Computational Optimization and Applications (2023) 86:1229–1246
https://doi.org/10.1007/s10589-023-00530-y

Extension of switch point algorithm to boundary-value problems

William W. Hager[1]

Received: 1 July 2023 / Accepted: 11 September 2023 / Published online: 30 September 2023

Abstract
In an earlier paper (https://doi.org/10.1137/21M1393315), the switch point algorithm was developed for solving optimal control problems whose solutions are either singular or bang-bang or both singular and bang-bang, and which possess a finite number of jump discontinuities in an optimal control at the points in time where the solution structure changes. The class of control problems that were considered had a given initial condition, but no terminal constraint. The theory is now extended to include problems with both initial and terminal constraints, a structure that often arises in boundary-value problems. Substantial changes to the theory are needed to handle this more general setting. Nonetheless, the derivative of the cost with respect to a switch point is again the jump in the Hamiltonian at the switch point.

Keywords Switch point algorithm · Singular control · Bang–bang control · Boundary-value problems

Mathematics Subject Classification 49M25 · 49M37 · 65K05 · 90C30

1 Introduction

An earlier paper [1] develops the Switch Point Algorithm for initial-value problems with bang-bang or singular solutions. This paper extends the algorithm to problems with terminal constraints. More precisely, we consider fixed terminal time control

July 1, 2023, revised September 8, 2023. The author gratefully acknowledges support by the National Science Foundation under grants 1819002 and 2031213, and by Office of Naval Research under grant N00014-22-1-2397.

✉ William W. Hager
hager@ufl.edu
http://people.clas.ufl.edu/hager/

1 Department of Mathematics, University of Florida, PO Box 118105, Gainesville, FL 32611-8105, USA

problems of the form

$$\min C(\mathbf{x}(T)) \quad \text{subject to } \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{u}(t) \in \mathcal{U}(t), \\ \mathbf{x}_I(0) = \mathbf{b}_I, \quad \mathbf{x}_E(T) = \mathbf{b}_E, \tag{1.1}$$

where $\mathbf{x} : [0, T] \to \mathbb{R}^n$ is absolutely continuous, $\mathbf{u} : [0, T] \to \mathbb{R}^m$ is essentially bounded, $C : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $\mathcal{U}(t)$ is a closed and bounded set for each $t \in [0, T]$, I and E are subsets of $\{1, 2, \ldots, n\}$, and $\mathbf{x}_I$ denotes the subvector of $\mathbf{x}$ associated with indices $i \in I$. The vectors $\mathbf{b}_I$ and $\mathbf{b}_E$ are given initial and terminal values for the state. It is assumed that $|I| + |E| = n$, where $|S|$ denotes the number of elements in a set S, and the dynamics $\mathbf{f}$ and the objective C are continuously differentiable. Here and throughout the paper, differential equations should hold almost everywhere on $[0, T]$. Problems of this form arise in boundary-value problems such as the fish harvesting problem in [32], which is also studied in the PhD thesis [6] of Summer Atkins.

With the notation given above, the paper [1] considered an initial value problem where $|I| = n$ and $|E| = 0$. In this special case, any $\mathbf{u}$ satisfying the control constraint is feasible, and the associated state is the solution to an initial value problem. When $|E| > 0$, components of the initial state corresponding to $i \in I^c$, the complement of I, are unknown. The nonspecified components of the initial state along with the control $\mathbf{u}$ must be chosen to satisfy the boundary condition $\mathbf{x}_E(T) = \mathbf{b}_E$. Due to the terminal constraint, the theory developed in [1] is no longer applicable.

The costate associated with (1.1) satisfies the linear differential equation

$$\dot{\mathbf{p}}(t) = -\mathbf{p}(t)\nabla_x \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{p}_J(0) = \mathbf{0}, \quad \mathbf{p}_F(T) = \nabla_F C(\mathbf{x}(T)), \tag{1.2}$$

where J and F denote the complements of I and E respectively, $\mathbf{p} : [0, T] \to \mathbb{R}^n$ is a row vector, the objective gradient $\nabla_F C$ is a row vector whose i-th component is the partial derivative of C with respect to x_i, $i \in F$, and $\nabla_x \mathbf{f}$ denotes the Jacobian of the dynamics with respect to $\mathbf{x}$. Due to the terminal constraint $\mathbf{x}_E(T) = \mathbf{b}_E$, the objective is only a function of $\mathbf{x}_F(T)$. Under the assumptions of the Pontryagin minimum principle, a local minimizer of (1.1) and the associated costate have the property that

$$H(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}(t)) = \inf\{H(\mathbf{x}(t), \mathbf{v}, \mathbf{p}(t)) : \mathbf{v} \in \mathcal{U}(t)\} \tag{1.3}$$

for almost every $t \in [0, T]$, where $H(\mathbf{x}, \mathbf{u}, \mathbf{p}) = \mathbf{p}\mathbf{f}(\mathbf{x}, \mathbf{u})$ is the Hamiltonian.

When the Hamiltonian is linear in the control and the feasible control set has the form

$$\mathcal{U}(t) = \{\mathbf{v} \in \mathbb{R}^m : \boldsymbol{\alpha}(t) \le \mathbf{v} \le \boldsymbol{\beta}(t)\},$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta} : [0, T] \to \mathbb{R}^m$, it is often possible to decompose $[0, T]$ into a finite number of disjoint subintervals (s_i, s_{i+1}), where $0 = s_0 < s_1 < \ldots < s_N = T$, and on each subinterval, each component of an optimal control is either singular or bang-bang. Moreover, by singular control theory [36], it is often possible to express the control in feedback form as $\mathbf{u}(t) = \boldsymbol{\phi}_i(\mathbf{x}(t), t)$ for all $t \in (s_i, s_{i+1})$ for some function $\boldsymbol{\phi}_i$ defined on a larger interval containing (s_i, s_{i+1}). In the Switch Point Algorithm, the original control problem is solved by optimizing over the choice of the s_i, $0 < i < N$. In other words, if $\mathbf{F}_i(\mathbf{x}, t) := \mathbf{f}(\mathbf{x}, \boldsymbol{\phi}_i(\mathbf{x}, t))$ and $\mathbf{F}(\mathbf{x}, t) := \mathbf{F}_i(\mathbf{x}, t)$ for all $t \in (s_i, s_{i+1})$, $0 \le i < N$, then (1.1) is replaced by the problem

$$\min_{\mathbf{s}} C(\mathbf{x}(T)) \quad \text{subject to} \quad \dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t), \quad \mathbf{x}_I(0) = \mathbf{b}_I, \quad \mathbf{x}_E(T) = \mathbf{b}_E. \tag{1.4}$$

In order to solve (1.4) efficiently, we develop an algorithm for computing the derivative of the objective with respect to a switch point. This formula allows us to utilize gradient, conjugate gradient, and quasi-Newton methods in the solution process. Let $C(\mathbf{s})$ denote the objective in (1.4) parameterized by the switch points s_i, $0 < i < N$, and suppose that $\mathbf{x}$ is feasible in (1.4). Under a smoothness assumption for each $\mathbf{F}_i$ and invertibility assumptions for submatrices of related fundamental matrices, we obtain the following formula:

$$\frac{\partial C}{\partial s_i}(\mathbf{s}) = H_{i-1}(\mathbf{x}(s_i), \mathbf{p}(s_i), s_i) - H_i(\mathbf{x}(s_i), \mathbf{p}(s_i), s_i), \quad 0 < i < N, \tag{1.5}$$

where $H_i(\mathbf{x}, \mathbf{p}, t) = \mathbf{p}\mathbf{F}_i(\mathbf{x}, t)$, and the row vector $\mathbf{p} : [0, T] \to \mathbb{R}^n$ is the solution to the linear differential equation

$$\dot{\mathbf{p}}(t) = -\mathbf{p}(t)\nabla_x \mathbf{F}(\mathbf{x}(t), t), \quad t \in [0, T], \quad \mathbf{p}_F(T) = \nabla_F C(\mathbf{x}(T)), \quad \mathbf{p}_J(0) = \mathbf{0}. \tag{1.6}$$

Hence, the derivative of the objective with respect to all the switch points can be computed from one integration of the state equation in (1.4), followed by one integration of the costate equation in (1.6). The formula (1.5) matches the formula given in [1, Thm. 2.4] in the case $|E| = 0$. Summer Atkins in her thesis [6] also obtains this formula in the special case of the fish harvesting problem. Since $\mathbf{F}$ could jump at s_i, the existence of the Jacobian in (1.6) is generally restricted to the open intervals (s_i, s_{i+1}), and the differential equation only needs to hold almost everywhere.

See the earlier paper [1] for a detailed survey of literature concerning bang-bang and singular control problems, which includes the papers [2–5, 8–11, 25, 27, 30, 31, 38, 39]. Note that [31] and [39] express the partial derivative of the objective with respect to the switch points in terms of the matrix of partial derivatives of each state variable with respect to each switch point, where the matrix is obtained by forward propagation using the system dynamics. We circumvent the evaluation of the matrix of partial derivatives by using the costate equation to directly compute the partial derivative of the objective with respect to all the switching points. One benefit of computing the matrix of partial derivatives of each state with respect to each switch point is that with marginal additional work, second-order optimality conditions can be checked [33, Chap. 7].

In more recent work [34], the authors develop a method for solving bang-bang and singular optimal control problem using adaptive Legendre–Gauss–Radau collocation [12, 13, 24, 28, 29, 35] in which the structure of the solution is first determined, and a regularization technique is used in the singular regions, while the switch points are treated as free parameters in the optimization. The gradient methods that might be used in conjunction with the derivatives provided in the current paper do not require regularization, however, as discussed in Sect. 7, a good starting guess for the switch points is needed to ensure convergence.

The paper is organized as follows. Section 2 provides an existence result for a system of nonlinear equations. This key result is the basis for a stability analysis of the boundary-value problem associated with (1.1). In Sect. 3, stability with respect the terminal boundary constraint is analyzed, while Sect. 4 analyzes stability with respect to a switch point. In Sect. 5, the results of the previous sections are combined to obtain the derivative formula (1.5). Section 6 discusses problems where a singular control depends on both state and costate. Finally, Sect. 7 explores numerical issues.

Notation and terminology. Throughout the paper, $\|\cdot\|$ is any norm on $\mathbb{R}^n$. The ball with center $\mathbf{c} \in \mathbb{R}^n$ and radius ρ is denoted $\mathcal{B}_\rho(\mathbf{c}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{c}\| \le \rho\}$. The expression $\mathcal{O}(\boldsymbol{\theta})$ denotes a quantity whose norm is bounded by $c\|\boldsymbol{\theta}\|$, with c is a constant that is independent of $\boldsymbol{\theta}$. The Jacobian of $\mathbf{f}(\mathbf{x}, \mathbf{u})$ with respect to $\mathbf{x}$ is denoted $\nabla_x \mathbf{f}(\mathbf{x}, \mathbf{u})$; its (i, j) element is $\partial f_i(\mathbf{x}, \mathbf{u})/\partial x_j$. For a real-valued function such as C, the gradient $\nabla C(\mathbf{x})$ is a row vector, while $\nabla_F C(\mathbf{x})$ is a row vector whose i-th component, $i \in F$, is the partial derivative of C with respect to x_i. For a vector $\mathbf{x} \in \mathbb{R}^n$ and a set $I \subset \{1, 2, \ldots, n\}$, $\mathbf{x}_I$ is the subvector consisting of elements x_i, $i \in I$. If $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a matrix, and $\mathcal{R}$ and $\mathcal{C}$ are subsets of the row and column indices respectively, then $\mathbf{A}_{\mathcal{R}\mathcal{C}}$ is the submatrix corresponding to rows in $\mathcal{R}$ and columns in $\mathcal{C}$. All vectors in the paper are column vectors except for the costate which is a row vector.

2 An existence result

In order to derive the formula (1.5) for the derivative of the objective with respect to a switch point, we first need to analyze the stability of the boundary-value problem in (1.1). This analysis is done using the proposition stated below. The proposition is a special case of a general theorem developed in a sequence of papers [15–21, 23]. The general result, formulated in a Banach space with set-valued maps, has broad application in the convergence analysis of numerical algorithms, as seen in papers such as [14, 20–22]. The special case stated here is for finite dimensional point-to-point maps which is sufficient for handling the analysis of (1.1). This result is closely related to Newton's method, a favorite topic of Asen L. Dontchev, whom we remember in this volume.

Proposition 2.1 *Suppose that* $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$ *is continuously differentiable in* $\mathcal{B}_r(\mathbf{0})$ *for some* $r > 0$*, and define* $\delta = \|\mathbf{g}(\mathbf{0})\|$*. Let* $\mathcal{L} \in \mathbb{R}^{n\times n}$ *be an invertible matrix with* $\gamma := \|\mathcal{L}^{-1}\|$ *and with the property that for some* $\epsilon > 0$,

$$\|\nabla g(\boldsymbol{\theta}) - \mathcal{L}\| \le \epsilon \;\; \textit{for all } \boldsymbol{\theta} \in \mathcal{B}_r(\mathbf{0}). \tag{2.1}$$

If $\epsilon\gamma < 1$ *and* $\delta \le r(1-\gamma\epsilon)/\gamma$, *then there exists a unique* $\boldsymbol{\theta} \in \mathcal{B}_r(\mathbf{0})$ *such that* $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$. *Moreover, we have the bound*

$$\|\boldsymbol{\theta}\| \le \frac{\delta\gamma}{1-\epsilon\gamma}. \tag{2.2}$$

3 Stability with respect to terminal constraint

In analyzing the differentiability of the objective in (1.4) with respect to a switch point, there is no loss in generality in focusing on the case $N = 2$, where there is a single switch point $s \in (0, T)$ and the dynamics switches from $\mathbf{F}_0$ to $\mathbf{F}_1$ at $t = s$:

$$\mathbf{F}(\mathbf{x}, t) = \mathbf{F}_0(\mathbf{x}, t) \text{ for all } t \in [0, s) \quad \text{and} \quad \mathbf{F}(\mathbf{x}, t) = \mathbf{F}_1(\mathbf{x}, t) \text{ for all } t \in (s, T].$$

It is assumed that there exists a feasible, absolutely continuous state $\mathbf{x}$ which satisfies the constraints of (1.4). That is, $\mathbf{x}$ satisfies

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t), \quad \mathbf{x}_I(0) = \mathbf{b}_I, \quad \mathbf{x}_E(T) = \mathbf{b}_E. \tag{3.1}$$

Throughout the paper, $\mathbf{x}$ denotes a solution to this problem. In this section, we focus on the following question: If the endpoint constraint $\mathbf{b}_E$ in (3.1) is changed to $\mathbf{b}_E + \boldsymbol{\pi}$, does there exist a solution $\mathbf{x}^\pi$ to the perturbed problem

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t), \quad \mathbf{x}_I(0) = \mathbf{b}_I, \quad \mathbf{x}_E(T) = \mathbf{b}_E + \boldsymbol{\pi}, \tag{3.2}$$

and is the solution change bounded in terms of $\|\boldsymbol{\pi}\|$? The following assumption is used in this analysis.

Dynamics Smoothness. For $\rho > 0$, define the tubes

$$\mathcal{T}_0 = \{(\boldsymbol{\chi}, t) : t \in [0, s + \rho] \text{ and } \boldsymbol{\chi} \in \mathcal{B}_\rho(\mathbf{x}(t))\},$$
$$\mathcal{T}_1 = \{(\boldsymbol{\chi}, t) : t \in [s - \rho, T] \text{ and } \boldsymbol{\chi} \in \mathcal{B}_\rho(\mathbf{x}(t))\}.$$

It is assumed that on $\mathcal{T}_j$, $j = 0$ or 1, $\mathbf{F}_j$ is continuously differentiable, while $\mathbf{F}_j(\boldsymbol{\chi}, t)$ is Lipschitz continuously differentiable in $\boldsymbol{\chi}$, uniformly in t, with Lipschitz constant L.

Let us define $\boldsymbol{\theta}^* = \mathbf{x}_J(0)$, and let us consider the initial-value problem

$$\dot{\mathbf{y}}(t) = \mathbf{F}(\mathbf{y}(t), t), \quad \mathbf{y}_I(0) = \mathbf{b}_I, \quad \mathbf{y}_J(0) = \boldsymbol{\theta}^* + \boldsymbol{\theta}. \tag{3.3}$$

For $\boldsymbol{\theta} = \mathbf{0}$, $\mathbf{y} = \mathbf{x}$, the solution of (3.1), since $\mathbf{y}_J(0) = \mathbf{x}_J(0)$. Under Dynamics Smoothness, it follows from [1, Cor. 2.3] that (3.3) has a solution $\mathbf{y}_\theta$ when $\|\boldsymbol{\theta}\|$ is sufficiently small, and we have the bound

$$\|\mathbf{y}_\theta(t) - \mathbf{x}(t)\| = \|\mathbf{y}_\theta(t) - \mathbf{y}_0(t)\| \le e^{Lt}\|\boldsymbol{\theta}\| \quad \text{for all } t \in [0, T]. \tag{3.4}$$

By the continuity of $\nabla_x \mathbf{F}_j$ on $\mathcal{T}_j$, for $j = 0$ or 1, it follows that there is a constant β such that

$$\|\nabla_x \mathbf{F}(\boldsymbol{\chi}, t)\| \le \beta \ \text{ for all } \ t \in [0, T] \ \text{ and } \ \boldsymbol{\chi} \in \mathcal{B}_\rho(\mathbf{x}(t)). \tag{3.5}$$

A sharper estimate for the difference $\mathbf{y} - \mathbf{x}$ is obtained from the solution $\mathbf{z}_\theta$ of the linearized problem

$$\dot{\mathbf{z}}(t) = \nabla_x \mathbf{F}(\mathbf{x}(t), t)\mathbf{z}(t), \quad \mathbf{z}_I(0) = \mathbf{0}, \quad \mathbf{z}_J(0) = \boldsymbol{\theta}. \tag{3.6}$$

Since $\nabla_x \mathbf{F}(\mathbf{x}(t), t)$ is continuous on $[0, s)$ and on $(s, T]$, the solution to the linear differential equation (3.6) has a bound

$$\mathbf{z}_\theta(t) = O(\boldsymbol{\theta}) \text{ for all } t \in [0, T]. \tag{3.7}$$

Define for all $t \in [0, T]$ and $\alpha \in [0, 1]$,

$$\boldsymbol{\delta}(t) = \mathbf{y}_\theta(t) - \mathbf{x}(t) - \mathbf{z}_\theta(t) \quad \text{and} \quad \mathbf{x}(\alpha, t) = \mathbf{x}(t) + \alpha(\mathbf{y}_\theta(t) - \mathbf{x}(t)). \tag{3.8}$$

Differentiating $\boldsymbol{\delta}$ and utilizing a Taylor expansion with integral remainder term, we obtain for all $t \in [0, T]$, $t \ne s$, $\dot{\boldsymbol{\delta}}(t) = \dot{\mathbf{y}}_\theta(t) - \dot{\mathbf{x}}(t) - \dot{\mathbf{z}}_\theta(t) =$

$$\begin{aligned}
&\mathbf{F}(\mathbf{y}_\theta(t), t) - \mathbf{F}(\mathbf{x}(t), t) - \nabla_x \mathbf{F}(\mathbf{x}(t), t)\mathbf{z}_\theta(t) = \\
&\left(\int_0^1 \nabla_x \mathbf{F}(\mathbf{x}(\alpha, t), t)\, d\alpha\right)(\mathbf{y}_\theta(t) - \mathbf{x}(t)) - \left(\int_0^1 \nabla_x \mathbf{F}(\mathbf{x}(t), t)\, d\alpha\right)\mathbf{z}_\theta(t) = \\
&\left(\int_0^1 [\nabla_x \mathbf{F}(\mathbf{x}(\alpha, t), t) - \nabla_x \mathbf{F}(\mathbf{x}(t), t)]\, d\alpha\right)\mathbf{z}_\theta(t) + \left(\int_0^1 \nabla_x \mathbf{F}(\mathbf{x}(\alpha, t), t)\, d\alpha\right)\boldsymbol{\delta}(t).
\end{aligned} \tag{3.9}$$

Take $\boldsymbol{\theta}$ in (3.4) small enough that $\mathbf{y}_\theta(t)$ lies in the tube around $\mathbf{x}(t)$ where $\nabla_x \mathbf{F}$ is Lipschitz continuous. If L is the Lipschitz constant for $\nabla_x \mathbf{F}$, then we have

$$\|\nabla_x \mathbf{F}(\mathbf{x}(\alpha, t), t) - \nabla_x \mathbf{F}(\mathbf{x}(t), t)\| \le \alpha L \|\mathbf{y}_\theta(t) - \mathbf{x}(t)\| = \mathcal{O}(\boldsymbol{\theta}) \tag{3.10}$$

by (3.4). Take the norm of each side of (3.9). On the right side of (3.9), the coefficient of $\mathbf{z}_\theta$ is $\mathcal{O}(\boldsymbol{\theta})$ by (3.10), while $\mathbf{z}_\theta$ is $\mathcal{O}(\boldsymbol{\theta})$ by (3.7). Since $\|\nabla_x \mathbf{F}(\mathbf{x}(\alpha, t), t)\| \le \beta$ for all $\alpha \in [0, 1]$ and $t \in [0, T]$ by (3.5), the right side of (3.9) has the bound $\mathcal{O}(\|\boldsymbol{\theta}\|^2) + \beta\|\boldsymbol{\delta}(t)\|$. On the left side, exploit the fact from [1, Lem. 2.1] that the derivative of a norm is bounded by the norm of the derivative to obtain

$$\frac{d\|\boldsymbol{\delta}(t)\|}{dt} \le \|\dot{\boldsymbol{\delta}}(t)\| \le \mathcal{O}\left(\|\boldsymbol{\theta}\|^2\right) + \beta\|\boldsymbol{\delta}(t)\|. \tag{3.11}$$

By the initial conditions for $\mathbf{y}_\theta$, $\mathbf{x}$, and $\mathbf{z}_\theta$ in (3.3), (3.1), and (3.6) respectively, $\boldsymbol{\delta}(0) = \mathbf{0}$. This observation, together with (3.11) and Gronwall's inequality yield

$$\|(\mathbf{y}_\theta - \mathbf{x}) - \mathbf{z}_\theta\| = \|\boldsymbol{\delta}(t)\| = \mathcal{O}\left(\|\boldsymbol{\theta}\|^2\right). \tag{3.12}$$

Thus $\mathbf{z}_\theta$ provides an $\mathcal{O}\left(\|\boldsymbol{\theta}\|^2\right)$ approximation to the difference $\mathbf{y}_\theta - \mathbf{x}$.

The linearized problem (3.6) plays a fundamental role in the stability analysis of (3.1). Finding a solution of the perturbed problem (3.2) is equivalent to finding the starting condition $\boldsymbol{\theta}$ in (3.3) with the property that $\mathbf{y}_\theta(T) = \mathbf{b}_E + \boldsymbol{\pi}$. Since $\mathbf{z}_\theta$ is a close approximation to $\mathbf{y}_\theta - \mathbf{x}$, we could choose $\boldsymbol{\theta}$ so that $\mathbf{z}_\theta(T) = \boldsymbol{\pi}$, in which case

$$\mathbf{y}_\theta(T) = \mathbf{x}(T) + \mathbf{z}_\theta(T) + \mathcal{O}\left(\|\boldsymbol{\theta}\|^2\right) = \mathbf{b}_E + \boldsymbol{\pi} + \mathcal{O}\left(\|\boldsymbol{\theta}\|^2\right).$$

Therefore, for this choice of $\boldsymbol{\theta}$, the solution of (3.3) satisfies the perturbed boundary condition to within $\mathcal{O}\left(\|\boldsymbol{\theta}\|^2\right)$.

The fundamental matrix $\boldsymbol{\Phi} : [0, T] \to \mathbb{R}^{n\times n}$ associated with the linear system $\dot{\mathbf{z}}(t) = \nabla_x \mathbf{F}(\mathbf{x}(t), t)\mathbf{z}(t)$ is the solution to the initial-value problem

$$\dot{\boldsymbol{\Phi}}(t) = \nabla_x \mathbf{F}(\mathbf{x}(t), t)\boldsymbol{\Phi}(t), \quad \boldsymbol{\Phi}(0) = \mathbf{I}, \tag{3.13}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. The solution $\mathbf{z}$ of the linearized problem (3.6) is equal to the fundamental matrix times the initial condition. Due to the special choice of the initial condition in (3.6), the $\boldsymbol{\theta}$ that yields $\mathbf{z}_E(T) = \boldsymbol{\pi}$ is the solution to the linear system of equations $\boldsymbol{\Phi}_{EJ}(T)\boldsymbol{\theta} = \boldsymbol{\pi}$, where $\boldsymbol{\Phi}_{EJ}$ represents the submatrix of $\boldsymbol{\Phi}$ associated with columns J and rows E. If this square submatrix is invertible, then $\boldsymbol{\theta} = \boldsymbol{\Phi}_{EJ}(T)^{-1}\boldsymbol{\pi}$. With these insights, we have the following result:

Lemma 3.1 *Suppose that $\boldsymbol{\Phi}_{EJ}(T)$ is invertible and let $\gamma = \|\boldsymbol{\Phi}_{EJ}^{-1}(T)\|$. For $\boldsymbol{\pi}$ in a neighborhood $\mathcal{N}$ of the origin, the perturbed boundary-value problem (3.2) has a solution $\mathbf{x}^\pi$ and*

$$\|\mathbf{x}_J^\pi(0) - \mathbf{x}_J(0)\| = \|\mathbf{x}_J^\pi(0) - \boldsymbol{\theta}^*\| \le c\|\boldsymbol{\pi}\| \textit{ for all } \boldsymbol{\pi} \in \mathcal{N}, \tag{3.14}$$

where c is a constant that approaches γ as $\|\boldsymbol{\pi}\|$ approaches $\mathbf{0}$.

Proof We apply Proposition 2.1 with $\mathcal{L} = \nabla \mathbf{g}(\mathbf{0})$, where $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{y}_{\theta E}(T) - \mathbf{b}_E - \boldsymbol{\pi}$ and $\mathbf{y}_\theta$ is the solution of (3.3). Both $\mathbf{b}_E$ and $\boldsymbol{\pi}$ are independent of $\boldsymbol{\theta}$ so their derivatives are $\mathbf{0}$. From the analysis in [37, Chap. 1.6], the derivative of $\mathbf{y}_{\theta E}(T)$ with respect to $\boldsymbol{\theta}$, evaluated at $\boldsymbol{\theta} = \mathbf{0}$ is $\mathcal{L} = \boldsymbol{\Phi}_{EJ}(T)$. Moreover, it follows from [37, Chap. 1.6] that $\nabla \mathbf{g}(\boldsymbol{\theta})$ is continuously differentiable at $\boldsymbol{\theta} = \mathbf{0}$. Choose ϵ small enough that $\epsilon\gamma < 1$ and then choose r small enough that (2.1) holds; by continuity of the derivative of $\mathbf{g}$ at $\boldsymbol{\theta} = \mathbf{0}$, (2.1) holds for r sufficiently small. Since $\mathbf{g}(\mathbf{0}) = \boldsymbol{\pi}$, we have $\delta = \|\boldsymbol{\pi}\|$. Choose $\|\boldsymbol{\pi}\|$ small enough that $\delta \le r(1-\gamma\epsilon)/\gamma$. Since all the requirements for (2.2) have now been satisfied, there exists a unique $\boldsymbol{\theta} \in \mathcal{B}_r(\mathbf{0})$ such that $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$, or equivalently, such that $\mathbf{y}_{\theta E}(T) = \mathbf{b}_E + \boldsymbol{\pi}$. By (2.2), $\|\boldsymbol{\theta}\| \le c\|\boldsymbol{\pi}\|$, where $c = \gamma/(1-\epsilon\gamma)$ is independent of $\boldsymbol{\pi}$. Since $\mathbf{y}_\theta$ satisfies both the initial and terminal conditions for $\mathbf{x}^\pi$ in (3.2), we can take $\mathbf{x}^\pi = \mathbf{y}_\theta$. At $t = 0$,

$$\mathbf{x}_J^\pi(0) = \mathbf{y}_{\theta J}(0) = \boldsymbol{\theta}^* + \boldsymbol{\theta},$$

which rearranges to give (3.14). As ϵ tends to zero, we can let r also approach zero, in which case the denominator in (2.2) tends to one and the ball containing the solution $\boldsymbol{\theta}$ to $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ tends to zero. □

4 Stability with respect to the switch point

In order to obtain the derivative of the objective in (1.4) with respect to the switch point, we need to analyze the effect of perturbations in the switch point s. Let $\mathbf{F}^+$ be defined by

$$\mathbf{F}^+(\mathbf{x}, t) = \begin{cases} \mathbf{F}_0(\mathbf{x}, t) & \text{for all } t \in [0, s + \Delta s), \\ \mathbf{F}_1(\mathbf{x}, t) & \text{for all } t \in (s + \Delta s, T], \end{cases}$$

where $|\Delta s| \le \rho$. Hence, $\mathbf{F}^+$ is the dynamics gotten by changing the switch point from s to $s + \Delta s$. The boundary-value problem associated with the perturbed switch point is

$$\dot{\mathbf{x}}(t) = \mathbf{F}^+(\mathbf{x}(t), t), \quad \mathbf{x}_I(0) = \mathbf{b}_I, \quad \mathbf{x}_E(T) = \mathbf{b}_E, \tag{4.1}$$

and a solution, if it exists, is denoted $\mathbf{x}^+$. The goal in this section is to show that when the invertibility condition of Lemma 3.1 holds, the perturbed problem (4.1) has a solution that is stable with respect to the perturbation Δs.

Let $\mathbf{y}_\theta^+$ denote the solution to the perturbed initial-value problem

$$\dot{\mathbf{y}}(t) = \mathbf{F}^+(\mathbf{y}(t), t), \quad \mathbf{y}_I(0) = \mathbf{b}_I, \quad \mathbf{y}_J(0) = \boldsymbol{\theta}^* + \boldsymbol{\theta}, \tag{4.2}$$

where $\boldsymbol{\theta}^* = \mathbf{x}_J(0)$. When $\boldsymbol{\theta} = \mathbf{0}$, we omit the $\boldsymbol{\theta}$ subscript on $\mathbf{y}_\theta^+$ so $\mathbf{y}^+ := \mathbf{y}_0^+$. Since $\mathbf{F}^+ = \mathbf{F}_0$ on $[0, s)$, assuming $\Delta s > 0$, it follows that

$$\mathbf{y}^+(t) = \mathbf{y}_0^+(t) = \mathbf{x}(t) \;\; \text{for all } t \in [0, s). \tag{4.3}$$

For $t \in (s, T]$, it is shown in [1, (2.12)–(2.14)] that

$$\|\mathbf{y}^+(t) - \mathbf{x}(t)\| = \mathcal{O}(\Delta s) \text{ on } (s, T], \;\; \text{which implies } \mathbf{y}_E^+(T) = \mathbf{b}_E - \boldsymbol{\pi} \tag{4.4}$$

for some $\boldsymbol{\pi} = \mathcal{O}(\Delta s)$ since $\mathbf{x}(T) = \mathbf{b}_E$. By (4.3) and (4.4), $\mathbf{y}^+$ lies inside the tubes around $\mathbf{x}$ given in Dynamic Smoothness when Δs is sufficiently small. Moreover, as in (3.4), it follows from Dynamics Smoothness and [1, Cor. 2.3] that (4.2) has a solution $\mathbf{y}_\theta^+$ when $|\Delta s| \le \rho$ and $\|\boldsymbol{\theta}\|$ is sufficiently small, and we have the bound

$$\|\mathbf{y}_\theta^+(t) - \mathbf{y}^+(t)\| \le e^{Lt}\|\boldsymbol{\theta}\| \quad \text{for all } t \in [0, T]. \tag{4.5}$$

Combine (4.3)–(4.5), and the triangle inequality to obtain

$$\|\mathbf{y}_\theta^+(t) - \mathbf{x}(t)\| = \mathcal{O}(\Delta s) + \mathcal{O}(\boldsymbol{\theta}) \quad \text{for all } t \in [0, T]. \tag{4.6}$$

Now let us consider whether a solution exists to (4.1), assuming a solution to the original system (3.1) exists when $\Delta s = 0$. As in the previous section, our approach is to focus on the initial-value problem (4.2) and try to choose $\boldsymbol{\theta}$ such that $\mathbf{y}_\theta^+ = \mathbf{x}^+$ is a solution of (4.1). In particular, if we choose $\boldsymbol{\theta}$ such that

$$\left(\mathbf{y}_\theta^+(T) - \mathbf{y}_0^+(T)\right)_E = \boldsymbol{\pi},$$

then combining this with (4.4) gives

$$\mathbf{y}_{\theta E}^{+} = \mathbf{y}_{E}^{+} + \boldsymbol{\pi} = \mathbf{b}_E - \boldsymbol{\pi} + \boldsymbol{\pi} = \mathbf{b}_E.$$

Thus $\mathbf{y}_{\theta}^{+}$ satisfies the same boundary conditions as those for a solution $\mathbf{x}^{+}$ of (4.1). With this insight, the following result is established:

Lemma 4.1 *If $\boldsymbol{\Phi}_{EJ}(T)$ is invertible, then for Δs in a neighborhood of 0, the problem (4.1), with perturbed switch point $s + \Delta s$, has a solution $\mathbf{x}^{+}$, and we have*

$$\|\mathbf{x}_J^{+}(0) - \mathbf{x}_J(0)\| = \|\mathbf{x}_J^{+}(0) - \boldsymbol{\theta}^{*}\| \le c|\Delta s| \text{ for all } \Delta s \text{ near } 0, \tag{4.7}$$

where c is a constant that is independent of Δs.

Proof The lemma is stated in terms of the fundamental matrix $\boldsymbol{\Phi}$ that arises in the unperturbed problem of Sect. 3, and which satisfies

$$\dot{\boldsymbol{\Phi}}(t) = \nabla_x \mathbf{F}(\mathbf{x}(t), t)\boldsymbol{\Phi}(t), \quad \boldsymbol{\Phi}(0) = \mathbf{I}.$$

If the proof technique of Lemma 3.1 is applied to the problem (4.1) with a perturbed switch point, then the associated fundamental matrix is the solution of

$$\dot{\boldsymbol{\Phi}}^{+}(t) = \nabla_x \mathbf{F}^{+}(\mathbf{y}^{+}(t), t)\boldsymbol{\Phi}^{+}(t), \quad \boldsymbol{\Phi}^{+}(0) = \mathbf{I}. \tag{4.8}$$

Since $\mathbf{x}(t) = \mathbf{y}_0^{+}(t) = \mathbf{y}^{+}(t)$ and $\mathbf{F}^{+} = \mathbf{F}$ on the interval $[0, s]$, it follows that $\boldsymbol{\Phi}^{+}(t) = \boldsymbol{\Phi}(t)$ on $[0, s]$. On the interval $[s, s+\Delta s]$, $\boldsymbol{\Phi}$ is associated with the dynamics $\mathbf{F}_1$ while $\boldsymbol{\Phi}^{+}$ is associated with the dynamics $\mathbf{F}_0$, so the fundamental matrices satisfy

$$\dot{\boldsymbol{\Phi}}(t) = \nabla_x \mathbf{F}_1(\mathbf{x}(t), t)\boldsymbol{\Phi}(t) \quad \text{and} \quad \dot{\boldsymbol{\Phi}}^{+}(t) = \nabla_x \mathbf{F}_0(\mathbf{y}^{+}(t), t)\boldsymbol{\Phi}^{+}(t) \quad \text{on } [s, s + \Delta s]$$

with the initial condition $\boldsymbol{\Phi}(s) = \boldsymbol{\Phi}^{+}(s)$. Since $\mathbf{F}_0$ and $\mathbf{F}_1$ are smooth and the starting conditions for $\boldsymbol{\Phi}(t)$ and $\boldsymbol{\Phi}^{+}(t)$ at $t = s$ are the same, it follows that the difference $\mathbf{D} = \boldsymbol{\Phi}^{+} - \boldsymbol{\Phi}$ satisfies $\|\mathbf{D}(s + \Delta s)\| = \mathcal{O}(\Delta s)$. On the interval $[s + \Delta s, T]$, the fundamental matrices satisfy

$$\dot{\boldsymbol{\Phi}}(t) = \nabla_x \mathbf{F}_1(\mathbf{x}(t), t)\boldsymbol{\Phi}(t) \quad \text{and} \quad \dot{\boldsymbol{\Phi}}^{+}(t) = \nabla_x \mathbf{F}_1(\mathbf{y}^{+}(t), t)\boldsymbol{\Phi}^{+}(t)$$

Subtracting the two equations, the difference $\mathbf{D}$ satisfies

$$\dot{\mathbf{D}}(t) = \nabla_x \mathbf{F}_1(\mathbf{y}^{+}(t), t)\mathbf{D}(t) + [\nabla_x \mathbf{F}_1(\mathbf{x}(t), t) - \nabla_x \mathbf{F}_1(\mathbf{y}^{+}(t), t)]\boldsymbol{\Phi}(t), \tag{4.9}$$

where $\mathbf{D}(s + \Delta s) = \mathcal{O}(\Delta s)$. Choose Δs small enough that $\mathbf{y}^{+}$ lies within the tubes associated with Dynamics Smoothness. Hence, (4.4), Dynamics Smoothness, and the Lipschitz property for $\nabla_x \mathbf{F}_1$ imply that the coefficient of $\boldsymbol{\Phi}$ in (4.9) is $\mathcal{O}(\Delta s)$. By the boundedness of $\mathbf{y}^{+}$ and $\boldsymbol{\Phi}$, it follows that the solution $\mathbf{D}$ of the linear equation (4.9) satisfies $\mathbf{D}(T) = \mathcal{O}(\Delta s)$. Since $\boldsymbol{\Phi}_{EJ}(T)$ is invertible by assumption, then so is

$\mathbf{\Phi}^+_{EJ}(T)$ for $|\Delta s|$ sufficiently small and $\mathbf{\Phi}^+_{EJ}(T)$ converges to $\mathbf{\Phi}_{EJ}(T)$ as Δs tends to zero. Let us take Δs small enough that $\|\mathbf{\Phi}^+_{EJ}(T)^{-1}\| \le \gamma^+ := 2\gamma$.

Observe that the analysis of $\mathbf{\Phi}$ and $\mathbf{\Phi}^+$ concern the case where $\boldsymbol{\theta} = \mathbf{0}$. Next, $\boldsymbol{\theta}$ is introduced into the analysis. Similar to the approach in the proof of Lemma 3.1, we take $\mathcal{L} = \nabla \mathbf{g}(\mathbf{0}) = \mathbf{\Phi}^+_{EJ}(T)$ where $\mathbf{\Phi}^+$ is the solution of (4.8), $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{y}^+_{\theta E}(T) - \mathbf{b}_E$, and $\mathbf{y}^+_\theta$ is the solution of (4.2). Note that $\nabla \mathbf{g}(\boldsymbol{\theta})$ is the EJ submatrix of $\mathbf{\Phi}^+_\theta(T)$ where

$$\dot{\mathbf{\Phi}}^+_\theta(t) = \nabla_x \mathbf{F}^+(\mathbf{y}^+_\theta(t), t)\mathbf{\Phi}^+_\theta(t), \quad \mathbf{\Phi}^+(0) = \mathbf{I}. \tag{4.10}$$

Subtract the equation (4.8) for $\mathbf{\Phi}^+$ from (4.10) to obtain an equation for the difference $\mathbf{D}^+ = \mathbf{\Phi}^+_\theta - \mathbf{\Phi}^+$:

$$\dot{\mathbf{D}}^+(t) = \nabla_x \mathbf{F}^+\left(\mathbf{y}^+_\theta(t), t\right)\mathbf{D}^+(t) + \left[\nabla_x \mathbf{F}^+\left(\mathbf{y}^+_\theta(t), t\right) - \nabla_x \mathbf{F}^+\left(\mathbf{y}^+(t), t\right)\right]\mathbf{\Phi}^+(t), \tag{4.11}$$

where $\mathbf{D}^+(0) = \mathbf{0}$. By the Lipschitz property for $\nabla_x \mathbf{F}_0$ and $\nabla_x \mathbf{F}_1$ and by (4.5), the coefficient of $\mathbf{\Phi}^+$ in (4.11) is $\mathcal{O}(\boldsymbol{\theta})$ when $|\Delta s| \le \rho$ and $\boldsymbol{\theta}$ is sufficiently small. Since $\mathbf{y}^+_\theta$ and $\mathbf{\Phi}^+$ are both uniformly bounded, it follows from (4.11) that $\|\mathbf{D}^+(T)\| = \mathcal{O}(\boldsymbol{\theta})$. In our context, the left side of (2.1) is

$$\|\nabla \mathbf{g}(\boldsymbol{\theta}) - \nabla \mathbf{g}(\mathbf{0})\| \le \|\mathbf{\Phi}^+_\theta(T) - \mathbf{\Phi}^+(T)\| = \|\mathbf{D}^+(T)\| \le c\|\boldsymbol{\theta}\|,$$

for some constant c independent of $\boldsymbol{\theta}$ and $|\Delta s| \le \rho$. Choose $\epsilon > 0$ such that $\epsilon\gamma^+ < 1$, and choose r small enough that $\|\mathbf{D}^+(T)\| \le \epsilon$ when $\|\boldsymbol{\theta}\| \le r$.

By (4.4), $\delta = \|\mathbf{g}(\mathbf{0})\| = \|\mathbf{y}^+_E(T) - \mathbf{b}_E\| = \mathcal{O}(\Delta s)$. Choose Δs smaller, if necessary, to ensure that $\delta \le r(1 - \gamma^+\epsilon)/\gamma^+$. Hence, by Proposition 2.1, there exists a unique $\boldsymbol{\theta} \in \mathcal{B}_r(\mathbf{0})$ such that $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$, or equivalently, such that $\mathbf{y}^+_{\theta E}(T) = \mathbf{b}_E$. Moreover, $\mathbf{x}^+ = \mathbf{y}^+_\theta$ is a solution of the perturbed problem (4.1) and $\|\boldsymbol{\theta}\| \le c|\Delta s|$ where $c = \gamma^+/\left(1 - \epsilon\gamma^+\right)$ by (2.2). The identity $\mathbf{x}^+ = \mathbf{y}^+_\theta$ implies that

$$\mathbf{x}^+_J(0) = \mathbf{y}^+_{\theta J}(0) = \boldsymbol{\theta}^* + \boldsymbol{\theta},$$

which rearranges to give (4.7) since $\boldsymbol{\theta} = \mathcal{O}(\Delta s)$. □

5 Objective derivative with respect to switch point

Lemmas 3.1 and 4.1 will be combined to establish the formula (1.5) for the derivative of the objective with respect to a switch point. Notice that this formula involves the costate $\mathbf{p}$, which must satisfy complementary boundary conditions to those of $\mathbf{x}$. Since the costate equation is linear, its solution can be expressed in terms of a fundamental matrix denoted $\mathbf{\Psi}$, the unique solution of the initial-value problem

$$\dot{\mathbf{\Psi}} = -\nabla_x \mathbf{F}(\mathbf{x}(t), t))^{\mathsf{T}}\mathbf{\Psi}(t), \quad \mathbf{\Psi}(0) = \mathbf{I}. \tag{5.1}$$

Since $\mathbf{p}_J(0) = \mathbf{0}$ while $\mathbf{p}_F(T) = \nabla_F C(\mathbf{x}(T))$, a solution to the costate equation exists when $\mathbf{\Psi}_{FI}(T)$ is invertible.

Theorem 5.1 *If Dynamics Smoothness holds, the objective C is continuously differentiable, and both $\mathbf{\Phi}_{EJ}(T)$ and $\mathbf{\Psi}_{FI}(T)$ are invertible, then*

$$\frac{\partial C}{\partial s}(s) = H_0(\mathbf{x}(s), \mathbf{p}(s), s) - H_1(\mathbf{x}(s), \mathbf{p}(s), s), \tag{5.2}$$

where $H_j(\mathbf{x}, \mathbf{p}, t) = \mathbf{p}\mathbf{F}_j(\mathbf{x}, t)$, $j = 0$ or 1, and the row vector $\mathbf{p} : [0, T] \to \mathbb{R}^n$ is the solution to the linear differential equation

$$\dot{\mathbf{p}}(t) = -\mathbf{p}(t)\nabla_x \mathbf{F}(\mathbf{x}(t), t), \quad t \in [0, T], \quad \mathbf{p}_F(T) = \nabla_F C(\mathbf{x}(T)), \quad \mathbf{p}_J(0) = \mathbf{0}. \tag{5.3}$$

Proof By Lemma 4.1, the problem with perturbed switch point has a solution $\mathbf{x}^+$ for Δs sufficiently small. Our goal is to evaluate the limit

$$\lim_{\Delta s \to 0} \frac{C(\mathbf{x}^+(T)) - C(\mathbf{x}(T))}{\Delta s}.$$

Let $\mathbf{y}_\theta^+$ be the solution of (4.2) associated with the solution $\mathbf{x}^+$ of (4.1); that is, $\mathbf{y}_\theta^+ = \mathbf{x}^+$. Let $\mathbf{Z}$ be the solution to the following linearized system:

$$\dot{\mathbf{Z}}(t) = \nabla_x \mathbf{F}_0(\mathbf{x}(t), t)\mathbf{Z}(t), \quad t \in [0, s), \quad \mathbf{Z}_I(0) = \mathbf{0}, \quad \mathbf{Z}_J(0) = \boldsymbol{\theta}, \tag{5.4}$$

$$\dot{\mathbf{Z}}(t) = \nabla_x \mathbf{F}_1(\mathbf{x}(t), t)\mathbf{Z}(t), \quad t \in (s + \Delta s, T], \tag{5.5}$$

where

$$\mathbf{Z}(s + \Delta s) = \mathbf{Z}(s) + \Delta s[\mathbf{F}_0(\mathbf{x}(s), s) - \mathbf{F}_1(\mathbf{x}(s), s)]. \tag{5.6}$$

There is a unique solution to (5.4)–(5.6) due to the linearity of the first two equations. Since $\boldsymbol{\theta} = \mathcal{O}(\Delta s)$ by Lemma 4.1 and the coefficient of $\mathbf{Z}$ in (5.4) is continuous, it follows that $\mathbf{Z}(t) = \mathcal{O}(\Delta s)$ for $t \in [0, s]$. Since $\mathbf{F}_0(\mathbf{x}(s), s)$ and $\mathbf{F}_1(\mathbf{x}(s), s)$ are both continuous for $t \in [s, s + \rho]$, $\|\mathbf{Z}(s + \Delta s)\| = \mathcal{O}(\Delta s)$. Finally, due to the linearity of (5.5), we have

$$\mathbf{Z}(t) = \mathcal{O}(\Delta s) \quad \text{for } t \in [0, s] \cup [s + \Delta s, T]. \tag{5.7}$$

The difference between $\mathbf{y}_\theta^+ - \mathbf{x}$ and $\mathbf{Z}$ can be analyzed as in Sect. 3 in terms of $\boldsymbol{\delta}(t) = \mathbf{y}_\theta^+(t) - \mathbf{x}(t) - \mathbf{Z}(t)$. By the initial conditions for $\mathbf{y}_\theta^+$, for $\mathbf{x} = \mathbf{y}_0$, and for $\mathbf{Z}$ in (4.2), (3.3), and (5.4) respectively, it follows that $\boldsymbol{\delta}(0) = \mathbf{0}$. Exactly the same expansions between (3.9) and (3.12) yield $\|\boldsymbol{\delta}(t)\| = \mathcal{O}\left(\|\boldsymbol{\theta}\|^2\right)$ for all $t \in [0, s]$. Moreover, from Lemma 4.1 and the fact that $\boldsymbol{\theta}$ is chosen such that $\mathbf{y}_\theta^+ = \mathbf{x}^+$, we have $\|\boldsymbol{\theta}\| \le c|\Delta s|$. Hence,

$$\|\boldsymbol{\delta}(t)\| = \mathcal{O}\left(|\Delta s|^2\right) \quad \text{on } [0, s]. \tag{5.8}$$

Now consider the interval $[s, s + \Delta s]$, $|\Delta s| \le \rho$. Since $\mathbf{x}^+$ and $\mathbf{x}$ are twice continuously differentiable on $(s, s + \Delta s)$, a Taylor expansion gives

$$\mathbf{x}^+(s + \Delta s) = \mathbf{x}^+(s) + \Delta s \mathbf{F}_0(\mathbf{x}^+(s), s) + \mathcal{O}\left(|\Delta s|^2\right), \tag{5.9}$$

$$\mathbf{x}(s+\Delta s)=\mathbf{x}(s)+\Delta s\mathbf{F}_1(\mathbf{x}(s),s)+\mathcal{O}\left(|\Delta s|^2\right). \tag{5.10}$$

Subtracting (5.10) and (5.6) from the (5.9) and referring to the definition of $\boldsymbol{\delta}$ yields

$$\boldsymbol{\delta}(s+\Delta s)=\boldsymbol{\delta}(s)+\Delta s\left[\mathbf{F}_0\left(\mathbf{x}^+(s),s\right)-\mathbf{F}_0(\mathbf{x}(s),s)\right]+\mathcal{O}\left(|\Delta s|^2\right). \tag{5.11}$$

By (4.6) and the fact established in Lemma 4.1 that $\mathbf{y}_\theta^+=\mathbf{x}^+$ with $\boldsymbol{\theta}=\mathcal{O}(\Delta s)$, we have $\|\mathbf{x}^+(s)-\mathbf{x}(s)\|=\mathcal{O}(\Delta s)$. Due to Dynamics Smoothness and the Lipschitz continuity of $\mathbf{F}_0$, and the fact from (5.8) that $\boldsymbol{\delta}(s)=\mathcal{O}(|\Delta s|^2)$, (5.11) implies that $\boldsymbol{\delta}(s+\Delta s)=\mathcal{O}\left(|\Delta s|^2\right)$.

The final interval $[s+\Delta s,T]$ is treated exactly as in the expansions (3.9)–(3.12) except that $\boldsymbol{\delta}(0)=\mathbf{0}$ in (3.12) should be replaced by $\boldsymbol{\delta}(s+\Delta s)=\mathcal{O}\left(|\Delta s|^2\right)$. Nonetheless, we have $\|\boldsymbol{\delta}(t)\|=\mathcal{O}\left(|\Delta s|^2\right)$ for all $t\in[s+\Delta s,T]$. In summary,

$$\|\boldsymbol{\delta}(t)\|=\mathcal{O}\left(|\Delta s|^2\right)\quad\text{for all } t\in[0,s]\cup[s+\Delta s,T]. \tag{5.12}$$

If $\mathbf{p}$ is the solution of (5.3), which exists by the invertibility assumption for $\boldsymbol{\Psi}_{FI}(T)$, and $\mathbf{Z}$ is the solution of (5.4)–(5.6), then we integrate over $[s+\Delta s,T]$ and then integrate by parts to obtain

$$\begin{aligned}
0&=\int_{s+\Delta s}^T \mathbf{p}(t)\Big[\nabla_x\mathbf{F}(\mathbf{x}(t),t)\mathbf{Z}(t)-\dot{\mathbf{Z}}(t)\Big]\,dt\\
&=\int_{s+\Delta s}^T\Big[\mathbf{p}(t)\nabla_x\mathbf{F}(\mathbf{x}(t),t)+\dot{\mathbf{p}}(t)\Big]\mathbf{Z}(t)\,dt-\mathbf{p}(T)\mathbf{Z}(T)+\mathbf{p}(s+\Delta s)\mathbf{Z}(s+\Delta s)\\
&=-\mathbf{p}_E(T)\mathbf{Z}_E(T)-\mathbf{p}_F(T)\mathbf{Z}_F(T)+\mathbf{p}(s+\Delta s)\mathbf{Z}(s+\Delta s)\\
&=-\mathbf{p}_E(T)\mathbf{Z}_E(T)-\nabla_F C(\mathbf{x}(T))\mathbf{Z}_F(T)\\
&\quad+\mathbf{p}(s+\Delta s)[\mathbf{Z}(s)+\Delta s(\mathbf{F}_0(\mathbf{x}(s),s)-\mathbf{F}_1(\mathbf{x}(s),s))],
\end{aligned} \tag{5.13}$$

where the integral in the second equality vanishes due to (5.3) and the last equality is due to (5.6). Similarly, an integral over $[0,s]$ yields

$$\begin{aligned}
0&=\int_0^s \mathbf{p}(t)\Big[\nabla_x\mathbf{F}(\mathbf{x}(t),t)\mathbf{Z}(t)-\dot{\mathbf{Z}}(t)\Big]\,dt\\
&=\int_0^s\Big[\mathbf{p}(t)\nabla_x\mathbf{F}(\mathbf{x}(t),t)+\dot{\mathbf{p}}(t)\Big]\mathbf{Z}(t)\,dt-\mathbf{p}(s)\mathbf{Z}(s)+\mathbf{p}(0)\mathbf{Z}(0)\\
&=-\mathbf{p}(s)\mathbf{Z}(s)
\end{aligned} \tag{5.14}$$

since $\mathbf{p}_J(0)=\mathbf{0}=\mathbf{Z}_I(0)$.

Since C is continuously differentiable at $\mathbf{x}(T)$, the mean-value theorem gives

$$C(\mathbf{x}^+(T))-C(\mathbf{x}(T))=\nabla_F C(\mathbf{x}_\Delta)\left[\mathbf{x}_F^+(T)-\mathbf{x}_F(T)\right], \tag{5.15}$$

where $\mathbf{x}_\Delta$ is a point on the line segment connecting $\mathbf{x}^+(T)$ and $\mathbf{x}(T)$. Add (5.13)–(5.15) and substitute

$$\mathbf{x}^+(T) - \mathbf{x}(T) = \mathbf{x}^+(T) - \mathbf{x}(T) - \mathbf{Z}(T) + \mathbf{Z}(T) = \boldsymbol{\delta}(T) + \mathbf{Z}(T)$$

to obtain $C(\mathbf{x}^+(T)) - C(\mathbf{x}(T)) =$

$$\nabla_F C(\mathbf{x}_\Delta)\boldsymbol{\delta}_F(T) + [\nabla_F C(\mathbf{x}_\Delta) - \nabla_F C(\mathbf{x}(T)]\mathbf{Z}_F(T) + [\mathbf{p}(s+\Delta s) - \mathbf{p}(s)]\mathbf{Z}(s) \\ -\mathbf{p}_E(T)\mathbf{Z}_E(T) + \Delta s\mathbf{p}(s+\Delta s)[\mathbf{F}_0(\mathbf{x}(s), s) - \mathbf{F}_1(\mathbf{x}(s), s)]. \tag{5.16}$$

Bounds are now obtained for each of the terms in (5.16). By (5.12), $\|\boldsymbol{\delta}(T)\| = \mathcal{O}\left(|\Delta s|^2\right)$ so $|\nabla_F C(\mathbf{x}_\Delta)\boldsymbol{\delta}_F(T)| = \mathcal{O}\left(|\Delta s|^2\right)$. Since the distance between $\mathbf{x}(t)$ and $\mathbf{x}^+(t) = \mathbf{y}_\theta^+(t)$ is $\mathcal{O}(\Delta s)$ by (4.6) and Lemma 4.1, the distance between $\mathbf{x}_\Delta$ and $\mathbf{x}(T)$ is also $\mathcal{O}(\Delta s)$. Since $\mathbf{Z}(T) = \mathcal{O}(\Delta s)$ by (5.7), it follows that $\mathbf{Z}_F(T) = \mathcal{O}(\Delta s)$, while the coefficient of $\mathbf{Z}_F$ tends to zero as Δs tends to zero. Similarly, $\mathbf{Z}(s) = \mathcal{O}(\Delta s)$ by (5.7), and the coefficient of $\mathbf{Z}(s)$ tends to 0 as $|\Delta s|$ tends to 0. Finally, since $\mathbf{x}_E^+(T) = \mathbf{x}_E(T) = \mathbf{b}_E$ and $\boldsymbol{\delta}(T) = \mathcal{O}\left(|\Delta s|^2\right)$, it follows that

$$\mathcal{O}\left(|\Delta s|^2\right) = \|\boldsymbol{\delta}_E(T)\| = \|\mathbf{x}_E^+(T) - \mathbf{x}_E(T) - \mathbf{Z}_E(T)\| = \|\mathbf{Z}_E(T)\|,$$

which implies that $\mathbf{p}_E(T)\mathbf{Z}_E(T) = \mathcal{O}\left(|\Delta s|^2\right)$. Divide (5.16) by Δs and let Δs tend to zero to obtain

$$\frac{\partial C}{\partial s}(s) = \lim_{\Delta s \to 0} \frac{C(\mathbf{x}^+(T)) - C(\mathbf{x}(T))}{\Delta s} = \mathbf{p}(s)[\mathbf{F}_0(\mathbf{x}(s), s) - \mathbf{F}_1(\mathbf{x}(s), s)],$$

which completes the proof. □

Theorem 5.1 requires the invertibility of two submatrices of two different fundamental matrices. It would seem that either of the matrices could be singular; consequently, the theory would not be applicable in certain degenerate situations. In particular, if the state $\mathbf{x}(t) \in \mathbb{R}^2$ and if the control problem is a boundary-value problem, rather than an initial-value problem, then $|I| = |J| = |E| = |F| = 1$. Hence, the critical submatrices of $\boldsymbol{\Phi}(T)$ or $\boldsymbol{\Psi}(T)$ would be one by one; that is, they would be scalars. It would seem that either of these off-diagonal scalars elements of $\boldsymbol{\Phi}(T)$ or $\boldsymbol{\Psi}(T)$ could be zero. The invertibility conditions associated with the fundamental matrices arise when we consider perturbations of the boundary conditions in either the state or costate equation. Situations where either of these boundary-value problems becomes infeasible after a small perturbation may be situations where these invertibility assumptions are violated.

6 Singular control depending on both state and costate

In this section, we consider the case where a singular control depends on both the state and costate. Suppose the $\mathbf{u}(t) = \boldsymbol{\phi}_i(\mathbf{x}(t), \mathbf{p}(t), t)$ for all $t \in [0, s)$ when $i = 0$ and for all $t \in (s, T]$ when $i = 1$. The state dynamics becomes $\mathbf{F}_i(\mathbf{x}, \mathbf{p}, t) = \mathbf{f}(\mathbf{x}, \boldsymbol{\phi}_i(\mathbf{x}, \mathbf{p}, t))$ for $t \in [0, s)$ when $i = 0$ or for $t \in (s, T]$ when $i = 1$. After replacing $\mathbf{u}$ by $\boldsymbol{\phi}_i(\mathbf{x}, \mathbf{p}, t)$ in the costate equation, we obtain the factor

$$\mathbf{F}_{ix}(x, p, t) = \nabla_x \mathbf{f}(\mathbf{x}, \mathbf{u})|_{\mathbf{u}=\boldsymbol{\phi}_i(\mathbf{x},\mathbf{p},t)},$$

for $t \in [0, s)$ when $i = 0$ and for $t \in (s, T]$ when $i = 1$. The state/costate coupled system is

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), \mathbf{p}(t), t), \quad \dot{\mathbf{p}}(t) = -\mathbf{p}(t)\mathbf{F}_x(\mathbf{x}(t), \mathbf{p}(t), t), \tag{6.1}$$

with the boundary conditions $\mathbf{x}_I(0) = \mathbf{b}_I$, $\mathbf{x}_E(T) = \mathbf{b}_E$, and $\mathbf{p}(0) = \mathbf{p}_0$, where $\mathbf{p}_0$ is to be determined. If $\mathbf{u}^*$ is a local minimizer for the control problem (1.1) and $(\mathbf{x}^*, \mathbf{p}^*)$ are the associated state and costate, then $(\mathbf{x}^*, \mathbf{p}^*)$ will satisfy both (6.1) and the boundary conditions, assuming $\mathbf{p}_0 = \mathbf{p}^*(0)$. From this perspective, we can think of the objective $C(\mathbf{x}(T))$ as being a function $C(s, \mathbf{p}_0)$ that depends on both the switching point and the starting value $\mathbf{p}_0$ for the costate. To solve this problem efficiently, we need a formula for the derivative of C with respect to either s or $\mathbf{p}_0$.

As in [1, Sect. 3], to compute the derivative with respect to s (holding $\mathbf{p}_0$ fixed), we view the state/costate pair $(\mathbf{x}, \mathbf{p})$ as a new generalized state. There are $2n$ boundary conditions, the initial condition $\mathbf{p}(0) = \mathbf{p}_0$ of dimension n, and the boundary conditions $\mathbf{x}_I(0) = \mathbf{b}_I$ and $\mathbf{x}_E(T) = \mathbf{b}_E$ of total dimension n. This fits the framework of (1.1), but with a state variable of dimension $2n$ and with $2n$ boundary conditions. The Dynamics Smoothness condition should then be replaced by a Generalized Dynamics Smoothness condition; in the Generalized condition, each occurrence of $\mathbf{x}$ is replaced by the pair $(\mathbf{x}, \mathbf{p})$. Associated with the generalized state, there is a generalized costate $(\mathbf{y}(t), \mathbf{z}(t)) \in \mathbb{R}^{2n}$, where $(\mathbf{y}, \mathbf{z})$ is a row vector. The generalized Hamiltonian is

$$\mathcal{H}_i(\mathbf{x}, \mathbf{p}, \mathbf{y}, \mathbf{z}, t) = \mathbf{y}\mathbf{F}_i(\mathbf{x}, \mathbf{p}, t) - \mathbf{p}\mathbf{F}_{ix}(\mathbf{x}, \mathbf{p}, t)\mathbf{z}^{\mathsf{T}}.$$

The generalized costate is the solution of the linear system of differential equations

$$\dot{\mathbf{y}}(t) = -\nabla_x \mathcal{H}(\mathbf{x}(t), \mathbf{p}(t), \mathbf{y}(t), \mathbf{z}(t), t), \;\; \mathbf{y}_J(0) = \mathbf{0}, \;\; \mathbf{y}_F(T) = \nabla_F C(\mathbf{x}(T)), \tag{6.2}$$
$$\dot{\mathbf{z}}(t) = -\nabla_p \mathcal{H}(\mathbf{x}(t), \mathbf{p}(t), \mathbf{y}(t), \mathbf{z}(t), t), \;\; \mathbf{z}(T) = \mathbf{0}, \tag{6.3}$$

where the boundary condition $\mathbf{z}(T) = \mathbf{0}$ is due to the fact that the objective does not depend on $\mathbf{p}(T)$. Therefore, under the assumptions of Theorem 5.1, but with Dynamics Smoothness replaced by Generalized Dynamics Smoothness, we have

$$\frac{\partial C}{\partial s}(s, \mathbf{p}_0) = \mathcal{H}_0(\mathbf{x}(s), \mathbf{p}(s), \mathbf{y}(s), \mathbf{z}(s), s) - \mathcal{H}_1(\mathbf{x}(s), \mathbf{p}(s), \mathbf{y}(s), \mathbf{z}(s), s), \tag{6.4}$$

where $(\mathbf{y}, \mathbf{z})$ satisfy (6.2)–(6.3). The formula for the derivative of the objective with respect to $\mathbf{p}_0$ is similar to that shown in [1, Sect. 3]; the only change is to replace the boundary conditions $\mathbf{x}(0) = \mathbf{x}_0$ and $\mathbf{y}(T) = \nabla C(\mathbf{x}(T))$ in [1, Sect. 3] by the boundary conditions $\mathbf{x}_I(0) = \mathbf{b}_I$, $\mathbf{x}_E(T) = \mathbf{b}_E$, $\mathbf{y}_J(0) = \mathbf{0}$, and $\mathbf{y}_F(T) = \nabla_F C(\mathbf{x}(T))$. The final formula is again

$$\frac{\partial C(\mathbf{s}, \mathbf{p}_0)}{\partial \mathbf{p}_0} = \mathbf{z}(0),$$

where $(\mathbf{y}, \mathbf{z})$ satisfies (6.2)–(6.3).

7 Algorithms

The derivative obtained in this paper is very useful when solving a singular control problem using a gradient-based optimization method, however, a good starting guess for the switching points is needed. One useful approach for generating an initial guess is to employ an Euler discretization with total variation regularization. Details on this regularization technique can be found in [1, Sect. 5] and [7]. A brief summary is as follows: The Euler discretized and regularized version of (1.1) is

$$\begin{aligned} &\min C\left(\mathbf{x}_N\right) + \rho \sum_{i=1}^{m} \sum_{j=1}^{N-1} |u_{ij} - u_{i,j-1}| \\ &\text{subject to } \ \mathbf{x}_{j+1} = \mathbf{x}_j + h\mathbf{f}\left(\mathbf{x}_j, \mathbf{u}_j\right), \ \mathbf{x}_{0I} = \mathbf{b}_I, \ \mathbf{x}_{NE} = \mathbf{b}_E, \ \mathbf{u}_j \in \mathcal{U}\left(t_j\right), \end{aligned} \tag{7.1}$$

where $0 \le j \le N-1$, $h = T/N$, $t_j = jh$, and N is the number of mesh intervals. The parameter ρ controls the strength of the TV regularization term, and as ρ increases, the oscillations in $\mathbf{u}$ should decrease. The nonsmooth problem (7.1) is equivalent to the smooth optimization problem

$$\begin{aligned} &\min C(\mathbf{x}_N) + \rho \sum_{i=1}^{m} \sum_{j=1}^{N-2} v_{ij} + w_{ij} \\ &\text{subject to } \ \mathbf{x}_{j+1} = \mathbf{x}_j + h\mathbf{f}\left(\mathbf{x}_j, \mathbf{u}_j\right), \ \mathbf{x}_{0I} = \mathbf{b}_I, \ \mathbf{x}_{NE} = \mathbf{b}_E, \\ &\quad \mathbf{u}_j \in \mathcal{U}\left(t_j\right), \ \mathbf{u}_{l+1} - \mathbf{u}_l = \mathbf{v}_l - \mathbf{w}_l, \ \mathbf{v}_l \ge \mathbf{0}, \ \mathbf{w}_l \ge \mathbf{0}, \end{aligned} \tag{7.2}$$

where $0 \le j \le N-1$ and $0 \le l \le N-2$. This sparse optimization problem is efficiently solved using the gradient-based algorithm in [26] when the control constraint set $\mathcal{U}\left(t_j\right)$ is a Cartesian product of intervals.

After obtaining an initial guess for the switch points, the formula for the derivative of the objective with respect to the switch points can be used to optimize the location of the switch points. For a problem with multiple switch points, the derivative with respect to all the switch points can be computed with two integrations. Here we focus on the ordinary state and costate (not the generalized case). Define $s_0 = 0$ and $s_N = T$, and let s_i denote the current estimate of the i-th switch point, $0 < i < N$. Assuming, for the moment, that the current state iterate $\mathbf{x}$ satisfies its boundary conditions, the first step in computing the derivative with respect to the switch points is to integrate both the state equation (3.1) and the equation (5.1) for the costate fundamental matrix $\boldsymbol{\Psi}$ across each of the intervals (s_i, s_{i+1}), $i = 0, 1, \ldots, N-1$. We save the values of the state on each interval (s_i, s_{i+1}) and the final fundamental matrix $\boldsymbol{\Psi}(T)$. The invertibility of $\boldsymbol{\Psi}_{FI}(T)$ can be checked at this point. Assuming it is invertible, the terminal condition $\mathbf{p}_E(T)$ for the costate is obtained:

$$\mathbf{p}_E(T) = \boldsymbol{\Psi}_{EI}(T)\boldsymbol{\Psi}_{FI}^{-1}(T)\nabla_F C(\mathbf{x}(T))^{\mathsf{T}}, \tag{7.3}$$

which is combined with $\mathbf{p}_F(T) = \nabla_F C(\mathbf{x}(T))$. Next, integrate the costate (5.3) backwards, starting from the the value for $\mathbf{p}(T)$ that was just determined, and across each of the intervals (s_i, s_{i+1}), $i = N-1, N-2, \ldots, 1$. The change in the Hamiltonian at each of the switch points gives the derivative of the objective with respect to the switch points, in accordance with Theorem 5.1.

After obtaining the gradient of the objective with respect to the switch points, most optimizers perform a line search along a search direction and move to a new set of switch points. To perform this line search, we need to determine the solution to the boundary-value problem associated with a new set of switch points, that are often near the original switch points. The techniques developed in Sects. 3 and 4 can be used to find the new solution of the boundary-value problem (4.1) associated with the new switch points, starting from a solution of the original boundary-value problem (3.1).

Consistent with the notation of Sect. 4, let $\boldsymbol{\theta}^*$ denote $\mathbf{x}_J(0)$, where $\mathbf{x}$ is the solution to the boundary-value problem (3.1) associated with the original switch points. The goal is to find $\boldsymbol{\theta}$ such that the solution to the problem (4.2) satisfies $\mathbf{y}_E(T) = \mathbf{b}_E$. Our initial guess could be $\boldsymbol{\theta}_0 = \mathbf{0}$. If $\boldsymbol{\theta}_k$ denotes the guess at iteration k, $\mathbf{y}^k$ is the solution $\mathbf{y}^+$ of (4.2) associated with $\boldsymbol{\theta} = \boldsymbol{\theta}_k$, and $\boldsymbol{\Phi}^k$ is the fundamental matrix $\boldsymbol{\Phi}^+$ in (4.8) associated with $\mathbf{y}^+ = \mathbf{y}^k$, then Newton's method for computing $\boldsymbol{\theta}$ is

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{\Phi}_{EJ}^k(T)^{-1}\left(\mathbf{y}_E^k(T) - \mathbf{b}_E\right). \tag{7.4}$$

While performing the Newton iteration, we can also check the invertibility of $\boldsymbol{\Phi}_{EJ}^k(T)$.

Note that two fundamental matrices arise in the computations, one matrix $\boldsymbol{\Phi}$ is associated with the state dynamics (3.13), and the other matrix $\boldsymbol{\Psi}$ is associated with the costate dynamics (5.1). In the formulas (7.3) and (7.4), only part of these fundamental matrices are required. In particular, we only need the part of $\boldsymbol{\Phi}$ associated with columns in J, and the part of $\boldsymbol{\Psi}$ associated with columns in I. Hence, the computation of the fundamental matrices can be streamlined by only computing the parts of $\boldsymbol{\Phi}$ or $\boldsymbol{\Psi}$ that are utilized.

8 Conclusions

The Switch Point Algorithm of [1] for an initial-value problem was extended to handle both initial and terminal boundary conditions. The formula for the derivative of the objective with respect to a switch point reduced to the change in the Hamiltonian across a switch point. This was the same formula obtained for an initial-value problem.

Nonetheless, significant modifications in the analysis were needed to handle terminal constraints. In particular, it was necessary to analyze the existence and stability of solutions to a boundary-value problem under perturbations in the terminal constraint or in the switch points; moreover, the invertibility of certain submatrices of fundamental matrices for the linearized state equation and for the costate equation were required.

Acknowledgements Many thanks to Christian Austin for pointing out Taylor's book [37] which provides in Chapter 1.6 a compact treatment of differentiability for the solution of a differential equation with respect to an initial condition.

Data availibility Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The author has no competing interests to declare that are relevant to the content of this article.

References

1. Aghaee, M., Hager, W.W.: The switch point algorithm. SIAM J. Control Optim. **59**, 2570–2593 (2021)
2. Aly, G.: The computation of optimal singular control. Int. J. Control **28**, 681–688 (1978)
3. Anderson, G.M.: An indirect numerical method for the solution of a class of optimal control problems with singular arcs. IEEE Trans. Automat. Control **17**, 363–365 (1972)
4. Andrés-Martínez, O., Biegler, L.T., Flores-Tlacuahuac, A.: An indirect approach for singular optimal control problems. Comput. Chem. Eng. **139**, 106923 (2020)
5. Andrés-Martínez, O., Flores-Tlacuahuac, A., Kameswaran, S., Biegler, L.T.: An efficient direct/indirect transcription approach for singular optimal control. Am. Inst. Chem. Eng. J. **65**, 937–946 (2019)
6. Atkins, S.: Regularization of singular control problems that arise in mathematical biology, PhD thesis, Department of Mathematics, University of Florida, Gainesville, FL (2021)
7. Atkins, S., Aghaee, M., Martcheva, M., Hager, W.: Solving singular control problems in mathematical biology using PASA. In: Tuncer, N., Martcheva, M., Prosper, O., Childs, L. (eds.) Computational and Mathematical Population Dynamics, pp. 319–419. World Scientific, Singapore (2023)
8. Betts, J.T.: Practical Methods for Optimal Control Using Nonlinear Programming, 2nd edn. SIAM, Philadelphia (2010)
9. Bryson, A.E., Ho, Y.-C.: Applied Optimal Control. Hemisphere Publishing, New York (1975)
10. Chen, W., Biegler, L.T.: Nested direct transcription optimization for singular optimal control problems. Am. Inst. Chem. Eng. J. **62**, 3611–3627 (2016)
11. Chen, W., Ren, Y., Zhang, G., Biegler, L.T.: A simultaneous approach for singular optimal control based on partial moving grid. Am. Inst. Chem. Eng. J. **65**, e16584 (2019)
12. Darby, C.L., Hager, W.W., Rao, A.V.: Direct trajectory optimization using a variable low-order adaptive pseudospectral method. J. Spacecr. Rockets **48**, 433–445 (2011)
13. Darby, C.L., Hager, W.W., Rao, A.V.: An hp-adaptive pseudospectral method for solving optimal control problems. Optim. Control Appl. Methods **32**, 476–502 (2011)
14. Dontchev, A., Hager, W.W., Poore, A., Yang, B.: Optimality, stability, and convergence in nonlinear control. Applied Math. Optim. **31**, 297–326 (1995)
15. Dontchev, A.L., Hager, W.W.: Lipschitzian stability in nonlinear control and optimization. SIAM J. Control Optim. **31**, 569–603 (1993)
16. Dontchev, A.L., Hager, W.W.: An inverse mapping theorem for set-valued maps. Proc. Am. Math. Soc. **121**, 481–489 (1994)
17. Dontchev, A.L., Hager, W.W.: Lipschitzian stability for state constrained nonlinear optimal control. SIAM J. Control Optim. **36**, 698–718 (1998)
18. Dontchev, A.L., Hager, W.W.: A new approach to Lipschitz continuity in state constrained optimal control. Syst. Control Lett. **35**, 137–143 (1998)

19. Dontchev, A.L., Hager, W.W.: The Euler approximation in state constrained optimal control. Math. Comput. **70**, 173–203 (2000)
20. Dontchev, A.L., Hager, W.W., Veliov, V.M.: Second-order Runge–Kutta approximations in constrained optimal control. SIAM J. Numer. Anal. **38**, 202–226 (2000)
21. Hager, W.W.: Multiplier methods for nonlinear optimal control. SIAM J. Numer. Anal. **27**, 1061–1080 (1990)
22. Hager, W.W.: Runge–Kutta methods in optimal control and the transformed adjoint system. Numer. Math. **87**, 247–282 (2000)
23. Hager, W. W.: Numerical analysis in optimal control. In: Hoffmann, K.-H., Lasiecka, I., Leugering, G., Sprekels, J., Tröltzsch, F. (eds.) International Series of Numerical Mathematics, , vol. 139, pp. 83–93. Birkhauser Verlag, Basel/Switzerland (2001)
24. Hager, W.W., Hou, H., Mohapatra, S., Rao, A.V., Wang, X.-S.: Convergence rate for a radau hp-collocation method applied to constrained optimal control. Comput. Optim. Appl. **74**, 274–314 (2019)
25. Hager, W.W., Rostamian, R.: Optimal coatings, bang-bang controls, and gradient techniques. Optim. Control Appl. Methods **8**, 1–20 (1987)
26. Hager, W.W., Zhang, H.: Algorithm 1035: a gradient-based implementation of the polyhedral active set algorithm. ACM Trans. Math. Softw. **49**(2), 20.1-20.13 (2023)
27. Jacobon, D.H., Gershwin, S.B., Lele, M.M.: Computation of optimal singular controls. IEEE Trans. Autom. Control **15**, 67–73 (1970)
28. Liu, F., Hager, W.W., Rao, A.V.: Adaptive mesh refinement method for optimal control using nonsmoothness detection and mesh size reduction. J. Franklin Inst. **352**, 4081–4106 (2015)
29. Liu, F., Hager, W.W., Rao, A.V.: Adaptive mesh refinement method for optimal control using decay rates of Legendre polynomial coefficients. IEEE Trans. Control Syst. Technol. **26**, 1475–1483 (2018)
30. Maurer, H.: Numerical solution of singular control problems using multiple shooting techniques. J. Optim. Theory Appl. **18**, 235–257 (1976)
31. Maurer, H., Büskens, C., Kim, J.-H.R., Kaya, C.Y.: Optimization methods for the verification of second order sufficient conditions for bang–bang controls. Optim. Control Appl. Methods **26**, 129–156 (2005)
32. Neubert, M.: Marine reserves and optimal harvesting. Ecol. Lett. **6**, 843–849 (2003)
33. Osmolovskii, N.P., Maurer, H.: Applications to Regular and Bang–Bang Control: Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control. SIAM Publications, Philadelphia (2012)
34. Pager, E.R., Rao, A.V.: Method for solving bang-bang and singular optimal control problems using adaptive Radau collocation. Comput. Optim. Appl. **81**, 857–887 (2022)
35. Patterson, M.A., Hager, W.W., Rao, A.V.: A *ph* mesh refinement method for optimal control. Optim. Control Appl. Methods **36**, 398–421 (2015)
36. Schättler, H., Ledzewicz, U.: Geometric Optimal Control. Springer, New York (2012)
37. Taylor, M.E.: Partial Differential Equations I. Basic Theory. Springer, New York (2011)
38. Vossen, G.: Numerische Lösungsmethoden, hinreichende Optimalitätsbedingungen und Sensitivitätsanalyse für optimale bang-bang und singuläre Steuerungen, PhD thesis, Universität Münster, Germany (2006)
39. Vossen, G.: Switching time optimization for bang–bang and singular controls. J. Optim. Theory Appl. **144**, 409–429 (2010)