2019 IEEE 58th Conference on Decision and Control (CDC)
Palais des Congrès et des Expositions Nice Acropolis
Nice, France, December 11-13, 2019

Consistent approximation of optimal control problems using Bernstein polynomials

Venanzio Cichella, Isaac Kaminer, Claire Walton, Naira Hovakimyan, António M. Pascoal

Abstract— **We present a direct method for the solution of nonlinear optimal control problems based on Bernstein polynomial approximations. We show, using a rigorous setting, that the proposed method yields consistent approximations of time continuous optimal control problems. We demonstrate that the method can also be used for the estimation of optimal control problems costates. This result leads to the formulation of the Covector Mapping Theorem for Bernstein polynomial approximation. Finally, we exploit the numerical and geometric properties of Bernstein polynomials, and illustrate the advantages of the method through numerical examples.**

I. Introduction

Optimal control problems that arise in most engineering applications are in general very complex. Finding a closed-form solution to these problems can be difficult or even impossible, and therefore they must be solved numerically. Direct methods, for example, are based on transcribing optimal control problems into nonlinear programming problems (NLPs) using some discretization scheme [1]–[4], and solving the latter by using ready-to-use NLP solvers (e.g. MATLAB, SNOPT, etc.). Pioneering work on direct methods includes that of Polak on the consistency of approximation theory [5], which provides a theoretical framework to assess the convergence properties of direct methods. Motivated by [5], a wide range of direct methods that use different discretization schemes have been developed, including Euler [5], Runge-Kutta [6], and Pseudospectral (PS) [7] methods.

Pseudospectral methods are the most popular direct methods nowadays, mainly due to their spectral (exponential) rate of convergence and their successful application to solve a wide range of optimization problems, e.g. [7]–[9]. However, as pointed out in [10]–[12], there is one salient disadvantage associated with these methods. When discretizing the state and/or the input, the constraints are enforced at the discretization nodes and satisfaction of constraints cannot be guaranteed in between the nodes. To mitigate this drawback, the order of approximation (number of nodes) can be increased; however, this leads to larger NLPs, which may become computationally expensive and too inefficient to solve. This problem does not limit itself to PS methods, but it is common to methods that are based on discretization. Pseudospectral methods also suffer from a drawback when dealing with non-smooth optimal control problems. This is mainly related to the Gibbs phenomenon [13], common to approximation methods based on orthogonal polynomials. This phenomenon, visible in the form of oscillations, reduces the accuracy of the approximation to first order away from discontinuities and to $\mathcal{O}(1)$ in the neighborhoods of jumps [14]. Several extensions of PS methods have been developed to deal with this disadvantage and lessen the effect of the Gibbs phenomenon (e.g. [15]–[17]). Some methods require the location of the discontinuities to be known a priori, which is often impractical or difficult. Other methods estimate these locations, which could result in inefficiency or ill conditioning of the discretized problem, especially when the number of discontinuities is large and unknown.

We propose a direct method based on Bernstein approximations. The latter have several nice properties. First, the approximants converge uniformly to the functions that they approximate, and so do their derivatives [18]. Moreover, Bernstein polynomials behave well even when the functions being approximated are non-smooth. As demonstrated in [19], the Gibbs phenomenon does not occur when approximating piecewise smooth, monotone functions with left and right derivatives at every point by Bernstein polynomials. Thus, the proposed method based on Bernstein approximation lends itself to problems that have discontinuous states and/or controls, e.g. bang-bang optimal control problems (see also [20]). Finally, due to their geometric properties, Bernstein polynomials yield computationally efficient algorithms for the computation of constraints for the whole time interval where optimization takes place, and not only at discretization points (see [18], [21], [22]). Hence, with the proposed approach the solutions can be guaranteed to be feasible and satisfy the constraints for all times, while retaining the computational efficiency of direct methods.

Bernstein approximations converge slower than other interpolation or approximation techniques. In fact, the approach proposed in the present paper is outperformed by, for example, PS methods in terms of convergence rate. This is not surprising, since the choice of nodes and the interpolating polynomials in the PS methods are dictated by approximation accuracy and convergence speed, while sacrificing satisfaction of constraints in between the nodes. On the other hand, our approach prioritizes constraint satisfaction at the expense

The authors acknowledge support from ONR (N00014-19-1-2106 and N00014-19-WX0-0155), NSF (NRI 1830639), AFOSR (FA9550-15-1-0518), the H2020 EU Marine Robotics RIN (Project ID 731103), and the FCT Project UID/EEA/5009/2013.

Venanzio Cichella is with the Department of Mechanical Engineering, University of Iowa, Iowa City, 52242 IA venanzio-cichella@uiowa.edu. Isaac Kaminer and Claire Walton are with the Department of Mechanical and Aerospace Engineering, Naval Postgraduate School, Monterey, CA 93940 {kaminer,clwalton1}@nps.edu. Naira Hovakimyan is with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 nhovakim@illinois.edu. António M. Pascoal is with the Institute for Systems and Robotics (ISR), Instituto Superior Tecnico (IST), Univ. Lisbon, Portugal. antoniog@isr.ist.utl.pt.

978-1-7281-1398-2/19/$31.00 ©2019 IEEE

of a slower convergence rate.

The paper is structured as follows: Section II presents preliminary mathematical results. Section III introduces the optimal control problem of interest and the proposed discretization method. Section IV demonstrates that the method yields consistent approximations to the optimal solution. Section V derives the Karush–Kuhn–Tucker (KKT) conditions associated with the NLP, and compares these conditions with the first order optimality conditions for the original optimal control problem. Numerical examples are discussed in Section VI. The paper ends with conclusions in Section VII.

II. Mathematical background

The Nth order Bernstein polynomial $\boldsymbol{x}_N : [0,1] \to \mathbb{R}^n$ is given by

$$\boldsymbol{x}_N(t) = \sum_{j=0}^{N} \bar{\boldsymbol{x}}_{j,N} b_{j,N}(t)\,, \qquad t \in [0,1]\,, \tag{1}$$

where $\bar{\boldsymbol{x}}_{0,N}, \ldots, \bar{\boldsymbol{x}}_{N,N}$ are Bernstein coefficients, and $b_{j,N}(t) = \binom{N}{j} t^j (1-t)^{N-j}$, $j \in \{0, \ldots, N\}$ is the Bernstein basis of degree N with $\binom{N}{j} = \frac{N!}{j!(N-j)!}$. The derivative and integral of $\boldsymbol{x}_N(t)$ are computed as

$$\dot{\boldsymbol{x}}_N(t) = N \sum_{j=0}^{N-1} (\bar{\boldsymbol{x}}_{j+1,N} - \bar{\boldsymbol{x}}_{j,N}) b_{j,N-1}(t)\,,$$

$$\int_0^1 \boldsymbol{x}_N(t) dt = w \sum_{j=0}^{N} \bar{\boldsymbol{x}}_{j,N}\,, \qquad w = \frac{1}{N+1}\,. \tag{2}$$

A vector valued function $\boldsymbol{x} : [0,1] \to \mathbb{R}^n$ can be approximated by the Nth order Bernstein polynomial $\boldsymbol{x}_N(t)$ computed as in (1) with $\bar{\boldsymbol{x}}_{j,N} = \boldsymbol{x}(t_j)$ and $t_j = \frac{j}{N}$, i.e.

$$\boldsymbol{x}_N(t) = \sum_{j=0}^{N} \boldsymbol{x}(t_j) b_{j,N}(t)\,, \qquad t_j = \frac{j}{N}\,. \tag{3}$$

Lemma 1 ([23], [24]): Let $\boldsymbol{x}(t) \in \mathcal{C}_n^0$ on $[0,1]$. Then, $\boldsymbol{x}_N(t)$ given by (3) satisfies

$$||\boldsymbol{x}_N(t) - \boldsymbol{x}(t)|| \leq C_0 W_x(N^{-\frac{1}{2}})\,, \quad \forall t \in [0,1],$$

where C_0 is a positive constant satisfying $C_0 < 5n/4$, and $W_x(\cdot)$ is the modulus of continuity of $\boldsymbol{x}(t)$ in $[0,1]$. If $\boldsymbol{x}(t) \in \mathcal{C}_n^1$, then

$$\|\dot{\boldsymbol{x}}_N(t) - \dot{\boldsymbol{x}}(t)\| \leq C_1 W_{x'}(N^{-\frac{1}{2}})\,, \quad \forall t \in [0,1],$$

where C_1 is a positive constant satisfying $C_1 < 9n/4$ and $W_{x'}(\cdot)$ is the modulus of continuity of $\dot{\boldsymbol{x}}(t)$ in $[0,1]$.

Lemma 2 ([25]): Assume $\boldsymbol{x}(t) \in \mathcal{C}_n^{r+2}$, $r \geq 0$. Then, $\boldsymbol{x}_N(t)$ computed as in (3) satisfies

$$||\boldsymbol{x}_N(t) - \boldsymbol{x}(t)|| \leq \frac{C_0}{N}, \;\ldots,\; ||\boldsymbol{x}_N^{(r)}(t) - \boldsymbol{x}^{(r)}(t)|| \leq \frac{C_r}{N}\,,$$

$\forall t \in [0,1]$, where $\boldsymbol{x}^{(r)}(t)$ denotes the rth derivative of $\boldsymbol{x}(t)$, and $C_0, \ldots, C_r$ are positive constants independent of N.

Lemma 3: If $\boldsymbol{x}(t) \in \mathcal{C}_n^0$ on $[0,1]$, then we have

$$\left\| \int_0^1 \boldsymbol{x}(t) dt - w \sum_{j=0}^{N} \boldsymbol{x}\left(\frac{j}{N}\right) \right\| \leq C_I W_x(N^{-\frac{1}{2}})\,,$$

with $w = \frac{1}{N+1}$, where $C_I > 0$ is independent of N. Moreover, if $\boldsymbol{x}(t) \in \mathcal{C}_n^2$, then

$$\left\| \int_0^1 \boldsymbol{x}(t) dt - w \sum_{j=0}^{N} \boldsymbol{x}\left(\frac{j}{N}\right) \right\| \leq \frac{C_I}{N}\,.$$

The Lemma above follows directly from Lemmas 1 and 2 and Equation (2).

III. Problem formulation

Consider the following optimal control problem:

Problem 1 (Problem P): Determine $\boldsymbol{x} : [0,1] \to \mathbb{R}^{n_x}$ and $\boldsymbol{u} : [0,1] \to \mathbb{R}^{n_u}$ that minimize

$$I(\boldsymbol{x}(t), \boldsymbol{u}(t)) = E(\boldsymbol{x}(0), \boldsymbol{x}(1)) + \int_0^1 F(\boldsymbol{x}(t), \boldsymbol{u}(t)) dt\,, \tag{4}$$

$$\text{subject to:} \quad \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t))\,, \quad \forall t \in [0,1], \tag{5}$$

$$\boldsymbol{e}(\boldsymbol{x}(0), \boldsymbol{x}(1)) = \boldsymbol{0}\,, \tag{6}$$

$$\boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{u}(t)) \leq \boldsymbol{0}\,, \quad \forall t \in [0,1]\,, \tag{7}$$

where $E : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}$ and $F : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ are the terminal and running costs, respectively, $\boldsymbol{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$ describes the system dynamics, $\boldsymbol{e} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_e}$ is the vector of boundary conditions, and $\boldsymbol{h} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_h}$ is the vector of state and input constraints.

The following assumptions hold:

Assumption 1: E, F, $\boldsymbol{f}$, $\boldsymbol{e}$, and $\boldsymbol{h}$ are Lipschitz continuous with respect to their arguments.

Assumption 2: Problem P admits optimal solutions $\boldsymbol{x}^*(t)$ and $\boldsymbol{u}^*(t)$ that satisfy $\boldsymbol{x}^*(t) \in \mathcal{C}_{n_x}^1$ and $\boldsymbol{u}^*(t) \in \mathcal{C}_{n_u}^0$.

In what follows we formulate a discretized version of Problem P, hereafter referred to as Problem P_N, where N denotes the *order of approximation*. First, consider the Nth order vectors of Bernstein polynomials

$$\boldsymbol{x}_N(t) = \sum_{j=0}^{N} \bar{\boldsymbol{x}}_{j,N} b_{j,N}(t),\; \boldsymbol{u}_N(t) = \sum_{j=0}^{N} \bar{\boldsymbol{u}}_{j,N} b_{j,N}(t), \tag{8}$$

with $\boldsymbol{x}_N : [0,1] \to \mathbb{R}^{n_x}$, $\boldsymbol{u}_N : [0,1] \to \mathbb{R}^{n_u}$, $\bar{\boldsymbol{x}}_{j,N} \in \mathbb{R}^{n_x}$ and $\bar{\boldsymbol{u}}_{j,N} \in \mathbb{R}^{n_u}$. Let $\bar{\boldsymbol{x}}_N \in \mathbb{R}^{n_x \times (N+1)}$ and $\bar{\boldsymbol{u}}_N \in \mathbb{R}^{n_u \times (N+1)}$ be defined as $\bar{\boldsymbol{x}}_N = [\bar{\boldsymbol{x}}_{0,N}\,, \ldots, \bar{\boldsymbol{x}}_{N,N}]$, $\bar{\boldsymbol{u}}_N = [\bar{\boldsymbol{u}}_{0,N}\,, \ldots, \bar{\boldsymbol{u}}_{N,N}]$. Let $t_j = \frac{j}{N}$, $j \in \{0, \ldots, N\}$. Then, Problem P_N can be stated as follows:

Problem 2 (Problem P_N): Determine $\bar{\boldsymbol{x}}_N$ and $\bar{\boldsymbol{u}}_N$ that minimize

$$I_N(\bar{\boldsymbol{x}}_N, \bar{\boldsymbol{u}}_N) = E(\boldsymbol{x}_N(0), \boldsymbol{x}_N(t_N)) + w \sum_{j=0}^{N} F(\boldsymbol{x}_N(t_j), \boldsymbol{u}_N(t_j))\,, \tag{9}$$

$$\text{subject to} \quad \|\dot{\boldsymbol{x}}_N(t_j) - \boldsymbol{f}(\boldsymbol{x}_N(t_j), \boldsymbol{u}_N(t_j))\| \leq \delta_P^N\,, \tag{10}$$

$$\boldsymbol{e}(\boldsymbol{x}_N(0), \boldsymbol{x}_N(t_N)) = \boldsymbol{0}\,, \tag{11}$$

$$\boldsymbol{h}(\boldsymbol{x}_N(t_j), \boldsymbol{u}_N(t_j)) \leq \delta_P^N \boldsymbol{1}\,, \tag{12}$$

$\forall j = 0, \ldots, N$, where $w = \frac{1}{N+1}$, and δ_P^N is a small positive number that depends on N and satisfies $\lim_{N\to\infty} \delta_P^N = 0$.

Remark 1: The *relaxation bound* δ_P^N is introduced to guarantee that Problem P_N has a feasible solution. It can be made arbitrarily small by choosing a sufficiently large order of approximation N.

IV. Feasibility and consistency of Problem P_N

The outcome of Problem P_N is a set of optimal Bernstein coefficients $\bar{\boldsymbol{x}}_N^*$ and $\bar{\boldsymbol{u}}_N^*$ that determine the vectors of Bernstein polynomials $\boldsymbol{x}_N^*(t)$ and $\boldsymbol{u}_N^*(t)$, i.e.

$$\boldsymbol{x}_N^*(t) = \sum_{j=0}^{N} \bar{\boldsymbol{x}}_{j,N}^* b_{j,N}(t),\ \boldsymbol{u}_N^*(t) = \sum_{j=0}^{N} \bar{\boldsymbol{u}}_{j,N}^* b_{j,N}(t)\,. \quad (13)$$

We now address the following: (i) existence of a feasible solution to Problem P_N, and (ii) convergence of the pair $(\boldsymbol{x}_N^*(t), \boldsymbol{u}_N^*(t))$ to the optimal solution of Problem P, given by $(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$. The main results of this section are summarized in Theorems 1 and 2 below.

Theorem 1 (Feasibility): Let

$$\delta_P^N = C_P \max\{W_{x'}(N^{-\frac{1}{2}}),\, W_x(N^{-\frac{1}{2}}),\, W_u(N^{-\frac{1}{2}})\}\,, \quad (14)$$

where C_P is a positive constant independent of N, and $W_{x'}(\cdot)$, $W_x(\cdot)$ and $W_u(\cdot)$ are the moduli of continuity of $\dot{\boldsymbol{x}}(t)$, $\boldsymbol{x}(t)$ and $\boldsymbol{u}(t)$, respectively. Then, Problem P_N is feasible for any arbitrary order of approximation $N \in \mathbb{Z}^+$.

Proof: Let $\boldsymbol{x}(t)$ and $\boldsymbol{u}(t)$ be a solution for Problem P, which exists by Assumption 2. Define

$$\bar{\boldsymbol{x}}_{k,N} = \boldsymbol{x}(t_k),\ \bar{\boldsymbol{u}}_{k,N} = \boldsymbol{u}(t_k),\ \forall k \in \{0, \ldots, N\}\,. \quad (15)$$

Next we show that the above polynomials satisfy the constraints in (10), (11) and (12), with δ_P^N defined in (14). Lemma 1 and Assumption 2 imply that

$$\begin{aligned} ||\boldsymbol{x}_N(t) - \boldsymbol{x}(t)|| &\le C_x W_x(N^{-\frac{1}{2}})\,, \\ ||\boldsymbol{u}_N(t) - \boldsymbol{u}(t)|| &\le C_u W_u(N^{-\frac{1}{2}})\,, \\ ||\dot{\boldsymbol{x}}_N(t) - \dot{\boldsymbol{x}}(t)|| &\le C_{x'} W_{x'}(N^{-\frac{1}{2}})\,, \end{aligned} \quad (16)$$

for all $t \in [0, 1]$, where $\boldsymbol{x}_N(t)$ and $\boldsymbol{u}_N(t)$ are computed as in Equation (8) with Bernstein coefficients given by Equation (15), $C_x < 5n_x/4$, $C_u < 5n_u/4$, and $C_{x'} < 9n_x/4$. Subtracting $\dot{\boldsymbol{x}}(t_k) - \boldsymbol{f}(\boldsymbol{x}(t_k), \boldsymbol{u}(t_k)) = 0$ from the left hand side of Equation (10) yields

$$\begin{aligned} ||\dot{\boldsymbol{x}}_N(t_k) - \boldsymbol{f}(\boldsymbol{x}_N(t_k), \boldsymbol{u}_N(t_k))|| &\le ||\dot{\boldsymbol{x}}_N(t_k) - \dot{\boldsymbol{x}}(t_k)|| \\ &+ ||\boldsymbol{f}(\boldsymbol{x}_N(t_k), \boldsymbol{u}_N(t_k)) - \boldsymbol{f}(\boldsymbol{x}(t_k), \boldsymbol{u}(t_k))||\,. \end{aligned}$$

Using Equation (16) and the fact that $\boldsymbol{f}$ is Lipschitz (see Assumption 1) with Lipschitz constant L_f, we obtain

$$\begin{aligned} ||\dot{\boldsymbol{x}}_N(t_k) - \boldsymbol{f}(\boldsymbol{x}_N(t_k), \boldsymbol{u}_N(t_k))|| &\le \\ &(C_{x'} + L_f(C_x + C_u))W_{\max}\,, \end{aligned}$$

with $W_{\max} = \max\{W_{x'}(N^{-\frac{1}{2}}), W_x(N^{-\frac{1}{2}}), W_u(N^{-\frac{1}{2}})\}$. Thus, the dynamic constraint in Equation (10) is satisfied with δ_P^N given by Equation (14). The satisfaction of constraints (11) and (12) follows similarly from Assumption 1.

Corollary 1: From the proof of Theorem 1 it follows that if $\boldsymbol{x}^*(t) \in \mathcal{C}^3_{n_x}$ and $\boldsymbol{u}^*(t) \in \mathcal{C}^2_{n_u}$, then Theorem 1 holds with $\delta_P^N = C_P N^{-1}$, where $C_P > 0$ is independent of N.

Theorem 2 (Consistency): Let $\{(\bar{\boldsymbol{x}}_N^*, \bar{\boldsymbol{u}}_N^*)\}_{N=N_1}^{\infty}$ be a sequence of optimal solutions to Problem P_N, and $\{(\boldsymbol{x}_N^*(t), \boldsymbol{u}_N^*(t))\}_{N=N_1}^{\infty}$ a sequence of Bernstein polynomials, given by (13). Assume $\{(\boldsymbol{x}_N^*(t), \boldsymbol{u}_N^*(t))\}_{N=N_1}^{\infty}$ has a uniform accumulation point, i.e.

$$\lim_{N\to\infty} (\boldsymbol{x}_N^*(t), \boldsymbol{u}_N^*(t)) = (\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))\,, \quad \forall t \in [0,1], \quad (17)$$

and assume $\dot{\boldsymbol{x}}^\infty(t)$ and $\boldsymbol{u}^\infty(t)$ are continuous on $[0, 1]$. Then, $(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))$ is an optimal solution for Problem P.

Proof: The proof is divided into three steps: (1) we prove that $(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))$ is a feasible solution to Problem P; (2) we show that $\lim_{N\to\infty} I_N(\bar{\boldsymbol{x}}_N^*, \bar{\boldsymbol{u}}_N^*) = I(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))$; (3) we prove that $(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))$ is an optimal solution of Problem P, i.e. $I(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t)) = I(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$.

Step (1)**.** First, we show by contradiction that $(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))$ satisfies the dynamic constraint of Problem P, $\dot{\boldsymbol{x}}^\infty(t) - \boldsymbol{f}(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t)) = \boldsymbol{0}$. Assume that the above equality does not hold. Then, there exists t', such that

$$||\dot{\boldsymbol{x}}^\infty(t') - \boldsymbol{f}(\boldsymbol{x}^\infty(t'), \boldsymbol{u}^\infty(t'))|| > 0\,. \quad (18)$$

Since the nodes $\{t_k\}_{k=0}^N$, $t_k = \frac{k}{N}$ are dense in $[0, 1]$, there exists a sequence of indices $\{k_N\}_{N=0}^\infty$ such that $\lim_{N\to\infty} t_{k_N} = t'$. Then, from continuity of $\dot{\boldsymbol{x}}^\infty(t)$, $\boldsymbol{x}^\infty(t)$ and $\boldsymbol{u}^\infty(t)$, the left hand side of Equation (18) satisfies

$$\begin{aligned} &||\dot{\boldsymbol{x}}^\infty(t') - \boldsymbol{f}(\boldsymbol{x}^\infty(t'), \boldsymbol{u}^\infty(t'))|| = \\ &\qquad \lim_{N\to\infty} ||\dot{\boldsymbol{x}}_N^*(t_{k_N}) - \boldsymbol{f}(\boldsymbol{x}_N^*(t_{k_N}), \boldsymbol{u}_N^*(t_{k_N}))||. \end{aligned}$$

However, the dynamic constraint in Problem P_N implies that $\lim_{N\to\infty} ||\dot{\boldsymbol{x}}_N^*(t_{k_N}) - \boldsymbol{f}(\boldsymbol{x}_N^*(t_{k_N}), \boldsymbol{u}_N^*(t_{k_N}))|| = 0$, which contradicts Equation (18), thus proving that $(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))$ satisfies the dynamic constraint in Equation (5). The equality and inequality constraints in (11) and (12) follow similarly.

Step (2)**.** We need to show that

$$\lim_{N\to\infty} w \sum_{j=0}^{N} F(\boldsymbol{x}_N^*(t_j), \boldsymbol{u}_N^*(t_j)) = \int_0^1 F(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))dt\,, \quad (19)$$

$$\lim_{N\to\infty} E(\boldsymbol{x}_N^*(0), \boldsymbol{x}_N^*(t_N)) = E(\boldsymbol{x}^\infty(0), \boldsymbol{x}^\infty(1))\,. \quad (20)$$

Using Lemma 3, together with the Lipschitz assumption on F (see Assumption 1) and the continuity of $\boldsymbol{x}^\infty(t)$ and $\boldsymbol{u}^\infty(t)$, we get $\lim_{N\to\infty} w \sum_{j=0}^{N} F(\boldsymbol{x}^\infty(t_j), \boldsymbol{u}^\infty(t_j)) = \int_0^1 F(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t))dt$. Finally, applying the convergence assumption given by Equation (17), the result in Equation (19) follows. Similarly, using the Lipschitz assumption on E, one can show that Equation (20) holds.

Step (3)**.** Let $\tilde{\bar{\boldsymbol{x}}}_{k,N} = \boldsymbol{x}^*(t_k)$, $\tilde{\bar{\boldsymbol{u}}}_{k,N} = \boldsymbol{u}^*(t_k)$, $\forall k \in \{1, \ldots, N\}$ and $\tilde{\bar{\boldsymbol{x}}}_N = [\tilde{\bar{\boldsymbol{x}}}_{0,N}, \ldots, \tilde{\bar{\boldsymbol{x}}}_{N,N}]$, $\tilde{\bar{\boldsymbol{u}}}_N = [\tilde{\bar{\boldsymbol{u}}}_{0,N}, \ldots, \tilde{\bar{\boldsymbol{u}}}_{N,N}]$. One can show that there exists N_1 such that for any $N \ge N_1$ the pair $(\tilde{\bar{\boldsymbol{x}}}_N, \tilde{\bar{\boldsymbol{u}}}_N)$ is a feasible solution of Problem P_N, and that

$$\lim_{N\to\infty} I_N(\tilde{\bar{\boldsymbol{x}}}_N, \tilde{\bar{\boldsymbol{u}}}_N) = I(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))\,. \quad (21)$$

We note that

$$\begin{aligned} I(\boldsymbol{x}^*(t),\boldsymbol{u}^*(t)) &\le I(\boldsymbol{x}^\infty(t),\boldsymbol{u}^\infty(t)) \\ &= \lim_{N\to\infty} I_N(\bar{\boldsymbol{x}}_N^*,\bar{\boldsymbol{u}}_N^*) \le \lim_{N\to\infty} I_N(\tilde{\bar{\boldsymbol{x}}}_N,\tilde{\bar{\boldsymbol{u}}}_N)\,, \end{aligned} \tag{22}$$

which gives $I(\boldsymbol{x}^*(t),\boldsymbol{u}^*(t)) = I(\boldsymbol{x}^\infty(t),\boldsymbol{u}^\infty(t))$.

V. Costate estimation for Problem P

Let $\boldsymbol{\lambda}(t) : [0,1] \to \mathbb{R}^{n_x}$ be the costate trajectory for Problem P, and let $\boldsymbol{\mu}(t) : [0,1] \to \mathbb{R}^{n_h}$ and $\boldsymbol{\nu} \in \mathbb{R}^{n_e}$ be the multipliers. Defining the Lagrangian of the Hamiltonian as

$$\begin{aligned} &\mathcal{L}(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\lambda}(t),\boldsymbol{\mu}(t)) = \\ &\qquad \mathcal{H}(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\lambda}(t)) + \boldsymbol{\mu}^\top(t)\boldsymbol{h}(\boldsymbol{x}(t),\boldsymbol{u}(t))\,, \end{aligned}$$

where $\mathcal{H}(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\lambda}(t)) = F(\boldsymbol{x}(t),\boldsymbol{u}(t)) + \boldsymbol{\lambda}^\top(t)\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t))$, the dual of Problem P can be formulated as follows [26]:

Problem 3 (Problem P_λ): Determine $\boldsymbol{x}(t)$, $\boldsymbol{u}(t)$, $\boldsymbol{\lambda}(t)$, $\boldsymbol{\mu}(t)$ and $\boldsymbol{\nu}$ that satisfy Equations (5), (6), (7) and

$$\boldsymbol{\mu}^\top(t)\boldsymbol{h}(\boldsymbol{x}(t),\boldsymbol{u}(t)) = 0\,, \quad \boldsymbol{\mu}(t) \ge 0\,, \tag{23}$$

$$\begin{aligned} \dot{\boldsymbol{\lambda}}^\top(t) = &-F_x(\boldsymbol{x}(t),\boldsymbol{u}(t)) - \boldsymbol{\lambda}^\top(t)\boldsymbol{f}_x(\boldsymbol{x}(t),\boldsymbol{u}(t)) \\ &- \boldsymbol{\mu}^\top(t)\boldsymbol{h}_x(\boldsymbol{x}(t),\boldsymbol{u}(t))\,, \end{aligned} \tag{24}$$

$$\boldsymbol{\lambda}^\top(0) = -\boldsymbol{\nu}^\top\boldsymbol{e}_{x(0)}(\boldsymbol{x}(0),\boldsymbol{x}(1)) - E_{x(0)}(\boldsymbol{x}(0),\boldsymbol{x}(1))\,, \tag{25}$$

$$\boldsymbol{\lambda}^\top(1) = \boldsymbol{\nu}^\top\boldsymbol{e}_{x(1)}(\boldsymbol{x}(0),\boldsymbol{x}(1)) + E_{x(1)}(\boldsymbol{x}(0),\boldsymbol{x}(1))\,, \tag{26}$$

$$\begin{aligned} &\boldsymbol{\lambda}^\top(t)\boldsymbol{f}_u(\boldsymbol{x}(t),\boldsymbol{u}(t)) + F_u(\boldsymbol{x}(t),\boldsymbol{u}(t)) \\ &\qquad + \boldsymbol{\mu}^\top(t)\boldsymbol{h}_u(\boldsymbol{x}(t),\boldsymbol{u}(t)) = 0\,, \end{aligned} \tag{27}$$

for all $t \in [0,1]$, where subscripts denote partial derivatives.

The following assumptions are imposed on Problem P_λ.

Assumption 3: $E, F, \boldsymbol{f}, \boldsymbol{e}$ and $\boldsymbol{h}$ are continuously differentiable with respect to their arguments, and their gradients are Lipschitz continuous over their domains of definition.

Assumption 4: Solutions $\boldsymbol{x}^*(t)$, $\boldsymbol{u}^*(t)$, $\boldsymbol{\lambda}^*(t)$, $\boldsymbol{\mu}^*(t)$ and $\boldsymbol{\nu}^*$ of Problem P_λ exist and satisfy $\boldsymbol{x}^*(t) \in \mathcal{C}^1_{n_x}$, $\boldsymbol{u}^*(t) \in \mathcal{C}^0_{n_u}$, $\boldsymbol{\lambda}^*(t) \in \mathcal{C}^1_{n_x}$ and $\boldsymbol{\mu}^*(t) \in \mathcal{C}^0_{n_h}$ in $[0,1]$.

Remark 2: Problem P_λ implicitly assumes the absence of pure state constraints in Problem P. If the inequality constraint in Equation (7) is independent of $\boldsymbol{u}(t)$, then the costate $\boldsymbol{\lambda}(t)$ must also satisfy the jump condition $\boldsymbol{\lambda}(t_e^-) = \boldsymbol{\lambda}(t_e^+) + \boldsymbol{h}^\top_{x(t_e)}\boldsymbol{\eta}$, where t_e is the entry or exit time into a constrained arc in which the inequality constraint is active, t_e^- and t_e^+ denote the left-hand side and right-hand side limits of the trajectory, respectively, and $\boldsymbol{\eta}$ is a covector [26].

Define the following Nth order Bernstein polynomials:

$$\boldsymbol{\lambda}_N(t) = \sum_{j=0}^N \bar{\boldsymbol{\lambda}}_{j,N} b_{j,N}(t),\ \boldsymbol{\mu}_N(t) = \sum_{j=0}^N \bar{\boldsymbol{\mu}}_{j,N} b_{j,N}(t), \tag{28}$$

with $\boldsymbol{\lambda}_N : [0,1] \to \mathbb{R}^{n_x}$, $\boldsymbol{\mu}_N : [0,1] \to \mathbb{R}^{n_h}$, $\bar{\boldsymbol{\lambda}}_{j,N} \in \mathbb{R}^{n_x}$ and $\bar{\boldsymbol{\mu}}_{j,N} \in \mathbb{R}^{n_h}$, and the vector $\bar{\boldsymbol{\nu}} \in \mathbb{R}^{n_e}$. Finally, let $\bar{\boldsymbol{\lambda}}_N = [\bar{\boldsymbol{\lambda}}_{0,N}\,,\ \dots\,,\ \bar{\boldsymbol{\lambda}}_{N,N}]$, $\bar{\boldsymbol{\mu}}_N = [\bar{\boldsymbol{\mu}}_{0,N}\,,\ \dots\,,\ \bar{\boldsymbol{\mu}}_{N,N}]$. With the above notation, the Lagrangian for Problem P_N can be written as

$$\begin{aligned} \mathcal{L}_N = &\, E(\boldsymbol{x}_N(0),\boldsymbol{x}_N(t_N)) + w\sum_{j=0}^N F(\boldsymbol{x}_N(t_j),\boldsymbol{u}_N(t_j)) \\ &+ \sum_{j=0}^N \boldsymbol{\lambda}_N^\top(t_j)(-\dot{\boldsymbol{x}}_N(t_j) + \boldsymbol{f}(\boldsymbol{x}_N(t_j),\boldsymbol{u}_N(t_j))) \\ &+ \sum_{j=0}^N \boldsymbol{\mu}_N^\top(t_j)\boldsymbol{h}(\boldsymbol{x}_N(t_j),\boldsymbol{u}_N(t_j)) + \bar{\boldsymbol{\nu}}^\top\boldsymbol{e}(\boldsymbol{x}_N(0),\boldsymbol{x}_N(t_N))\,. \end{aligned}$$

The dual of Problem P_N can now be stated as follows:

Problem 4 (Problem $P_{N\lambda}$): Determine $\bar{\boldsymbol{x}}_N$, $\bar{\boldsymbol{u}}_N$, $\bar{\boldsymbol{\lambda}}_N$, $\bar{\boldsymbol{\mu}}_N$ and $\bar{\boldsymbol{\nu}}$ that satisfy the primal feasibility conditions, namely Equations (10), (11) and (12), the complementary slackness and dual feasibility conditions

$$\begin{aligned} &\left\|\boldsymbol{\mu}_N^\top(t_k)\boldsymbol{h}(\boldsymbol{x}_N(t_k),\boldsymbol{u}_N(t_k))\right\| \le N^{-1}\delta_D^N\,, \\ &\boldsymbol{\mu}_N(t_k) \ge -N^{-1}\delta_D^N\boldsymbol{1}\,, \qquad \forall k = 0,\dots,N\,, \end{aligned} \tag{29}$$

and the stationarity conditions

$$\left\|\frac{\partial \mathcal{L}_N}{\partial \bar{\boldsymbol{x}}_{k,N}}\right\| \le \delta_D^N\,, \quad \left\|\frac{\partial \mathcal{L}_N}{\partial \bar{\boldsymbol{u}}_{k,N}}\right\| \le \delta_D^N\,, \tag{30}$$

$\forall k = 0,\dots,N$, where δ_D^N is a small positive number that depends on N and satisfies $\lim_{N\to\infty}\delta_D^N = 0$.

At this point one might expect results similar to the ones in Section IV, i.e. feasibility (Theorem 1) and consistency (Theorem 2). Nevertheless, similarly to most results on costate estimation [7], [27], [28], this is not the case, and additional conditions must be added to Equations (10)-(12), (29) and (30) in order to obtain consistent approximations of the solutions of Problem P_λ. These conditions, often referred to as *closure conditions* in the literature, are given as follows:

$$\begin{aligned} &\left\|\frac{\boldsymbol{\lambda}_N^\top(0)}{w} + \bar{\boldsymbol{\nu}}^\top\boldsymbol{e}_{x(0)}(\boldsymbol{x}_N(0),\boldsymbol{x}_N(t_N))\right. \\ &\qquad\qquad \left. + E_{x(0)}(\boldsymbol{x}_N(0),\boldsymbol{x}_N(t_N))\right\| \le \delta_D^N \end{aligned} \tag{31}$$

$$\begin{aligned} &\left\|\frac{\boldsymbol{\lambda}_N^\top(t_N)}{w} - \bar{\boldsymbol{\nu}}^\top\boldsymbol{e}_{x(1)}(\boldsymbol{x}_N(0),\boldsymbol{x}_N(t_N))\right. \\ &\qquad\qquad \left. - E_{x(1)}(\boldsymbol{x}_N(0),\boldsymbol{x}_N(t_N))\right\| \le \delta_D^N\,. \end{aligned} \tag{32}$$

In other words, the closure conditions are constraints that must be added to Problem $P_{N\lambda}$ so that the solution of this problem approximates the solution of Problem P_λ. We notice that the conditions given above are discrete approximations of the conditions given by Equations (25) and (26). With this setup, we define the following problem:

Problem 5 (Problem $P^{clos}_{N\lambda}$): Determine $\bar{\boldsymbol{x}}_N$, $\bar{\boldsymbol{u}}_N$, $\bar{\boldsymbol{\lambda}}_N$, $\bar{\boldsymbol{\mu}}_N$ and $\bar{\boldsymbol{\nu}}$ that satisfy the primal feasibility conditions, namely Equations (10), (11) and (12), the complementary slackness and dual feasibility conditions (29), the stationarity conditions (30), and the closure conditions (31) and (32).

The solution of Problem $P^{clos}_{N\lambda}$ presents a set of optimal Bernstein coefficients $\bar{\boldsymbol{x}}_N^*$, $\bar{\boldsymbol{u}}_N^*$, $\bar{\boldsymbol{\lambda}}_N^*$, $\bar{\boldsymbol{\mu}}_N^*$ (which determine the Bernstein polynomials $\boldsymbol{x}_N^*(t)$, $\boldsymbol{u}_N^*(t)$, $\boldsymbol{\lambda}_N^*(t)$ and $\boldsymbol{\mu}_N^*(t)$) and a vector $\bar{\boldsymbol{\nu}}^*$.

In what follows we investigate the ability of the solutions of Problem $P^{clos}_{N\lambda}$ to approximate the solutions of Problem P_λ. The main results of this section are summarized below.

Theorem 3 (Feasibility): Let

$$\delta_D^N = C_D \max\{\delta_P^N\,,\, W_{\lambda'}(N^{-\frac{1}{2}})\,,\, W_{\lambda}(N^{-\frac{1}{2}})\,,\, W_{\mu}(N^{-\frac{1}{2}})\}\,, \tag{33}$$

where C_D is a positive constant independent of N, δ_P^N was defined in Equation (14), and $W_{\lambda'}(\cdot)$, $W_{\lambda}(\cdot)$, and $W_{\mu}(\cdot)$ are the moduli of continuity of $\dot{\boldsymbol{\lambda}}(t)$, $\boldsymbol{\lambda}(t)$ and $\boldsymbol{\mu}(t)$, respectively. Then, Problem $P_{N\lambda}^{clos}$ is feasible for arbitrary order of approximation $N \in \mathbb{Z}^+$.

Corollary 2: If solutions $\boldsymbol{x}^*(t)$, $\boldsymbol{u}^*(t)$, $\boldsymbol{\lambda}^*(t)$, $\boldsymbol{\mu}^*(t)$ and $\boldsymbol{\nu}^*$ of Problem P_λ exist and satisfy $\dot{\boldsymbol{x}}^*(t) \in \mathcal{C}^2_{n_x}$, $\boldsymbol{u}^*(t) \in \mathcal{C}^2_{n_u}$, $\dot{\boldsymbol{\lambda}}^*(t) \in \mathcal{C}^2_{n_x}$, and $\boldsymbol{\mu}^*(t) \in \mathcal{C}^2_{n_h}$ in $[0,1]$, then Theorem 3 holds with $\delta_P^N = C_P N^{-1}$ and $\delta_D^N = C_D N^{-1}$, where C_P and C_D are positive constants independent of N.

Theorem 4 (Consistency): Let $\{(\bar{\boldsymbol{x}}_N^*, \bar{\boldsymbol{u}}_u^*, \bar{\boldsymbol{\lambda}}_N^*, \bar{\boldsymbol{\mu}}_N^*, \bar{\boldsymbol{\nu}}^*)\}_{N=N_1}^{\infty}$ be a sequence of solutions of Problem $P_{N\lambda}^{\text{clos}}$. Consider the sequence of transformed solutions $\{(\bar{\boldsymbol{x}}_N^*, \bar{\boldsymbol{u}}_N^*, \tilde{\bar{\boldsymbol{\lambda}}}_N^*, \tilde{\bar{\boldsymbol{\mu}}}_N^*, \bar{\boldsymbol{\nu}}^*)\}_{N=N_1}^{\infty}$, with $\tilde{\bar{\boldsymbol{\lambda}}}_{j,N}^* = \frac{\bar{\boldsymbol{\lambda}}_{j,N}^*}{w}$, $\tilde{\bar{\boldsymbol{\mu}}}_{j,N}^* = \frac{\bar{\boldsymbol{\lambda}}_{j,N}^*}{w}$, and the corresponding polynomial approximation $\{(\boldsymbol{x}_N^*(t), \boldsymbol{u}_N^*(t), \tilde{\boldsymbol{\lambda}}_N^*(t), \tilde{\boldsymbol{\mu}}_N^*(t), \bar{\boldsymbol{\nu}}^*)\}_{N=N_1}^{\infty}$. Assume that

$$\lim_{N\to\infty} (\boldsymbol{x}_N^*(t), \boldsymbol{u}_N^*(t), \tilde{\boldsymbol{\lambda}}_N^*(t), \tilde{\boldsymbol{\mu}}_N^*(t), \bar{\boldsymbol{\nu}}^*) = (\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t), \tilde{\boldsymbol{\lambda}}^\infty(t), \tilde{\boldsymbol{\mu}}^\infty(t), \bar{\boldsymbol{\nu}}^\infty)\,, \quad \forall t \in [0,1],$$

and assume $\dot{\boldsymbol{x}}^\infty(t)$, $\boldsymbol{u}^\infty(t)$, $\dot{\tilde{\boldsymbol{\lambda}}}^\infty(t)$ and $\tilde{\boldsymbol{\mu}}^\infty(t)$ are continuous on $[0,1]$. Then, $(\boldsymbol{x}^\infty(t), \boldsymbol{u}^\infty(t), \tilde{\boldsymbol{\lambda}}^\infty(t), \tilde{\boldsymbol{\mu}}^\infty(t), \bar{\boldsymbol{\nu}}^\infty)$ is a solution of Problem P_λ.

Proof: The proofs of Theorems 3 and 4 and Corollary 2 can be found in [29].

Theorem 5 (Covector Mapping Theorem): Under the same assumptions of Theorems 3 and 4, when $N \to \infty$, the *covector mapping*

$$\boldsymbol{x}_N^*(t) \mapsto \boldsymbol{x}^*(t)\,, \quad \boldsymbol{u}_N^*(t) \mapsto \boldsymbol{u}^*(t)\,,$$

$$\frac{\boldsymbol{\lambda}_N^*}{w} \mapsto \boldsymbol{\lambda}^*(t)\,, \quad \frac{\boldsymbol{\mu}_N^*(t)}{w} \mapsto \boldsymbol{\mu}^*(t)\,, \quad \bar{\boldsymbol{\nu}}^* \mapsto \boldsymbol{\nu}^*$$

is a bijective mapping between the solution of Problem $P_{N\lambda}^{\text{clos}}$ and the solution of Problem P_λ.

Proof: The result follows directly from Theorems 3 and 4.

VI. Numerical Examples

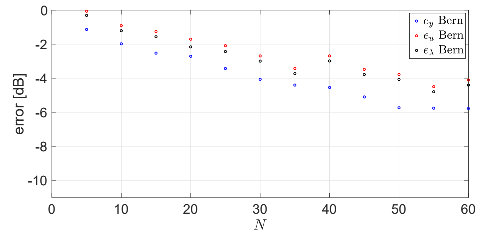
The numerical results presented in this section are obtained using MATLAB's built in *fmincon* function.

Example 1 (see [27]): Determine $y : [0,5] \to \mathbb{R}$ and $u : [0,5] \to \mathbb{R}$ that minimize

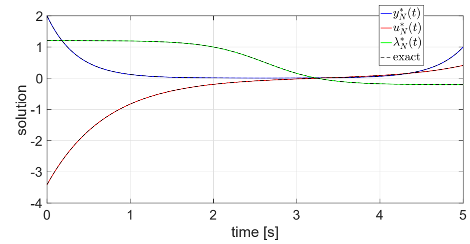
$$I(y(t), u(t)) = \frac{1}{2}\int_0^5 (y(t) + u^2(t))dt\,,$$

$$\text{subject to} \quad \dot{y}(t) = 2y(t) + 2u(t)\sqrt{y(t)}\,, \quad \forall t \in [0,5],$$
$$y(0) = 2\,, \quad y(5) = 1$$

The above example was solved using the Bernstein approximation method for orders $N = 5, 10, \ldots, 55, 60$. Similar to



(a) Error in Bernstein approximation method.



(b) Solutions using the Bernstein approximation method for order $N = 40$.

Fig. 1: Results for Example 1.

[27], we define the following errors

$$e_y = \max_{k=0,\ldots,N} \log_{10} |y_N^*(t_k) - y^*(t_k))|\,,$$
$$e_u = \max_{k=0,\ldots,N} \log_{10} |u_N^*(t_k) - u^*(t_k))|\,,$$
$$e_\lambda = \max_{k=0,\ldots,N} \log_{10} |\lambda_N^*(t_k) - \lambda^*(t_k))|\,,$$

where $y_N^*(t_k)$, $u_N^*(t_k)$ and $\lambda_N^*(t_k)$ are the state, input and costate evaluated at the time nodes $t_k = 5k/N$, $k = 0, \ldots, N$, while $y^*(t_k)$, $u^*(t_k)$ and $\lambda^*(t_k)$ are the exact solutions. Figure 1a illustrates the convergence of the above errors to zero as the order of approximation increases. Figure 1b shows the state, input, and costate obtained using the Bernstein approximation method for order $N = 40$ alongside the exact solution.

Example 2 (bang-bang): Determine $y : [0,2] \to \mathbb{R}$ and $u : [0,2] \to \mathbb{R}$ that minimize

$$I(y(t), u(t)) = \int_0^2 (3u(t) - 2y(t))dt\,, \tag{34}$$

$$\text{subject to} \quad \dot{y}(t) = y(t) + u(t)\,, \quad \forall t \in [0,2], \tag{35}$$
$$y(0) = 4\,, \quad y(2) = 39.392 \tag{36}$$
$$0 \le u(t) \le 2 \quad \forall t \in [0,2]\,. \tag{37}$$

The optimal control for the above example is:

$$u^*(t) = \begin{cases} 2 & 0 \le t \le 1.096 \\ 0 & 1.096 \le t \le 2. \end{cases}$$

Example 2 is solved using the Bernstein approximation method for orders of approximation $N = 10, 15, 30, 55$. To impose the input constraints in Equation (37) we use the convex hull property of Bernstein polynomials [18]. Namely, letting $u_N(t) = \sum_{j=0}^N \bar{u}_{j,N} b_{j,N}(t)$ be the Bernstein polynomial approximating the control input, the saturation constraints are imposed on its Bernstein coefficients, $\bar{u}_{j,N}$,

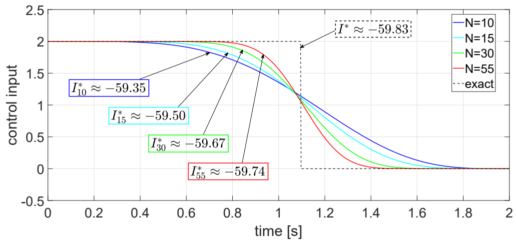


Fig. 2: Solution to Example 2.

$j = 0, \ldots, N$. This ensures satisfaction of the constraints for the whole trajectory $u_N(t)$, $\forall t \geq 0$. The results are illustrated in Figure 2. It can be noted that the solutions and the optimal costs resulting from the Bernstein approximation method converge to the exact ones, and that the approximated control inputs behave nicely despite the discontinuity. There are no jumps in the neighborhood of the discontinuities (the Gibbs phenomenon does not occur [19]), and the exact value of the discontinuity ($t = 1.096$s) is detected with reasonable accuracy even for low orders of approximation. The reader is referred to [14]–[17], [30], where the performance of PS methods when dealing with bang-bang optimal control problems are discussed.

VII. Conclusion

We proposed a numerical method to approximate nonlinear optimal control problems by nonlinear programming (NLPs) using Bernstein polynomials. A rigorous analysis was provided that shows convergence of the solution of the NLP to the solution of the continuous-time problem. Conditions were derived under which the Karush-Kuhn-Tucker multipliers of the NLP converge to the costates of the optimal control problem. This led to the formulation of the Covector Mapping Theorem for Bernstein approximations, enabling numerical computation of the costates. The theoretical findings were validated through numerical examples, and the advantages of the method were discussed.

References

[1] A. V. Rao, "A survey of numerical methods for optimal control," *Advances in the Astronautical Sciences*, vol. 135, no. 1, pp. 497–528, 2009.

[2] J. T. Betts, *Practical methods for optimal control and estimation using nonlinear programming*. SIAM, 2010.

[3] ——, "Survey of numerical methods for trajectory optimization," *Journal of guidance, control, and dynamics*, vol. 21, no. 2, pp. 193–207, 1998.

[4] B. A. Conway, "A survey of methods available for the numerical optimization of continuous dynamic systems," *Journal of Optimization Theory and Applications*, vol. 152, no. 2, pp. 271–306, 2012.

[5] E. Polak, "Optimization: Algorithms and consistent approximations," 1997.

[6] A. Schwartz and E. Polak, "Consistent approximations for optimal control problems based on runge–kutta integration," *SIAM Journal on Control and Optimization*, vol. 34, no. 4, pp. 1235–1269, 1996.

[7] Q. Gong, I. M. Ross, W. Kang, and F. Fahroo, "Connections between the covector mapping theorem and convergence of pseudospectral methods for optimal control," *Computational Optimization and Applications*, vol. 41, no. 3, pp. 307–335, 2008.

[8] F. Fahroo and I. M. Ross, "On discrete-time optimality conditions for pseudospectral methods," in *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 2006, p. 6304.

[9] K. Bollino, L. R. Lewis, P. Sekhavat, and I. M. Ross, "Pseudospectral optimal control: a clear road for autonomous intelligent path planning," in *AIAA Infotech@ Aerospace 2007 Conference and Exhibit*, 2007, p. 2831.

[10] Y. Chen, M. Cutler, and J. P. How, "Decoupled multiagent path planning via incremental sequential convex programming," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 5954–5961.

[11] F. Augugliaro, A. P. Schoellig, and R. D'Andrea, "Generation of collision-free trajectories for a quadrocopter fleet: A sequential convex programming approach," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 1917–1922.

[12] V. Cichella, I. Kaminer, C. Walton, and N. Hovakimyan, "Optimal motion planning for differentially flat systems using Bernstein approximation," *IEEE Control Systems Letters*, vol. 2, no. 1, pp. 181–186, 2018.

[13] E. Hewitt and R. E. Hewitt, "The Gibbs-Wilbraham phenomenon: an episode in fourier analysis," *Archive for history of Exact Sciences*, vol. 21, no. 2, pp. 129–160, 1979.

[14] A. Gelb and J. Tanner, "Robust reprojection methods for the resolution of the gibbs phenomenon," *Applied and Computational Harmonic Analysis*, vol. 20, no. 1, pp. 3–25, 2006.

[15] E. Tohidi, A. Pasban, A. Kilicman, and S. L. Noghabi, "An efficient pseudospectral method for solving a class of nonlinear optimal control problems," in *Abstract and Applied Analysis*, vol. 2013. Hindawi, 2013.

[16] I. M. Ross and F. Fahroo, "Pseudospectral knotting methods for solving nonsmooth optimal control problems," *Journal of Guidance, Control, and Dynamics*, vol. 27, no. 3, pp. 397–405, 2004.

[17] C. L. Darby, W. W. Hager, and A. V. Rao, "An hp-adaptive pseudospectral method for solving optimal control problems," *Optimal Control Applications and Methods*, vol. 32, no. 4, pp. 476–502, 2011.

[18] R. T. Farouki, "The Bernstein polynomial basis: a centennial retrospective," *Computer Aided Geometric Design*, vol. 29, no. 6, pp. 379–419, 2012.

[19] H. Gzyl and J. L. Palacios, "On the approximation properties of Bernstein polynomials via probabilistic tools," *Boletın de la Asociación Matemática Venezolana*, vol. 10, no. 1, pp. 5–13, 2003.

[20] L. A. Ricciardi and M. Vasile, "Direct transcription of optimal control problems with finite elements on Bernstein basis," *Journal of Guidance, Control and Dynamics*, 2018.

[21] J.-W. Chang, Y.-K. Choi, M.-S. Kim, and W. Wang, "Computation of the minimum distance between two Bézier curves/surfaces," *Computers & Graphics*, vol. 35, no. 3, pp. 677–684, 2011.

[22] C. Kielas-Jensen and V. Cichella, "BeBOT: Bernstein polynomial toolkit for trajectory generation," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, accepted.

[23] P. Sikkema, "Der wert einiger konstanten in der theorie der approximation mit bernstein-polynomen," *Numerische Mathematik*, vol. 3, no. 1, pp. 107–116, 1961.

[24] M. J. D. Powell, *Approximation theory and methods*. Cambridge university press, 1981.

[25] M. S. Floater, "On the convergence of derivatives of Bernstein approximation," *Journal of Approximation Theory*, vol. 134, no. 1, pp. 130–135, 2005.

[26] R. F. Hartl, S. P. Sethi, and R. G. Vickson, "A survey of the maximum principles for optimal control problems with state constraints," *SIAM review*, vol. 37, no. 2, pp. 181–218, 1995.

[27] D. Garg, M. A. Patterson, C. Francolin, C. L. Darby, G. T. Huntington, W. W. Hager, and A. V. Rao, "Direct trajectory optimization and costate estimation of finite-horizon and infinite-horizon optimal control problems using a radau pseudospectral method," *Computational Optimization and Applications*, vol. 49, no. 2, pp. 335–358, 2011.

[28] B. Singh, R. Bhattacharya, and S. R. Vadali, "Verification of optimality and costate estimation using hilbert space projection," *Journal of guidance, control, and dynamics*, vol. 32, no. 4, pp. 1345–1355, 2009.

[29] V. Cichella, I. Kaminer, C. Walton, N. Hovakimyan, and A. Pascoal, "Bernstein approximation of optimal control problems," *arXiv preprint arXiv:1812.06132*, 2018.

[30] W. Kang, Q. Gong, and I. M. Ross, "Convergence of Pseudospectral Methods for a Class of Discontinuous Optimal Control," in *2005 IEEE 44th Conference on Decision and Control (CDC)*, Seville, Spain, 2005, pp. 2799–2804.