Shape Restricted Smoothing Splines via Constrained Optimal Control and Nonsmooth Newton's Methods

Jinglai Shen* and Teresa M. Lebair†

January 24, 2014

Abstract

Shape restricted smoothing splines receive growing attention, motivated by a wide range of important applications in science and engineering. In this paper, we consider smoothing splines subject to general linear dynamics and control constraints, and formulate them as finite-horizon constrained linear optimal control problems with unknown initial state and control. By exploring techniques from functional and variational analysis, optimality conditions are developed in terms of variational inequalities. Due to the control constraints, the optimality conditions give rise to a nonsmooth B-differentiable equation of an optimal initial condition, whose unique solution completely determines the shape restricted smoothing spline. A modified nonsmooth Newton's algorithm with line search is exploited to solve this equation; detailed convergence analysis of the proposed algorithm is presented. In particular, using techniques from nonsmooth analysis, polyhedral theory, and switching systems, we show the global convergence of the algorithm when a shape restricted smoothing spline is subject to a general polyhedral control constraint.

1 Introduction

Spline models are extensively studied in approximation theory, numerical analysis, and statistics with broad applications in science and engineering. Informally speaking, a univariate spline model gives a piecewise polynomial curve that "best" fits a given set of data. Such a spline can be attained via efficient numerical algorithms and enjoys many favorable analytical and statistical properties [6]. A number of variations and extensions have been developed, for instance, penalized polynomial splines [32] and smoothing splines [38]. Specifically, the smoothing spline model is a smooth function $f : [0, 1] \to \mathbb{R}$ in a suitable function space that minimizes the following objective functional:

$$\frac{1}{n}\sum_{i=1}^{n}\Big(f(t_i)-y_i\Big)^2+\lambda\int_0^1\Big(f^{(m)}(t)\Big)^2dt, \tag{1}$$

where y_i are data points at $t_i \in [0, 1]$, $i = 1, \ldots, n$, $f^{(m)}$ denotes the m-th derivative of f, and $\lambda > 0$ is a penalty parameter that characterizes a tradeoff between data fidelity and smoothness of f. We refer the interested reader to [38] and the references therein for extensive discussions on statistical properties of smoothing splines.

From a control systems point of view, the smoothing spline model (1) is closely related to the finite-horizon linear quadratic optimal control problem by treating $f^{(m)}$ as a control input [9]. This

*Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21250, U.S.A. Email: shenj@umbc.edu. This research is supported by NSF grants CMMI-1030246 and DMS-1042967.

†Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21250, U.S.A. Email: ei44375@umbc.edu.

has led to a highly interesting spline model defined by a linear control system, which is called a *control theoretic spline* [9]. It is shown in [9] and the references therein, e.g., [13, 14, 18, 36, 40], that a number of smoothing, interpolation, and path planning problems can be incorporated into this paradigm and studied using control theory and optimization techniques on Hilbert spaces with efficient numerical schemes. Other relevant approaches include control theoretic wavelets [11].

Despite significant progress and many important results for *unconstrained* spline models, various models of biological, engineering and economic systems contain functions whose shape and/or dynamics are subject to *inequality* constraints, e.g., the monotone and convex constraints. Examples include monotone regulatory functions in genetic networks [33] and a shape restricted function in an attitude control system [26]. Other applications are found in reliability engineering (e.g., survival/hazard functions), medicine (e.g., dose-response curves), finance (e.g., option/delivery price), and astronomy (e.g., galaxy mass functions). Incorporating the knowledge of shape constraints into a construction procedure improves estimation efficiency and accuracy [25]. This has raised surging interest in the study of constrained splines and shape restricted estimation in statistics and other related fields; see some recent results [19, 32, 34, 35, 39].

The present paper focuses on shape restricted smoothing splines formulated as a constrained linear optimal control problem with unknown initial state and control. Two types of constraints arise in this framework: (i) control constraints; and (ii) state constraints. General state constraints are very difficult to approach both analytically and numerically, except in a few simple cases [7, 9, 12, 14]. Being more tractable, control constraints receive increasing attention, as a variety of shape constraints, which may be imposed on derivatives of a function, can be easily formulated as control constraints. It should be noted that a control constrained optimal control problem is inherently nonsmooth, and thus is considerably different from a classical (unconstrained) linear optimal control problem. Nevertheless, most of the current literature focuses on relatively simpler linear dynamics and special control constraints, e.g., [9, 14, 18, 19, 37], and many critical questions remain open in analysis and computation when general dynamics and control constraints are taken into account. For example, a widely used approach in the literature concentrates on shape restricted smoothing splines whose linear dynamics are defined by certain nilpotent matrices, and whose control constraints are given by a cone in $\mathbb{R}$ [8, 17]. In this case, the smoothing spline is a piecewise continuous polynomial with a known degree. Hence computation of the smoothing spline boils down to determining parameters of a polynomial on each interval, which can be further reduced to a quadratic or semidefinite program that attains efficient algorithms [7, 9, 18]. However, this approach fails to deal with general linear dynamics and control constraints, since the solution form of a general shape restricted smoothing spline is unknown *a priori*. Therefore new tools are needed to handle complex dynamics and general control constraint induced nonsmoothness.

In this paper, we develop new analytical and numerical results for smoothing splines subject to general dynamics and control constraints by using optimal control and nonsmooth optimization techniques. The major contributions of the paper are summarized as follows:

1. Optimal control formulation and analysis. By using the Hilbert space method and convex and variational techniques, optimality conditions are established for the constrained optimal control problem of the shape restricted smoothing spline in the form of variational inequalities. These optimality conditions yield a nonsmooth equation of an optimal initial condition; the unique solution of this equation completely determines an optimal control and thus the desired smoothing spline.

2. Numerical computation and convergence analysis. To solve the indicated equation of an optimal initial condition, we show the B-differentiability and other nonsmooth properties of this equation. A modified nonsmooth Newton's algorithm with line search is invoked to solve the equation. This algorithm does not require knowing the solution form of a smoothing spline *a priori*. A major part of the paper is devoted to convergence analysis of this algorithm. By using various

techniques from nonsmooth analysis, polyhedral theory, and piecewise affine switching systems, we establish the global convergence of the proposed algorithm for a general polyhedral control constraint under mild technical conditions.

The paper is organized as follows. In Section 2, we formulate a shape restricted smoothing spline as a constrained optimal control problem with optimality conditions developed in Section 3. A nonsmooth Newton's algorithm for the smoothing spline is proposed in Section 4; its convergence analysis and numerical results are presented in Section 5 and Section 6 respectively. Finally, conclusions are made in Section 7.

2 Shape Restricted Smoothing Splines: Constrained Optimal Control Formulation

Consider the linear control system on $\mathbb{R}^\ell$ subject to control constraint:

$$\dot{x} = Ax + Bu, \qquad y = Cx, \tag{2}$$

where $A \in \mathbb{R}^{\ell\times\ell}$, $B \in \mathbb{R}^{\ell\times m}$, and $C \in \mathbb{R}^{p\times\ell}$. Let $\Omega \subseteq \mathbb{R}^m$ be a closed convex set. The control constraint is given by $u \in L_2([0,1];\mathbb{R}^m)$ and $u(t) \in \Omega$ for almost all $t \in [0,1]$, where $L_2([0,1];\mathbb{R}^m)$ is the space of square $\mathbb{R}^m$-valued (Lebesgue) integrable functions. We denote this constrained linear control system by $\Sigma(A,B,C,\Omega)$. Define the set of permissible controls, which is clearly convex:

$$\mathcal{W} := \Big\{ u \in L_2([0,1];\mathbb{R}^m) \,|\, u(t) \in \Omega, \text{ a.e. } [0,1] \Big\}.$$

Let the underlying function $f : [0,1] \to \mathbb{R}^p$ be the output $f(t) := Cx(t)$ for an absolutely continuous trajectory $x(t)$ of $\Sigma(A,B,C,\Omega)$, which can be completely determined by its initial state and control. This leads to the following (generalized) regression problem on the interval $[0,1]$:

$$y_i \;=\; f(t_i) + \varepsilon_i, \qquad i = 0,1,\ldots,n, \tag{3}$$

where t_i's are the pre-specified design points with $0 = t_0 < t_1 < \cdots < t_n = 1$, $y_i \in \mathbb{R}^p$ are samples, and $\varepsilon_i \in \mathbb{R}^p$ are noise or errors. Given the sample observation $(t_i, y_i)_{i=0}^n$, and let $w_i > 0, i = 1,\ldots,n$ be such that $\sum_{i=1}^n w_i = 1$ (e.g., $w_i = t_i - t_{i-1}$). Define the cost functional

$$J := \sum_{i=1}^n w_i \big\| y_i - Cx(t_i) \big\|_2^2 + \lambda \int_0^1 \|u(t)\|_2^2 dt, \tag{4}$$

where $\lambda > 0$ is the penalty parameter. The goal of a shape restricted smoothing spline is to find an absolutely continuous trajectory $x(t)$ (which is determined by its initial state and control) that minimizes the cost functional J subject to the dynamics of the linear control system $\Sigma(A,B,C,\Omega)$ (2) and the control constraint $u \in \mathcal{W}$.

Remark 2.1. Let $R \in \mathbb{R}^{m\times m}$ be a symmetric positive definite matrix. A more general cost functional

$$J := \sum_{i=1}^n w_i \big\| y_i - Cx(t_i) \big\|_2^2 + \lambda \int_0^1 u^T(t) R u(t) dt \tag{5}$$

may be considered. However, a suitable control transformation will yield an equivalent problem defined by the cost functional (4). In fact, let $R = P^T P$ for an invertible matrix P. Let $v(t) = Pu(t)$, $\Omega' = P\Omega$, and $\mathcal{W}' := \big\{ v \in L_2([0,1];\mathbb{R}^m) \,|\, v(t) \in \Omega', \text{ a.e. } [0,1] \big\}$. Clearly, Ω' remains closed and convex, and $\mathcal{W}'$ is still convex. Therefore the constrained optimal control problem defined by (5) for the linear control system $\Sigma(A,B,C,\Omega)$ is equivalent to that defined by (4) with u replaced by v for the linear system $\Sigma(A, BP^{-1}, C, \Omega')$ subject to the constraint $(v, x_0) \in \mathcal{W}' \times \mathbb{R}^\ell$.

Example 2.1. The constrained linear control model (2) covers a wide range of estimation problems subject to shape and/or dynamical constraints. For instance, the standard monotone regression problem is a special case of the model (2) by letting the scalars $A = 0$, $B = C = 1$, and $\Omega = \mathbb{R}_+$. Another case is the convex regression, for which

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{R}^2, \quad C^T = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathbb{R}^2, \quad \Omega = \mathbb{R}_+.$$

3 Optimality Conditions of Shape Restricted Smoothing Splines

This section develops optimality conditions of the finite-horizon constrained optimal control problem (4) using the Hilbert space techniques. We first introduce the following functions $P_i : [0, 1] \to \mathbb{R}^{p\times m}$ inspired by [9]:

$$P_i(t) := \begin{cases} Ce^{A(t_i - t)}B, & \text{if } \ t \in [0, t_i] \\ 0, & \text{if } \ t > t_i \end{cases}, \quad i = 1, \dots, n.$$

Hence,

$$f(t_i) = Cx(t_i) = Ce^{At_i}x_0 + \int_0^1 P_i(t)u(t)dt, \quad i = 1, \dots, n,$$

where x_0 denotes the initial state of $x(t)$. Define the set $\mathcal{P} := \mathcal{W} \times \mathbb{R}^\ell$. It is easy to verify that $\mathcal{P}$ is convex. The constrained optimal control problem is formulated as

$$\inf_{(u,x_0)\in\mathcal{P}} J(u, x_0), \tag{6}$$

where $J : \mathcal{P} \to \mathbb{R}_+$ is given by

$$J(u, x_0) := \sum_{i=1}^n w_i \left\| Ce^{At_i}x_0 + \int_0^1 P_i(t)u(t)dt - y_i \right\|_2^2 + \lambda \int_0^1 \|u(t)\|_2^2 dt.$$

For given design points $\{t_i\}_{i=1}^n$ in $[0, 1]$, we introduce the following condition:

$$\mathbf{H}.\mathbf{1}: \quad \text{rank} \begin{pmatrix} Ce^{At_1} \\ Ce^{At_2} \\ \vdots \\ Ce^{At_n} \end{pmatrix} = \ell.$$

It is easy to see, via $t_i \in [0, 1]$ for all i, that if (C, A) is an observable pair, then the condition **H.1** holds for all sufficiently large n. Under this condition, the existence and uniqueness of an optimal solution can be shown via standard arguments in functional analysis, e.g., [2, 15, 16]. We present its proof in the following theorem for self-containment.

Theorem 3.1. *Suppose $\{(t_i, y_i)\}$, $\{w_i\}$, and $\lambda > 0$ are given. Under the condition* **H.1***, the optimization problem (6) has a unique optimal solution $(u_*, x_0^*) \in \mathcal{P}$.*

Proof. Consider the Hilbert space $L_2([0, 1]; \mathbb{R}^m) \times \mathbb{R}^\ell$ endowed with the inner product $\langle (u, x), (v, z) \rangle := \int_0^1 u^T(t)v(t)dt + x^T z$ for any $(u, x), (v, z) \in L_2([0, 1]; \mathbb{R}^m) \times \mathbb{R}^\ell$. Its induced norm satisfies $\|(u, x)\|^2 := \|u\|_{L_2}^2 + \|x\|_2^2$, where $\|u\|_{L_2} := \Big(\int_0^1 u^T(t)u(t)dt \Big)^{1/2}$ for any $u \in L_2([0, 1]; \mathbb{R}^m)$ and the latter $\| \cdot \|_2$ is the Euclidean norm on $\mathbb{R}^\ell$. The following properties of $J : L_2([0, 1]; \mathbb{R}^m) \times \mathbb{R}^\ell \to \mathbb{R}_+$ can be easily verified via the positive definiteness of the matrix $\sum_{i=1}^n w_i(Ce^{At_i})^T(Ce^{At_i}) \in \mathbb{R}^{\ell\times\ell}$ due to **H.1**:

(i) J is coercive, i.e., for any sequence $\{(u_k, x_k)\}$ with $\|(u_k, x_k)\| \to \infty$ as $k \to \infty$, $J(u_k, x_k) \to \infty$ as $k \to \infty$.

(ii) J is strictly convex, i.e., for any $(u, x), (v, z) \in L_2([0, 1]; \mathbb{R}^m) \times \mathbb{R}^\ell$, $J(\alpha(u, x) + (1 - \alpha)(v, z)) < \alpha J(u, x) + (1 - \alpha)J(v, z)$, $\forall\ \alpha \in (0, 1)$.

Pick an arbitrary $(\widetilde{u}, \widetilde{x}) \in \mathcal{P}$ and define the level set $\mathcal{S} := \{(u, x) \in \mathcal{P} : J(u, x) \leq J(\widetilde{u}, \widetilde{x})\}$. Due to the convexity and the coercive property of J, $\mathcal{S}$ is a convex and (L_2-norm) bounded set in $L_2([0, 1]; \mathbb{R}^m) \times \mathbb{R}^\ell$. Since the Hilbert space $L_2([0, 1]; \mathbb{R}^m) \times \mathbb{R}^\ell$ is reflexive and self dual, it follows from Banach-Alaoglu Theorem [16] that an arbitrary sequence $\{(u_k, x_k)\}$ in $\mathcal{S}$ with $u_k \in \mathcal{W}$ and $x_k \in \mathbb{R}^\ell$ has a subsequence $\{(u'_k, x'_k)\}$ that attains a weak*, thus weak, limit $(u_*, x^*) \in L_2([0, 1]; \mathbb{R}^m) \times \mathbb{R}^\ell$. Clearly, $x^* \in \mathbb{R}^\ell$. Without loss of generality, we assume that for each u'_k, $u'_k(t) \in \Omega$ for all $t \in [0, 1]$. Therefore, $Ce^{At_i}x'_k + \int_0^1 P_i(t)u'_k(t)dt$ converges to $Ce^{At_i}x^* + \int_0^1 P_i(t)u_*(t)dt$ for each i.

Next we show that $u_* \in \mathcal{W}$ via the closedness and convexity of Ω. In view of the weak convergence of (u'_k) to u_*, it follows from Mazur's Lemma [24, Lemma 10.19] that there exists a sequence of convex combinations of (u'_k), denoted by (v_k), that converges to u_* strongly in $L_2([0, 1]; \mathbb{R}^m)$, i.e., for each k, there exist an integer $p_k \geq k$ and real numbers $\lambda_{k,j} \geq 0, k \leq j \leq p_k$ with $\sum_{j=k}^{p_k} \lambda_{k,j} = 1$ such that $v_k = \sum_{j=k}^{p_k} \lambda_{k,j}u_k$, and $\|v_k - u_*\|_{L_2} \to 0$ as $k \to \infty$. Since each $u_k(t) \in \Omega, \forall\, t \in [0, 1]$, the same holds for each v_k via the convexity of Ω. Furthermore, due to the strong convergence of (v_k) to u_* (i.e., in the L_2-norm), (v_k) converges to u_* in measure [3, pp. 69], and hence has a subsequence that converges to u_* pointwise almost everywhere on $[0, 1]$ (cf. [3, Theorem 7.6] or [15, Theorem 5.2]). Since Ω is closed, $u_*(t) \in \Omega$ for almost all $t \in [0, 1]$. This shows that $u_* \in \mathcal{W}$.

Furthermore, by using the (L_2-norm) boundedness of (u'_k) and the triangle inequality for the L_2-norm, it is easy to show that for any $\eta > 0$, there exists $K \in \mathbb{N}$ such that $\|u_*\|^2_{L_2} \leq \|u'_k\|^2_{L_2} + \eta, \forall\ k \geq K$. These results imply that for any $\varepsilon > 0$, $J(u_*, x^*) \leq J(u'_k, x'_k) + \varepsilon$ for all k sufficiently large. Consequently, $J(u_*, x^*) \leq J(\widetilde{u}, \widetilde{x})$ such that $(u_*, x^*) \in \mathcal{S}$. This thus shows that $\mathcal{S}$ is sequentially compact. In view of (strong) continuity of J, we see that a global optimal solution exists on $\mathcal{S}$ [16, Section 5.10, Theorem 2], and thus on $\mathcal{P}$. Moreover, since J is strictly convex and the set $\mathcal{P}$ is convex, the optimal solution (u_*, x_0^*) must be unique. □

The next result provides the necessary and sufficient optimality conditions in terms of variational inequalities. Toward this end, we introduce the *variational inequality* (VI) on $\mathbb{R}^n$. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a function and $\mathcal{K}$ be a closed convex set in $\mathbb{R}^n$. The variational inequality defined by $\mathcal{K}$ and F, denoted by VI($\mathcal{K}, F$), is to find a vector $z^* \in \mathcal{K}$ such that $\langle z - z^*, F(z^*)\rangle \geq 0$ for all $z \in \mathcal{K}$. In what follows, we let SOL($\mathcal{K}, F$) denote the solution set of VI($\mathcal{K}, F$). It should be pointed out that a VI is a highly nonlinear and nonsmooth problem, even if F is linear and $\mathcal{K}$ is polyhedral. See the monograph [10] and the references therein for comprehensive discussions. In what follows, given a closed convex set $\mathcal{S}$ in $\mathbb{R}^n$, we use $\Pi_{\mathcal{S}}(z)$ to denote the Euclidean projection of $z \in \mathbb{R}^n$ onto $\mathcal{S}$ [10]. It is known that $\Pi_{\mathcal{S}}(\cdot)$ is Lipschitz continuous on $\mathbb{R}^n$ with the Lipschitz constant $L = 1$ when the Euclidean norm is used (i.e., the non-expansive property of $\Pi_{\mathcal{S}}$).

Theorem 3.2. *The pair $(u_*, x_0^*) \in \mathcal{P}$ is an optimal solution to (6) if and only if the following two conditions hold:*

$$u_*(t) \ = \ \Pi_\Omega\big(- G(t, u_*(t), x_0^*)/\lambda\big), \quad a.e.\ \ [0, 1], \tag{7}$$

$$L(u_*, x_0^*) \ = \ 0, \tag{8}$$

where

$$G(t, u_*(t), x_0^*) := \sum_{i=1}^n w_i P_i^T(t)\Big(Ce^{At_i}x_0^* + \int_0^1 P_i(t)u_*(t)dt - y_i\Big), \tag{9}$$

and

$$L(u_*, x_0^*) := \sum_{i=1}^{n} w_i \big(Ce^{At_i}\big)^T \Big(Ce^{At_i} x_0^* + \int_0^1 P_i(t) u_*(t) dt - y_i\Big).$$

Proof. Let $(u', x') \in \mathcal{P}$ be arbitrary. Due to the convexity of $\mathcal{P}$, $(u_*, x_0^*) + \varepsilon[(u', x') - (u_*, x_0^*)] \in \mathcal{P}$ for all $\varepsilon \in [0, 1]$. Further, since (u_*, x_0^*) is a global optimizer, we have $J((u_*, x_0^*) + \varepsilon[(u', x') - (u_*, x_0^*)]) \geq J(u_*, x_0^*)$ for all $\varepsilon \in [0, 1]$. Therefore

$$\begin{aligned} 0 \;\leq\; & \lim_{\varepsilon \downarrow 0} \frac{J((u_*, x_0^*) + \varepsilon[(u', x') - (u_*, x_0^*)]) - J(u_*, x_0^*)}{\varepsilon} \\ =\; & 2\Bigg[\sum_{i=1}^{n} w_i \left\langle Ce^{At_i} x_0^* + \int_0^1 P_i(t) u_*(t) dt - y_i,\; Ce^{At_i}(x' - x_0^*) + \int_0^1 P_i(t)\Big(u'(t) - u_*(t)\Big) dt \right\rangle \\ & \quad + \lambda \int_0^1 u_*(t)^T \Big(u'(t) - u_*(t)\Big) dt\Bigg]. \end{aligned}$$

This thus yields the necessary optimality condition: for all $(u', x') \in \mathcal{P}$,

$$\begin{aligned} \sum_{i=1}^{n} w_i \left\langle Ce^{At_i} x_0^* + \int_0^1 P_i(t) u_*(t) dt - y_i,\; Ce^{At_i}(x' - x_0^*) + \int_0^1 P_i(t)\Big(u'(t) - u_*(t)\Big) dt \right\rangle & \\ + \lambda \int_0^1 u_*(t)^T \Big(u'(t) - u_*(t)\Big) dt \;\geq\; 0. & \qquad (10) \end{aligned}$$

This condition is also sufficient in light of the following inequality due to the convexity of J:

$$J(u', x') - J(u_*, x_0^*) \;\;\geq\;\; \lim_{\varepsilon \downarrow 0} \frac{J((u_*, x_0^*) + \varepsilon[(u', x') - (u_*, x_0^*)]) - J(u_*, x_0^*)}{\varepsilon}, \quad \forall\, (u', x') \in \mathcal{P}.$$

We now show that the optimality condition (10) is equivalent to

$$\Big\langle v(t) - u_*(t),\; \lambda u_*(t) + G(t, u_*(t), x_0^*) \Big\rangle \;\geq\; 0, \text{ a.e. } [0, 1], \; \forall\, v \in \mathcal{W}, \qquad (11)$$

where $G(t, u_*(t), x_0^*)$ is given in (9), and

$$\Big\langle L(x_0^*, u_*),\; x' - x_0^* \Big\rangle \;\geq\; 0, \quad \forall\, x' \in \mathbb{R}^\ell. \qquad (12)$$

Clearly, if (11) and (12) hold, then (10) holds. Conversely, by setting $u' = u_*$, we have from (10) that

$$\sum_{i=1}^{n} w_i \left\langle Ce^{At_i} x_0^* + \int_0^1 P_i(t) u_*(t) dt - y_i,\; Ce^{At_i}(x' - x_0^*) \right\rangle \geq 0, \quad \forall\; x' \in \mathbb{R}^\ell.$$

Since x' is arbitrary in $\mathbb{R}^\ell$, this yields (12) and thus (8). Furthermore, the condition (10) is reduced to

$$\int_0^1 \Big\langle u'(t) - u_*(t),\; \lambda u_*(t) + G(t, u_*(t), x_0^*) \Big\rangle dt \;\geq\; 0, \quad \forall\; u' \in \mathcal{W}.$$

Let $\widetilde{G}(t, u_*(t), x_0^*) := \lambda u_*(t) + G(t, u_*(t), x_0^*)$. Since $\widetilde{G} \in L_2([0, 1]; \mathbb{R}^m)$, $u' \in L_2([0, 1]; \mathbb{R}^m)$ and Ω is closed and convex, it follows from [22, Section 2.1] that the above integral inequality is equivalent to the variational inequality (11), which is further equivalent to $u_*(t) \in \mathrm{SOL}\big(\Omega, \widetilde{G}(t, \cdot, x_0^*)\big)$, a.e. $[0, 1]$. Hence, for almost all $t \in [0, 1]$,

$$\Big\langle w - u_*(t),\; u_*(t) + G(t, u_*(t), x_0^*)/\lambda \Big\rangle \;\geq\; 0, \quad \forall\; w \in \Omega.$$

This shows $u_*(t) = \Pi_\Omega(-G(t, u_*(t), x_0^*)/\lambda)$ a.e. $[0, 1]$. □

In what follows, we further develop the optimal control solution for the shape restricted smoothing spline. Let $\widehat{f}(t, x_0^*)$ denote the shape restricted smoothing spline for the given $\{y_i\}$, i.e.,

$$\widehat{f}(t, x_0^*) := Ce^{At}x_0^* + \int_0^1 Ce^{A(t-s)}Bu_*(s, x_0^*)ds,$$

where u_* is the optimal control, and x_0^* is the optimal initial state.

Corollary 3.1. *The shape restricted smoothing spline $\widehat{f}(t, x_0^*)$ satisfies*

$$\sum_{i=1}^{n} w_i \big(Ce^{At_i}\big)^T \big(\widehat{f}(t_i, x_0^*) - y_i\big) = 0, \tag{13}$$

and $G(t, u_*(t), x_0^*)$ in (9) is given by

$$G(t, u_*(t), x_0^*) = \begin{cases} 0, & \forall\ t \in [0, t_1) \\ -\displaystyle\sum_{i=1}^{k} w_i \big(Ce^{A(t_i - t)}B\big)^T \big(\widehat{f}(t_i, x_0^*) - y_i\big), & \forall\ t \in [t_k, t_{k+1}),\ k = 1, \dots, n-1 \end{cases} \tag{14}$$

Proof. Note that $\widehat{f}(t_i, x_0^*) = Ce^{At_i}x_0^* + \int_0^1 P_i(s)u_*(s, x_0^*)ds$ for $i = 1, \dots, n$. In light of (8) and the definition of $L(u_*, x_0^*)$, we obtain (13). To establish (14), letting $\mathbf{I}_S$ denote the indicator function of a set S, we see from (13) that

$$\begin{aligned} &\sum_{i=1}^{n} w_i \Big(Ce^{At_i}\Big)^T \Big(\widehat{f}(t_i, x_0^*) - y_i\Big)\mathbf{I}_{[0,t_i]} \\ &= \begin{cases} 0, & \forall\ t \in [0, t_1) \\ -\displaystyle\sum_{i=1}^{k} w_i \big(Ce^{At_i}\big)^T \big(\widehat{f}(t_i, x_0^*) - y_i\big), & \forall\ t \in [t_k, t_{k+1}),\ k = 1, \dots, n-1 \end{cases}. \end{aligned} \tag{15}$$

Moreover, it follows from (9) and the definition of P_i that

$$\begin{aligned} G(t, u_*(t), x_0^*) &= \sum_{i=1}^{n} w_i \Big(Ce^{A(t_i - t)}B \cdot \mathbf{I}_{[0,t_i]}\Big)^T \Big(\widehat{f}(t_i, x_0^*) - y_i\Big) \\ &= \sum_{i=1}^{n} w_i \Big(Ce^{At_i}e^{-At}B\Big)^T \cdot \mathbf{I}_{[0,t_i]} \cdot \Big(\widehat{f}(t_i, x_0^*) - y_i\Big) \\ &= \Big(e^{-At}B\Big)^T \sum_{i=1}^{n} w_i \Big(Ce^{At_i}\Big)^T \Big(\widehat{f}(t_i, x_0^*) - y_i\Big)\mathbf{I}_{[0,t_i]}. \end{aligned}$$

By virtue of this and (15), we obtain (14). □

This corollary shows that if x_0^* is known, then the smoothing spline $\widehat{f}(t, x_0^*)$ can be determined inductively. In fact, on each interval $[t_k, t_{k+1})$, $G(t, u_*(t), x_0^*)$, and thus $u_*(t)$, depends on $\widehat{f}(t_i, x_0^*)$ with $t_i \le t_k$ only. This property will be exploited for numerical computation of the shape constrained smoothing spline in Section 4.

We mention a few special cases of VIs which are of particular interest as follows. Let $\perp$ denote the orthogonality of two vectors in $\mathbb{R}^n$, i.e., $a \perp b$ implies $a^Tb = 0$. If $\mathcal{K}$ is the closed convex cone $\mathcal{C}$, then $z^* \in \text{SOL}(\mathcal{K}, F)$ if and only if $\mathcal{C} \ni z^* \perp F(z^*) \in \mathcal{C}^*$, where $\mathcal{C}^*$ is the dual cone of $\mathcal{C}$. In particular, if $\mathcal{K}$ is the nonnegative orthant $\mathbb{R}^n_+$, then $z^* \in \text{SOL}(\mathcal{K}, F)$ if and only if $0 \le z^* \perp F(z^*) \ge 0$, where the latter is called the *complementarity problem* (cf. [5, 10] for details). Especially, when $F(z)$ is affine, i.e., $F(z) = Mz+q$ for a square matrix M and a vector q, then the complementarity problem becomes the linear complementarity problem (LCP). Another special case of significant interest is when $\mathcal{K}$ is a polyhedron, namely, $\mathcal{K} = \{z \in \mathbb{R}^n \,|\, Dz \geq b,\ Ez = d\}$, where $D \in \mathbb{R}^{r\times n}$, $E \in \mathbb{R}^{q\times n}$, and $b \in \mathbb{R}^r$, $d \in \mathbb{R}^q$. In this case, it is well known that $z_* \in \mathrm{SOL}(\mathcal{K}, F)$ if and only if there exist multipliers $\chi \in \mathbb{R}^r, \mu \in \mathbb{R}^q$ such that $F(z_*) - D^T\chi + E^T\mu = 0,\ 0 \leq \chi \perp Dz_* - b \geq 0,\ Ez_* - d = 0$ [10, Proposition 1.2.1].

Along with the above results, we obtain the following optimality condition for u_* in term of a complementarity problem when Ω is polyhedral.

Proposition 3.1. *Let $\Omega = \{w \in \mathbb{R}^m \,|\, Dw \geq b\}$ be a (nonempty) polyhedron with $D \in \mathbb{R}^{r\times m}$ and $b \in \mathbb{R}^r$. Then*

$$u_*(t) \;=\; \Big[- G(t, u_*(t), x_0^*) + D^T\chi(G(t, u_*(t), x_0^*))\Big]/\lambda, \qquad a.e.\ \ [0,1],$$

where $D^T\chi : \mathbb{R}^m \to \mathbb{R}^m$ is a continuous piecewise affine function defined by the solution of the linear complementarity problem:

$$0 \leq \chi \perp \lambda^{-1}DD^T\chi - \lambda^{-1}Dz - b \geq 0.$$

Proof. It follows from $u_* \in \mathrm{SOL}(\Omega, \widetilde{G}(t, \cdot, x_0^*))$ a.e. $[0, 1]$, where $\widetilde{G}(t, u_*(t), x_0^*) = \lambda u_* + G(u_*(t), x^*)$, and the above discussions that u_* is the optimal control if and only if for almost all $t \in [0, 1]$, there exists $\chi \in \mathbb{R}^r$ such that

$$\lambda u_*(t) + G(u_*(t), x_0^*) - D^T\chi = 0, \qquad \text{and} \qquad 0 \leq \chi \perp Du_*(t) - b \geq 0.$$

This is equivalent to the linear complementarity problem

$$0 \leq \chi \perp \lambda^{-1}DD^T\chi - \lambda^{-1}z - b \geq 0, \tag{16}$$

where $z := G(u_*(t), x_0^*)$. Due to the positive semidefinite plus (PSD-plus) structure [31], it follows from complementarity theory [5] that for any $z \in \mathbb{R}^m$, the LCP (16) has a solution $\chi(z)$ and $D^T\chi(z)$ is unique. To be self-contained, we present this argument below.

We first show that the LCP (16) is feasible for any z, i.e., there exists χ such that $\chi \geq 0$ and $(\lambda)^{-1}DD^T\chi - \lambda^{-1}Dz - b \geq 0$. Suppose not. Then it follows from the Theorem of the Alternative (or a version of Farkas' Lemma, e.g., [5, Theorem 2.7.9]) that there exists a vector v such that

$$v^T\lambda^{-1}DD^T \leq 0, \quad v \geq 0, \quad v^T\big[\lambda^{-1}Dz + b\big] > 0.$$

Clearly the first two inequalities show that $D^Tv = 0$. Together with the last inequality, we also have $v^Tb > 0$. Since the polyhedron Ω is nonempty, there exists w such that $Dw - b \geq 0$. In view of $v \geq 0$, we have $v^T(Dw - b) \geq 0$. On the other hand, $v^T(Dw - b) = -v^Tb < 0$, a contradiction. This shows the feasibility of the LCP. Moreover, since DD^T is positive semidefinite, the feasibility leads to the solvability of the LCP [5, Theorem 3.1.2]. The uniqueness of $D^T\chi$ follows from the standard sign reverse argument in the LCP theory. In fact, for a fixed z, let χ^1, χ^2 be two solutions. Since $(\chi^1 - \chi^2)_i\big(\lambda^{-1}DD^T(\chi^1 - \chi^2)\big)_i \leq 0, \forall\ i = 1, \ldots, r$, we have $(\chi^1 - \chi^2)^T\lambda^{-1}DD^T(\chi^1 - \chi^2) \leq 0$, yielding $D^T\chi^1 = D^T\chi^2$. This shows the uniqueness of $D^T\chi(z)$ and also implies that $D^T\chi(\cdot)$ is continuous and piecewise affine. □

4 Computation of Shape Restricted Smoothing Splines via Nonsmooth Newton's Method

In this section, we develop a numerical algorithm for the computation of a general class of shape restricted smoothing splines based on a nonsmooth Newton's method. As indicated below Corollary 3.1, in order to determine the smoothing spline $\widehat{f}(t, x_0^*)$, it suffices to find the optimal initial state x_0^*, since once x_0^* is known, u_* and $\widehat{f}$ can be computed recursively. In fact, it follows from Corollary 3.1 that $\widehat{f}(t, x_0^*)$ is given by

$$\widehat{f}(t, x_0^*) \;=\; Ce^{At}x_0^* + \int_0^t Ce^{A(t-s)}Bu_*(s, x_0^*)ds, \tag{17}$$

where

$$u_*(t, x_0^*) \;=\; \begin{cases} \Pi_\Omega(0), & \forall\, t \in [0, t_1) \\ \Pi_\Omega\Big(\lambda^{-1}\sum_{i=1}^{k} w_i\big(Ce^{A(t_i-t)}B\big)^T\big(\widehat{f}(t_i, x_0^*) - y_i\big)\Big), & \forall\, t \in [t_k, t_{k+1}),\ k = 1, \dots, n-1 \end{cases}$$

and $\widehat{f}(t, x_0^*)$ satisfies

$$H_{y,n}(x_0^*) \;:=\; \sum_{i=1}^{n} w_i\Big(Ce^{At_i}\Big)^T\Big(\widehat{f}(t_i, x_0^*) - y_i\Big) \;=\; 0. \tag{18}$$

To compute the optimal initial state x_0^*, we consider the equation $H_{y,n}(z) = 0$, where $\widehat{f}(t, z)$ in $H_{y,n}(z)$ is defined by (17) with x_0^* replaced by z. The following lemma is a direct consequence of Theorem 3.1 and the definition of $\widehat{f}$.

Lemma 4.1. *For any given $\{(t_i, y_i)\}$, $\{w_i\}$, and $\lambda > 0$, the equation $H_{y,n}(z) = 0$ defined above has a unique solution, which corresponds to the optimal initial state x_0^* of the smoothing spline $\widehat{f}(t, x_0^*)$.*

It should be noted that the function $H_{y,n} : \mathbb{R}^\ell \to \mathbb{R}^\ell$ is nonsmooth in general, due to the constraint induced nonsmoothness of $u_*(t, x_0^*)$ in x_0^*. However, the following proposition shows the B(ouligand)-differentiability of $\widehat{f}(t, z)$ in z [10, Section 3.1]. Recall that a function $G : \mathbb{R}^\ell \to \mathbb{R}^\ell$ is B-differentiable if it is Lipschitz continuous and directionally differentiable on $\mathbb{R}^\ell$, namely, for any $z \in \mathbb{R}^\ell$ and any direction vector $d \in \mathbb{R}^\ell$, the following (one-sided) directional derivative exists

$$G'(z; d) := \lim_{\tau \downarrow 0} \frac{G(z + \tau d) - G(z)}{\tau}.$$

Proposition 4.1. *Assume that $\Pi_\Omega : \mathbb{R}^m \to \mathbb{R}^m$ is directionally differentiable on $\mathbb{R}^m$. For any given $\{(t_i, y_i)\}$, $\{w_i\}$, $\lambda > 0$, and $z \in \mathbb{R}^\ell$, the function $\widehat{f}(t, z)$ is B-differentiable in z for any fixed $t \in [0, 1]$.*

Proof. We prove the B-differentiability of $\widehat{f}(t, z)$ in z by induction on the intervals $[t_k, t_{k+1}], k = 0, 1, \dots, n-1$. Consider $t \in [0, t_1]$ first. Since $u_*(t, z) = \Pi_\Omega(0), \forall\, t \in [0, t_1)$ and $\widehat{f}(t, z)$ is continuous in t, $\widehat{f}(t, z) = Ce^{At}z + \int_0^t Ce^{A(t-s)}B\,\Pi_\Omega(0)ds, \forall\, t \in [0, t_1]$, which is clearly Lipschitz continuous and directionally differentiable, thus B-differentiable, in z for any fixed $t \in [0, t_1]$.

Now assume that $\widehat{f}(t, \cdot)$ is B-differentiable for all $t \in [0, t_1] \cup \cdots \cup [t_{k-1}, t_k]$, and consider the interval $[t_k, t_{k+1}]$. Note that for any $t \in [t_k, t_{k+1})$, the optimal control

$$u_*(t, z) \;=\; \Pi_\Omega\left(\lambda^{-1}\sum_{i=1}^{k} w_i\Big(Ce^{A(t_i-t)}B\Big)^T\Big(\widehat{f}(t_i, z) - y_i\Big)\right). \tag{19}$$

Since the functions $\Pi_\Omega(\cdot)$ and $\widehat{f}(t_i, \cdot), i = 1, \ldots, k$ are all B-differentiable, it follows from [10, Proposition 3.1.6] that the composition given in $u_*(t, z)$ remains B-differentiable in z for each fixed $t \in [t_k, t_{k+1})$. For a given direction vector $d \in \mathbb{R}^\ell$ and a given $\tau \geq 0$, $u_*(t, z + \tau d)$ is continuous in t on $[t_k, t_{k+1})$. Therefore, $u_*(t, z + \tau d)$ is (Borel) measurable on $[t_k, t_{k+1})$ for any fixed τ and d. Since

$$u'_*(t, z; d) = \lim_{\tau \downarrow 0} \frac{u_*(t, z + \tau d) - u_*(t, z)}{\tau},$$

the function $u'_*(t, z; d)$ is also (Borel) measurable on $[t_k, t_{k+1})$ for any fixed z and d [3, Corollary 2.10 or Corollary 5.9]. Moreover, it is easy to show that $u'_*(t, z; d)$ is bounded on the interval $[t_k, t_{k+1})$, i.e., there exists $\varrho_k > 0$ such that $\|u'_*(t, z; d)\| \leq \varrho_k$ for all $t \in [t_k, t_{k+1})$. This shows that $u'_*(t, z; d)$ is (Lebesgue) integrable in t on $[t_k, t_{k+1}]$. In view of this and the Lebesgue Dominated Convergence Theorem [3, Theorem 5.6], we have $\widehat{f}'(t, z; d) = Ce^{At}d + \int_0^t Ce^{A(t-s)}Bu'_*(s, z; d)ds$ for all $t \in [t_k, t_{k+1}]$. This shows that $\widehat{f}(t, \cdot)$ is directionally differentiable for each $t \in [t_k, t_{k+1}]$. Furthermore, since $\|\Pi_\Omega(z) - \Pi_\Omega(z')\|_2 \leq \|z - z'\|_2$ for all $z, z' \in \mathbb{R}^\ell$, and $u_*(t, z)$ depends on finitely many $\widehat{f}(t_i, z)$ on the interval $[t_j, t_{j+1})$ with $j = 1, \ldots, k$ (cf. (19)), it can be shown that for each $j = 1, \ldots, k$, there exists a uniform Lipschitz constant $L_j > 0$ (independent of t) such that for any $t \in [t_j, t_{j+1})$, $\|u_*(t, z) - u_*(t, z')\|_2 \leq L_j\|z - z'\|_2$ for all $z, z' \in \mathbb{R}^\ell$. In view of $\widehat{f}(t, z) = Ce^{At}z + \int_0^t Ce^{A(t-s)}Bu_*(s, z)ds$, the continuity of $\widehat{f}$ in t, and the induction hypothesis, we deduce the Lipschitz continuity of $\widehat{f}(t, \cdot)$ for each fixed $t \in [t_k, t_{k+1}]$. Therefore, the proposition follows by the induction principle. □

Clearly, the assumption of global directional differentiability of the Euclidean projector Π_Ω is critical to Proposition 4.1. In what follows, we identify a few important cases where this assumption holds. One of the most important cases is when Ω is polyhedral. In this case, as shown in Proposition 3.1, $\Pi_\Omega(\cdot)$ is a continuous piecewise affine function, and its directional derivative is given by a piecewise linear function of a direction vector d (cf. [10, Section 4.1] or [27]). When Ω is non-polyhedral, we consider a finitely generated convex set, i.e., $\Omega = \{w \in \mathbb{R}^m \,|\, G(w) \leq 0\}$, where $G : \mathbb{R}^m \to \mathbb{R}^{p_1}$ is such that each component function G_i is twice continuously differentiable and convex for $i = 1, \ldots, p_1$. It is known that if, for each $w \in \mathbb{R}^m$, the set Ω satisfies either sequentially bounded constraint qualification (SBCQ) or constant rank constraint qualification (CRCQ) at $\Pi_\Omega(w)$, then Π_Ω is directionally differentiable; see [10, Sections 4.4-4.5] for detailed discussions.

More differential properties can be obtained for $\widehat{f}(t, z)$. Motivated by [23, Theorem 8], we consider semismoothness. A function $G : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *semismooth* at $z_* \in \mathbb{R}^n$ [10] if G is B-differentiable at all points in a neighborhood of z_* and satisfies

$$\lim_{z_* \neq z \to z_*} \frac{G'(z; z - z_*) - G'(z_*; z - z_*)}{\|z - z_*\|} = 0.$$

Semismooth functions play an important role in nonsmooth analysis and optimization; see [10] and the references therein for details.

Lemma 4.2. *Assume that $\Pi_\Omega : \mathbb{R}^m \to \mathbb{R}^m$ is directionally differentiable on $\mathbb{R}^m$. For any given $\{(t_i, y_i)\}$, $\{w_i\}$, $\lambda > 0$, and $z \in \mathbb{R}^\ell$, if $u_*(t, \cdot)$ is semismooth at z for each fixed $t \in [0, 1]$, so is $\widehat{f}(t, \cdot)$.*

Proof. Fix $\{(t_i, y_i)\}$, $\{w_i\}$, $\lambda > 0$, and $z \in \mathbb{R}^\ell$. It suffices to prove that $\widehat{x}(t, \cdot)$ is semismooth at z for each fixed $t \in [0, 1]$, where $\widehat{x}(t, z)$ satisfies the ODE: $\dot{x}(t) = Ax(t) + Bu_*(t, z)$, $t \in [0, 1]$ with $x(0) = z$. It follows from the proof of Proposition 4.1 that $\widehat{x}(t, z)$ is B-differential in z on $[0, 1]$ and for a given $d \in \mathbb{R}^\ell$ and $t \in [0, 1]$,

$$\widehat{x}'(t, z; d) = e^{At}d + \int_0^t e^{A(t-s)}Bu'_*(s, z; d)ds.$$

Algorithm 1 Modified Nonsmooth Newton's Method with Line Search

```
Choose scalars $\beta \in (0,1)$ and $\gamma \in (0, \frac{1}{2})$;
Initialize $k = 0$ and choose an initial vector $z^0 \in \mathbb{R}^\ell$ such that $q(t, v_j(z^0))$ is non-degenerate on each $[t_j, t_{j+1}]$;
repeat
  $k \leftarrow k + 1$;
  Find a direction vector $d^k$ such that $H_{y,n}(z^{k-1}) + H'_{y,n}(z^{k-1}; d^k) = 0$;
  Let $m_k$ be the first nonnegative integer $m$ for which $g(z^{k-1}) - g(z^{k-1} + \beta_k^m d^k) \geq -\gamma\beta_k^m g'(z^{k-1}; d^k)$;
  $z^k \leftarrow z^{k-1} + \beta^{m_k} d^k$;
  if $q(t, v_j(z^k))$ is degenerate on some $[t_j, t_{j+1}]$ then
    Choose $d' \in \mathbb{R}^\ell$ with sufficiently small $\|d'\| > 0$ such that $q(t, v_j(z^k + d'))$ is non-degenerate on each $[t_j, t_{j+1}]$;
    $z^k \leftarrow z^k + d'$;
  end if
until $g(z^k)$ is sufficiently small
return $z^k$
```

In view of this, it is easy to verify that for a fixed $t \in [0, 1]$ and any $\widetilde{z} \in \mathbb{R}^\ell$,

$$\widehat{x}'(t, \widetilde{z}; \widetilde{z} - z) - \widehat{x}'(t, z; \widetilde{z} - z) \ = \ \int_0^t e^{A(t-s)} B\Big(u'_*(s, \widetilde{z}; \widetilde{z} - z) - u'_*(s, z; z - \widetilde{z})\Big) ds.$$

By the semismoothness of $u_*(s, \cdot)$ at z, we have for each fixed $s \in [0, t]$,

$$\lim_{z \neq \widetilde{z} \to z} \frac{u'_*(s, \widetilde{z}; \widetilde{z} - z) - u'_*(s, z; \widetilde{z} - z)}{\|\widetilde{z} - z\|} = 0.$$

Furthermore, it is shown in the proof of Proposition 4.1 that $u'_*(s, \widetilde{z}; \widetilde{z} - z)$ and $u'_*(s, z; \widetilde{z} - z)$ are Lebesgue integrable and bounded on $[0, 1]$. Therefore, it follows from Lebesgue Dominated Convergence Theorem [3, Theorem 5.6] that

$$\lim_{z \neq \widetilde{z} \to z} \frac{\widehat{x}'(t, \widetilde{z}; \widetilde{z} - z) - \widehat{x}'(t, z; \widetilde{z} - z)}{\|\widetilde{z} - z\|} = 0.$$

This shows that $\widehat{x}(t, z)$ is semismooth at z for each $t \in [0, 1]$. □

Proposition 4.2. *If Π_Ω is semimsooth at any point in $\mathbb{R}^m$, then for any $z \in \mathbb{R}^\ell$, $\widehat{f}(t, \cdot)$ is semismooth at z for each $t \in [0, 1]$. In particular, this holds true if Ω is polyhedral.*

Proof. Note that semimsoothness implies B-differentiability. Furthermore, $\widehat{f}(t, \cdot)$ is clearly semismooth in z on $[0, t_1]$. Now assume that $\widehat{f}(t, z)$ is semismooth in z for all $t \in [0, t_k]$. By the induction hypothesis and (19), we see that for any fixed $t \in [t_k, t_{k+1}]$, $u_*(t, z)$ is a composition of Π_Ω and a semismooth function of z. It follows from [10, Proposition 7.4.4] that $u_*(t, \cdot)$ is semismooth at z for any $t \in [t_k, t_{k+1}]$. In light of Lemma 4.2, $\widehat{f}(t, \cdot)$ is semismooth at z on $[t_k, t_{k+1}]$ and on $[0, t_{k+1}]$. By the induction principle, $\widehat{f}(t, \cdot)$ is semismooth at z for each $t \in [0, 1]$. Finally, since Ω is polyhedral, Π_Ω is continuous piecewise affine and hence (strongly) semismooth [10, Proposition 7.4.7]. □

It follows from the above results that $H_{y,n}$ is a vector-valued B-differentiable function (under the assumption that $\Pi_\Omega(\cdot)$ is directionally differentiable). To solve the equation $H_{y,n}(z) = 0$, we consider a nonsmooth Newton's method with line search in [20]; its (unique) solution is the optimal initial state x_0^* that completely determines the smoothing spline $\widehat{f}(t, x_0^*)$. It should be mentioned that if $\widehat{f}$ is semismooth, other nonsmooth Newton's methods may be applied [10]. However, these methods require computing limiting Jacobian, which is usually numerically expensive. Hence, we consider the nonsmooth Newton's method method in [20], which only requires computing directional derivatives. This method is suitably modified to deal with dynamics induced issues in the paper.

To describe the nonsmooth Newton's algorithm, we introduce the merit function $g : \mathbb{R}^\ell \to \mathbb{R}_+$:

$$g(z) := \frac{1}{2} H_{y,n}^T(z) H_{y,n}(z).$$

Thus g is B-differentiable and $g'(z; d) = H_{y,n}^T(z) H'_{y,n}(z; d)$. The numerical procedure of the modified nonsmooth Newton's method is described in Algorithm 1. Detailed discussions of this algorithm and convergence analysis are given in Section 5.

5 Convergence Analysis of the Nonsmooth Newton's Method

In this section, we study the global convergence of the proposed modified nonsmooth Newton's method. We present some technical preliminaries first. Recall that the recession cone of a closed convex set $\mathcal{K}$ in $\mathbb{R}^n$ is defined by $\mathcal{K}^\infty := \{d \in \mathbb{R}^n \,|\, x + \mu d \in \mathcal{K}, \forall\, \mu \geq 0\}$ for some $x \in \mathcal{K}$. It is known (cf. [1]) that in a finite dimensional space such as $\mathbb{R}^n$, $\mathcal{K}^\infty$ is equivalent to the asymptotic cone of $\mathcal{K}$, where the latter is defined by

$$\left\{ d \in \mathbb{R}^n \,|\, \text{ there exist } 0 < \mu_k \to \infty, x_k \in \mathcal{K} \text{ such that } \lim_{k\to\infty} \frac{x_k}{\mu_k} = d \right\}.$$

Furthermore, $\mathcal{K}^\infty$ is a closed convex cone, and $\mathcal{K}$ is bounded if and only if $\mathcal{K}^\infty = \{0\}$. More equivalent definitions and properties of recession cones can be found in [1, Proposition 2.1.5]. We give a lemma pertaining to the Euclidean projection onto a recession cone as follows.

Lemma 5.1. *Let Ω be a closed convex set in $\mathbb{R}^m$, let (v_k) be a sequence in $\mathbb{R}^m$, and let (μ_k) be a positive real sequence such that $\lim_{k\to\infty} \mu_k = \infty$ and $\lim_{k\to\infty} \frac{v_k}{\mu_k} = d$ for some $d \in \mathbb{R}^m$. Then*

$$\lim_{k\to\infty} \frac{\Pi_\Omega(v_k)}{\mu_k} = \Pi_{\Omega^\infty}(d),$$

where Ω^∞ is the recession cone of Ω.

Proof. It follows from a similar argument as in [10, Lemma 6.3.13] that

$$\lim_{\mu\to\infty} \frac{\Pi_\Omega(\mu d)}{\mu} = \Pi_{\Omega^\infty}(d).$$

Therefore, it suffices to show $\lim_{k\to\infty} \frac{\Pi_\Omega(v_k)}{\mu_k} = \lim_{k\to\infty} \frac{\Pi_\Omega(\mu_k d)}{\mu_k}$. Without loss of generality, we let the vector norm $\|\cdot\|$ be the Euclidean norm. By virtue of the non-expansive property of the Euclidean projector with respect to the Euclidean norm, we have

$$\frac{\big\|\Pi_\Omega(v_k) - \Pi_\Omega(\mu_k d)\big\|}{\mu_k} \leq \frac{\big\|v_k - \mu_k d\big\|}{\mu_k} = \left\| \frac{v_k}{\mu_k} - d \right\| \longrightarrow 0, \quad \text{ as } k \to \infty.$$

This shows the equivalence of the two limits, and hence completes the proof. □

With the help of this lemma, we establish a boundedness result for level sets defined by $H_{y,n}$. For a given $z^* \in \mathbb{R}^\ell$, define the level set $\mathcal{S}_{z^*} := \{z \in \mathbb{R}^\ell \,:\, \|H_{y,n}(z)\| \leq \|H_{y,n}(z^*)\|\}$.

Proposition 5.1. *Let Ω be a closed convex set in $\mathbb{R}^m$. Given any $\{(t_i, y_i)\}$, $\{w_i\}$, $\lambda > 0$, and $z^* \in \mathbb{R}^\ell$, the level set $\mathcal{S}_{z^*}$ is bounded.*

Proof. We prove the boundedness of $\mathcal{S}_{z^*}$ by contradiction. Suppose not. Then there exists a sequence (z_k) in $\mathcal{S}_{z^*}$ such that $\|z_k\| \to \infty$ as $k \to \infty$. Without loss of generality, we may assume that $(z_k/\|z_k\|)$ converges to $v^* \in \mathbb{R}^\ell$ with $\|v^*\| = 1$ by taking a suitable subsequence of (z_k) if necessary. Define the functions $\widetilde{f} : [0,1] \times \mathbb{R}^\ell \to \mathbb{R}^p$ and $\widetilde{u}_* : [0,1] \times \mathbb{R}^\ell \to \mathbb{R}^m$ as:

$$\widetilde{f}(t,z) \ := \ Ce^{At}z + \int_0^t Ce^{A(t-s)} B \widetilde{u}_*(s,z)\, ds,$$

and

$$\widetilde{u}_*(t,z) \ := \ \begin{cases} \Pi_{\Omega^\infty}(0), & \forall\, t \in [0, t_1) \\ \Pi_{\Omega^\infty}\Big(\lambda^{-1} \displaystyle\sum_{i=1}^{k} w_i \big(Ce^{A(t_i - t)}B\big)^T \widetilde{f}(t_i, z)\Big), & \forall\, t \in [t_k, t_{k+1}), \ k = 1, \ldots, n-1 \end{cases}$$

where Ω^∞ is the recession cone of Ω. Note that $\widetilde{f}$ can be treated as the shape restricted smoothing spline obtained from the linear control system $\Sigma(A, B, C, \Omega^\infty)$ for the given $\widetilde{y} := (\widetilde{y}_i)_{i=1}^n = 0$, namely, the control constraint set Ω is replaced by its recession cone Ω^∞ and y by the zero vector.

We claim that for each fixed $t \in [0, 1]$,

$$\lim_{k\to\infty} \frac{\widehat{f}(t, z_k)}{\|z_k\|} \ = \ \widetilde{f}(t, v^*).$$

We prove this claim by induction on the intervals $[t_j, t_{j+1}]$ for $j = 0, 1, \ldots, n-1$.

Consider the interval $[0, t_1]$ first. Recall that $u_*(t, z_k) = \Pi_\Omega(0), \forall\, t \in [0, t_1)$ such that $\widehat{f}(t, z_k) = Ce^{At}z_k + \int_0^t Ce^{A(t-s)}B\Pi_\Omega(0)ds$ for all $t \in [0, t_1]$. Hence, in view of $\Pi_{\Omega^\infty}(0) = 0$ such that $\widetilde{u}_*(t, v^*) = 0$ and $\widetilde{f}(t, v^*) = Ce^{At}v^*$ for all $t \in [0, t_1]$, we have, for each fixed $t \in [0, t_1]$,

$$\lim_{k\to\infty} \frac{\widehat{f}(t, z_k)}{\|z_k\|} \ = \ \lim_{k\to\infty} Ce^{At} \frac{z_k}{\|z_k\|} \ = \ Ce^{At}v^* \ = \ \widetilde{f}(t, v^*).$$

Now suppose the claim holds true for all $t \in [0, t_j]$ with $j \in \{1, \ldots, n-2\}$, and consider $[t_j, t_{j+1}]$. Note that for each $t \in [t_j, t_{j+1})$,

$$u_*(t,z) \ = \ \Pi_\Omega\left(\lambda^{-1} \sum_{i=1}^{j} w_i \big(Ce^{A(t_i-t)}B\big)^T \big(\widehat{f}(t_i, z) - y_i\big)\right).$$

By the induction hypothesis and the boundedness of $Ce^{A(t_i-t)}B$ on $[t_j, t_{j+1}]$ for all $i = 1, \ldots, j$, we have, for each fixed $t \in [t_j, t_{j+1})$,

$$\lim_{k\to\infty} \frac{\lambda^{-1} \sum_{i=1}^{j} w_i \big(Ce^{A(t_i-t)}B\big)^T \big(\widehat{f}(t_i, z) - y_i\big)}{\|z_k\|} \ = \ \lambda^{-1} \sum_{i=1}^{j} w_i \big(Ce^{A(t_i-t)}B\big)^T \widetilde{f}(t_i, v^*).$$

By Lemma 5.1, we further have, for each fixed $t \in [t_s, t_{s+1})$ with $s \in \{1, \ldots, j\}$,

$$\begin{aligned} \lim_{k\to\infty} \frac{u_*(t, z_k)}{\|z_k\|} \ &= \ \lim_{k\to\infty} \frac{\Pi_\Omega\Big(\lambda^{-1} \sum_{i=1}^{s} w_i \big(Ce^{A(t_i-t)}B\big)^T \big(\widehat{f}(t_i, z) - y_i\big)\Big)}{\|z_k\|} \\ &= \ \Pi_{\Omega^\infty}\left(\lambda^{-1} \sum_{i=1}^{s} w_i \big(Ce^{A(t_i-t)}B\big)^T \widetilde{f}(t_i, v^*)\right) \\ &= \ \widetilde{u}_*(t, v^*). \end{aligned}$$

Clearly, $\widetilde{u}_*(\cdot, v^*)$ is Lebesgue integrable and uniformly bounded on $[t_j, t_{j+1}]$. Therefore, for each fixed $t \in [t_j, t_{j+1}]$,

$$\begin{aligned}
\lim_{k\to\infty} \frac{\widehat{f}(t, z_k)}{\|z_k\|} &= \lim_{k\to\infty} \frac{Ce^{At} z_k + \displaystyle\int_0^t Ce^{A(t-s)} B u_*(s, z_k)\, ds}{\|z_k\|} \\
&= \lim_{k\to\infty} \frac{Ce^{At} z_k}{\|z_k\|} + \int_0^t Ce^{A(t-s)} B \left(\lim_{k\to\infty} \frac{u_*(s, z_k)}{\|z_k\|} \right) ds \\
&= Ce^{At} v^* + \int_0^t Ce^{A(t-s)} B \widetilde{u}_*(s, v^*)\, ds \\
&= \widetilde{f}(t, v^*),
\end{aligned}$$

where the second equality follows from the Lebesgue Dominated Convergence Theorem [3, Theorem 5.6]. This establishes the claim by the induction principle.

In light of the claim and the definition of $H_{y,n}$ in (18), we hence have

$$\lim_{k\to\infty} \frac{H_{y,n}(z_k)}{\|z_k\|} \;=\; \sum_{i=1}^{n} w_i \big(Ce^{At_i}\big)^T \widetilde{f}(t_i, v^*) \;=\; \widetilde{H}_{\widetilde{y},n}(v^*)\big|_{\widetilde{y}=0},$$

where $\widetilde{H}_{\widetilde{y},n} : \mathbb{R}^\ell \to \mathbb{R}^\ell$ with $\widetilde{y} = (\widetilde{y}_i)_{i=1}^n$ is defined by

$$\widetilde{H}_{\widetilde{y},n}(z) \;:=\; \sum_{i=1}^{n} w_i \big(Ce^{At_i}\big)^T \Big(\widetilde{f}(t_i, z) - \widetilde{y}_i \Big).$$

Since the smoothing spline $\widetilde{f}$ is obtained from the linear control system $\Sigma(A, B, C, \Omega^\infty)$, and the recession cone Ω^∞ contains the zero vector, it is easy to verify that when $\widetilde{y} = 0$, the optimal solution pair $(\widetilde{u}_*, \widetilde{x}_0^*)$ for $\widetilde{f}(t, \widetilde{x}_0^*)$ is $\widetilde{x}_0^* = 0$ and $\widetilde{u}_*(t, \widetilde{x}_0^*) \equiv 0$ on $[0, 1]$ (such that $\widetilde{f}(t, \widetilde{x}_0^*) \equiv 0$ on $[0, 1]$). Based on Lemma 4.1, we deduce that the equation $\widetilde{H}_{0,n}(z) = 0$ has a unique solution $z = 0$. Since $v^* \neq 0$, we must have $\widetilde{H}_{0,n}(v^*) \neq 0$. Consequently,

$$\lim_{k\to\infty} \frac{\|H_{y,n}(z_k)\|}{\|z_k\|} \;=\; \big\|\widetilde{H}_{0,n}(v^*)\big\| > 0.$$

This shows that $\|H_{y,n}(z)\|$ is unbounded on $\mathcal{S}_{z^*}$, which yields a contradiction. □

Next we show that under a certain order condition on w_i and λ, a directional vector d can always be found for the equation $H_{y,n}(z) + H'_{y,n}(z; d) = 0$ for any $z \in \mathbb{R}^\ell$. This result, along with the boundedness of $\mathcal{S}_{z^*}$ shown in Proposition 5.1, paves the way for the global convergence of Algorithm 1 [20, Theorem 4]. In particular, we shall focus on the case where the control constraint set Ω is polyhedral for the following reasons: (i) the class of polyhedral Ω is already very broad and includes a number of important applications; (ii) since a closed convex set is the intersection of all closed half-spaces containing it, it can be approximated by a polyhedron with good precision; (iii) when Ω is polyhedral, Π_Ω is globally B-differentiable while this is not the case for a non-polyhedral Ω, unless certain constrained qualification is imposed *globally*. Furthermore, for a non-polyhedral Ω, the directional derivatives of Π_Ω are much more difficult to characterize and compute.

Let $\Omega = \{w \in \mathbb{R}^m \,|\, Dw \geq b\}$ be a polyhedron with $D \in \mathbb{R}^{r\times m}$ and $b \in \mathbb{R}^r$. Proposition 3.1 shows that $\Pi_\Omega : \mathbb{R}^m \to \mathbb{R}^m$ is a (Lipschitz) continuous and piecewise affine (PA) function. It follows from (19) that for $t \in [t_k, t_{k+1})$ with $k = 1, 2, \ldots, n-1$, $Bu_*(t, z) \;=\; B\Pi_\Omega\big(B^T e^{-A^T t} v_k(z)\big)$, where

$$v_k(z) \;:=\; \lambda^{-1} \sum_{i=1}^{k} w_i \big(Ce^{At_i}\big)^T \Big(\widehat{f}(t_i, z) - y_i \Big) \in \mathbb{R}^\ell. \tag{20}$$

Define the function $F : \mathbb{R}^\ell \to \mathbb{R}^\ell$ as $F := B \circ \Pi_\Omega \circ B^T$, which is also Lipschitz continuous and piecewise affine. It follows from the theory of piecewise smooth functions (e.g., [27]) that such a function admits an appealing geometric structure for its domain, which provides an alternative representation of the function. Specifically, let Ξ be a finite family of polyhedra $\{\mathcal{X}_i\}_{i=1}^{m_*}$, where each $\mathcal{X}_i := \{\, v \in \mathbb{R}^\ell \mid G_i v \ge h_i \,\}$ for a matrix G_i and a vector h_i. We call Ξ a *polyhedral subdivision* of $\mathbb{R}^\ell$ [10, 27] if

(a) the union of all polyhedra in Ξ is equal to $\mathbb{R}^\ell$, i.e., $\bigcup_{i=1}^{m_*} \mathcal{X}_i = \mathbb{R}^\ell$,

(b) each polyhedron in Ξ has a nonempty interior (thus is of dimension ℓ), and

(c) the intersection of any two polyhedra in Ξ is either empty or a common proper face of both polyhedra, i.e., $\mathcal{X}_i \cap \mathcal{X}_j \neq \emptyset \implies \big[\, \mathcal{X}_i \cap \mathcal{X}_j = \mathcal{X}_i \cap \{v \,|\, (G_i v - h_i)_\alpha = 0\} = \mathcal{X}_j \cap \{v \,|\, (G_j v - h_j)_\beta = 0\}$ for nonempty index sets α and β with $\mathcal{X}_i \cap \{v \,|\, (G_i v - h_i)_\alpha = 0\} \neq \mathcal{X}_i$ and $\mathcal{X}_j \cap \{v \,|\, (G_j v - h_j)_\beta = 0\} \neq \mathcal{X}_j \,\big]$.

For a Lipschitz PA function F, one can always find a polyhedral subdivision of $\mathbb{R}^\ell$ and finitely many affine functions $g_i(v) \equiv E_i v + l_i$ such that F coincides with one of the g_i's on each polyhedron in Ξ [10, Proposition 4.2.1] or [27]. Therefore, an alternative representation of F is given by

$$F(v) \,=\, E_i v + l_i, \quad \forall\, v \in \mathcal{X}_i, \qquad i = 1, \ldots, m_*,$$

and $v \in \mathcal{X}_i \cap \mathcal{X}_j \Longrightarrow E_i v + l_i = E_j v + l_j$.

Given $v \in \mathbb{R}^\ell$, define the index set $\mathcal{I}(v) := \{i \,|\, v \in \mathcal{X}_i\}$. Moreover, given a direction vector $\widetilde{d} \in \mathbb{R}^\ell$, there exists $j \in \mathcal{I}(v)$ (dependent on $\widetilde{d}$) such that $F'(v; \widetilde{d}) = E_j \widetilde{d}$. (A more precise characterization of the directional derivative of the Euclidean projection is defined by the critical cone [10, Theorem 4.1.1], which shows that for a fixed v, $F'(v; \widetilde{d})$ is continuous and piecewise linear (PL) in $\widetilde{d}$.) In view of this and (19), we have that, for each fixed $t \in [t_k, t_{k+1})$ with $k = 1, \ldots, n-1$, there exists $j \in \mathcal{I}(e^{-A^T t} v_k(z))$ (dependent on d) such that

$$B u_*'(t, z; d) \,=\, E_j e^{-A^T t} v_k'(z; d), \qquad \text{where} \quad v_k'(z; d) = \lambda^{-1} \sum_{i=1}^{k} w_i \big(C e^{A t_i}\big)^T \widehat{f}'(t_i, z; d). \tag{21}$$

For each fixed t, the matrix E_j not only depends on z, which is usually known, but also depends on the direction vector d that is unknown *a priori* in a numerical algorithm. This leads to great complexity and difficulty in solving the equation $H_{y,n}(z) + H_{y,n}'(z; d) = 0$ for a given z, where d is the unknown. In what follows, we identify an important case where $e^{-A^T t} v_k(z)$ is in the interior of some polyhedron $\mathcal{X}_j$ such that the matrix E_j relies on z (and t) only but is independent of d.

For notational convenience, define

$$q(t, v) \,:=\, e^{-A^T t} v, \quad v \in \mathbb{R}^\ell,$$

which satisfies the linear ODE: $\dot{q}(t, v) = -A^T q(t, v)$. For a polyhedron $\mathcal{X}_i = \{v \,|\, G_i v \ge h_i\}$ in Ξ, define $\mathcal{Y}_i := \{v \in \mathbb{R}^\ell \,|\, (G_i v - h_i, G_i(-A^T)v, G_i(-A^T)^2 v, \ldots, G_i(-A^T)^\ell v) \succcurlyeq 0\}$, where $\succcurlyeq$ denotes the lexicographical nonnegative order. For a given $v \in \mathbb{R}^\ell$, let the index set $\mathcal{J}(v) := \{i \,|\, v \in \mathcal{Y}_i\}$. Clearly, $\mathcal{J}(v) \subseteq \mathcal{I}(v)$ for any v. Furthermore, given t_*, there exist $\varepsilon > 0$ and $\mathcal{X}_i$ such that $q(t, v) \in \mathcal{X}_i$ for all $t \in [t_*, t_* + \varepsilon]$ if and only if $q(t_*, v) \in \mathcal{Y}_i$. We introduce more notions below.

Definition 5.1. Let $q(t, v)$ and a time t_* be given. If $\mathcal{J}(q(t_*, v)) \neq \mathcal{I}(q(t_*, v))$, then we call t_* *a critical time* along $q(t, v)$ and its corresponding state $q(t_*, v)$ a *critical state*. Furthermore, if there exist $\varepsilon > 0$ and a polyhedron $\mathcal{X}_i$ in Ξ such that $q(t, v) \in \mathcal{X}_i, \forall\, t \in [t_* - \varepsilon, t_* + \varepsilon]$, then we call t_* a *non-switching-time* along $q(t, v)$; otherwise, we call t_* a *switching time* along $q(t, v)$.

It is known that a switching time must be a critical time but not vice versa [28]. Furthermore, a critical state must be on the boundary of a polyhedron in Ξ. The following result, which is a direct consequence of [28, Proposition 7], presents an extension of the so-called *non-Zenoness* of piecewise affine or linear systems (e.g., [4, 21, 29, 30]).

Proposition 5.2. *Consider $q(t,v)$ and a compact time interval $[t_*, t_*+T]$ where $T>0$. Then there are finitely many critical times on $[t_*, t_*+T]$ along $q(t,v)$. Particularly, there exists a partition $t_* = \widehat{t}_0 < \widehat{t}_1 < \cdots < \widehat{t}_{M-1} < \widehat{t}_M = t_* + T$ such that for each $i = 0, 1, \ldots, M-1$, $\mathcal{I}(q(t,v)) = \mathcal{J}(q(t,v)) = \mathcal{J}(q(t',v)), \forall\, t \in (\widehat{t}_i, \widehat{t}_{i+1})$ for any $t' \in (\widehat{t}_i, \widehat{t}_{i+1})$.*

It follows from the above proposition that for any given v, there are finitely many critical times on the compact time interval $[t_k, t_{k+1}]$ along $q(t,v)$, where $k \geq 1$. We call $q(t,v)$ *non-degenerate* on $[t_k, t_{k+1}]$ if, for any two consecutive critical times $\widehat{t}_j$ and $\widehat{t}_{j+1}$ on $[t_k, t_{k+1}]$, there exists an index i_* (dependent on $(\widehat{t}_j, \widehat{t}_{j+1})$) such that $\mathcal{I}(q(t,v)) = \{i_*\}$ for all $t \in (\widehat{t}_j, \widehat{t}_{j+1})$. In other words, $q(t,v)$ is non-degenerate if it is in the interior of some polyhedron of Ξ on the entire $(\widehat{t}_j, \widehat{t}_{j+1})$.

We introduce more notation and assumptions. First, it is clear that there exist constants $\rho_1 > 0$ and $\rho_2 > 0$ such that $\|Ce^{A(t-s)}\|_\infty \leq \rho_1$ for all $t, s \in [0,1]$ and $\max_{i \in \{1,\ldots,m_*\}} \|E_i\|_\infty \leq \rho_2$. Moreover, we assume that

H.**2** there exist constants $\rho_t > 0, \mu \geq \nu > 0$ such that for all n,

$$\max_{0 \leq i \leq n-1} |t_{i+1} - t_i| \leq \frac{\rho_t}{n}, \qquad \frac{\nu}{n} \leq w_i \leq \frac{\mu}{n}, \quad \forall\, i.$$

Theorem 5.1. *Let Ω be a polyhedron in $\mathbb{R}^m$. Assume that $\mathbf{H.1}-\mathbf{H.2}$ hold and $\lambda \geq \mu^2 \rho_1^2 \rho_2 \rho_t/(4\nu)$. Given $z \in \mathbb{R}^\ell$, let $v_k(z)$ be defined as in (20). Suppose that $q(t, v_k(z)) = e^{-A^T t} v_k(z)$ is non-degenerate on $[t_k, t_{k+1}]$ for each $k = 1, 2\ldots, n-1$. Then there exists a unique direction vector $d \in \mathbb{R}^\ell$ satisfying $H_{y,n}(z) + H'_{y,n}(z; d) = 0$.*

Proof. It follows from the non-degeneracy of $q(t, v_k(z))$ and Proposition 5.2 that, for the given z and each $[t_k, t_{k+1}]$ with $k = 1, \ldots, n-1$, there exists a partition $t_k = \widehat{t}_{k,0} < \widehat{t}_{k,1} < \cdots < \widehat{t}_{k,M_k-1} < \widehat{t}_{k,M_k} = t_{k+1}$ such that for each $j = 0, \ldots, M_k - 1$, $q(t, v_k(z))$ is in the interior of some polyhedron of Ξ for all $t \in (\widehat{t}_{k,j}, \widehat{t}_{k,j+1})$. It is easy to show via the continuity of $\widehat{f}(t,z)$ in z that for each open interval $(\widehat{t}_{k,j}, \widehat{t}_{k,j+1})$, there exists a matrix $E_{k,j} \in \{E_1, \ldots, E_q\}$ such that for all $t \in (\widehat{t}_{k,j}, \widehat{t}_{k,j+1})$, $Bu'_*(t,z;d) = E_{k,j} e^{-A^T t} v'_k(z;d)$. Letting $\widetilde{w}_i := w_i/\lambda, i = 1, \ldots, n$ and by virtue of (21), we have, for $r \geq k+1$,

$$\begin{aligned}
\int_{t_k}^{t_{k+1}} Ce^{A(t_r - s)} Bu'_*(s,z;d)ds &= \int_{t_k}^{t_{k+1}} Ce^{A(t_r-s)} \left(\sum_{j=0}^{M_k-1} E_{k,j} \cdot \mathbf{I}_{[\widehat{t}_{k,j}, \widehat{t}_{k,j+1}]} \right) e^{-A^T s} v'_k(z;d) ds \\
&= \int_{t_k}^{t_{k+1}} Ce^{A(t_r-s)} \left(\sum_{j=0}^{M_k-1} E_{k,j} \cdot \mathbf{I}_{[\widehat{t}_{k,j}, \widehat{t}_{k,j+1}]} \right) e^{-A^T s} \sum_{i=1}^{k} \widetilde{w}_i \big(Ce^{At_i}\big)^T \widehat{f}'(t_i, z; d) ds \\
&= \sum_{i=1}^{k} \widetilde{w}_i \left\{ \int_{t_k}^{t_{k+1}} Ce^{A(t_r-s)} \left(\sum_{j=0}^{M_k-1} E_{k,j} \cdot \mathbf{I}_{[\widehat{t}_{k,j}, \widehat{t}_{k,j+1}]} \right) \big(Ce^{A(t_i-s)}\big)^T ds \right\} \widehat{f}'(t_i, z; d) \\
&= \sum_{i=1}^{k} \widetilde{w}_i \, V_{(r,k,i),z} \, \widehat{f}'(t_i, z; d),
\end{aligned}$$

where, for each $i = 1, \ldots, k$,

$$V_{(r,k,i),z} := \int_{t_k}^{t_{k+1}} Ce^{A(t_r-s)} \left(\sum_{j=0}^{M_k-1} E_{k,j} \cdot \mathbf{I}_{[\widehat{t}_{k,j}, \widehat{t}_{k,j+1}]} \right) \big(Ce^{A(t_i-s)}\big)^T ds \in \mathbb{R}^{p \times p}.$$

Note that for a fixed triple (r, k, i), $V_{(r,k,i),z}$ depends on z only and $r > k \geq i \geq 1$. For $r > i \geq 1$, define $W_{(r,i),z} := \widetilde{w}_i \sum_{j=i}^{r-1} V_{(r,j,i),z}$. Therefore, for each $k = 1, \ldots, n-1$,

$$\begin{aligned} \widehat{f}'(t_{k+1}, z; d) &= Ce^{At_{k+1}}d + \sum_{j=1}^{k} \int_{t_j}^{t_{j+1}} Ce^{A(t_{k+1}-s)} B u'_*(s, z; d) ds \\ &= Ce^{At_{k+1}}d + \sum_{j=1}^{k} \sum_{i=1}^{j} \widetilde{w}_i \, V_{(k+1,j,i),z} \, \widehat{f}'(t_i, z; d) = Ce^{At_{k+1}}d + \sum_{i=1}^{k} W_{(k+1,i),z} \, \widehat{f}'(t_i, z; d). \end{aligned}$$

In what follows, we drop z in the subscript of W for notational simplicity. In view of $\widehat{f}'(t_1, z; d) = Ce^{At_1}d$, it can be shown via induction that for each $k = 2, \ldots, n$,

$$\begin{aligned} \widehat{f}'(t_k, z; d) &= Ce^{At_k}d + W_{(k,k-1)}Ce^{At_{k-1}}d + \Big(W_{(k,k-2)} + W_{(k,k-1)}W_{(k-1,k-2)}\Big)Ce^{At_{k-2}}d \\ &\quad + \quad \cdots \quad \cdots \quad + \quad \cdots \quad \cdots \\ &\quad + \Bigg(W_{(k,1)} + W_{(k,s)}W_{(s,1)} + \sum_{s_1=3}^{k-1} \sum_{s_2=2}^{s_1-1} W_{(k,s_1)}W_{(s_1,s_2)}W_{(s_2,1)} + \cdots \\ &\qquad \cdots \quad \cdots \cdots + W_{(k,k-1)}W_{(k-1,k-2)} \cdots W_{(3,2)}W_{(2,1)}\Bigg)Ce^{At_1}d \\ &= \sum_{j=1}^{k} \widetilde{W}_{(k,j)} \, Ce^{At_j} d, \end{aligned}$$

where the matrices $\widetilde{W}_{(k,j)}$ of order p are defined in terms of $W_{(k,s)}$ as shown above.

For a given $r \in \{1, \ldots, n\}$, define

$$\mathbf{C}_r := \begin{pmatrix} Ce^{At_1} \\ Ce^{At_2} \\ \vdots \\ Ce^{At_r} \end{pmatrix} \in \mathbb{R}^{rp \times \ell}, \ \mathbf{W}_r := \operatorname{diag}(w_1 I_p, \ldots, w_r I_p) \begin{bmatrix} I_p & & & & \\ \widetilde{W}_{(2,1)} & I_p & & & \\ \widetilde{W}_{(3,1)} & \widetilde{W}_{(3,2)} & I_p & & \\ \vdots & \vdots & \ddots & \ddots & \\ \widetilde{W}_{(r,1)} & \widetilde{W}_{(r,2)} & \cdots & \widetilde{W}_{(r,r-1)} & I_p \end{bmatrix} \in \mathbb{R}^{rp \times rp}, \tag{22}$$

where I_p is the identity matrix of order p and $\mathbf{W}_r$ depends on z but is independent of d. Hence, the directional derivative of $\sum_{i=1}^{r} w_i (Ce^{At_i})^T \big(\widehat{f}(t_i, z) - y_i\big)$ along the direction vector d is given by

$$\sum_{i=1}^{r} w_i \big(Ce^{At_i}\big)^T \widehat{f}'(t_i, z; d) = \mathbf{C}_r^T \, \mathbf{W}_r \, \mathbf{C}_r \, d.$$

Clearly, $\mathbf{W}_r$ is invertible for any r, and it can be easily verified via the property of $\widetilde{W}_{(k,j)}$ that

$$\mathbf{W}_r^{-1} = \begin{bmatrix} I_p & & & & \\ -W_{(2,1)} & I_p & & & \\ -W_{(3,1)} & -W_{(3,2)} & I_p & & \\ \vdots & \vdots & \ddots & \ddots & \\ -W_{(r,1)} & -W_{(r,2)} & \cdots & -W_{(r,r-1)} & I_p \end{bmatrix} \operatorname{diag}(w_1^{-1} I_p, \ldots, w_r^{-1} I_p).$$

Moreover, define the symmetric matrix

$$\mathbf{V}_r \,:=\, \frac{1}{2}\Big(\mathbf{W}_r^{-1} + \big(\mathbf{W}_r^{-1}\big)^T\Big) \,=\, \begin{bmatrix} \frac{I_p}{w_1} & -\frac{W_{(2,1)}}{2w_1} & -\frac{W_{(3,1)}}{2w_1} & \cdots & -\frac{W_{(r,1)}}{2w_1} \\ -\frac{W_{(2,1)}}{2w_1} & \frac{I_p}{w_2} & -\frac{W_{(3,2)}}{2w_2} & \cdots & -\frac{W_{(r,2)}}{2w_2} \\ -\frac{W_{(3,1)}}{2w_1} & -\frac{W_{(3,2)}}{2w_2} & \frac{I_p}{w_3} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\frac{W_{(r,r-1)}}{2w_{r-1}} \\ -\frac{W_{(r,1)}}{2w_1} & -\frac{W_{(r,2)}}{2w_2} & \cdots & -\frac{W_{(r,r-1)}}{2w_{r-1}} & \frac{I_p}{w_r} \end{bmatrix}.$$

It follows from assumption **H.2** that $\max_i \widetilde{w}_i \le \mu/(\lambda n)$ and that for any $1 \le j < k$,

$$\|W_{(k,j)}\|_\infty \,\le\, \frac{\mu}{\lambda\, n}\int_{t_j}^{t_k} \rho_1^2 \rho_2 d\tau \,\le\, \frac{\mu \rho_1^2 \rho_2}{\lambda\, n} \cdot \frac{\rho_t (k-j)}{n}.$$

Furthermore, we deduce from **H.2** that $\max_{i,j} \frac{w_i}{w_j} \le \mu/\nu$. Therefore, for any fixed $k = 1, \dots, n$,

$$\begin{aligned} w_k \left(\sum_{i=1}^{k-1} \Big\| \frac{W_{(k,i)}}{2w_i} \Big\|_\infty + \sum_{i=k+1}^{n} \Big\| \frac{W_{(i,k)}}{2w_k} \Big\|_\infty \right) \;&\le\; \frac{\mu^2 \rho_1^2 \rho_2 \rho_t}{2\lambda\nu\, n^2} \left(\sum_{i=1}^{k-1}(k-i) + \sum_{i=k+1}^{n}(i-k) \right) \\ &\le\; \frac{\mu^2 \rho_1^2 \rho_2 \rho_t}{2\lambda\nu\, n^2} \sum_{i=1}^{n-1} i \;\le\; \frac{\mu^2 \rho_1^2 \rho_2 \rho_t \,(n-1)}{4\lambda\nu\, n} \;<\; 1, \end{aligned}$$

where the last inequality follows from the assumption on λ. This implies that the symmetric matrix $\mathbf{V}_r$ is strictly diagonally dominant for any r. Hence, for each r, $\mathbf{V}_r$ is positive definite, so are $\mathbf{W}_r^{-1}$ and $\mathbf{W}_r$ (although not symmetric).

Finally, note that $H'_{y,n}(z; d) \,=\, \mathbf{C}_n^T \,\mathbf{W}_n \,\mathbf{C}_n \, d$, and $\mathbf{C}_n$ has full column rank, in light of the assumption **H.1**. Consequently, $\mathbf{C}_n^T \,\mathbf{W}_n \,\mathbf{C}_n$ is positive definite such that the unique direction vector $d = -\big(\mathbf{C}_n^T \,\mathbf{W}_n \,\mathbf{C}_n\big)^{-1} H_{y,n}(z)$ solves the equation $H_{y,n}(z) + H'_{y,n}(z; d) = 0$. □

The above result relies on the critical non-degenerate property of $q(t, v_k(z))$. In what follows, we consider the case where $q(t, v_k(z))$ is degenerate on some sub-interval of $[t_k, t_{k+1}]$. Geometrically, this implies that the trajectory of $q(t, v_k(z))$ travels on a face of a polyhedron in Ξ for some time. It shall be shown that under mild assumptions, a suitable small perturbation of z will lead to a non-degenerate trajectory. Recall that each polyhedron $\mathcal{X}_i$ in the polyhedral subdivision Ξ is defined by the matrix $G_i \in \mathbb{R}^{m_i}$ and the vector $h_i \in \mathbb{R}^{m_i}$. Since each $\mathcal{X}_i$ has non-empty interior, we assume, without loss of generality, that for each $j = 1, \dots, m_i$, the set $\{v \in \mathcal{X}_i \,|\, (G_i v - h_i)_j = 0\}$ represents a (unique) facet of $\mathcal{X}_i$ (i.e., a $(\ell - 1)$-dimensional face of $\mathcal{X}_i$) [27, Proposition 2.1.3], where $(G_i)_{j\bullet}$ denotes the jth row of G_i and satisfies $\|(G_i)_{j\bullet}^T\|_2 = 1$.

Proposition 5.3. *Let Ω be a polyhedron in $\mathbb{R}^m$. For a given $z \in \mathbb{R}^\ell$, suppose that $q(t, v_k(z))$ is degenerate on the interval $[t_k, t_{k+1}]$ for some $k \in \{1, \dots, n-1\}$, where $v_k(z)$ is defined in (20). Assume that (C, A) is an observable pair,* **H.1** *−* **H.2** *hold, and $\lambda \ge \mu^2 \rho_1^2 \rho_2 \rho_t/(4\nu)$. Then for any $\varepsilon > 0$, there exists $d \in \mathbb{R}^\ell$ with $0 < \|d\| \le \varepsilon$ such that $q(t, v_k(z + d))$ is non-degenerate on $[t_k, t_{k+1}]$ for each $k = 1, \dots, n-1$.*

Proof. Fix $\varepsilon > 0$. Define the set of vector-scalar pairs that represent all the facets of the polyhedra in Ξ:

$$\mathcal{S} \,:=\, \Big\{ \big((G_i)_{j\bullet}^T, (h_i)_j \big) \,|\, i = 1, \dots, m_*, \; j = 1, \dots, m_i \Big\}.$$

Note that if $q(t, v_k(z))$ is degenerate on $[t_k, t_{k+1}]$ for some k, then there exist a pair $(g, \alpha) \in \mathcal{S}$ and an open subinterval $\mathcal{T} \subset [t_k, t_{k+1}]$ such that $g^T q(t, v_k(z)) - \alpha = 0$ for all $t \in \mathcal{T}$, which is further equivalent to $g^T q(t, v_k(z)) - \alpha = 0$ for all $t \in [t_k, t_{k+1}]$ in view of $q(t, v_k(z)) = e^{-A^T t} v_k(z)$. In view of this, we define for each $k \in \{1, \ldots, n-1\}$, $\mathcal{S}_{z,k,D} := \{(g, \alpha) \in \mathcal{S} \,|\, g^T q(t, v_k(z)) = \alpha, \, \forall\, t \in [t_k, t_{k+1}]\}$.

Let k_1 be the smallest k such that $\mathcal{S}_{z,k,D}$ is nonempty (or equivalently $q(t, v_k(z))$ is degenerate on $[t_k, t_{k+1}]$). Clearly, $k \geq 1$. Since $q(t, v_k(z))$ is non-degenerate on $[t_k, t_{k+1}]$ for each $k = 1, \ldots, k_1 - 1$, it follows from a similar argument in the proof of Theorem 5.1 that $v'_{k_1}(z; d) = \lambda^{-1} \mathbf{C}^T_{k_1} \mathbf{W}_{k_1,z} \mathbf{C}_{k_1} d$, where we write $\mathbf{W}_{k_1}$ as $\mathbf{W}_{k_1,z}$ to emphasize its dependence on z (but independent of d). Consider the following two cases:

(i) $(g, \alpha) \in \mathcal{S}_{z,k_1,D}$. It follows from the B-differentiability of $v_{k_1}(\cdot)$ that $q(t, v_{k_1}(z + d)) = q(t, v_{k_1}(z)) + q\big(t, v'_{k_1}(z; d) + o(\|d\|)\big)$ for each $t \in [t_{k_1}, t_{k_1+1}]$. Therefore, using the fact that $\|g\|_2 = 1$, we have for each $t \in [t_{k_1}, t_{k_1+1}]$,

$$g^T q(t, v_{k_1}(z+d)) - \alpha = \big(g^T q(t, v_{k_1}(z)) - \alpha\big) + g^T q\big(t, v'_{k_1}(z; d) + o(\|d\|)\big) = g^T q\big(t, v'_{k_1}(z; d)\big) + o(\|d\|).$$

Furthermore, under **H.2** and the assumption on λ, it is shown in Theorem 5.1 that $\mathbf{W}_{k_1,z}$ is positive definite. Let the observability matrix

$$V_g := \begin{pmatrix} g^T \\ g^T A^T \\ \vdots \\ g^T (A^T)^{\ell-1} \end{pmatrix} \in \mathbb{R}^{\ell \times \ell}.$$

Since $g^T q(t, v'_{k_1}(z; d)) = g^T e^{-A^T t} \mathbf{C}^T_{k_1} \mathbf{W}_{k_1,z} \mathbf{C}_{k_2} d$, we see that $g^T q(t, v'_{k_1}(z; d))$ is nonvanishing on $[t_{k_1}, t_{k_1+1}]$ if and only if $d \notin \text{Ker}(V_g \mathbf{C}^T_{k_1} \mathbf{W}_{k_1,z} \mathbf{C}_{k_1})$. Since g is nonzero, $\mathbf{W}_{k_1,z}$ is positive definite, and (C, A) is an observable pair, it is easy to show that $V_g \mathbf{C}^T_{k_1} \mathbf{W}_{k_1,z} \mathbf{C}_{k_1} \neq 0$ such that $\text{Ker}(V_g \mathbf{C}^T_{k_1} \mathbf{W}_{k_1,z} \mathbf{C}_{k_1})$ is a proper subspace of $\mathbb{R}^\ell$. Hence there exists a scalar $\tau > 0$ such that for any $d \notin \text{Ker}(V_g \mathbf{C}^T_{k_1} \mathbf{W}_{k_1,z} \mathbf{C}_{k_1})$ with $0 < \|d\| \leq \tau$, $g^T q(t, v'_{k_1}(z; d))$, and thus $g^T q(t, v_{k_1}(z + d)) - \alpha$, is nonvanishing on $[t_{k_1}, t_{k_1+1}]$, which further implies that $g^T q(t, v_{k_1}(z + d)) - \alpha$ has at most finitely many zeros on $[t_{k_1}, t_{k_1+1}]$.

(ii) $(g, \alpha) \in \mathcal{S} \setminus \mathcal{S}_{z,k_1,D}$. This means that there exists $t_* \in [t_{k_1}, t_{k_1+1}]$ such that $g^T q(t_*, v_{k_1}(z + d)) - \alpha \neq 0$. Due to the continuity of $v_{k_1}(z)$, we see that there exists $\tau > 0$ such that if $\|d\| \leq \tau$, then $g^T q(t_*, v_{k_1}(z + d)) - \alpha \neq 0$, which also implies that $g^T q(t, v_{k_1}(z + d)) - \alpha$ has at most finitely many zeros on $[t_{k_1}, t_{k_1+1}]$. Similarly, we see that for each $k = 1, \ldots, k_1 - 1$ and each $(g, \alpha) \in \mathcal{S}$, $g^T q(t, v_k(z + d)) - \alpha$ has at most finitely many zeros on $[t_{k_1}, t_{k_1+1}]$.

By virtue of the finiteness of $\mathcal{S}$ and the above results, we obtain a finite union of proper subspaces of $\mathbb{R}^\ell$ denoted by S and a constant $\eta > 0$ such that for each $(g, \alpha) \in \mathcal{S}$ and any $d \notin S$ with $0 < \|d\| \leq \eta$, $g^T q(t, v_k(z+d)) - \alpha$ has at most finitely many zeros on $[t_k, t_{k+1}]$ for each $k = 1, \ldots, k_1$. Since $g^T q(t, v_k(z + d)) - \alpha \neq 0$ for all but finitely many times in $[t_k, t_{k+1}]$ with $k = 1, \ldots, k_1$ for all $(g, \alpha) \in \mathcal{S}$, we conclude that except finitely many times in $[t_k, t_{k+1}]$, $q(t, v_k(z + d))$ must be in the interior of some polyhedron in Ξ at each $t \in [t_k, t_{k+1}]$, where $k = 1, \ldots, k_1$. This shows that $q(t, v_k(z + d))$ is non-degenerate on $[t_k, t_{k+1}]$ for each $k = 1, \ldots, k_1$. In particular, we can choose a nonzero vector d^1 with $\|d^1\| \leq \varepsilon/n$ satisfying this condition.

Now define $\widetilde{z}^1 := z + d^1$, and let k_2 be the smallest k such that $\mathcal{S}_{\widetilde{z}^1,k,D}$ is nonempty. Clearly, $k_2 \geq k_1 + 1$. By replacing z by $\widetilde{z}^1$ in the preceding proof, we deduce via a similar argument that there exists a nonzero vector d^2 with $\|d^2\| \leq \min(\varepsilon/n, \|d^1\|/4)$ such that $q(t, v_k(\widetilde{z}^1 + d^2))$ is non-degenerate on $[t_k, t_{k+1}]$ for each $k = 1, \ldots, k_2$. Continuing this process and using induction, we obtain at most $(n-1)$ nonzero vectors d^j with $\|d^j\| \leq \min(\varepsilon/n, \|d^1\|/2^j)$ for $j \geq 2$ and $d^* := \sum_j d^j$ such that $q(t, v_k(z+d^*))$ is non-degenerate on $[t_k, t_{k+1}]$ for each $k = 1, \ldots, n-1$. Obviously $\|d^*\| \leq \varepsilon$. Furthermore, by virue of $\| \sum_{j \geq 2} d^j \| \leq \|d^1\|/2$, we conclude that $d^* \neq 0$. □

Finally, we establish the global convergence of Algorithm 1 under suitable assumptions.

Theorem 5.2. *Let Ω be a polyhedron in $\mathbb{R}^m$. If (C, A) is an observable pair, the assumptions in Theorem 5.1 hold, and $\liminf_k \beta^{m_k} > 0$, then the sequence (z^k) generated by Algorithm 1 has an accumulation point that is a solution to the equation $H_{y,n}(z) = 0$.*

Proof. Let (z^k) be a sequence generated by Algorithm 1 from an initial vector $z^0 \in \mathbb{R}^\ell$; the existence of (z^k) is due to Theorem 5.1 and Proposition 5.3. Without loss of generality, we assume that $H_{y,n}(z^k) \neq 0$ for each k. Letting d' be the perturbation vector in the algorithm in case of degeneracy, we have $g(z^{k-1}) - g(z^k - d') \geq \sigma\beta^{m_k}\|H_{y,n}(z^k - d')\|_2^2$. Since $\|d'\|$ can be arbitrarily small and $H_{y,n}$ and g are continuous, it follows from a similar argument as in the proof of [20, Theorem 4] that $g(z^{k-1}) - g(z^k) \geq \sigma\beta^{m_k}\Big(\|H_{y,n}(z^k)\|_2^2 + o(\|H_{y,n}(z^k)\|_2^2)\Big)$. Hence, $\big(g(z^k)\big)$ is a nonnegative and strictly decreasing sequence. This also shows, in view of Proposition 5.1, that the sequence (z^k) is bounded and thus has an accumulation point. Furthermore, $\big(g(z^k)\big)$ converges and $\lim_{k\to\infty}(\beta^{m_k}\|H_{y,n}(z^k)\|_2^2 + \varepsilon_k) = 0$, where each $|\varepsilon_k|$ is arbitrarily small by choosing small $\|d'\|$. (For example, $|\varepsilon_k|$ can be of order $o(\|H_{y,n}(z^k)\|_2^2)$ by choosing a suitable d'.) Hence, if $\liminf_k \beta^{m_k} > 0$, then an accumulation point of (z^k) is a desired solution to the B-differentiable equation $H_{y,n}(z) = 0$. □

6 Numerical Examples

In this section, three nontrivial numerical examples are given to demonstrate performance of the shape restricted smoothing spline and the proposed nonsmooth Newton's method. In each example, the underlying true function $f : [0, 1] \to \mathbb{R}$ is defined by $A \in \mathbb{R}^{2\times 2}$, $B = (0, 1)^T$, $C = (1, 0)$, a true initial state x_0, and a true control function $u \in L_2([0, 1], \mathbb{R})$ with the control constraint set Ω. The sample data (y_i) is generated by $y_i = f(t_i) + \varepsilon_i$, where (ε_i) is iid zero mean random error with variance σ^2. The weights w_i are chosen as $w_i = \frac{1}{n}$ for each $i = 1, \ldots, n$ in all cases. Furthermore, different choices of possibly non-equally spaced design points (t_i) are considered in order to illustrate flexibility of the proposed algorithm.

In what follows, the true underlying function f, the corresponding matrix A and true control function u, the choice of design points t_i, the true initial state x_0, initial guess of the initial condition z^0 in the algorithm, the variance σ, and the penalty parameter λ are given for each example. For notational convenience, we use the indicator function $\mathbf{I}_S$ for a set S below.

Example 6.1. The convex constraint with non-equally spaced designed points:
$f(t) = \big(\frac{4}{3}t^3 - t + 1\big)\cdot \mathbf{I}_{[0,\frac{1}{2})} + \big(-\frac{8}{3}t^3 + 6t^2 - 4t + \frac{3}{2}\big)\cdot \mathbf{I}_{[\frac{1}{2},\frac{3}{4})} + \big(\frac{1}{2}t + \frac{3}{8}\big)\cdot \mathbf{I}_{[\frac{3}{4},1]}$,
$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\quad x_0 = (1, -1)^T, \quad u(t) = 8t\cdot \mathbf{I}_{[0,\frac{1}{2})} + (12 - 16t)\cdot \mathbf{I}_{[\frac{1}{2},\frac{3}{4})}$,
$\Omega = [0, \infty), \quad z^0 = (2, 3)^T, \quad \sigma = 0.1, \quad \lambda = 10^{-4}$, and the design points

$$(t_i)_{i=0}^n = \Big\{0, \frac{1}{2n}, \ldots, \frac{1}{20}, \frac{1}{20} + \frac{4}{3n}, \ldots, \frac{9}{20}, \frac{9}{20} + \frac{1}{2n}, \ldots, \frac{11}{20}, \frac{11}{20} + \frac{1}{2n}, \ldots, \frac{19}{20}, \frac{19}{20} + \frac{1}{2n}, \ldots, 1\Big\}.$$

Example 6.2. The unbounded control constraint with non-equally spaced designed points:

$$f(t) = \begin{cases} 11.610t(e^{-t} + e^{-2t}) - 27.219e^{-t} + 25.219e^{-2t} + 2 & \text{if } t \in [0, \frac{1}{4}) \\ -6.234e^{-t} + 3.257e^{-2t} + 3 & \text{if } t \in [\frac{1}{4}, \frac{1}{2}) \\ -11.610t(e^{-t} + e^{-2t}) + 18.222e^{-t} - 21.692e^{-2t} + 3 & \text{if } t \in [\frac{1}{2}, \frac{3}{4}) \\ -3.345e^{-t} + 1.306e^{-2t} + 2 & \text{if } t \in [\frac{3}{4}, 1] \end{cases}$$

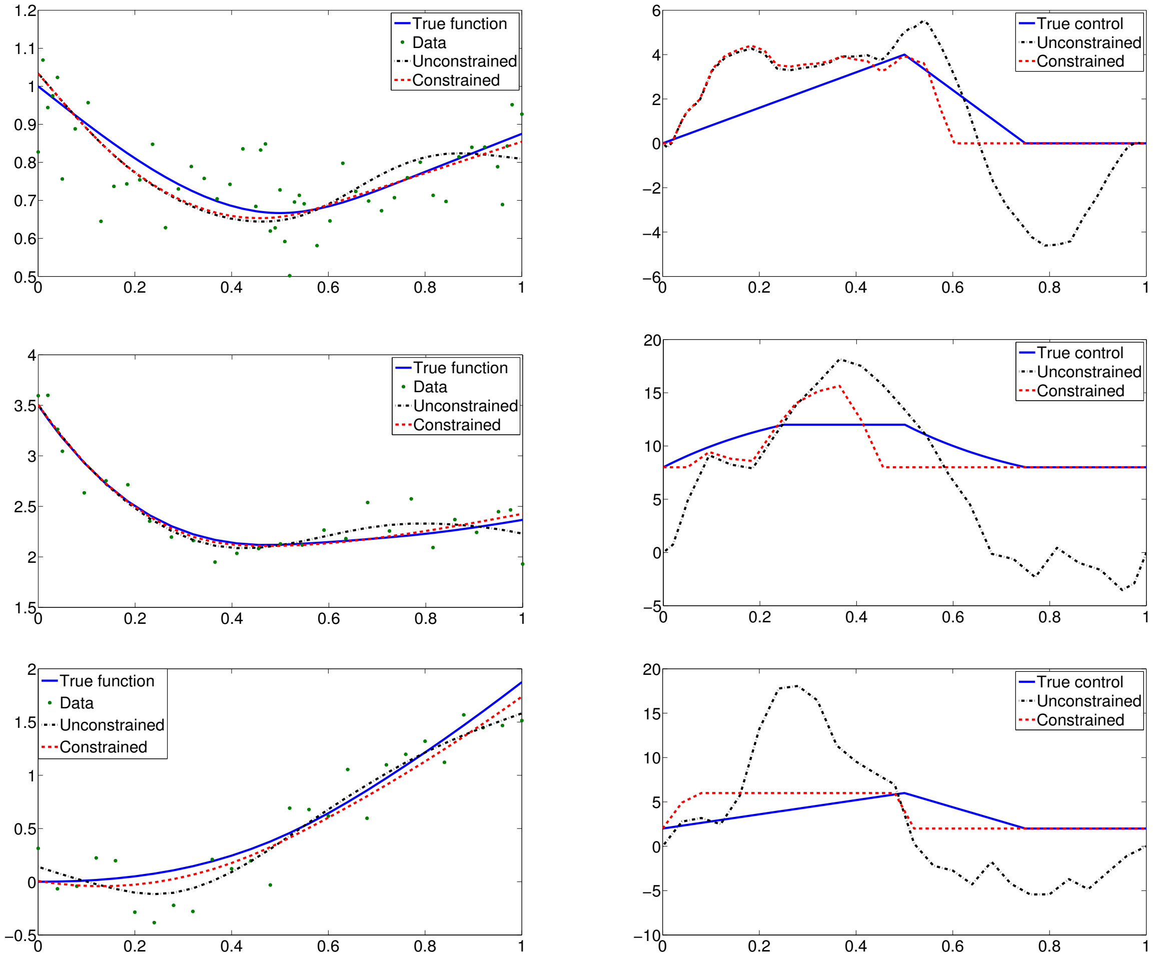

Figure 1: Left column: spline performance of Examples 6.1 (top), 6.2 (middle), and 6.3 (bottom); right column: the corresponding control performance of Examples 6.1–6.3.

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}, \quad x_0 = (7/2, -7)^T, \quad u(t) = \begin{cases} 23.219(e^{-t} - e^{-2t}) + 8 & \text{if } t \in [0, \frac{1}{4}) \\ 12 & \text{if } t \in [\frac{1}{4}, \frac{1}{2}) \\ -38.282e^{-t} + 63.117e^{-2t} + 6 & \text{if } t \in [\frac{1}{2}, \frac{3}{4}) \\ 8 & \text{if } t \in [\frac{3}{4}, 1] \end{cases}$$

$\Omega = [8, \infty)$, $z^0 = (0, 1/2)^T$, $\sigma = 0.2$, $\lambda = 10^{-4}$, and the design points

$$(t_i)_{i=0}^n = \Big\{0, \frac{1}{2n}, \frac{2}{2n}, \dots, \frac{1}{20}, \frac{1}{20} + \frac{9}{8n}, \dots, \frac{19}{20}, \frac{19}{20} + \frac{1}{2n}, \dots, 1\Big\}.$$

Example 6.3. The bounded control constraint with equally spaced designed points:
$f(t) = \big(\frac{4}{3}t^3 + t^2\big) \cdot \mathbf{I}_{[0,\frac{1}{2})} + \big(-\frac{8}{3}t^3 + 7t^2 - 3t + \frac{1}{2}\big) \cdot \mathbf{I}_{[\frac{1}{2},\frac{3}{4})} + \big(t^2 + \frac{3}{2}t - \frac{5}{8}\big) \cdot \mathbf{I}_{[\frac{3}{4},1]}$,
$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $x_0 = (0, 0)^T$, $u(t) = (8t + 2) \cdot \mathbf{I}_{[0,\frac{1}{2})} + (14 - 16t) \cdot \mathbf{I}_{[\frac{1}{2},\frac{3}{4})} + 2 \cdot \mathbf{I}_{[\frac{3}{4},1]}$,
$\Omega = [2, 6]$, $z^0 = (2, 3)^T$, $\sigma = 0.3$, $\lambda = 10^{-4}$, and the equally spaced design points $t_i = \frac{i}{n}$.

Table 1: Performance of Shape Restricted (constr.) Splines vs. Unconstrained Splines

		$\|f-\widehat{f}\|_{L_2}$		$\|f-\widehat{f}\|_{L_\infty}$		$\|x_0-\widehat{x}_0\|_2$	
Example	sample size	constr.	unconstr.	constr.	unconstr.	constr.	unconstr.
Ex. 6.1	$n = 25$	0.00696	0.00723	0.06809	0.07216	0.25985	0.30825
	$n = 50$	0.00351	0.00362	0.04971	0.05218	0.19141	0.22549
	$n = 100$	0.00177	0.00180	0.03487	0.03588	0.14021	0.15958
Ex. 6.2	$n = 25$	0.01302	0.01492	0.12639	0.15609	0.76778	1.45583
	$n = 50$	0.00704	0.00791	0.09998	0.12474	0.70899	1.41832
	$n = 100$	0.00387	0.00436	0.08048	0.10519	0.75410	1.54277
Ex. 6.3	$n = 25$	0.01728	0.02138	0.16761	0.22974	0.44519	0.97093
	$n = 50$	0.00912	0.01074	0.13525	0.16891	0.36184	0.67901
	$n = 100$	0.00463	0.00531	0.09601	0.12063	0.31549	0.61803

The proposed nonsmooth Newton's algorithm is used to compute the shape restricted smoothing splines for the three examples. In all cases, we choose $\beta = 0.25$ and $\gamma = 0.1$ in Algorithm 1 with the terminating tolerance as 10^{-6}. The numerical results for Example 6.1 with $n = 50$, Example 6.2 with $n = 25$, and Example 6.3 with $n = 25$ are displayed in Figure 1. For comparison, the unconstrained smoothing splines are also shown in Figure 1. The number of iterations for numerical convergence of the proposed nonsmooth Newton's algorithm ranges from a single digit to 160 with the median between 9 and 34 (depending on system parameters, sample data and size, and initial state guesses). It is observed that the proposed nonsmooth Newton's algorithm converges superlinearly overall.

To further compare the performance of shape restricted smoothing splines and unconstrained smoothing splines, simulations were run 200 times, and the average performance over these simulations was recorded in each case. Three performance metrics are considered, namely, the L_2-norm, the L_∞-norm, and the 2-norm of the difference between the true and computed initial conditions. Table 1 summarizes the spline performance of the two splines for different sample sizes, where $\widehat{f}$ denotes the computed smoothing splines and $\widehat{x}_0$ denotes the computed initial condition. It is seen in all of the above examples that the shape restricted smoothing spline usually outperforms its unconstrained counterpart. It should be pointed out that the performance of shape restricted smoothing splines critically depends on the penalty parameter λ, the weights w_i, the control constraint set Ω, and the function class that the true function belongs to. However, detailed discussions of performance issues are beyond the scope of the current paper and will be addressed in future.

7 Conclusion

Shape restricted smoothing splines subject to general linear dynamics and control constraints are studied in this paper. Such a constrained smoothing spline is formulated as a finite-horizon constrained optimal control problem with unknown initial state and control. Optimality conditions are derived using the Hilbert space methods and variational techniques. To compute the shape restricted smoothing spline, the optimality conditions are converted to a nonsmooth B-differentiable equation, and a modified nonsmooth Newton's algorithm with line search is proposed to solve the equation. Detailed convergence analysis of this algorithm is given, and numerical examples show the effectiveness of the proposed algorithm.

References

[1] A. Auslender and M. Teboulle. *Asymptotic Cones and Functions in Optimization and Variational Inequalities*. Springer-Verlag, Heidelberg, 2002.

[2] A.V. Balakrishnan. *Introduction to Optimization in a Hilbert Space*. Lecture Notes in Operations Research and Mathematical Systems, Vol. 42, Springer-Verlag, 1971.

[3] R.G. Bartle. *The Elements of Integration and Lebesgue Measure*. John Wiley & Sons, Inc., 1995.

[4] M.K. Çamlibel, J.S. Pang, and J. Shen. Conewise linear systems: non-Zenoness and observability. *SIAM Journal on Control and Optimization*, Vol. 45(5), pp. 1769–1800, 2006.

[5] R.W. Cottle, J.S. Pang, and R.E. Stone. *The Linear Complementarity Problem*, Academic Press Inc., (Cambridge 1992).

[6] C. De Boor. *A Practical Guide to Splines*. Springer, 2001.

[7] A.L. Dontchev, H.-D. Qi, L. Qi, and H. Yin. A Newton method for shape-preserving spline interpolation. *SIAM Journal on Optimization*, Vol. 13(2), pp. 588–602, 2002.

[8] A.L. Dontchev, H.-D. Qi, and L. Qi. Quadratic convergence of Newton's method for convex interpolation and smoothing. *Constructive Approximation*, Vol. 19, pp. 123–143, 2003.

[9] M. Egerstedt and C. Martin. *Control Theoretic Splines*. Princeton University Press, 2010.

[10] F. Facchinei and J.S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer-Verlag, 2003.

[11] A. Girard. Towards a multiresolution approach to linear control. *IEEE Trans. on Automatic Control*, Vol. 51(8), pp. 1261–1270, 2006.

[12] L. Han, M.K. Camlibel, J.-S. Pang, and W.P.M.H. Heemels. A unified numerical scheme for linear-quadratic optimal control problems with joint control and state constraints. *Optimization Methods and Software*, Vol. 27(4-5), pp. 761–799, 2012.

[13] H. Kano, M. Egerstedt, H. Nakata, and C.F. Martin. B-splines and control theory. *Applied Mathematics and Computation*, Vol. 145, pp. 265–288, 2003.

[14] H. Kano, H. Fujioka, and C.F. Martin. Optimal smoothing spline with constraints on its derivatives. *Proc. of the 49nd IEEE Conf. on Decision and Control*, pp. 6785–6790, 2010.

[15] S. Lang. *Real and Functional Analysis*. Springer-Verlage, 3rd Edition, 1993.

[16] D.G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons Inc., 1969.

[17] C.A. Micchelli and F.I. Utreras. Smoothing and interploation in a convex subset of a Hilbert space. *SIAM Journal on Scientific and Statistical Computing*, Vol. 9(4), pp. 728–746, 1988.

[18] M. Nagahara, C.F. Martin, and Y. Yamamoto. Quadratic programming for monotone control theoretical splines. *Proc. of the SICE 2010 Annual Conference*, pp. 531–534, 2010.

[19] J. Pal and M. Woodroofe. Large sample properties of shape restricted regression estimators with smoothness adjustments. *Statistica Sinica*, Vol. 17, pp. 1601–1616, 2007.

[20] J.S. Pang. Newton's method for B-differentiable equations. *Mathematics of Operations Research*, Vol. 15, pp. 311–341, 1990.

[21] J.S. Pang and J. Shen. Strongly regular differential variational systems. *IEEE Trans. on Automatic Control*, Vol. 52(2), pp. 242–255, 2007.

[22] J.S. Pang and D. Stewart. Differential variational inequalities. *Mathematical Programming Series A*, Vol. 113, pp. 345–424, 2008.

[23] J.S. Pang and D. Stewart. Solution dependence on initial conditions in differential variational inequalities. *Mathematical Programming Series B*, Vol. 116, pp. 429–460, 2009.

[24] M. Renardy and R.C. Rogers. *An Introduction to Partial Differential Equations*. Springer, 2nd Edition, 2004.

[25] T. Robertson, F.T. Wright, and R.L. Dykstra. *Order Restricted Statistical Inference*. John Wiley & Sons Ltd., 1988.

[26] A.K. Sanyal, M. Chellappa, J.L. Valk, J. Shen, J. Ahmed, and D.S. Bernstein. Globally convergent adaptive tracking of spacecraft angular velocity with inertia identification. *Proc. of the 42nd IEEE Conf. on Decision and Control*, pp. 2704–2709, 2003.

[27] S. Scholtes. Introduction to piecewise differentiable equations. Habilitation thesis, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe (1994).

[28] J. Shen. Observability analysis of conewise linear systems via directional derivative and positive invariance techniques. *Automatica*, Vol. 46(5), pp. 843–851, 2010.

[29] J. Shen. Robust non-Zenoness of piecewise analytic systems with applications to complementarity systems. *Proc. of 2010 American Control Conference*, pp. 148–153, Baltimore, MD, 2010.

[30] J. Shen and J.S. Pang. Linear complementarity systems: Zeno states. *SIAM Journal on Control and Optimization*, Vol. 44, pp. 1040–1066, 2005.

[31] J. Shen and J.S. Pang. Linear complementarity systems with singleton properties: non-Zenoness. *Proc. of 2007 American Control Conference*, pp. 2769–2774, New York, 2007.

[32] J. Shen and X. Wang. Estimation of monotone functions via P-splines: A constrained dynamical optimization approach. *SIAM Journal on Control and Optimization*, Vol. 49(2), pp. 646–671, 2011.

[33] J. Shen and X. Wang. Estimation of shape constrained functions in dynamical systems and its application to genetic networks. *Proc. of American Control Conference*, pp. 5948–5953, Baltimore, 2010.

[34] J. Shen and X. Wang. A constrained optimal control approach to smoothing splines. *Proc. of the 50th IEEE Conf. on Decision and Control*, pp. 1729–1734, Orlando, FL, 2011.

[35] J. Shen and X. Wang. Convex regression via penalized splines: a complementarity approach. *Proc. of American Control Conference*, pp. 332–337, Montreal, Canada, 2012.

[36] S. Sun, M.B. Egerstedt, and C.F. Martin. Control theoretic smoothing splines. *IEEE Trans. on Automatic Control*, Vol. 45(12), pp. 2271–2279, 2000.

[37] C. Tantiyaswasdikul and M. Woodroofe. Isotonic smoothing splines under sequential designs. *Journal of Statistical Planning and Inference*, Vol. 38, pp. 75-88, 1994.

[38] G. Wahba. *Spline Models for Observational Data*. Philadelphia: SIAM, 1990.

[39] X. Wang and J. Shen. Uniform convergence and rate adaptive estimation of convex functions via constrained optimization. *SIAM Journal on Control and Optimization*, Vol. 51(4), pp. 2753–2787, 2013.

[40] Y. Zhou, M. Egerstedt, and C.F. Martin. Hilbert space methods for control theoretic splines: a unified treatment. *Communication in Information and Systems*. Vol. 6(1), pp. 55–82, 2006.