

Article

Consistency of Approximation of Bernstein Polynomial-Based Direct Methods for Optimal Control

Venanzio Cichella ^{1,*}, Isaac Kaminer ², Claire Walton ³, Naira Hovakimyan ⁴ and António Pascoal ⁵¹ Department of Mechanical Engineering, University of Iowa, Iowa City, IA 52242, USA² Department of Mechanical and Aerospace Engineering, Naval Postgraduate School, Monterey, CA 93940, USA³ Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX 78249, USA⁴ Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA⁵ Institute for Systems and Robotics (ISR), Instituto Superior Tecnico (IST), University of Lisbon, 1049-001 Lisbon, Portugal

* Correspondence: venanzio-cichella@uiowa.edu

Abstract: Bernstein polynomial approximation of continuous function has a slower rate of convergence compared to other approximation methods. “The fact seems to have precluded any numerical application of Bernstein polynomials from having been made. Perhaps they will find application when the properties of the approximant in the large are of more importance than the closeness of the approximation.”—remarked P.J. Davis in his 1963 book, *Interpolation and Approximation*. This paper presents a direct approximation method for nonlinear optimal control problems with mixed input and state constraints based on Bernstein polynomial approximation. We provide a rigorous analysis showing that the proposed method yields consistent approximations of time-continuous optimal control problems and can be used for costate estimation of the optimal control problems. This result leads to the formulation of the Covector Mapping Theorem for Bernstein polynomial approximation. Finally, we explore the numerical and geometric properties of Bernstein polynomials, and illustrate the advantages of the proposed approximation method through several numerical examples.

Keywords: numerical optimal control; Bernstein polynomials; Bezier curves

Citation: Cichella, V.; Kaminer, I.; Walton, C.; Hovakimyan, N.; Pascoal, A. Consistency of Approximation of Bernstein Polynomial-Based Direct Methods for Optimal Control. *Machines* **2022**, *10*, 1132. <https://doi.org/10.3390/machines10121132>

Academic Editor: Dan Zhang

Received: 28 September 2022

Accepted: 23 November 2022

Published: 28 November 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Motion planning plays an important role in enabling robotic systems to accomplish tasks assigned to them autonomously, safely and reliably. Over the past decades, many approaches to generating trajectories have been proposed. Examples include bug algorithms, artificial potential functions, roadmap path planners, cell decomposition methods, and optimal control-based trajectory generation. The reader is referred to [1–8] and references therein for detailed discussions and comparisons of these methods. Each technique has different advantages and disadvantages, and is best suited to certain types of problems. Motion planning based on optimal control, i.e., optimal motion planning, is particularly suitable for applications that require the trajectory to optimize some costs while guaranteeing satisfaction of a complex set of vehicle and problem constraints. These applications include multi-robot road search [9], coordinated tracking [10], optimal and constrained formation control [11], and adversarial swarm defense [12].

Optimal control problems that arise from robotics and motion-planning 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. Numerical methods include indirect and direct methods [13]. Indirect methods solve the problems by converting them into boundary value problems. Then, the solutions are found by

solving systems of differential equations. On the other hand, direct methods are based on transcribing optimal control problems into nonlinear programming problems (NLPs) using a discretization scheme [6,13–15]. These NLPs can be solved using ready-to-use NLP solvers (e.g., MATLAB, SNOPT, etc.) and do not require calculation of costate and adjoint variables, as indirect methods do.

A wide range of direct methods that use different discretization schemes have been developed, including direct single shootings, direct multiple shooting and direct collocation methods [6,13–15]. The software packages that implement some of these methods (e.g., PSOPT [16], NLOptControl [17], GPOPS II [18], PROPT [19], DIDO [20] and CasADi [21]) are particularly relevant; some of these have been applied successfully to solve a wide range of real-world problems [22–28]. Theoretical results in the literature on direct methods include those related to consistency of approximation theory; see [29], which provides a framework to assess the convergence properties of Euler and Range–Kutta discretization schemes. Motivated by the consistency of approximation theory, direct methods that use different discretization schemes have been developed, including Pseudospectral methods based on Legendre, Chebyshev and Lagrange polynomials [28]. One drawback of direct methods is that the costate of the original optimal control problem cannot be readily obtained from the approximated solution. Nevertheless, in several applications—such as motion planning and control for safety-critical robotic systems—the knowledge of the costate is important because it allows for the evaluation of the fulfillment of necessary conditions of optimality. This evaluation, in turn, provides important insights into the validity and optimality of the solution. Therefore, approaches for obtaining estimates of the costate from direct methods have been proposed in the literature on direct collocation [30–32] and direct shooting [33].

In [34] we presented a direct method based on Bernstein polynomials. We showed that the geometric properties of these polynomials allow for the implementation of efficient algorithms for the computation of state and input constraints, which are particularly useful for motion planning and trajectory generation applications [35,36]. Additional works that exploit the properties of Bernstein polynomials for nonlinear optimal control can be found in [37–41]. Furthermore, in [42] we used the approximation properties of Bernstein polynomials to derive consistency and convergence results for the proposed direct method. In the present paper, we propose an approximation scheme for primal and dual optimal control problems based on Bernstein polynomials. In particular, we propose an approach to approximate the costate of a general non-linear optimal control problem of Bolza type using the Lagrange multipliers of the Bernstein polynomial-based discrete approximation. We derive transformations that relate the Lagrange multipliers of the nonlinear programming problem to the costate of the original optimal control problem. These transformations are often referred to as *covector mapping* in the literature on direct methods for optimal control [28,29,43]. Finally, we demonstrate uniform convergence properties of the method.

The paper is structured as follows: in Section 2, we present the notation and the mathematical results, which will be used later in the paper. Section 3 introduces the optimal control problem of interest and some related assumptions, and presents the approximation method based on Bernstein approximation that approximates the optimal control problem into an NLP. In Section 4 we derive the Karush–Kuhn–Tucker (KKT) conditions associated with the NLP. Section 5 compares these conditions to the first-order optimality conditions for the original optimal control problem and states the Covector Mapping Theorem for Bernstein approximation. Numerical examples are discussed in Section 6, while Section 7 highlights the significance of the theoretical findings applied to a specific multi-robot simulation scenario, namely optimal defense against swarm attacks. The paper ends with conclusions in Section 8.

2. Notation and Mathematical Background

Vector-valued functions are denoted by bold letters, $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]^\top$, while vectors are denoted by bold letters with an upper bar, $\bar{\mathbf{x}} = [x_1, \dots, x_n]^\top \in \mathbb{R}^n$. The symbol

\mathcal{C}^r denotes the space of functions with r continuous derivatives. \mathcal{C}_n^r denotes the space of n -vector valued functions in \mathcal{C}^r . $\|\cdot\|$ denotes the Euclidean norm, $\|\bar{x}\| = \sqrt{x_1^2 + \dots + x_n^2}$.

The Bernstein basis polynomials of degree N are defined as

$$b_{j,N}(t) = \binom{N}{j} t^j (1-t)^{N-j}, \quad t \in [0, 1],$$

for $j = 0, \dots, N$, with $\binom{N}{j} = \frac{N!}{j!(N-j)!}$. A N th-order Bernstein polynomial $x_N : [0, 1] \rightarrow \mathbb{R}$ is a linear combination of $N + 1$ Bernstein basis polynomials of order N , i.e.,

$$x_N(t) = \sum_{j=0}^N \bar{x}_j b_{j,N}(t), \quad t \in [0, 1],$$

where $\bar{x}_j \in \mathbb{R}$, $j = 0, \dots, N$, are referred to as Bernstein coefficients. For the sake of generality, and with a slight abuse of terminology, in this paper, we extend the definition of a Bernstein polynomial given above to a vector of N th-order polynomials $\mathbf{x}_N : [0, 1] \rightarrow \mathbb{R}^n$ expressed in the following form

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

where $\bar{\mathbf{x}}_{0,N}, \dots, \bar{\mathbf{x}}_{N,N} \in \mathbb{R}^n$.

In what follows, we provide a review of numerical properties of Bernstein polynomials that are used throughout this paper. The derivative and integral of a Bernstein polynomial $\mathbf{x}_N(t)$ can be easily computed as

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

and

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

respectively.

Bernstein polynomials can be used to approximate smooth functions. Consider a n -vector valued function $\mathbf{x} : [0, 1] \rightarrow \mathbb{R}^n$. The N th order Bernstein approximation of $\mathbf{x}(t)$ is a vector of Bernstein polynomials $\mathbf{x}_N(t)$ computed as in (1) with $\bar{\mathbf{x}}_{j,N} = \mathbf{x}(t_j)$ and $t_j = \frac{j}{N}$ for all $j = 0, \dots, N$. Namely,

$$\mathbf{x}(t) \approx \mathbf{x}_N(t) = \sum_{j=0}^N \mathbf{x}(t_j) b_{j,N}(t), \quad t_j = \frac{j}{N}. \quad (3)$$

The following results hold for Bernstein approximations.

Lemma 1 (Uniform convergence of Bernstein approximation). *Let $\mathbf{x}(t) \in \mathcal{C}_n^0$ on $[0, 1]$, and let $\mathbf{x}_N(t)$ be computed as in Equation (3). Then, for arbitrary order of approximation $N \in \mathbb{Z}^+$, the Bernstein approximation $\mathbf{x}_N(t)$ satisfies*

$$\|\mathbf{x}_N(t) - \mathbf{x}(t)\| \leq C_0 W_x(N^{-\frac{1}{2}}),$$

where C_0 is a positive constant satisfying $C_0 < 5n/4$, and $W_x(\cdot)$ is the modulus of continuity of $\mathbf{x}(t)$ in $[0, 1]$ [44–46]. Moreover, if $\mathbf{x}(t) \in \mathcal{C}_n^1$, then

$$\|\dot{\mathbf{x}}_N(t) - \dot{\mathbf{x}}(t)\| \leq C_1 W_{x'}(N^{-\frac{1}{2}}),$$

where C_1 is a positive constant satisfying $C_1 < 9n/4$ and $W_{x'}(\cdot)$ is the modulus of continuity of $\dot{\mathbf{x}}(t)$ in $[0, 1]$ [47].

Lemma 2 ([48]). Assume $\mathbf{x}(t) \in C_n^{r+2}$, $r \geq 0$, and let $\mathbf{x}_N(t)$ be computed as in Equation (3). Let $\mathbf{x}^{(r)}(t)$ denote the r th derivative of $\mathbf{x}(t)$. Then, the following inequalities hold for all $t \in [0, 1]$:

$$\begin{aligned} \|\mathbf{x}_N(t) - \mathbf{x}(t)\| &\leq \frac{C_0}{N}, \\ &\vdots \\ \|\mathbf{x}_N^{(r)}(t) - \mathbf{x}^{(r)}(t)\| &\leq \frac{C_r}{N}, \end{aligned}$$

where C_0, \dots, C_r are independent of N .

Lemma 3. If $\mathbf{x}(t) \in C_n^0$ on $[0, 1]$, then we have

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

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

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

The Lemma above follows directly from Lemmas 1 and 2 and Equation (2). The following property of Bernstein polynomials is relevant to this paper.

Property 1 (End point values). The Bernstein polynomial given by Equation (1) satisfies $\mathbf{x}_N(0) = \bar{\mathbf{x}}_{0,N}$ and $\mathbf{x}_N(1) = \bar{\mathbf{x}}_{N,N}$.

3. Problem Formulation

This paper considers the following optimal control problem:

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

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

subject to

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \forall t \in [0, 1], \tag{5}$$

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

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

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

Next, we formulate a discretized version of Problem P, here referred to as Problem P_N , where N denotes the order of approximation. This requires that we approximate the input and state functions, the cost function, the system dynamics and the equality and inequality constraints in Problem P. First, consider the following N th-order vectors of Bernstein polynomials:

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

with $\mathbf{x}_N : [0, 1] \rightarrow \mathbb{R}^{n_x}$, $\mathbf{u}_N : [0, 1] \rightarrow \mathbb{R}^{n_u}$, $\bar{\mathbf{x}}_{j,N} \in \mathbb{R}^{n_x}$ and $\bar{\mathbf{u}}_{j,N} \in \mathbb{R}^{n_u}$. Let $\bar{\mathbf{x}}_N \in \mathbb{R}^{n_x \times (N+1)}$ and $\bar{\mathbf{u}}_N \in \mathbb{R}^{n_u \times (N+1)}$ be defined as

$$\bar{\mathbf{x}}_N = [\bar{\mathbf{x}}_{0,N}, \dots, \bar{\mathbf{x}}_{N,N}], \quad \bar{\mathbf{u}}_N = [\bar{\mathbf{u}}_{0,N}, \dots, \bar{\mathbf{u}}_{N,N}].$$

Let $0 = t_0 < t_1 < \dots < t_N = 1$ be a set of equidistant time nodes, i.e., $t_j = \frac{j}{N}$. Then, Problem P_N can be stated as follows:

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

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

subject to

$$\|\dot{\mathbf{x}}_N(t_j) - \mathbf{f}(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))\| \leq \delta_P^N, \quad \forall j = 0, \dots, N, \quad (10)$$

$$\mathbf{e}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) = \mathbf{0}, \quad (11)$$

$$\mathbf{h}(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j)) \leq \delta_P^N \mathbf{1}, \quad \forall j = 0, \dots, N, \quad (12)$$

where $w = \frac{1}{N+1}$, and δ_P^N is a small positive number that depends on N and converges uniformly to 0, i.e., $\lim_{N \rightarrow \infty} \delta_P^N = 0$.

Remark 1. Compared to the constraints of Problem P , the dynamic and inequality constraints given by Equations (10) and (12) are relaxed. Motivated by previous work on consistency of approximation theory [29], the bound δ_P^N , referred to as relaxation bound, is introduced to guarantee that Problem P_N has a feasible solution. As will become clear later, the relaxation bound can be made arbitrarily small by choosing a sufficiently large order of approximation N . Furthermore, note that when $N \rightarrow \infty$, the right-hand sides of Equations (10) and (12) are equal to zero, i.e., the difference between the constraints imposed by Problems P and P_N vanishes.

Remark 2. The outcome of Problem P_N is a set of optimal Bernstein coefficients $\bar{\mathbf{x}}_N^*$ and $\bar{\mathbf{u}}_N^*$ that determine the vectors of Bernstein polynomials $\mathbf{x}_N^*(t)$ and $\mathbf{u}_N^*(t)$, i.e.,

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

In our previous work, see [42], we provide theoretical results demonstrating: (i) the existence of a feasible solution to Problem P_N , and (ii) the convergence of the pair $(\mathbf{x}_N^*(t), \mathbf{u}_N^*(t))$ to the optimal solution of Problem P , given by $(\mathbf{x}^*(t), \mathbf{u}^*(t))$. Nevertheless, the present paper focuses on the existence and convergence of the estimates of the costates of Problem P , which are introduced next.

4. Costate Estimation for Problem P

4.1. First-Order Optimality Conditions of Problem P

We start by deriving the first-order necessary conditions for Problem P . Let $\boldsymbol{\lambda}(t) : [0, 1] \rightarrow \mathbb{R}^{n_x}$ be the costate trajectory, and let $\boldsymbol{\mu}(t) : [0, 1] \rightarrow \mathbb{R}^{n_h}$ and $\mathbf{v} \in \mathbb{R}^{n_e}$ be the multipliers. By defining the Lagrangian of the Hamiltonian (also known as the D-form [49]) as

$$\mathcal{L}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = \mathcal{H}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t)) + \boldsymbol{\mu}^\top(t) \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t)),$$

where the Hamiltonian \mathcal{H} is given by

$$\mathcal{H}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t)) = F(\mathbf{x}(t), \mathbf{u}(t)) + \boldsymbol{\lambda}^\top(t) \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)),$$

the dual of Problem P can be formulated as follows [49].

Problem 3 (Problem P_λ). Determine $\mathbf{x}(t)$, $\mathbf{u}(t)$, $\boldsymbol{\lambda}(t)$, $\boldsymbol{\mu}(t)$ and \mathbf{v} that for all $t \in [0, 1]$ satisfy Equations (5)–(7) and

$$\boldsymbol{\mu}^\top(t)\mathbf{h}(\mathbf{x}(t), \mathbf{u}(t)) = 0, \quad \boldsymbol{\mu}(t) \geq 0, \tag{14}$$

$$\dot{\boldsymbol{\lambda}}^\top(t) + \mathcal{L}_x(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = 0, \tag{15}$$

$$\boldsymbol{\lambda}^\top(0) = -\mathbf{v}^\top \mathbf{e}_{x(0)}(\mathbf{x}(0), \mathbf{x}(1)) - E_{x(0)}(\mathbf{x}(0), \mathbf{x}(1)), \tag{16}$$

$$\boldsymbol{\lambda}^\top(1) = \mathbf{v}^\top \mathbf{e}_{x(1)}(\mathbf{x}(0), \mathbf{x}(1)) + E_{x(1)}(\mathbf{x}(0), \mathbf{x}(1)), \tag{17}$$

$$\mathcal{L}_u(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = 0. \tag{18}$$

In the above problem, subscripts are used to denote partial derivatives, e.g., $F_x(\mathbf{x}, \mathbf{u}) = \frac{\partial}{\partial \mathbf{x}}F(\mathbf{x}, \mathbf{u})$.

The following assumptions are imposed onto Problem P_λ .

Assumption 1. $E, F, \mathbf{f}, \mathbf{e}$ and \mathbf{h} are continuously differentiable with respect to their arguments, and their gradients are Lipschitz continuous over the domain.

Assumption 2. Solutions $\mathbf{x}^*(t)$, $\mathbf{u}^*(t)$, $\boldsymbol{\lambda}^*(t)$, $\boldsymbol{\mu}^*(t)$ and \mathbf{v}^* of Problem P_λ exist and satisfy $\mathbf{x}^*(t) \in \mathcal{C}_{n_x}^1$, $\mathbf{u}^*(t) \in \mathcal{C}_{n_u}^0$, $\boldsymbol{\lambda}^*(t) \in \mathcal{C}_{n_x}^1$ and $\boldsymbol{\mu}^*(t) \in \mathcal{C}_{n_h}^0$ in $[0, 1]$.

Remark 3. Notice that Problem P_λ implicitly assumes the absence of pure state constraints in Problem P . If the inequality constraint in Equation (7) is independent of $\mathbf{u}(t)$, then the costate $\boldsymbol{\lambda}(t)$ must also satisfy the following jump condition [49]:

$$\boldsymbol{\lambda}(t_e^-) = \boldsymbol{\lambda}(t_e^+) + \mathbf{h}_{x(t_e)}^\top \boldsymbol{\eta},$$

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 constant covector. For simplicity, the theoretical results that will be presented in Section 5 do not consider the jump conditions above, i.e., the inequality constraints are dependent on $\mathbf{u}(t)$. Nevertheless, numerical examples will be presented in Section 6, showing the applicability of the discretization method to pure state-constrained problems.

4.2. KKT Conditions of Problem P_N

Now, we derive the necessary conditions of Problem P_N . Let us introduce the following N th-order Bernstein polynomials:

$$\boldsymbol{\lambda}_N(t) = \sum_{j=0}^N \bar{\boldsymbol{\lambda}}_{j,N} b_{j,N}(t), \quad \boldsymbol{\mu}_N(t) = \sum_{j=0}^N \bar{\boldsymbol{\mu}}_{j,N} b_{j,N}(t), \tag{19}$$

with $\boldsymbol{\lambda}_N : [0, 1] \rightarrow \mathbb{R}^{n_x}$, $\boldsymbol{\mu}_N : [0, 1] \rightarrow \mathbb{R}^{n_h}$, $\bar{\boldsymbol{\lambda}}_{j,N} \in \mathbb{R}^{n_x}$ and $\bar{\boldsymbol{\mu}}_{j,N} \in \mathbb{R}^{n_h}$, and the vector $\bar{\mathbf{v}} \in \mathbb{R}^{n_e}$. Finally, let $\bar{\boldsymbol{\lambda}}_N \in \mathbb{R}^{n_x \times (N+1)}$ and $\bar{\boldsymbol{\mu}}_N \in \mathbb{R}^{n_h \times (N+1)}$ be defined as

$$\bar{\boldsymbol{\lambda}}_N = [\bar{\boldsymbol{\lambda}}_{0,N}, \dots, \bar{\boldsymbol{\lambda}}_{N,N}], \quad \bar{\boldsymbol{\mu}}_N = [\bar{\boldsymbol{\mu}}_{0,N}, \dots, \bar{\boldsymbol{\mu}}_{N,N}].$$

With the above notation, the Lagrangian for problem P_N can be written as

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

Then, the duality of Problem P_N can be stated as follows:

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

$$\begin{aligned} \left\| \mu_N^\top(t_k) \mathbf{h}(\mathbf{x}_N(t_k), \mathbf{u}_N(t_k)) \right\| &\leq N^{-1} \delta_D^N, \\ \mu_N(t_k) &\geq -N^{-1} \delta_D^N \mathbf{1}, \quad \forall k = 0, \dots, N, \end{aligned} \quad (20)$$

and the stationarity conditions

$$\left\| \frac{\partial \mathcal{L}_N}{\partial \bar{x}_{k,N}} \right\| \leq \delta_D^N, \quad \left\| \frac{\partial \mathcal{L}_N}{\partial \bar{u}_{k,N}} \right\| \leq \delta_D^N, \quad \forall k = 0, \dots, N, \quad (21)$$

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

At this point, similarly to most results on costate estimation [50–52], we introduce additional conditions that must be added to Equations (10)–(12), (20) and (21) 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:

$$\left\| \frac{\lambda_N^\top(0)}{w} + \bar{v}^\top \mathbf{e}_{x(0)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) + E_{x(0)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) \right\| \leq \delta_D^N, \quad (22)$$

$$\left\| \frac{\lambda_N^\top(t_N)}{w} - \bar{v}^\top \mathbf{e}_{x(1)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) - E_{x(1)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) \right\| \leq \delta_D^N. \quad (23)$$

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 (16) and (17). With this setup, we define the following problem:

Problem 5 (Problem $P_{N\lambda}^{clos}$). Determine \bar{x}_N , \bar{u}_N , $\bar{\lambda}_N$, $\bar{\mu}_N$ and \bar{v} that satisfy the primal feasibility conditions, namely Equations (10)–(12), the complementary slackness and dual feasibility conditions (20), the stationarity conditions (21), and the closure conditions (22) and (23).

The solution of Problem $P_{N\lambda}^{clos}$ presents a set of optimal Bernstein coefficients \bar{x}_N^* , \bar{u}_N^* , $\bar{\lambda}_N^*$, $\bar{\mu}_N^*$ (which determine the Bernstein polynomials $x_N^*(t)$, $u_N^*(t)$, $\lambda_N^*(t)$ and $\mu_N^*(t)$) and a vector \bar{v}^* .

5. Feasibility and Consistency of Problem $P_{N\lambda}^{clos}$

The objective of this section is to investigate the ability of the solutions of Problem $P_{N\lambda}^{clos}$ to approximate the solutions of Problem P_λ . In what follows, we first show the existence of a solution to Problem $P_{N\lambda}^{clos}$ (feasibility). Second, we investigate the convergence properties of this solution as $N \rightarrow \infty$ (consistency). Third, by combining these two results, we finally formulate the *covector mapping theorem* for Bernstein approximations, which provides a map between the solution of Problem $P_{N\lambda}^{clos}$ and the solution of Problem P_λ . The main results of this section are reported in the three theorems below and summarized in Figure 1.

Theorem 1 (Feasibility). Let

$$\delta_D^N = C_D \max\{W_{x'}(N^{-\frac{1}{2}}), W_x(N^{-\frac{1}{2}}), W_u(N^{-\frac{1}{2}}), W_{\lambda'}(N^{-\frac{1}{2}}), W_\lambda(N^{-\frac{1}{2}}), W_\mu(N^{-\frac{1}{2}})\}, \quad (24)$$

$$\delta_P^N = C_P \max\{W_{x'}(N^{-\frac{1}{2}}), W_x(N^{-\frac{1}{2}}), W_u(N^{-\frac{1}{2}})\}, \quad (25)$$

where C_D and C_P are positive constants independent of N , and $W_{x'}(\cdot)$, $W_x(\cdot)$, $W_u(\cdot)$, $W_{\lambda'}(\cdot)$, $W_\lambda(\cdot)$ and $W_\mu(\cdot)$ are the moduli of continuity of $\dot{x}(t)$, $x(t)$, $u(t)$, $\dot{\lambda}(t)$, $\lambda(t)$ and $\mu(t)$, respectively. Then Problem $P_{N\lambda}^{clos}$ is feasible for arbitrary order of approximation $N \in \mathbb{Z}^+$.

Proof. This proof follows by constructing a solution for Problem $P_{N\lambda}^{clos}$, with δ_D^N given by Equation (24). To this end, let $x(t)$, $u(t)$, $\lambda(t)$, $\mu(t)$ and v be a solution of Problem P_λ , which exists by Assumption 2, and define

$$\bar{x}_{j,N} = x(t_j), \quad \bar{u}_{j,N} = u(t_j), \tag{26}$$

$$\bar{\lambda}_{j,N} = w\lambda(t_j), \quad \bar{\mu}_{j,N} = w\mu(t_j), \quad \bar{v} = v, \tag{27}$$

for all $j = 0, \dots, N$, $t_j = \frac{j}{N}$, $w = \frac{1}{N+1}$, with corresponding Bernstein polynomials given by

$$\begin{aligned} x_N(t) &= \sum_{j=0}^N \bar{x}_{j,N} b_{j,N}(t), & u_N(t) &= \sum_{j=0}^N \bar{u}_{j,N} b_{j,N}(t), \\ \lambda_N(t) &= \sum_{j=0}^N \bar{\lambda}_{j,N} b_{j,N}(t), & \mu_N(t) &= \sum_{j=0}^N \bar{\mu}_{j,N} b_{j,N}(t). \end{aligned} \tag{28}$$

The remainder of this proof shows that $x_N(t)$, $u_N(t)$, $\lambda_N(t)$, $\mu_N(t)$ and \bar{v} given above satisfy Equations (20)–(23). The satisfaction of Equations (10)–(12) can be demonstrated using a proof similar to the one of [42], and is thus omitted. We start by defining the Bernstein coefficients $\tilde{\lambda}_{j,N}$ and $\tilde{\mu}_{j,N}$ as follows

$$\tilde{\lambda}_{j,N} = \frac{\bar{\lambda}_{j,N}}{w}, \quad \tilde{\mu}_{j,N} = \frac{\bar{\mu}_{j,N}}{w}, \tag{29}$$

with corresponding Bernstein polynomials given by

$$\tilde{\lambda}_N(t) = \sum_{j=0}^N \tilde{\lambda}_{j,N} b_{j,N}(t), \quad \tilde{\mu}_N(t) = \sum_{j=0}^N \tilde{\mu}_{j,N} b_{j,N}(t).$$

Notice that

$$\tilde{\lambda}_N(t) = \frac{\lambda_N(t)}{w}, \quad \tilde{\mu}_N(t) = \frac{\mu_N(t)}{w}. \tag{30}$$

Combining Equations (26), (27) and (29) and using Assumption 2 and Lemma 1, we get

$$\begin{aligned} \|x_N(t) - x(t)\| &\leq C_x W_x(N^{-\frac{1}{2}}), \\ \|u_N(t) - u(t)\| &\leq C_u W_u(N^{-\frac{1}{2}}), \\ \|\dot{x}_N(t) - \dot{x}(t)\| &\leq C_{x'} W_{x'}(N^{-\frac{1}{2}}), \\ \|\tilde{\lambda}_N(t) - \lambda(t)\| &\leq C_\lambda W_\lambda(N^{-\frac{1}{2}}), \\ \|\tilde{\mu}_N(t) - \mu(t)\| &\leq C_\mu W_\mu(N^{-\frac{1}{2}}), \\ \|\dot{\tilde{\lambda}}_N(t) - \dot{\lambda}(t)\| &\leq C_{\lambda'} W_{\lambda'}(N^{-\frac{1}{2}}), \end{aligned} \tag{31}$$

where $C_\lambda < \frac{5n_x}{4}$, $C_\mu < \frac{5n_u}{4}$, $C_{\lambda'} < \frac{9n_x}{4}$ and $W_\lambda(\cdot)$, $W_\mu(\cdot)$ and $W_{\lambda'}(\cdot)$ are the moduli of continuity of $\lambda(t)$, $\mu(t)$ and $\dot{\lambda}(t)$, respectively.

Now, we show that the bound in Equation (20) is satisfied. Using Equation (30), and adding and subtracting $w(\mu^\top(t_k)h(x_N(t_k), u_N(t_k)) + \mu^\top(t_k)h(x(t_k), u(t_k)))$, we get

$$\begin{aligned} \|\mu_N^\top(t_k)h(x_N(t_k), u_N(t_k))\| &= \|w\tilde{\mu}_N^\top(t_k)h(x_N(t_k), u_N(t_k))\| \\ &\leq w\|(\tilde{\mu}_N^\top(t_k) - \mu^\top(t_k))h(x_N(t_k), u_N(t_k))\| \\ &\quad + w\|\mu^\top(t_k)h(x(t_k), u(t_k))\| \\ &\quad + w\|\mu^\top(t_k)(h(x_N(t_k), u_N(t_k)) - h(x(t_k), u(t_k)))\| \end{aligned}$$

Using Equation (14), the above inequality reduces to

$$\begin{aligned} \|\boldsymbol{\mu}_N^\top(t_k)\mathbf{h}(\mathbf{x}_N(t_k), \mathbf{u}_N(t_k))\| &\leq w\|(\tilde{\boldsymbol{\mu}}_N^\top(t_k) - \boldsymbol{\mu}^\top(t_k))\mathbf{h}(\mathbf{x}_N(t_k), \mathbf{u}_N(t_k))\| \\ &\quad + w\|\boldsymbol{\mu}^\top(t_k)(\mathbf{h}(\mathbf{x}_N(t_k), \mathbf{u}_N(t_k)) - \mathbf{h}(\mathbf{x}(t_k), \mathbf{u}(t_k)))\| \\ &\leq w\|\mathbf{h}(\mathbf{x}_N(t_k), \mathbf{u}_N(t_k))\|C_\mu W_\mu(N^{-\frac{1}{2}}) \\ &\quad + w\|\boldsymbol{\mu}^\top(t_k)\|L_h(C_x W_x(N^{-\frac{1}{2}}) + C_u W_u(N^{-\frac{1}{2}})), \end{aligned}$$

where we used the bounds in Equation (31) together with the Lipschitz assumption on \mathbf{h} (see Assumptions 1). Finally, from using Assumptions 1 and 2, it follows that \mathbf{h} and $\boldsymbol{\mu}$ are bounded on $[0, 1]$ with bounds h_{\max} and μ_{\max} , respectively. Therefore, we get

$$\|\boldsymbol{\mu}_N^\top(t_k)\mathbf{h}(\mathbf{x}_N(t_k), \mathbf{u}_N(t_k))\| \leq w[h_{\max}C_\mu W_\mu(N^{-\frac{1}{2}}) + \mu_{\max}L_h(C_x W_x(N^{-\frac{1}{2}}) + C_u W_u(N^{-\frac{1}{2}}))],$$

which implies that the bound in Equation (20) is satisfied with δ_D^N given by Equation (24) and $C_D > h_{\max}C_\mu + \mu_{\max}L_h(C_x + C_u)$. Similarly,

$$\boldsymbol{\mu}_N(t_k) = w\tilde{\boldsymbol{\mu}}_N(t_k) \geq w\boldsymbol{\mu}(t_k) - w\|\boldsymbol{\mu}(t_k) - \tilde{\boldsymbol{\mu}}_N(t_k)\|\mathbf{1} - N^{-1}C_\mu W_\mu(N^{-\frac{1}{2}})\mathbf{1},$$

which proves that Equation (20) holds.

Now, consider the left equation in (21). For $k = 0$, we have

$$\begin{aligned} \left\| \frac{\partial \mathcal{L}_N}{\partial \bar{\mathbf{x}}_{0,N}} \right\| &= \left\| E_{x(0)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) \right. \\ &\quad + w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) \\ &\quad + \sum_{j=0}^N \boldsymbol{\lambda}_N^\top(t_j)(f_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) - \dot{b}_{0,N}(t_j)) \\ &\quad + \sum_{j=0}^N \boldsymbol{\mu}_N^\top(t_j)\mathbf{h}_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) \\ &\quad \left. + \bar{\mathbf{v}}^\top e_{x(0)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) \right\|. \end{aligned} \tag{32}$$

Substituting $w\tilde{\boldsymbol{\lambda}}_N(t_j) = \boldsymbol{\lambda}_N(t_j)$ and $w\tilde{\boldsymbol{\mu}}_N(t_j) = \boldsymbol{\mu}_N(t_j)$, the equation above can be written as

$$\begin{aligned} \left\| \frac{\partial \mathcal{L}_N}{\partial \bar{\mathbf{x}}_{0,N}} \right\| &= \left\| E_{x(0)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) \right. \\ &\quad + w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) \\ &\quad + w \sum_{j=0}^N \tilde{\boldsymbol{\lambda}}_N^\top(t_j)(f_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) - \dot{b}_{0,N}(t_j)) \\ &\quad + w \sum_{j=0}^N \tilde{\boldsymbol{\mu}}_N^\top(t_j)\mathbf{h}_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) \\ &\quad \left. + \bar{\mathbf{v}}^\top e_{x(0)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) \right\|. \end{aligned} \tag{33}$$

Notice that the following inequalities are satisfied:

$$\begin{aligned} \left\| w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) - \int_0^1 F_x(\mathbf{x}(t), \mathbf{u}(t))b_{0,N}(t)dt \right\| \\ \leq \bar{C}_1(N^{-\frac{1}{2}} + W_x(N^{-\frac{1}{2}}) + W_u(N^{-\frac{1}{2}})), \end{aligned} \tag{34a}$$

$$\left\| w \sum_{j=0}^N \tilde{\lambda}_N^\top(t_j) b_{0,N}(t_j) - \int_0^1 \lambda^\top(t) b_{0,N}(t) dt \right\| \leq \bar{C}_2(N^{-\frac{1}{2}} + W_\lambda(N^{-\frac{1}{2}})), \tag{34b}$$

$$\begin{aligned} & \left\| w \sum_{j=0}^N \tilde{\lambda}_N^\top(t_j) f_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j)) b_{0,N}(t_j) - \int_0^1 \lambda^\top(t) f_x(\mathbf{x}(t), \mathbf{u}(t)) b_{0,N}(t) dt \right\| \\ & \leq \bar{C}_3(N^{-\frac{1}{2}} + W_\lambda(N^{-\frac{1}{2}}) + W_x(N^{-\frac{1}{2}}) + W_u(N^{-\frac{1}{2}})), \end{aligned} \tag{34c}$$

$$\begin{aligned} & \left\| w \sum_{j=0}^N \tilde{\mu}_N^\top(t_j) h_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j)) b_{0,N}(t_j) - \int_0^1 \mu^\top(t) h_x(\mathbf{x}(t), \mathbf{u}(t)) b_{0,N}(t) dt \right\| \\ & \leq \bar{C}_4(N^{-\frac{1}{2}} + W_\mu(N^{-\frac{1}{2}}) + W_x(N^{-\frac{1}{2}}) + W_u(N^{-\frac{1}{2}})), \end{aligned} \tag{34d}$$

for some positive $\bar{C}_1, \bar{C}_2, \bar{C}_3$ and \bar{C}_4 independent of N . Proof of the above inequalities is given in Appendix A. Then, the combination of Equations (33) and (34) yields the following inequality

$$\begin{aligned} \left\| \frac{\partial \mathcal{L}_N}{\partial \bar{\mathbf{x}}_{0,N}} \right\| & \leq \left\| E_{x(0)}(\mathbf{x}(0), \mathbf{x}(1)) + \int_0^1 F_x(\mathbf{x}(t), \mathbf{u}(t)) b_{0,N}(t) dt \right. \\ & \quad - \int_0^1 \lambda^\top(t) \dot{b}_{0,N}(t) dt \\ & \quad + \int_0^1 \lambda^\top(t) f_x(\mathbf{x}(t), \mathbf{u}(t)) b_{0,N}(t) dt \\ & \quad + \int_0^1 \mu^\top(t) h_x(\mathbf{x}(t), \mathbf{u}(t)) b_{0,N}(t) dt \\ & \quad \left. + \bar{\mathbf{v}}^\top e_{x(0)}(\mathbf{x}(0), \mathbf{x}(1)) \right\| \\ & \quad + \bar{C} \max\{N^{-\frac{1}{2}}, W_x(N^{-\frac{1}{2}}), W_u(N^{-\frac{1}{2}}), W_\lambda(N^{-\frac{1}{2}}), W_\mu(N^{-\frac{1}{2}})\}, \end{aligned} \tag{35}$$

with $\bar{C} \geq 4 \max\{\bar{C}_1, \bar{C}_2, \bar{C}_3, \bar{C}_4\}$. Using integration by parts, we have $\int_0^1 \lambda^\top(t) \dot{b}_{0,N}(t) dt = -\int_0^1 \dot{\lambda}^\top(t) b_{0,N}(t) dt + [\lambda^\top(t) b_{0,N}(t)]_0^1$. Thus, since $b_{0,N}(0) = 1, b_{N,N}(0) = 0$, the above inequality becomes

$$\begin{aligned} \left\| \frac{\partial \mathcal{L}_N}{\partial \bar{\mathbf{x}}_{0,N}} \right\| & \leq \left\| E_{x(0)}(\mathbf{x}(0), \mathbf{x}(1)) + \lambda^\top(0) + \bar{\mathbf{v}}^\top e_{x(0)}(\mathbf{x}(0), \mathbf{x}(1)) \right. \\ & \quad + \int_0^1 \left(\dot{\lambda}^\top(t) + F_x(\mathbf{x}(t), \mathbf{u}(t)) + \lambda^\top(t) f_x(\mathbf{x}(t), \mathbf{u}(t)) \right. \\ & \quad \left. \left. + \mu^\top(t) h_x(\mathbf{x}(t), \mathbf{u}(t)) \right) b_{0,N}(t) dt \right\| \\ & \quad + \bar{C} \max\{N^{-\frac{1}{2}}, W_x(N^{-\frac{1}{2}}), W_u(N^{-\frac{1}{2}}), W_\lambda(N^{-\frac{1}{2}}), W_\mu(N^{-\frac{1}{2}})\}. \end{aligned} \tag{36}$$

Finally, using Equations (15) and (16), the above inequality reduces to the left condition in Equation (21) for $k = 0$, with δ_D^N given by Equation (24) and $C_D \geq \bar{C}$. The same condition for $k = 1, \dots, N$ can be shown to be satisfied using an identical argument. The stationarity condition in the right of Equation (21) can also be verified similarly, and the computations are thus omitted. To show that the closure condition (22) is satisfied, we use the definitions in Equations (26) and (27) together with the end point values property of Bernstein polynomials, Property 1 in Section 2, which gives

$$\begin{aligned} & \left\| \frac{\lambda_N^\top(0)}{w} + \bar{\mathbf{v}}^\top e_{x(0)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) + E_{x(0)}(\mathbf{x}_N(0), \mathbf{x}_N(t_N)) \right\| \\ & \leq \left\| \lambda^\top(0) + \bar{\mathbf{v}}^\top e_{x(0)}(\mathbf{x}(0), \mathbf{x}(1)) + E_{x(0)}(\mathbf{x}(0), \mathbf{x}(1)) \right\| = 0, \end{aligned}$$

where the last equality follows from Equation (16). An identical argument can be used to show that the closure condition (23) holds, thus completing the proof of Theorem 1.

Corollary 1. *If solutions $\mathbf{x}^*(t)$, $\mathbf{u}^*(t)$, $\boldsymbol{\lambda}^*(t)$, $\boldsymbol{\mu}^*(t)$ and \mathbf{v}^* of Problem P_λ exist and satisfy $\dot{\mathbf{x}}^*(t) \in \mathcal{C}_{n_x}^2$, $\mathbf{u}^*(t) \in \mathcal{C}_{n_u}^2$, $\dot{\boldsymbol{\lambda}}^*(t) \in \mathcal{C}_{n_\lambda}^2$, and $\boldsymbol{\mu}^*(t) \in \mathcal{C}_{n_\mu}^2$ in $[0, 1]$, then Theorem 1 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 the order of approximation, N .*

□

Proof. The proof of Corollary 1 follows easily by applying Lemma 2 to the proof of Theorem 1.

Remark 4. *We notice that for arbitrarily small scalar $\epsilon_D > 0$, there exists N_1 such that for all $N \geq N_1$, we have $\delta_D^N \leq \epsilon_D$; i.e., the relaxation bound in Problem $P_{N\lambda}^{clos}$ can be made arbitrarily small by choosing sufficiently large N .*

Theorem 2 (Consistency). *Let $\{(\bar{\mathbf{x}}_N^*, \bar{\mathbf{u}}_N^*, \bar{\boldsymbol{\lambda}}_N^*, \bar{\boldsymbol{\mu}}_N^*, \bar{\mathbf{v}}^*)\}_{N=N_1}^\infty$ be a sequence of solutions of Problem $P_{N\lambda}^{clos}$. Consider the sequence of transformed solutions $\{(\tilde{\mathbf{x}}_N^*, \tilde{\mathbf{u}}_N^*, \tilde{\boldsymbol{\lambda}}_N^*, \tilde{\boldsymbol{\mu}}_N^*, \tilde{\mathbf{v}}^*)\}_{N=N_1}^\infty$, with*

$$\tilde{\boldsymbol{\lambda}}_{j,N}^* = \frac{\bar{\boldsymbol{\lambda}}_{j,N}^*}{w}, \quad \tilde{\boldsymbol{\mu}}_{j,N}^* = \frac{\bar{\boldsymbol{\mu}}_{j,N}^*}{w}, \tag{37}$$

and the corresponding polynomial approximation $\{(\mathbf{x}_N^(t), \mathbf{u}_N^*(t), \boldsymbol{\lambda}_N^*(t), \boldsymbol{\mu}_N^*(t), \mathbf{v}^*)\}_{N=N_1}^\infty$. Assume that the latter has a uniform accumulation point, i.e.,*

$$\lim_{N \rightarrow \infty} (\mathbf{x}_N^*(t), \mathbf{u}_N^*(t), \boldsymbol{\lambda}_N^*(t), \boldsymbol{\mu}_N^*(t), \mathbf{v}^*) = (\mathbf{x}^\infty(t), \mathbf{u}^\infty(t), \boldsymbol{\lambda}^\infty(t), \boldsymbol{\mu}^\infty(t), \mathbf{v}^\infty), \quad \forall t \in [0, 1],$$

and assume $\dot{\mathbf{x}}^\infty(t)$, $\mathbf{u}^\infty(t)$, $\dot{\boldsymbol{\lambda}}^\infty(t)$ and $\tilde{\boldsymbol{\mu}}^\infty(t)$ are continuous on $[0, 1]$. Then,

$$(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t), \boldsymbol{\lambda}^\infty(t), \tilde{\boldsymbol{\mu}}^\infty(t), \mathbf{v}^\infty)$$

is a solution of Problem P_λ .

□

Proof. The objective is to show that $\mathbf{x}^\infty(t), \mathbf{u}^\infty(t), \boldsymbol{\lambda}^\infty(t), \tilde{\boldsymbol{\mu}}^\infty(t)$ and \mathbf{v}^∞ satisfy Equations (5)–(7) and (14)–(18). The satisfaction of Equations (5)–(7) has been demonstrated in ([42] [Proof of Theorem 2]). We start by showing Equation (14), and we do so using a proof by contradiction. Assume that $\mathbf{x}^\infty(t), \mathbf{u}^\infty(t), \tilde{\boldsymbol{\mu}}^\infty(t)$ do not satisfy Equation (14). Then, there exists $t' \in [0, 1]$, such that

$$\|\tilde{\boldsymbol{\mu}}^{\infty\top}(t') \mathbf{h}(\mathbf{x}^\infty(t'), \mathbf{u}^\infty(t'))\| > 0. \tag{38}$$

Since the nodes $\{t_k\}_{k=0}^N$ are dense in $[0, 1]$, there exists a sequence of indices $\{k_N\}_{N=0}^\infty$ such that

$$\lim_{N \rightarrow \infty} t_{k_N} = t',$$

which implies

$$\begin{aligned} \lim_{N \rightarrow \infty} \|\tilde{\boldsymbol{\mu}}^\infty(t') - \tilde{\boldsymbol{\mu}}^\infty(t_{k_N})\| &= 0, \\ \lim_{N \rightarrow \infty} \|\mathbf{x}^\infty(t') - \mathbf{x}^\infty(t_{k_N})\| &= 0, \\ \lim_{N \rightarrow \infty} \|\mathbf{u}^\infty(t') - \mathbf{u}^\infty(t_{k_N})\| &= 0. \end{aligned}$$

Then, we have

$$\begin{aligned} \|\tilde{\mu}^{\infty\top}(t')\mathbf{h}(\mathbf{x}^\infty(t'), \mathbf{u}^\infty(t'))\| &\leq \lim_{N \rightarrow \infty} \|(\tilde{\mu}_N^{*\top}(t') - \tilde{\mu}_N^{*\top}(t_{k_N}))\mathbf{h}(\mathbf{x}_N^*(t'), \mathbf{u}_N^*(t'))\| \\ &\quad + \lim_{N \rightarrow \infty} \|\tilde{\mu}_N^{*\top}(t_{k_N})(\mathbf{h}(\mathbf{x}_N^*(t'), \mathbf{u}_N^*(t')) \\ &\quad \quad - \mathbf{h}(\mathbf{x}_N^*(t_{k_N}), \mathbf{u}_N^*(t_{k_N})))\| \\ &\quad + \lim_{N \rightarrow \infty} \|\tilde{\mu}_N^{*\top}(t_{k_N})\mathbf{h}(\mathbf{x}_N^*(t_{k_N}), \mathbf{u}_N^*(t_{k_N}))\| \\ &= \lim_{N \rightarrow \infty} \frac{1}{w} \|\mu_N^{*\top}(t_{k_N})\mathbf{h}(\mathbf{x}_N^*(t_{k_N}), \mathbf{u}_N^*(t_{k_N}))\| = 0, \end{aligned}$$

where we used Equation (20). This contradicts Equation (38). Similarly, we can show that $\tilde{\mu}^\infty(t) \geq 0$, thus proving that $\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)$ and $\tilde{\mu}^\infty(t)$ satisfy Equation (14).

Furthermore, we notice that if $\mathbf{x}^\infty(t), \mathbf{u}^\infty(t), \tilde{\lambda}^\infty(t), \tilde{\mu}^\infty(t)$ and \tilde{v}^∞ satisfy Equations (21)–(23), then the following holds for all $k = 0, \dots, N$:

$$\begin{aligned} \|\tilde{\lambda}^{\infty\top}(0) + \tilde{v}^{\infty\top} \mathbf{e}_{x(0)}(\mathbf{x}^\infty(0), \mathbf{x}^\infty(1)) + E_{x(0)}(\mathbf{x}^\infty(0), \mathbf{x}^\infty(1))\| &= 0, \\ \|\lambda^{\infty\top}(1) - \tilde{v}^{\infty\top} \mathbf{e}_{x(1)}(\mathbf{x}^\infty(0), \mathbf{x}^\infty(1)) - E_{x(1)}(\mathbf{x}^\infty(0), \mathbf{x}^\infty(1))\| &= 0, \\ \left\| \int_0^1 \left[\dot{\lambda}^{\infty\top}(t) + F_x(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) + \tilde{\lambda}^{\infty\top}(t) f_x(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) \right. \right. \\ &\quad \left. \left. + \tilde{\mu}^{\infty\top}(t) \mathbf{h}_x(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) \right] b_{k,N}(t) dt \right\| = 0, \\ \left\| \int_0^1 \left[F_u(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) + \tilde{\lambda}^{\infty\top}(t) f_u(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) \right. \right. \\ &\quad \left. \left. + \tilde{\mu}^{\infty\top}(t) \mathbf{h}_u(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) \right] b_{k,N}(t) dt \right\| = 0. \end{aligned}$$

Since $\{b_{k,N}(t)\}_{k=0}^N$ is a linearly independent basis set, the last two equations above imply

$$\begin{aligned} \|\dot{\lambda}^{\infty\top}(t) + F_x(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) + \tilde{\lambda}^{\infty\top}(t) f_x(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) + \tilde{\mu}^{\infty\top}(t) \mathbf{h}_x(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t))\| &= 0, \\ \|F_u(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) + \tilde{\lambda}^{\infty\top}(t) f_u(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t)) + \tilde{\mu}^{\infty\top}(t) \mathbf{h}_u(\mathbf{x}^\infty(t), \mathbf{u}^\infty(t))\| &= 0, \end{aligned}$$

for all $t \in [0, 1]$. This proves that $\mathbf{x}^\infty(t), \mathbf{u}^\infty(t), \tilde{\lambda}^\infty(t), \tilde{\mu}^\infty(t)$ and $\tilde{v}^\infty(t)$ satisfy Equations (15)–(18). □

Theorem 3 (Covector Mapping Theorem). *Under the same assumptions of Theorems 1 and 2, when $N \rightarrow \infty$, the covector mapping*

$$\begin{aligned} \mathbf{x}_N^*(t) \mapsto \mathbf{x}^*(t), \quad \mathbf{u}_N^*(t) \mapsto \mathbf{u}^*(t), \\ \frac{\lambda_N^*(t)}{w} \mapsto \lambda^*(t), \quad \frac{\mu_N^*(t)}{w} \mapsto \mu^*(t), \quad \tilde{v}^* \mapsto v^* \end{aligned} \tag{39}$$

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

Proof. The above result follows directly from Theorems 1 and 2. In fact, if

$$\{\mathbf{x}^*(t), \mathbf{u}^*(t), \lambda^*(t), \mu^*(t), v^*\}$$

is a solution to Problem P_λ , which exists by Assumption 2, then from Theorem 1, it follows that $\{\mathbf{x}^*(t), \mathbf{u}^*(t), w\lambda^*(t), w\mu^*(t), v^*\}$ is a solution to Problem $P_{N\lambda}^{clos}$ (see Equations (26)–(28)). Conversely, by using Equation (37), a solution

$$(\mathbf{x}_N^*(t), \mathbf{u}_N^*(t), \boldsymbol{\lambda}_N^*(t), \boldsymbol{\mu}_N^*(t), \bar{\mathbf{v}}^*)$$

that solves Problem $P_{N\lambda}^{\text{clos}}$ provides a solution

$$(\mathbf{x}_N^*(t), \mathbf{u}_N^*(t), \frac{\boldsymbol{\lambda}_N^*(t)}{w}, \frac{\boldsymbol{\mu}_N^*(t)}{w}, \bar{\mathbf{v}}^*) = (\mathbf{x}_N^*(t), \mathbf{u}_N^*(t), \tilde{\boldsymbol{\lambda}}_N^*(t), \tilde{\boldsymbol{\mu}}_N^*(t), \bar{\mathbf{v}}^*)$$

that converges to a solution to Problem P_λ (see Theorem 2). \square

Remark 5. Define the Hamiltonian approximation

$$\mathcal{H}_N(t) = F(\mathbf{x}_N^*(t), \mathbf{u}_N^*(t)) + \frac{\boldsymbol{\lambda}_N^*(t)}{w} \mathbf{f}(\mathbf{x}_N^*(t), \mathbf{u}_N^*(t)),$$

then, Theorem 3 implies $\lim_{N \rightarrow \infty} \mathcal{H}_N(t) = \mathcal{H}(t)$.

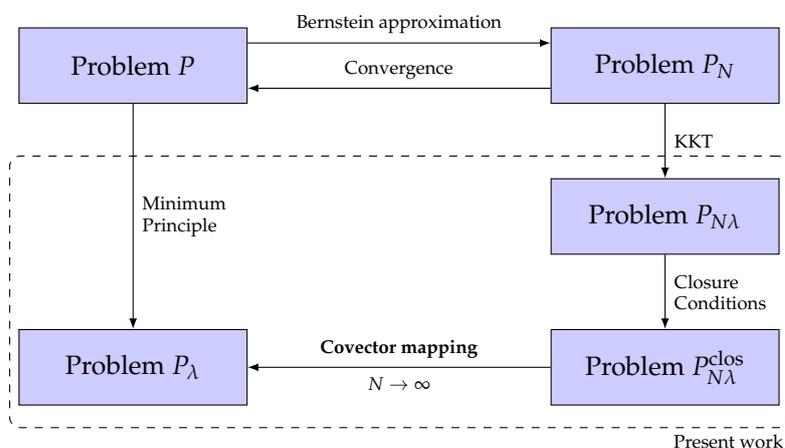


Figure 1. Diagram of the covector mapping principle for Bernstein approximation. The solution to Problem $P_{N\lambda}^{\text{clos}}$ converges to that of Problem P_λ as $N \rightarrow \infty$.

6. Numerical Examples

6.1. Example 1: 1D Minimum Time Problem

The first example we consider is the classical minimum time problem for a double integrator plant.

$$\min_u J = t_f,$$

subject to

$$\begin{aligned} \dot{x}_1 &= x_2, & \dot{x}_2 &= u, & x_1(0) &= 1, & x_2(0) &= -1 \\ x_1(t_f) &= x_2(t_f) &= 0, \\ |u(t)| &\leq 1, \forall t \in [0, t_f]. \end{aligned}$$

The analytical solution to this problem is well known: the optimal control input is bang-bang and the Hamiltonian along the optimal trajectories is equal to -1 [53]. Figure 2 includes plots of the state and control trajectories for this problem for $N = 45$, as well as a graph of the analytical solution. It is clear that the Bernstein polynomial solution captures the switching time precisely and closely approximates the optimal control input solution. Figure 3 includes the plot of the Hamiltonian approximation \mathcal{H}_N computed using Covector Mapping Theorem. As expected, it is equal to -1 with the exception of a small bump around the switching time.

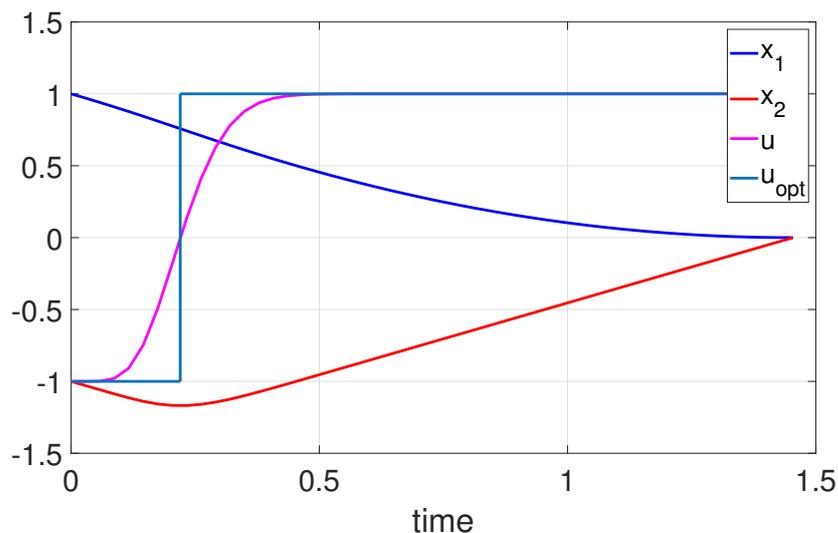


Figure 2. Example 1: state trajectories $x_1(t)$ and $x_2(t)$, and control input $u(t)$. The control input approximates the optimal solution (bang-bang).

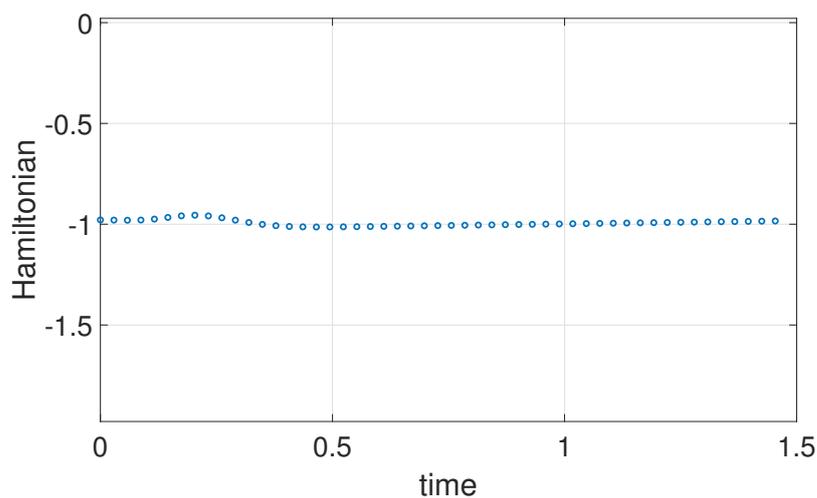


Figure 3. Example 1: the approximated Hamiltonian converges to the Hamiltonian of the problem, i.e., -1 . See Theorem 3 and Remark 5.

6.2. Example 2: 3D Minimum Time Problem

In this example, we consider a minimum-time problem for a simplified 3D model of a multi-rotor drone. The vehicle is asked to reach the origin in minimum time from a given initial condition with all the control inputs bounded by ± 1 . Unlike Example 1, there is no known analytical solution to this problem. However, we know that the Hamiltonian along the optimal trajectories is equal to -1 [53].

$$\min_u J = t_f,$$

subject to

$$\begin{aligned} \dot{x}_1 &= x_4, & \dot{x}_2 &= x_5, & \dot{x}_3 &= x_6, \\ \dot{x}_4 &= u_1, & \dot{x}_5 &= u_2, & \dot{x}_6 &= -g + u_3, \\ x_{10} &= 1 & x_{20} &= 2, & x_{30} &= 3, \\ x_{40} &= -1, & x_{50} &= -1, & x_{60} &= -1, \\ x_1(t_f) &= x_2(t_f) = x_3(t_f) = 0, \\ x_4(t_f) &= x_5(t_f) = x_6(t_f) = 0, \end{aligned}$$

$$|u_1(t)| \leq 1, \quad |u_2(t)| \leq 1, \quad |-g + u_3(t)| \leq 1, \quad \forall t \in [0, t_f].$$

Figure 4 shows the 3D plot of the position of the vehicle. The vehicle clearly reaches the origin from the given initial condition. Figure 5 includes graphs of the control inputs. They satisfy the ± 10 bound imposed in the problem formulation. Finally, Figure 6 shows the plot of the Hamiltonian approximation $\mathcal{H}_N, N = 45$. It is equal to -1 , as predicted by theory, thus indicating that the solution obtained is indeed close to optimal.

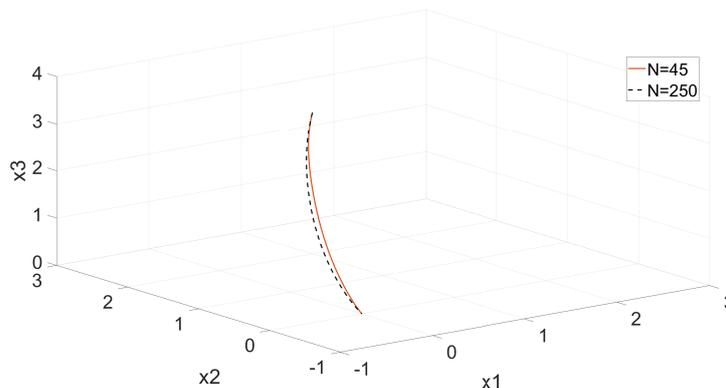


Figure 4. Example 2: 3D position plot, i.e., $p = [x_1, x_2, x_3]$. The solid line represents the solution obtained with $N = 45$, while the dashed line depicts the (near optimal) solution obtained with $N = 250$.

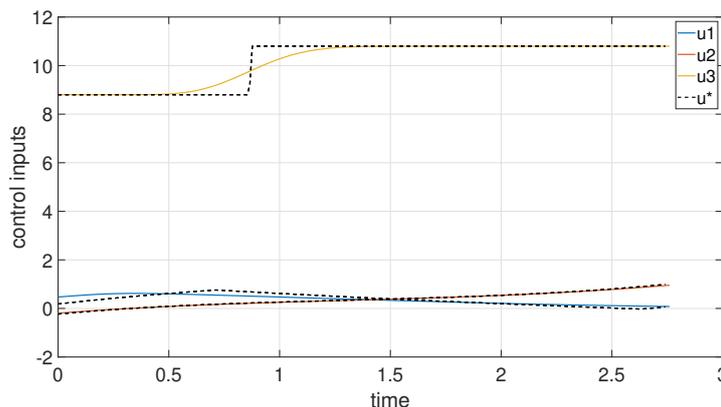


Figure 5. Example 2: Control inputs, i.e., vehicle’s acceleration along the three axis. The solid lines represent the solution obtained with $N = 45$, while the dashed lines depict the (near optimal) solution obtained with $N = 250$.

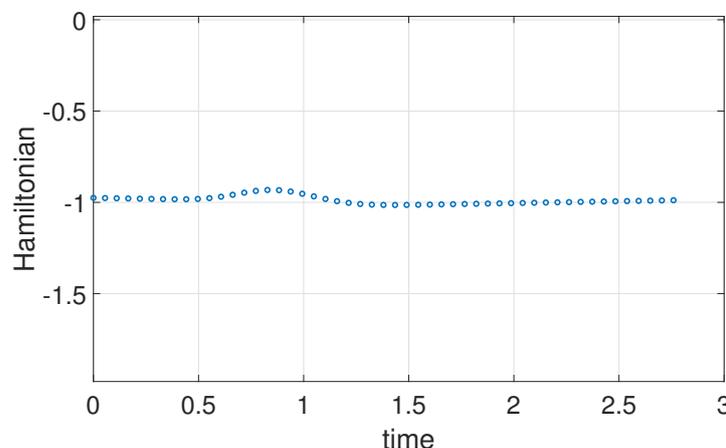


Figure 6. Example 2: the approximated Hamiltonian converges to the Hamiltonian of the problem, i.e., -1 . See Theorem 3 and Remark 5.

7. Defense against a Swarm Attack

The numerical analysis presented here involves a scenario in which an enemy swarm is attempting to destroy an high-value unit (HVU). The HVU is defended by a number of defending agents whose trajectories are optimized to maximize the probability of the HVU survival. The attacking agents dynamics are defined using Leonard swarm dynamics model [54]. A virtual leader is guiding the attacking swarm towards the HVU’s position. The attacking and defending agents are equipped with weapons systems which allow them to inflict damage on each other. The attacking agents inflict damage on the defending agents and try to destroy the HVU. The defending agents inflict damage on the attacking agents and attempt to destroy them or herd them away.

Attacking agent $i \in \{1, \dots, N_A\}$ has position $x_i(t) \in R^3$ and defending agent $k \in \{1, \dots, N_D\}$ has position $s_k(t) \in R^3$. The equations of motion for attacker i is

$$\ddot{x}_i = \sum_{j \neq i}^N \frac{f_I(x_{ij})}{\|x_{ij}\|} x_{ij} + \sum_{k=1}^M \frac{f_d(s_{ik})}{\|s_{ik}\|} s_{ik} + K \frac{h_i}{\|h_i\|} - b\dot{x}_i, \tag{40}$$

for $i = 1 \dots N$. There are four terms in this equation, representing: (1) attractive and repulsive forces $f_I(x_{ij})$ from other attacking agents j , where $x_{ij} = x_i - x_j$ is the distance between attackers i and j ; (2) a constant “virtual leader” force with magnitude K pulling them toward the HVU’s position, where $h_i = h - x_i$ and h is the position of the HVU; (3) purely repulsive forces $f_d(s_{ik})$ due to defending agents, where $s_{ik} = x_i - s_k$ is the distance between attacker i and defender k ; and (4) a damping force proportional to the \dot{x}_i .

For the mathematical forms of f_I and f_d , we have chosen the Leonard model [54], i.e., f_I and f_d can be written as gradients of a scalar potential functions that depends only on x_{ij} and s_{ik} . The force f_I is repulsive when $\|x_{ij}\| \leq d_0$, attractive when $d_0 < \|x_{ij}\| \leq d_1$ and zero when $\|x_{ij}\| > d_1$. For f_d , we only keep the repulsive term (since attackers should not be attracted to defenders), i.e., f_d is repulsive when $\|s_{ik}\| \leq s_0$ and zero when $\|s_{ik}\| > s_0$.

Defending agent i ’s dynamics are given by

$$\ddot{s}_i = u_i, \quad i = 1 \dots N_D, \quad s_i(t), u_i(t) \in R^3, \tag{41}$$

where the absolute value of each element of u_i , ($|u_{ij}|, j = 1, 2, 3$) is bounded by $u_{\max} = 1$.

Mutual attrition model: for hostile swarm engagements, agents are equipped with some weapons systems. The likelihood of destruction of an agent depends on its position (how close it has come to enemies) as well as the positions of those enemy agents, since each agent’s ability to inflict damage is contingent on its own survival. To model this mutual attrition, we use a damage function to track the probability that defender k is destroyed by a shot from attacker i , and vice versa. We choose a cumulative normal distribution function, Φ , to model the damage function [55]. Next, we define (i) the attrition rate at which attacker i is destroyed due to defender k , d_{ik}^{att} , (ii) the attrition rate of defender k due to attacker i , d_{ki}^{def} , and (iii) the attrition rate of the HVU, d_i^{hvu} , as follows:

$$d_{ik}^{\text{att}} = \lambda_d \Phi\left(\frac{\|s_{ik}\|^2}{\sigma_d}\right), \quad d_{ki}^{\text{def}} = \lambda_a \Phi\left(\frac{\|s_{ik}\|^2}{\sigma_a}\right), \quad d_i^{\text{hvu}} = \lambda_a \Phi\left(\frac{\|h_i\|^2}{\sigma_a}\right). \tag{42}$$

In the above equation σ_d is a defender Poisson parameter that corresponds to the range of the defenders’ weapons, λ_d is a defender Poisson parameter that corresponds to the defenders’ rate of fire, σ_a is an attacker Poisson parameter that corresponds to the attackers’ range and λ_a is an attacker Poisson parameter that corresponds to the attackers’ rate of fire. The probability of defender k destroying attacker i during a time interval of duration Δt is weighted by the current survival probability of defender k , $P_k^d(t)$. Thus, the probability that defender k will destroy attacker i during a given time interval $[t, t + \Delta t]$ is $P_k^d(t) d_{ik}^{\text{att}} \Delta t$. Assuming independence (i.e., defenders do not coordinate their fire), the expression $\prod_k^M (1 - [d_{ik}^{\text{att}} P_k^d(t)] \Delta t)$ represents the probability that i th attacker would survive

during a time interval $[t, t + \Delta t]$ due to attrition from all defenders. Therefore, the survival probability $Q_i(t + \Delta t)$ of attacker i is governed by

$$Q_i(t + \Delta t) = Q_i(t) \prod_k^{N_D} (1 - [d_{ik}^{\text{att}} P_k^d(t)] \Delta t), \quad (43)$$

where we assumed that probabilities of attacker i survival $Q_i(t_1)$, $Q_i(t_2)$ are independent for any t_1, t_2 . Similarly, the survival probability $P_k^d(t)$ of defender k and the survival probability $P(t)$ of the HVU are governed by

$$\begin{aligned} P_k^d(t + \Delta t) &= P_k^d(t) \prod_i^{N_A} (1 - [d_{ki}^{\text{def}} Q_i(t)] \Delta t), \\ P(t + \Delta t) &= P(t) \prod_k^{N_A} (1 - [d_k^{\text{hvu}} Q_k(t)] \Delta t). \end{aligned} \quad (44)$$

Initial conditions are set to $Q_i(0) = P(0) = P_k^d(0) = 1$ for all agents and the HVU.

Further rearranging Equations (43) and (44) and letting $\Delta t \rightarrow 0$, as derived in [56] we obtain:

$$\begin{aligned} \dot{Q}_i(t) &= -Q_i(t) \sum_k^{N_D} (1 - [d_{ik}^{\text{att}} P_k^d(t)]), \\ \dot{P}_k^d(t) &= -P_k^d(t) \sum_i^{N_A} (1 - [d_{ki}^{\text{def}} Q_i(t)]), \end{aligned} \quad (45)$$

$$\dot{P}(t) = -P(t) \sum_k^{N_A} (1 - [d_k^{\text{hvu}} Q_k(t)]). \quad (46)$$

The optimal control problem at hand can be expressed as Problem P by properly rescaling the time variable, i.e., $\tau = t/t_f$; see Equations (4)–(7). In particular, we seek to maximize the probability of HVU survival at the terminal time $t = t_f$ ($\tau = 1$), i.e., minimize $I = 1 - P(t_f)$. The system's state $x(t)$ includes the attacker and the defender positions and velocities, as well as probabilities of the attacker and defender survivals and the probability of the HVU survival:

$$\begin{aligned} x &= [x_1^T, \dots, x_{N_A}^T, (v_1^x)^T, \dots, (v_{N_A}^x)^T, s_1^T, \dots, s_{N_D}^T, \\ & (v_1^s)^T, \dots, (v_{N_D}^s)^T, Q_1, \dots, Q_{N_A}, P_1^d, \dots, P_{N_D}^d, P]^T, \end{aligned} \quad (47)$$

where v_i^x is the velocity of the i -th attacker and v_k^s is the velocity of the k -th defender. The control input vector is defined by stacking accelerations of each defender $u = [u_1^T, \dots, u_{N_D}^T]^T$. Using definitions of the system's state and control inputs the system dynamics function $f(\cdot, \cdot)$ is given by concatenation of Equations (40), (41), (45), and (46). Finally, the function $h(\cdot, \cdot)$ in our case becomes a function of defender control inputs only and includes $\pm u_{\max}$.

Figure 7 shows results for an optimization with one defender protecting an HVU against a swarm of five attackers for $N = 45$. The HVU is at the origin and the defender trajectory is color purple. The defender has a 50% larger weapons range ($\sigma_d/\sigma_a = 1.5$), as well as double the fire rate with respect to the attackers ($\lambda_d/\lambda_a = 2$). The defender initially herds the attackers on his right away from HVU than approaches the HVU and similarly herds the attackers to his left away from the HVU. Figure 8 illustrates the control inputs that drive the motion of the defender. Unlike minimum time problems, e.g., the previous example, in this case, the inequality constraints on the control input are never active. Figure 9 shows a sequence of the Hamiltonian approximations $\mathcal{H}_N, N = 5, \dots, 45$. The sequence clearly converges to zero, indicating that the final numerical solution for $N = 45$ is indeed a close approximation of the true optimal solution.

The reader is referred to [57,58] for additional numerical examples.

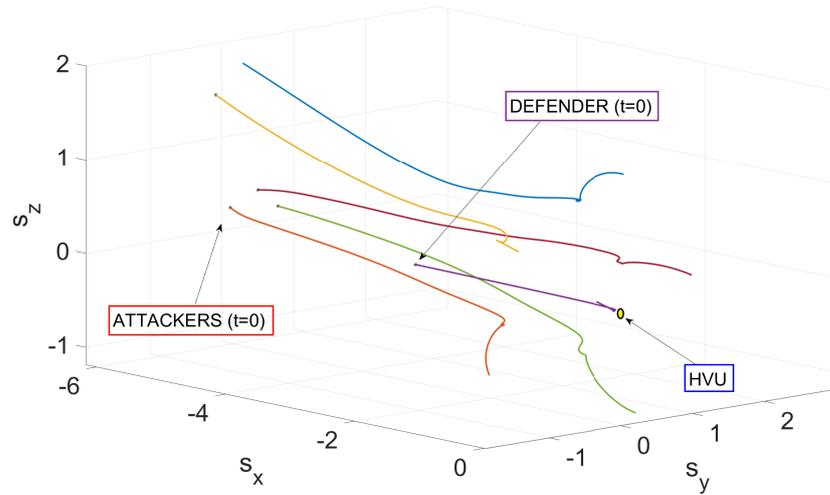


Figure 7. Defense against swarm attack. The plot shows optimal trajectory of one defender (purple) protecting a high-value unit (positioned at the origin) against five attackers.

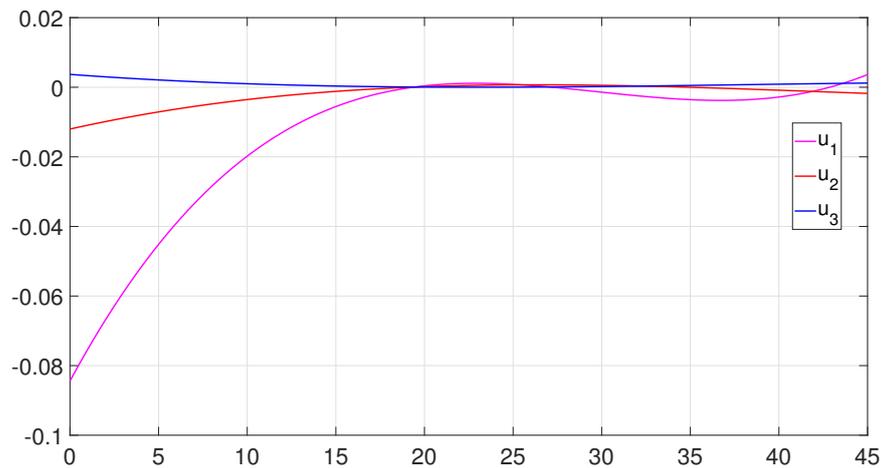


Figure 8. Defense against swarm attack. The plot shows the time history of the control input.

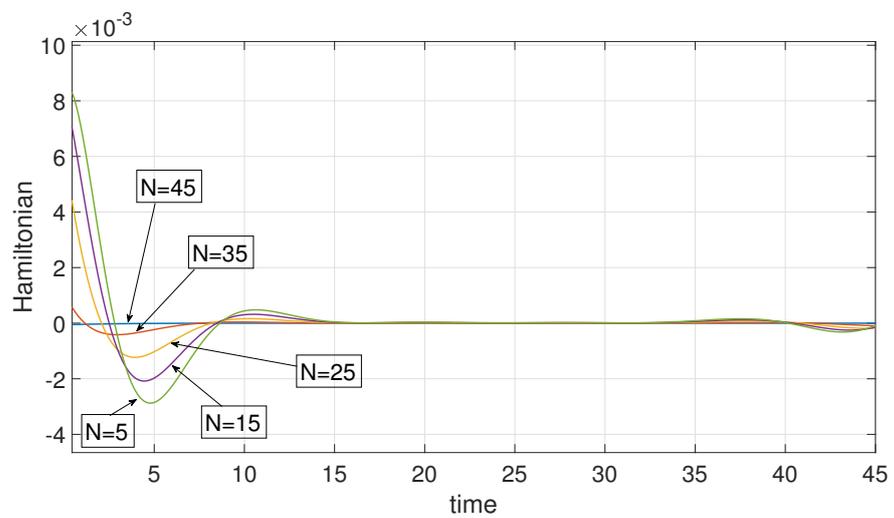


Figure 9. Hamiltonian Convergence. The plot shows a sequence of the Hamiltonian approximations $\mathcal{H}_N, N = 5, \dots, 45$, indicating that the numerical solution converges to the true optimal solution.

8. Conclusions

This paper proposed a numerical method for costate estimation of nonlinear constrained optimal control problems using Bernstein polynomials. A rigorous analysis is provided that shows convergence of the costate estimates to the dual variables of the continuous-time problem. To this end, a set of conditions are 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 approximation. The theoretical findings are validated through several numerical examples.

Author Contributions: Conceptualization, V.C., I.K., A.P., C.W.; methodology, V.C., I.K., A.P., C.W.; software, V.C., I.K.; validation, V.C., I.K., C.W.; investigation, V.C., I.K., A.P., C.W.; resources, V.C., I.K., A.P., N.H.; writing—original draft preparation, V.C., I.K.; writing—review and editing, V.C., I.K.; supervision, V.C., I.K., A.P., C.W., N.H.; project administration, V.C.; funding acquisition, V.C., I.K., A.P., N.H. All authors have read and agreed to the published version of the manuscript.

Funding: Venanzio Cichella was supported by Amazon, by the Office of Naval Research (grants N000141912106 and N000142112091), and by the National Science Foundation (grant 2136298). Antonio Pascoal was supported by H2020-EU.1.2.2—FET Proactive RAMONES (grant GA 101017808) and LARSyS-FCT (grant UIDB/50009/2020). Isaac Kaminer was supported by the Office of Naval Research (grants N0001421WX01974, N0001419WX00155, N0001422WX01906) and NPS CRUSER. Naira Hovakimyan was supported by NASA ULI (grant #80NSSC22M0070) and NSF RI (grant 2133656).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Equation (34)

Let us focus on Equation (34a). Adding and subtracting $\int_0^1 F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))b_{0,N}(t)dt$, we have

$$\begin{aligned} & \left\| w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) - \int_0^1 F_x(\mathbf{x}(t), \mathbf{u}(t))b_{0,N}(t)dt \right\| \\ & \leq \left\| w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) - \int_0^1 F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))b_{0,N}(t)dt \right. \\ & \quad \left. + \int_0^1 F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))b_{0,N}(t)dt - \int_0^1 F_x(\mathbf{x}(t), \mathbf{u}(t))b_{0,N}(t)dt \right\| \quad (\text{A1}) \\ & \leq \left\| w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) - \int_0^1 F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))b_{0,N}(t)dt \right\| \\ & \quad + \left\| \int_0^1 F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))b_{0,N}(t)dt - \int_0^1 F_x(\mathbf{x}(t), \mathbf{u}(t))b_{0,N}(t)dt \right\|. \end{aligned}$$

Using Lemma 3 and continuity of $F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))$ and $b_{0,N}(t)$, the first term on the right-hand side of the inequality above satisfies

$$\left\| w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j))b_{0,N}(t_j) - \int_0^1 F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))b_{0,N}(t)dt \right\| \leq C_I W_{F_x b_{0,N}}(N^{-\frac{1}{2}}),$$

where $W_{F_x b_{0,N}}(\cdot)$ is used to denote the modulus of continuity of the product

$$F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))b_{0,N}(t),$$

with $F_x(\mathbf{x}_N(t), \mathbf{u}_N(t))$ being a bounded function due to its continuity over a bounded domain. Denote its bound as $F_{x,\max}$. Notice that $b_{0,N}(t)$ is bounded, as $\max_{t \in [0,1]} b_{0,N}(t) \leq 1$. Then, using the properties of the modulus of continuity, we get

$$\begin{aligned} & \left\| w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j)) b_{0,N}(t_j) - \int_0^1 F_x(\mathbf{x}_N(t), \mathbf{u}_N(t)) b_{0,N}(t) dt \right\| \\ & \leq C_I F_{x,\max} W_{b_{0,N}}(N^{-\frac{1}{2}}) + C_I W_{F_x}(N^{-\frac{1}{2}}) \\ & \leq C_I F_{x,\max} N^{-\frac{1}{2}} + C_I W_{F_x}(N^{-\frac{1}{2}}), \end{aligned} \quad (\text{A2})$$

where $W_{F_x}(\cdot)$ is the modulus of continuity of F_x , and C_I is a positive constant independent of N . Furthermore, we have

$$\begin{aligned} & \left\| \int_0^1 F_x(\mathbf{x}_N(t), \mathbf{u}_N(t)) b_{0,N}(t) dt - \int_0^1 F_x(\mathbf{x}(t), \mathbf{u}(t)) b_{0,N}(t) dt \right\| \\ & \leq \int_0^1 \left\| F_x(\mathbf{x}_N(t), \mathbf{u}_N(t)) b_{0,N}(t) - F_x(\mathbf{x}(t), \mathbf{u}(t)) b_{0,N}(t) \right\| dt \\ & \leq L_{F_x} (C_x W_x(N^{-\frac{1}{2}}) + C_u W_u(N^{-\frac{1}{2}})), \end{aligned} \quad (\text{A3})$$

where L_{F_x} is the Lipschitz constant of F_x , $C_x < 5n_x/4$, $C_u < 5n_u/4$, and $W_x(\cdot)$ and $W_u(\cdot)$ are the moduli of continuity of \mathbf{x} and \mathbf{u} , respectively. Combining Equations (A2) and (A3) with Equation (A1) yields

$$\begin{aligned} & \left\| w \sum_{j=0}^N F_x(\mathbf{x}_N(t_j), \mathbf{u}_N(t_j)) b_{0,N}(t_j) - \int_0^1 F_x(\mathbf{x}(t), \mathbf{u}(t)) b_{0,N}(t) dt \right\| \\ & \leq C_I F_{x,\max} N^{-\frac{1}{2}} + C_I W_{F_x}(N^{-\frac{1}{2}}) + L_{F_x} (C_x W_x(N^{-\frac{1}{2}}) + C_u W_u(N^{-\frac{1}{2}})), \end{aligned}$$

which proves the bound in Equation (34a). The bounds in Equation (34b–d) follow easily using an identical argument.

References

- Ng, J.; Bräunl, T. Performance comparison of bug navigation algorithms. *J. Intell. Robot. Syst.* **2007**, *50*, 73–84. [\[CrossRef\]](#)
- Khatib, O. Real-time obstacle avoidance for manipulators and mobile robots. In *Autonomous Robot Vehicles*; Springer: Berlin/Heidelberg, Germany, 1986; pp. 396–404.
- Siegwart, R.; Nourbakhsh, I.R.; Scaramuzza, D. *Introduction to Autonomous Mobile Robots*; MIT Press: Cambridge, MA, USA, 2011.
- Choset, H.M. *Principles of Robot Motion: Theory, Algorithms, and Implementation*; MIT Press: Cambridge, MA, USA, 2005.
- Latombe, J.C. *Robot Motion Planning*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; Volume 124.
- Betts, J.T. Survey of numerical methods for trajectory optimization. *J. Guid. Control Dyn.* **1998**, *21*, 193–207. [\[CrossRef\]](#)
- Von Stryk, O.; Bulirsch, R. Direct and indirect methods for trajectory optimization. *Ann. Oper. Res.* **1992**, *37*, 357–373. [\[CrossRef\]](#)
- LaValle, S.M. *Planning Algorithms*; Cambridge University Press: Cambridge, UK, 2006.
- Cichella, V. Cooperative Autonomous Systems: Motion Planning and Coordinated Tracking Control for Multi-Vehicle Missions. Ph.D. Thesis. University of Illinois at Urbana-Champaign, Champaign, IL, USA, 2018.
- Cichella, V.; Kaminer, I.; Dobrokhodov, V.; Xargay, E.; Choe, R.; Hovakimyan, N.; Aguiar, A.P.; Pascoal, A.M. Cooperative Path-Following of Multiple Multirotors over Time-Varying Networks. *IEEE Trans. Autom. Sci. Eng.* **2015**, *12*, 945–957. [\[CrossRef\]](#)
- Sun, X.; Cassandras, C.G. Optimal dynamic formation control of multi-agent systems in constrained environments. *Automatica* **2016**, *73*, 169–179. [\[CrossRef\]](#)
- Walton, C.; Kaminer, I.; Gong, Q.; Clark, A.H.; Tsatsanifos, T. Defense against adversarial swarms with parameter uncertainty. *Sensors* **2022**, *22*, 4773. [\[CrossRef\]](#)
- Rao, A.V. A survey of numerical methods for optimal control. *Adv. Astronaut. Sci.* **2009**, *135*, 497–528.
- Betts, J.T. *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*; SIAM: Philadelphia, PA, USA, 2010.
- Conway, B.A. A survey of methods available for the numerical optimization of continuous dynamic systems. *J. Optim. Theory Appl.* **2012**, *152*, 271–306. [\[CrossRef\]](#)
- Becerra, V.M. Solving complex optimal control problems at no cost with PSOPT. In Proceedings of the 2010 IEEE International Symposium on Computer-Aided Control System Design, Yokohama, Japan, 8–10 September 2010; pp. 1391–1396.
- Febbo, H.; Jayakumar, P.; Stein, J.L.; Ersal, T. NLOptControl: A modeling language for solving optimal control problems. *arXiv* **2020**, arXiv:2003.00142.

18. Patterson, M.A.; Rao, A.V. GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using hp-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming. *ACM Trans. Math. Softw. (TOMS)* **2014**, *41*, 1–37. [[CrossRef](#)]
19. Rutquist, P.E.; Edvall, M.M. Propt—Matlab Optimal Control Software. Tomlab Optimization Inc.: Washington, DC, USA, 2010.
20. Ross, I.M. Enhancements to the DIDO Optimal Control Toolbox. *arXiv* **2020**. arXiv:2004.13112.
21. Andersson, J.A.; Gillis, J.; Horn, G.; Rawlings, J.B.; Diehl, M. CasADi: A software framework for nonlinear optimization and optimal control. *Math. Program. Comput.* **2019**, *11*, 1–36. [[CrossRef](#)]
22. Fahroo, F.; Ross, I.M. On discrete-time optimality conditions for pseudospectral methods. In Proceedings of the AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Keystone, CO, USA, 21–24 August 2006; p. 6304.
23. Bollino, K.; Lewis, L.R.; Sekhavat, P.; Ross, I.M. Pseudospectral optimal control: A clear road for autonomous intelligent path planning. In Proceedings of the AIAA Infotech@ Aerospace 2007 Conference and Exhibit, Rohnert Park, CA, USA, 7–10 May 2007; p. 2831.
24. Gong, Q.; Lewis, R.; Ross, M. Pseudospectral motion planning for autonomous vehicles. *J. Guid. Control. Dyn.* **2009**, *32*, 1039–1045. [[CrossRef](#)]
25. Bedrossian, N.S.; Bhatt, S.; Kang, W.; Ross, I.M. Zero-propellant maneuver guidance. *IEEE Control. Syst.* **2009**, *29*.
26. Bollino, K.; Lewis, L.R. Collision-free multi-UAV optimal path planning and cooperative control for tactical applications. In Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit, Honolulu, HI, USA, 18–21 August 2008; p. 7134.
27. Bedrossian, N.; Bhatt, S.; Lammers, M.; Nguyen, L.; Zhang, Y. First Ever Flight Demonstration of Zero Propellant Maneuver (TM) Attitude Control Concept. In Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit, Hilton Head, SC, USA, 20–23 August 2007; p. 76734.
28. Ross, I.M.; Karpenko, M. A review of pseudospectral optimal control: From theory to flight. *Annu. Rev. Control.* **2012**, *36*, 182–197. [[CrossRef](#)]
29. Polak, E. *Optimization: Algorithms and Consistent Approximations*; Springer: Berlin, Germany, 1997.
30. Fahroo, F.; Ross, I.M. Costate estimation by a Legendre pseudospectral method. *J. Guid. Control Dyn.* **2001**, *24*, 270–277. [[CrossRef](#)]
31. Darby, C.L.; Garg, D.; Rao, A.V. Costate estimation using multiple-interval pseudospectral methods. *J. Spacecr. Rocket.* **2011**, *48*, 856–866. [[CrossRef](#)]
32. Hager, W.W. Runge-Kutta methods in optimal control and the transformed adjoint system. *Numer. Math.* **2000**, *87*, 247–282. [[CrossRef](#)]
33. Grimm, W.; Markl, A. Adjoint estimation from a direct multiple shooting method. *J. Optim. Theory Appl.* **1997**, *92*, 263–283. [[CrossRef](#)]
34. Cichella, V.; Kaminer, I.; Walton, C.; Hovakimyan, N. Optimal Motion Planning for Differentially Flat Systems Using Bernstein Approximation. *IEEE Control Syst. Lett.* **2018**, *2*, 181–186. [[CrossRef](#)]
35. Kielas-Jensen, C.; Cichella, V. BeBOT: Bernstein polynomial toolkit for trajectory generation. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Venetian Macao, Macau, 3–8 November 2019; pp. 3288–3293.
36. Kielas-Jensen, C.; Cichella, V.; Berry, T.; Kaminer, I.; Walton, C.; Pascoal, A. Bernstein Polynomial-Based Method for Solving Optimal Trajectory Generation Problems. *Sensors* **2022**, *22*, 1869. [[CrossRef](#)] [[PubMed](#)]
37. Ricciardi, L.A.; Vasile, M. Direct transcription of optimal control problems with finite elements on Bernstein basis. *J. Guid. Control Dyn.* **2018**. [[CrossRef](#)]
38. Choe, R.; Puig-Navarro, J.; Cichella, V.; Xargay, E.; Hovakimyan, N. Cooperative Trajectory Generation Using Pythagorean Hodograph Bézier Curves. *J. Guid. Control Dyn.* **2016**, *39*, 1744–1763. [[CrossRef](#)]
39. Ghomanjani, F.; Farahi, M.H. Optimal control of switched systems based on Bezier control points. *Int. J. Intell. Syst. Appl.* **2012**, *4*, 16. [[CrossRef](#)]
40. Huo, M.; Yang, L.; Peng, N.; Zhao, C.; Feng, W.; Yu, Z.; Qi, N. Fast costate estimation for indirect trajectory optimization using Bezier-curve-based shaping approach. *Aerosp. Sci. Technol.* **2022**, *126*, 107582. [[CrossRef](#)]
41. Zhao, Z.; Kumar, M. Split-bernstein approach to chance-constrained optimal control. *J. Guid. Control Dyn.* **2017**, *40*, 2782–2795. [[CrossRef](#)]
42. Cichella, V.; Kaminer, I.; Walton, C.; Hovakimyan, N.; Pascoal, A.M. Optimal Multi-Vehicle Motion Planning using Bernstein Approximants. *IEEE Trans. Autom. Control* **2020**.. [[CrossRef](#)]
43. Schwartz, A.; Polak, E. Consistent approximations for optimal control problems based on Runge–Kutta integration. *SIAM J. Control Optim.* **1996**, *34*, 1235–1269. [[CrossRef](#)]
44. Bojanic, R.; Cheng, F. Rate of convergence of Bernstein polynomials for functions with derivatives of bounded variation. *J. Math. Anal. Appl.* **1989**, *141*, 136–151. [[CrossRef](#)]
45. Popoviciu, T. Sur l’approximation des fonctions convexes d’ordre supérieur. *Mathematica* **1935**, *10*, 49–54.
46. Sikkema, P. Der wert einiger konstanten in der theorie der approximation mit Bernstein-Polynomen. *Numer. Math.* **1961**, *3*, 107–116. [[CrossRef](#)]
47. Powell, M.J.D. *Approximation Theory and Methods*; Cambridge University Press: Cambridge, UK, 1981.
48. Floater, M.S. On the convergence of derivatives of Bernstein approximation. *J. Approx. Theory* **2005**, *134*, 130–135. [[CrossRef](#)]
49. Hartl, R.F.; Sethi, S.P.; Vickson, R.G. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev.* **1995**, *37*, 181–218. [[CrossRef](#)]

50. Garg, D.; Patterson, M.A.; Francolin, C.; Darby, C.L.; Huntington, G.T.; Hager, W.W.; Rao, A.V. Direct trajectory optimization and costate estimation of finite-horizon and infinite-horizon optimal control problems using a Radau pseudospectral method. *Comput. Optim. Appl.* **2011**, *49*, 335–358. [[CrossRef](#)]
51. Gong, Q.; Ross, I.M.; Kang, W.; Fahroo, F. Connections between the covector mapping theorem and convergence of pseudospectral methods for optimal control. *Comput. Optim. Appl.* **2008**, *41*, 307–335. [[CrossRef](#)]
52. Singh, B.; Bhattacharya, R.; Vadali, S.R. Verification of optimality and costate estimation using Hilbert space projection. *J. Guid. Control Dyn.* **2009**, *32*, 1345–1355. [[CrossRef](#)]
53. Kirk. *Optimal Control Theory: An Introduction*; Prentice-Hall: Hoboken, NJ, USA, 1970.
54. Ogren, P.; Fiorelli, E.; Leonard, N.E. Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment. *IEEE Trans. Autom. Control* **2004**, *49*, 1292–1302. [[CrossRef](#)]
55. Washburn, A.; Kress, M. *Combat Modeling*; Springer: Berlin/Heidelberg, Germany, 2009.
56. Walton, C.; Lambrianides, P.; Kaminer, I.; Royset, J.; Gong, Q. Optimal motion planning in rapid-fire combat situations with attacker uncertainty. *Naval Res. Logist.* **2018**, *65*, 101–119. [[CrossRef](#)]
57. Cichella, V.; Kaminer, I.; Walton, C.; Hovakimyan, N.; Pascoal, A. Bernstein approximation of optimal control problems. *arXiv* **2018**, arXiv:1812.06132.
58. Cichella, V.; Kaminer, I.; Walton, C.; Hovakimyan, N.; Pascoal, A.M. Consistent approximation of optimal control problems using Bernstein polynomials. In Proceedings of the 2019 IEEE 58th Conference on Decision and Control (CDC), Nice, France, 11–13 December 2019; pp. 4292–4297.