigma(u^r - u_d) + v^r \\
+ S^*_f \left[ \nabla f(x^r(T)) + \nabla g(x^r(T))^\top \lambda^r \\
- \sum_{i=1}^k (\bar{\mu}^r_i \nabla G_i(x^r(T)) + \bar{\nu}^r_i \nabla H_i(x^r(T))) \\
+ \sum_{i \in I^{0+}(\bar{u})} \xi^r_i H_i(x^r(T)) \nabla G_i(x^r(T)) \\
+ \sum_{i \in I^{0+}(\bar{u})} \xi^r_i G_i(x^r(T)) \nabla H_i(x^r(T)) \right].
\]
Suppose that the sequence \( \{ (\lambda^r, \bar{\mu}^r, \bar{\nu}^r, \xi^r_{I^{0+}(\bar{u})}) \} \) is unbounded. Let us define
\[
(\bar{\lambda}^r, \bar{\mu}^r, \bar{\nu}^r, \bar{\xi}^r) := \frac{(\lambda^r, \bar{\mu}^r, \bar{\nu}^r, \xi^r_{I^{0+}(\bar{u})})}{\left\| (\lambda^r, \bar{\mu}^r, \bar{\nu}^r, \xi^r_{I^{0+}(\bar{u})}) \right\|_2}.
\]
Then the sequence \( \{ (\bar{\lambda}^r, \bar{\mu}^r, \bar{\nu}^r, \bar{\xi}^r) \} \) possesses a convergent subsequence (w.l.o.g. we use the same index) with nonvanishing limit \( (\bar{\lambda}, \bar{\mu}, \bar{\nu}, \bar{\xi}) \). We multiply (47) with
\[
\left\| (\lambda^r, \bar{\mu}^r, \bar{\nu}^r, \xi^r_{I^{0+}(\bar{u})}) \right\|_2^{-1}
\]
and take the limit \( r \to \infty \) in \( L^2(0, T)^m \). Lemma 2.3 yields the existence of \( \bar{v} \in \mathcal{N}_S(\bar{u}) \) which satisfies
\[
0 = \bar{v} + S^*_f \left[ \nabla g(\bar{x}(T))^\top \bar{\lambda} - \sum_{i=1}^k \bar{\mu}_i \nabla G_i(\bar{x}(T)) + \bar{\nu}_i \nabla H_i(\bar{x}(T)) \right].
\]
Furthermore, we obtain
\[
\begin{align*}
\hat{\lambda} &\geq 0 \land \forall t \in I^0(\bar{u}): \hat{\lambda}_t = 0, \\
\forall i \in I^{+0}(\bar{u}): \hat{\mu}_i &= 0, \\
\forall i \in I^{0+}(\bar{u}): \hat{\nu}_i &= 0,
\end{align*}
\] (49)
and \((\hat{\lambda}, \hat{\mu}, \hat{\nu}) \neq 0\) from the choice of the multipliers, see (46) and the proof of [21, Theorem 3.1] for the details.

We invoke Lemmas 2.4 and 7.1 in order to find some \(g \in L^2(0,T)^q\) which satisfies (22d), (22e), and \(\hat{\nu} = (S^* \circ C^* + D^*)[\hat{g}]\). We deduce
\[
0 = (S^* \circ C^* + D^*)[\hat{g}] + S^*_T \left[ \nabla g(\bar{x}(T)) \right] - \sum_{i=1}^k \left[ \hat{\mu}_i \nabla G_i(\bar{x}(T)) + \hat{\nu}_i \nabla H_i(\bar{x}(T)) \right]
\] from (48). Due to Lemma 7.2, the operator \((C \circ S + D, S_T)\) is surjective which is why its adjoint is injective. Especially, we obtain
\[
0 = \nabla g(\bar{x}(T))^\top \hat{\lambda} - \sum_{i=1}^k \left[ \hat{\mu}_i \nabla G_i(\bar{x}(T)) + \hat{\nu}_i \nabla H_i(\bar{x}(T)) \right]
\] from above. However, due to (49) and \((\hat{\lambda}, \hat{\mu}, \hat{\nu}) \neq 0\), this is a contradiction to the postulated constraint qualification (43).

Hence, the sequence \(\{ (\lambda^r, \vec{\mu}^r, \vec{\nu}^r, \xi^r_{i=0(\bar{u})} \cup I^{+0(\bar{u})}) \} \) is bounded and converges w.l.o.g. to \((\lambda, \mu, \nu, \xi)\). Taking the limit \(r \to \infty\) in (47) w.r.t. the \(L^2(0,T)^m\)-norm and invoking Lemma 2.3 once more, we find \(v \in \mathcal{N}_S(\bar{u})\) which satisfies
\[
0 = S^*[E[\bar{x}] - x_d] + \sigma(\bar{u} - u_d) + \nu + S^*_T \left[ \nabla f(\bar{x}(T)) + \nabla g(\bar{x}(T))^\top \lambda \right. \\
\left. - \sum_{i=1}^k \left[ \mu_i \nabla G_i(\bar{x}(T)) + \nu_i \nabla H_i(\bar{x}(T)) \right] \right]
\] as well as
\[
\lambda \geq 0 \land \forall t \notin I^0(\bar{u}): \lambda_t = 0, \\
\forall i \in I^{+0}(\bar{u}): \mu_i = 0, \\
\forall i \in I^{0+}(\bar{u}): \nu_i = 0.
\]
We invoke Lemmas 2.4 and 7.1 in order to find \(g \in L^2(0,T)^q\) which satisfies (22d), (22e), and solves \(v = (S^* \circ C^* + D^*)[\hat{g}]\). Putting this into (50) and applying Lemma 4.4, we obtain
\[
0 = \begin{align*}
S^*[E[\bar{x}] - x_d + C^*[\hat{g}] + \sigma(\bar{u} - u_d) + D^*[\hat{g}] \\
+ \hat{f}'(\bar{u}) + \hat{g}'(\bar{u})[\lambda] - \sum_{i=1}^k \left[ \mu_i \hat{G}_i'(\bar{u}) + \nu_i \hat{H}_i'(\bar{u}) \right]
\end{align*}
\] Following the proof of Proposition 6.1, \((\bar{x}, \bar{u})\) is a W-stationary point of \((\text{OCTCC})\).

Now, suppose that this point is not C-stationary. Then we find \(i \in I^{00}(\bar{u})\) such that \(\mu_i \nu_t < 0\) holds true. If \(\mu_i < 0 \land \nu_i > 0\) hold, we have \(\mu_i \nu_t < 0\) and, hence, \(\xi^t_i > 0\) for sufficiently large \(r \in \mathbb{N}\). This leads to \(\nu_i^t \leq 0\) for sufficiently large \(r \in \mathbb{N}\) and, thus, to \(\nu_i \leq 0\) which is a contradiction. On the other hand, if \(\mu_i > 0 \land \nu_i < 0\) are satisfied, then \(\nu_i^t > 0\) and, hence, \(\xi^t_i > 0\) is valid for sufficiently large \(r \in \mathbb{N}\). This leads to \(\nu_i^t \geq 0\) for sufficiently large \(r \in \mathbb{N}\), i.e. \(\nu_i \geq 0\) which is again a contradiction. Consequently, \((\bar{x}, \bar{u})\) is a C-stationary point of \((\text{OCTCC})\). \(\square\)

The following remark directly follows from Theorem 7.3.

**Remark 8.1** Let \((\bar{x}, \bar{u}) \in H^1(0,T)^n \times L^2(0,T)^m\) be a locally optimal solution of \((\text{OCTCC})\) and let the assumptions of Theorem 8.1 hold. Then \((\bar{x}, \bar{u})\) is an M-stationary point of \((\text{OCTCC})\).

In order to guarantee that locally optimal solutions of \((\text{OCTCC}(\alpha^r))\) satisfy the corresponding KKT conditions presented in Lemma 8.1, a constraint qualification is needed. For that purpose, we finally state the following important observation which is inspired by [21, Theorem 3.2]. The technical proof parallels the validation of the corresponding result in [21] and, thus, is only presented in Appendix C for the sake of completeness.
Lemma 8.2 Let Assumption 7.1 be valid and let \( \{\alpha^r\} \subseteq \mathbb{R}^+ \) be a scalar sequence monotonically decreasing to zero. Let \((\bar{x}, \bar{u}) \in H^1(0, T)^n \times L^2(0, T)^m\) be a feasible point of \((\text{OCTCC})\) where the constraint qualification \((43)\) holds, and assume that the matrix \((28)\) possesses full row rank \(n\). Then there is a neighborhood \(U\) of \(\bar{u}\) such that for sufficiently large \(r \in \mathbb{N}\), KRZCQ holds at all feasible points \(u \in U\) of \((44)\).

Let the assumptions of Lemma 8.2 be satisfied. Note that the KKT conditions of \((44)\) equal the KKT conditions of \((\text{OCTCC}(\alpha^r))\). Thus, if for sufficiently large \(r \in \mathbb{N}\), \((x^r, u^r)\) is a locally optimal solution of \((44)\) such that \(u^r \in U\) holds true, then the KKT conditions presented in Lemma 8.1 are valid since KRZCQ holds at \(u^r\) for the surrogate problem \((44)\).

A Vector-valued function spaces

From [1, Theorem 6.3] we have that \(H^1(0, T)\) is compactly embedded into \(L^2(0, T)\) and \(C[0, T]\), respectively. In this section, we will justify that \(H^1(0, T)^n\) is compactly embedded into \(L^2(0, T)^n\) and \(C[0, T]^n\), respectively.

Subsequently, the components of a vector function \(x\) which maps to \(\mathbb{R}^n\) will be denoted by \(x_1, \ldots, x_n\).

Lemma A.1 We have

\[
\forall x \in C[0, T]^n; \quad \frac{1}{n\sqrt{n}} \sum_{j=1}^{n} \|x_j\|_{C[0, T]} \leq \|x\|_{C[0, T]^n} \leq \sum_{j=1}^{n} \|x_j\|_{C[0, T]},
\]

\[
\forall x \in L^2(0, T)^n; \quad \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \|x_j\|_{L^2(0, T)} \leq \|x\|_{L^2(0, T)^n} \leq \sum_{j=1}^{n} \|x_j\|_{L^2(0, T)},
\]

\[
\forall x \in H^1(0, T)^n; \quad \frac{1}{\sqrt{2n}} \sum_{j=1}^{n} \|x_j\|_{H^1(0, T)} \leq \|x\|_{H^1(0, T)^n} \leq \sqrt{2} \sum_{j=1}^{n} \|x_j\|_{H^1(0, T)}.
\]

Proof Recall that for nonnegative real numbers \(a_1, \ldots, a_n\), we have

\[
\frac{1}{\sqrt{n}} \sum_{j=1}^{n} a_j \leq \left( \sum_{j=1}^{n} a_j^2 \right)^{1/2} \leq \sum_{j=1}^{n} a_j.
\]

Choose a function \(x \in C[0, T]^n\). By definition, its components \(x_1, \ldots, x_n\) belong to \(C[0, T]\). We obtain

\[
\sum_{j=1}^{n} \|x_j\|_{C[0, T]} \leq n \sup_{t \in [0, T]} \|x(t)\|_1 \leq n\sqrt{n} \sup_{t \in [0, T]} \|x(t)\|_2 = n\sqrt{n} \|x\|_{C[0, T]^n}
\]

and

\[
\|x\|_{C[0, T]^n} = \sup_{t \in [0, T]} \|x(t)\|_2 \leq \sup_{t \in [0, T]} \|x(t)\|_1 \leq \sum_{j=1}^{n} \sup_{t \in [0, T]} \|x_j(t)\| = \sum_{j=1}^{n} \|x_j\|_{C[0, T]}
\]

which yields the first assertion of the lemma.

Now, choose \(x \in L^2(0, T)^n\) arbitrarily. Again, the definition of this space implies \(x_1, \ldots, x_n \in L^2(0, T)\). It is easily seen that

\[
\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \|x_j\|_{L^2(0, T)} \leq \left( \sum_{j=1}^{n} \|x_j\|_{L^2(0, T)}^2 \right)^{1/2} = \left( \int_0^T \sum_{j=1}^{n} x_j^2(t) \ dt \right)^{1/2} = \|x\|_{L^2(0, T)^n}
\]

and

\[
\|x\|_{L^2(0, T)^n} = \left( \sum_{j=1}^{n} \|x_j\|_{L^2(0, T)}^2 \right)^{1/2} \leq \sum_{j=1}^{n} \|x_j\|_{L^2(0, T)}
\]

hold true. Thus, the second assertion of the lemma is valid.

Finally, choose \(x \in H^1(0, T)^n\). By definition, we have \(x_1, \ldots, x_n \in H^1(0, T)\). Using the lemma’s second assertion, we obtain

\[
\sum_{j=1}^{n} \|x_j\|_{H^1(0, T)} = \sum_{j=1}^{n} \left( \|x_j\|_{L^2(0, T)} + \|\dot{x}_j\|_{L^2(0, T)} \right)^{1/2}
\]

\[
\leq \sum_{j=1}^{n} \left( \|x_j\|_{L^2(0, T)^n} + \|\dot{x}_j\|_{L^2(0, T)^n} \right)
\]

\[
\leq \sqrt{n} \left( \|x\|_{L^2(0, T)^n} + \|\dot{x}\|_{L^2(0, T)^n} \right)
\]

\[
\leq \sqrt{2n} \left( \|x\|_{L^2(0, T)^n} + \|\dot{x}\|_{L^2(0, T)^n} \right)^{1/2} = \sqrt{2n} \|x\|_{H^1(0, T)^n}.
\]
Additionally, it is possible to show
\[
\|x\|_{H^1(0,T)^n} \leq \|x\|_{L^2(0,T)^n} + \|\dot{x}\|_{L^2(0,T)^n} \leq \sum_{j=1}^{n} \left( \|x_j\|_{L^2(0,T)} + \|\dot{x}_j\|_{L^2(0,T)} \right) \\
\leq \sqrt{2} \sum_{j=1}^{n} \left( \|x_j\|_{L^2(0,T)}^2 + \|\dot{x}_j\|_{L^2(0,T)}^2 \right)^{\frac{1}{2}} = \sqrt{2} \sum_{j=1}^{n} \|x_j\|_{H^1(0,T)}.
\]

Combining the last two observations, the third assertion of the lemma is obtained. This completes the proof. □

We apply Lemma A.1 in order to obtain our desired result.

**Theorem A.1** We have $H^1(0,T)^n \hookrightarrow L^2(0,T)^n$ and $H^1(0,T)^n \hookrightarrow C[0,T]^n$. Additionally, these embeddings are compact.

**Proof** As mentioned earlier, we have that the embeddings $H^1(0,T) \hookrightarrow L^2(0,T)$ and $H^1(0,T) \hookrightarrow C[0,T]$ are compact, see [1, Theorem 6.3]. Using the estimates from Lemma A.1, $H^1(0,T)^n \hookrightarrow L^2(0,T)^n$ and $H^1(0,T)^n \hookrightarrow C[0,T]^n$ follow directly.

Choose a bounded sequence $\{x_k\} \subseteq H^1(0,T)^n$. Due to Lemma A.1, the sequences $\{x_{j,k}\} \subseteq H^1(0,T)$ are bounded for any $j = 1, \ldots, n$. Exploiting the compactness of $H^1(0,T) \hookrightarrow L^2(0,T)$ (of $H^1(0,T) \hookrightarrow C[0,T]$), we can extract a subsequence $\{x_{j,k}\}$ from $\{x_k\}$ such that $\{x_{j,k}\}$ converges w.r.t. the norm in $L^2(0,T)$ (w.r.t. the norm in $C[0,T]$). We rephrase this argument for each of the components $2, \ldots, n$ to obtain a subsequence of $\{x_k\}$ whose components all converge w.r.t. the norm in $L^2(0,T)$ (w.r.t. the norm in $C[0,T]$). Invoking Lemma A.1 once more, this subsequence of $\{x_k\}$ converges w.r.t. the $L^2(0,T)^n$-norm (w.r.t. the $C[0,T]^n$-norm) as well. Thus, the embedding $H^1(0,T)^n \hookrightarrow L^2(0,T)^n$ (the embedding $H^1(0,T)^n \hookrightarrow C[0,T]^n$) is compact as well. This completes the proof. □

**B Continuity of the solution operator**

Here, we show that the linear solution operator $S : L^2(0,T)^m \rightarrow H^1(0,T)^n$ of the linear ODE (13) is continuous.

First, we set
\[
A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\
\vdots & \ddots & \vdots \\
a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \quad B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,m} \\
\vdots & \ddots & \vdots \\
b_{n,1} & \cdots & b_{n,m} \end{pmatrix}
\]

and define
\[
\alpha := \max \{|a_{i,j}| : i, j \in \{1, \ldots, n\}\}, \\
\beta := \max \{|b_{i,l}| : i \in \{1, \ldots, n\}, l \in \{1, \ldots, m\}\}.
\]

Let $\Phi$ denote the matrix function which solves (14). Note that $\Phi$ is continuous on $[0,T]$. Thus, the function $\phi : [0,T]^2 \rightarrow \mathbb{R}$ defined by
\[
\phi(t,s) := \max \left\{ \sum_{j=1}^{n} \sum_{k=1}^{m} \Phi_{i,j}(t) \Phi_{i,k}^{-1}(s) b_{j,l} : i \in \{1, \ldots, n\}, l \in \{1, \ldots, m\} \right\}
\]
for any $(t,s) \in [0,T]^2$ is continuous as well and, thus, attains a maximum which we denote by $\gamma$.

For arbitrary $u \in L^2(0,T)^m$, we denote by $|u| \in L^1(0,T)$ the componentwise absolute value function of $u$, and $u_1, \ldots, u_m \in L^2(0,T)$ represent the component functions of $u$. From (16), Hölder’s inequality, and Lemma A.1 we obtain
\[
\|S[u]\|_{L^2(0,T)^n} = \int_0^T \int_0^T \|\phi(t,s) Bh(s) ds\|_2^2 dt \\
\leq \int_0^T \|\gamma \int_0^T E[u(s)] ds\|_2^2 dt = nT\gamma^2 \left( \sum_{l=1}^{m} \int_0^T |u_l(s)| ds \right)^2 \\
\leq nT\gamma^2 \left( \sum_{l=1}^{m} \int_0^T \|u_l\|_{L^2(0,T)} \right)^2 = nT\gamma^2 \left( \sum_{l=1}^{m} \int_0^T \|u_l\|_{L^2(0,T)} \right)^2 \\
\leq mnT^2\gamma^2 \|u\|_{L^2(0,T)^m}^2.
\]

By definition of $S$, the weak derivative of $S[u]$ is given by $AS[u] + Bu$ and we have
\[
\|AS[u] + Bu\|_{L^2(0,T)^n} \leq \|AS[u]\|_{L^2(0,T)^n} + \|Bu\|_{L^2(0,T)^n}
\]
by the triangle inequality. Obviously, we obtain
\[
\|Bu\|_{L^2(0,T)^n}^2 = \int_0^T \|Bu(t)\|_2^2 dt \leq \int_0^T \|\beta E[u(s)]\|_2^2 dt \\
= n\beta^2 \int_0^T \left( \sum_{l=1}^{m} |u_l(t)| \right)^2 dt \leq mn\beta^2 \int_0^T \left( \sum_{l=1}^{m} u_l^2(t) \right) dt \\
= mn\beta^2 \int_0^T \|u(t)\|_2^2 dt = mn\beta^2 \|u\|_{L^2(0,T)^m}^2.
\]
and, analogously,
\[
\|\mathbf{A}s[u]\|_{L^2(0,T)^n} \leq n^2 \alpha^2 \|s[u]\|_{L^2(0,T)^n} \leq mn^3 T^2 \alpha^2 \gamma^2 \|u\|_{L^2(0,T)^m}^2
\]
is derived. This yields
\[
\|\mathbf{A}s[u] + Bu\|_{L^2(0,T)^n} \leq \sqrt{mn}(nT\alpha\gamma + \beta)\|u\|_{L^2(0,T)^m}.
\]
Consequently, we have
\[
\|s[u]\|_{H^1(0,T)^n}^2 = \left(\|s[u]\|^2_{L^2(0,T)^n} + \|\mathbf{A}s[u] + Bu\|^2_{L^2(0,T)^n}\right)^{\frac{1}{2}} \leq \left(mnT^2 \alpha^2 + mn(nT\alpha\gamma + \beta)^2\right)^{\frac{1}{2}} \|u\|_{L^2(0,T)^m}.
\]
i.e. \(s\) is continuous from \(L^2(0,T)^m\) to \(H^1(0,T)^n\).

C Proof of Lemma 8.2

Here, we present the proof of Lemma 8.2.

Proof As shown in the proof of [21, Theorem 3.2], there is an \(\varepsilon > 0\) such that for all \(z \in \mathbb{U}^\theta_{\mathbb{R}^n}(\bar{x}(T))\), the constraint qualification
\[
0 = \nabla g(z)^\top - \sum_{k=1}^K \left[ \mu_k \nabla G_i(z) + \nu_i \nabla H_i(z) \right]
\]
\[
\lambda \geq 0 \land \forall i \notin I^0(\bar{u}): \lambda_i = 0,
\]
\[
\forall i \in I^0(\bar{u}): \mu_i = 0,
\]
\[
\forall i \in I^0^+(\bar{u}): \nu_i = 0
\]
is satisfied. On the other hand, there is \(\delta > 0\) such that for all \(z \in \mathbb{U}^\theta_{\mathbb{R}^n}(\bar{x}(T))\) and \(\alpha > 0\) sufficiently small, we have
\[
I^0(z) \subseteq I^0(\bar{u}),
\]
\[
I^0(z) \cup (I^0(z,\alpha) \cap I^0^+(\bar{u})) \cup (I^0(z,\alpha) \cap I^0^0(\bar{u})) \subseteq I^0^+(\bar{u}) \cup I^0^0(\bar{u}),
\]
\[
I^0(z) \cup (I^0(z,\alpha) \cap I^0^+(\bar{u})) \cup (I^0(z,\alpha) \cap I^0^0(\bar{u})) \subseteq I^0^+(\bar{u}) \cup I^0^0(\bar{u})
\]
and
\[
\forall i \in I^0^0(\bar{u}): G_i(z) > 0,
\]
\[
\forall i \in I^0^+(\bar{u}): H_i(z) > 0.
\]
Here, we used the intuitive definitions
\[
I^0(z) := \{i \in \{1, \ldots, K\} | g_i(z) = 0\},
\]
\[
I^0(z) := \{i \in \{1, \ldots, K\} | G_i(z) = 0\},
\]
\[
I^0(\bar{u}) := \{i \in \{1, \ldots, K\} | H_i(z) = 0\},
\]
\[
I^0(z,\alpha) := \{i \in \{1, \ldots, K\} | G_i(z)H_i(z) = \alpha\}.
\]
Defining \(\theta := \min \{\delta, \varepsilon\}\) yields that for any \(z \in \mathbb{U}^\theta_{\mathbb{R}^n}(\bar{x}(T))\) and sufficiently small \(\alpha > 0\), we have
\[
0 = \nabla g(z)^\top - \sum_{i \in I^0(z,\alpha) \cap (I^0(z,\alpha) \cap I^0(\bar{u}))} \mu_i \nabla G_i(z)
\]
\[
- \sum_{i \in I^0(z,\alpha) \cap (I^0^+(\bar{u}) \cup I^0^0(\bar{u}))} \xi_i [H_i(z)\nabla G_i(z) + G_i(z)\nabla H_i(z)],
\]
\[
\lambda \geq 0 \land \forall i \notin I^0(z): \lambda_i = 0,
\]
\[
\forall i \notin I^0(z) \cup (I^0(z,\alpha) \cap I^0^+(\bar{u})): \mu_i = 0,
\]
\[
\forall i \notin I^0(z) \cup (I^0(z,\alpha) \cap I^0^0(\bar{u})): \nu_i = 0,
\]
\[
\forall i \notin I^0(z,\alpha) \cap (I^0^+(\bar{u}) \cup I^0^0(\bar{u})): \xi_i = 0
\]
is satisfied. This leads to
\[
0 = \nabla g(z)^\top - \sum_{i=1}^K [\mu_i \nabla G_i(z) + \nu_i \nabla H_i(z)]
\]
\[
+ \sum_{k=1}^K \xi_k [H_k(z)\nabla G_k(z) + G_k(z)\nabla H_k(z)],
\]
\[
\lambda \geq 0 \land \forall i \notin I^0(z): \lambda_i = 0,
\]
\[
\mu \geq 0 \land \forall i \notin I^0(z): \mu_i = 0,
\]
\[
\nu \geq 0 \land \forall i \notin I^0(z): \nu_i = 0,
\]
\[
\xi \geq 0 \land \forall i \notin I^0(z,\alpha): \xi_i = 0
\]
for all $z \in U^{n}_{\tilde{g}}(\tilde{x}(T))$ and sufficiently small $\alpha > 0$.
Since we have $S_T \in L[L^2(0,T)^m,\mathbb{R}^n]$, see Section 4, there is a constant $\gamma > 0$ such that

$$\forall u \in L^2(0,T)^m: \quad \|S_T[u]\|_2 \leq \gamma\|u\|_{L^2(0,T)^m}$$

is valid. Hence, for an arbitrary point $u \in U^{m-1}_{L^2(0,T)^m}(\tilde{u})$ feasible to (44), where $r \in N$ is sufficiently large, and $x := S[u]$, the system

$$0 = \nabla g(x(T))\lambda + \sum_{i=1}^{k} [\mu_i \nabla G_i(x(T)) + \nu_i \nabla H_i(x(T))]$$

$$\lambda \geq 0 \land \forall i \notin I^p(u): \lambda_i = 0,$n
$$\tilde{\mu} \leq 0 \land \forall i \notin I^G(u): \tilde{\mu}_i = 0,$n
$$\tilde{\nu} \leq 0 \land \forall i \notin I^H(u): \tilde{\nu}_i = 0,$n
$$\xi \geq 0 \land \forall i \notin I^{GH}(u,\alpha^r): \xi_i = 0$$

possesses only the trivial solution $(\lambda, \tilde{\mu}, \tilde{\nu}, \xi) = (0,0,0,0)$. Note that we have

$$\lambda \geq 0 \land \forall i \notin I^p(u): \lambda_i = 0 \iff \lambda \in T_{-R_{\tilde{g}}^0} \cdot (g(x(T)))^\circ,$n
$$\tilde{\mu} \leq 0 \land \forall i \notin I^G(u): \tilde{\mu}_i = 0 \iff \tilde{\mu} \in T_{-R_{\tilde{g}}^0} \cdot (G(x(T)))^\circ,$n
$$\tilde{\nu} \leq 0 \land \forall i \notin I^H(u): \tilde{\nu}_i = 0 \iff \tilde{\nu} \in T_{-R_{\tilde{g}}^0} \cdot (H(x(T)))^\circ,$n

as well as

$$\xi_i \geq 0 \land \forall i \notin I^{GH}(u,\alpha^r): \xi_i = 0 \iff \xi_i \in T_{-R_{\tilde{g}}^0} \cdot (G_i(x(T))H_i(x(T)) - \alpha^r)^\circ$$

for all $i = 1, \ldots, k$. Now, a similar argumentation as we used in the proof of Lemma 7.3 implies the validity of the constraint qualification

$$(\sigma \circ \tilde{S} + \tilde{d}, \tilde{g}^I(u), \tilde{G}(u), \tilde{H}(u), \tilde{\Gamma}(u)[L^2(0,T)^m] - K(u))$$

$$= L^2(0,T)^n \times \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^k.$$n

Here, $\tilde{\Gamma}: L^2(0,T)^m \rightarrow \mathbb{R}^k$ is the mapping which possesses the component functions $\tilde{\Gamma}_1, \ldots, \tilde{\Gamma}_k: L^2(0,T)^m \rightarrow \mathbb{R}$ defined by

$$\forall i \in \{1, \ldots, k\} \forall u \in L^2(0,T)^m: \quad \tilde{\Gamma}_i(u) := \tilde{G}_i(u)\tilde{H}_i(u) - \alpha^r,$n

and $K(u) \subseteq L^2(0,T)^n \times \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^k$ is the convex cone given below:

$$K(u) := \{0\} \times R_{-R_{\tilde{g}}^0} \times (\tilde{g}(u)) \times R_{-R_{\tilde{g}}^0} \times (\tilde{G}(u)) \times R_{-R_{\tilde{g}}^0} \times (\tilde{H}(u)) \times R_{-R_{\tilde{g}}^0} \times (\tilde{\Gamma}(u)).$$

Clearly, the above constraint qualification implies that KRZCQ for (44) at $u$ is satisfied. \boxproof

References