Title

On the convergence of nonlinear optimal control using pseudospectral methods for feedback linearizable systems

Journal
International Journal of Robust and Nonlinear Control, 17(14)

ISSN
1049-8923

Authors
Kang, W
Gong, Q
Ross, IM
et al.

Publication Date
2007-09-25

DOI
10.1002/rnc.1166

Peer reviewed

INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
Int. J. Robust Nonlinear Control 2007; **17**:1251–1277
Published online 3 January 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/rnc.1166

On the convergence of nonlinear optimal control using pseudospectral methods for feedback linearizable systems[‡]

W. Kang[1,*,†], Q. Gong[2], I. M. Ross[2] and F. Fahroo[1]

[1] *Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93943, U.S.A.*
[2] *Department of Mechanical and Astronautical Engineering, Naval Postgraduate School, Monterey, CA 93943, U.S.A.*

SUMMARY

We consider the optimal control of feedback linearizable dynamic systems subject to mixed state and control constraints. In contrast to the existing results, the optimal controller addressed in this paper is allowed to be discontinuous. This generalization requires a substantial modification to the existing convergence analysis in terms of both the framework as well as the notion of convergence around points of discontinuity. Although the nonlinear system is assumed to be feedback linearizable, the optimal control does not necessarily linearize the dynamics. Such problems frequently arise in astronautical applications where stringent performance requirements demand optimality over feedback linearizing controls. We prove that a sequence of solutions obtained using the Legendre pseudospectral method converges to the optimal solution of the continuous-time problem under mild conditions. Published in 2007 by John Wiley & Sons, Ltd.

Received 13 February 2006; Revised 10 October 2006; Accepted 13 November 2006

KEY WORDS: pseudospectral methods; optimal control; feedback linearizable systems

1. INTRODUCTION

As a result of significant progress in large-scale computational algorithms and nonlinear programming, the so-called direct computational methods have become the industry standard for solving nonlinear optimal control problems [1, 2], particularly in aerospace applications [3, 4].

*Correspondence to: W. Kang, Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93943, U.S.A.
† E-mail: wkang@nps.edu

Contract/grant sponsor: Air Force Office of Scientific Research; contract/grant number: F1ATA0-60-6-2G002
Contract/grant sponsor: NPS, the Secretary of the Air Force

Published in 2007 by John Wiley & Sons, Ltd.

In simple terms, in a direct method, the continuous-time problem of optimal control is discretized, and the resulting discretized optimization problem is solved by nonlinear programming algorithms. In mathematical terms, this approach can be categorized as numerical functional analysis which implies that a number of issues pertaining to convergence of the approximation and the convergence of the algorithm need to be addressed [5]. This paper addresses the former issue, and in particular, the convergence of the approximations under pseudospectral (PS) discretization.

In the 1990s, PS methods were introduced for solving general nonlinear optimal control problems with constraints [6–9]; and since then, have gained considerable attention [3, 4, 10–15]. Over the last decade, the PS methods have been used to solve a broad class of industrial-strength optimal control problems, for instance, low-thrust orbit transfers [10], impulsive orbit transfers [14], ascent guidance [12, 13], etc. As a result of its successful applications at NASA, the latest version of the OTIS software package [16] has the Legendre PS method as a problem solving option. Further details on NASA's plans are described at: http://trajectory.grc.nasa.gov/projects/lowthrust.shtml. In addition, the commercially available software package, DIDO [17], uses PS methods exclusively for solving optimal control problems.

The popularity of PS methods motivates us to study a number of fundamental problems such as feasibility, convergence, and the rate of convergence. It has been proved [18] that PS methods offer a convergence rate that is faster than any polynomial rate for the approximation of analytic functions. This property can also be numerically demonstrated with regards to PS methods for control [9, 19]. Furthermore, PS methods provide Eulerian-like simplicity; thus, for a given error bound, a PS method generates a significantly smaller-scale optimization problem when compared to the traditional discretization methods, such as Euler and Runge–Kutta. This property is particularly attractive for control applications as it places real-time computation within easy reach of modern computational power [20–22]. It has also been demonstrated that PS methods also offer a ready approach to exploiting differential-geometric properties of a control system such as convexity and differential flatness [19, 20]. Despite its versatility and simplicity, a PS approach masks a wide range of deeply theoretical issues that lie at the intersection of approximation theory and control theory. For example, does the discretized problem always have a feasible solution? Does the discretized optimal solution converge to the continuous-time optimal solution? The answers to these fundamental questions are yet to be found because this is a relatively young area of research and many problems are still widely open. Nonetheless, some notable results have been proved. For instance, in [23, 24] a detailed relationship between the necessary conditions of the continuous-time optimal control problem and the Karush–Kuhn–Tucker (KKT) condition of the discrete optimization problem is revealed. In [25], the feasibility of the PS discretization is proved with relaxed inequality constraints for fixed relaxation margins. In [26], the feasibility and convergence results are proved for feedback linearizable nonlinear systems. In this paper, the relaxation margin of the constraints approaches zero and the convergence theorem is proved in a way similar to the theory of consistent approximations [27].

In [26], the feasibility of the PS discretization and a set of sufficient conditions for the convergence of the approximated optimal control are proved based on a key assumption that the optimal controller is at least continuous. Unfortunately, for many optimal control problems this assumption is not valid, especially when the control input is constrained. In this situation, the optimal controller is likely to be discontinuous, such as a bang-bang control. In this paper we extend the results in [26] to a more general case that includes discontinuous optimal controls.

Due to the lack of smoothness in the optimal control, the proof is much more involved than the one given in [26]. It is well known that an analysis of discontinuous controllers is a very challenging problem [28]. As far as the discretization is concerned, the existence of discontinuity in controllers raises fundamental problems in approximation theory. In this paper, we prove feasibility and the convergence results for the Legendre PS method when the controls are discontinuous and the dynamics is in a feedback linearizable form. We assume the dynamic system can be written in a normal form. It permits a modification of the standard PS method [7, 24] in a manner that is similar to dynamic inversion. That is, we seek polynomial approximations of the state trajectories while the controls are determined by an exact satisfaction of dynamics. This modification of a PS method permits us to prove sufficient conditions for the feasibility and convergence of the PS discretizations of discontinuous controllers. Furthermore, our method allows one to easily incorporate state and control constraints including mixed state and control constraints. Note that we do not linearize the dynamics by feedback control; rather, we find the optimal control for a generic cost function and this optimal control is not necessarily smooth. Such problems are particularly common in astronautical applications where stringent performance requirements demand that the control be optimal rather than feasible as implied by the linearizing control. We show that, under mild conditions, the PS discretized optimization problem always has a feasible solution even for discontinuous control input. Furthermore, sufficient conditions are derived for the numerical solution to converge to the solution of the original continuous-time constrained optimal control problem.

The paper is organized as follows: in Section 2, we briefly present the PS discretization method for constrained nonlinear optimal control problems. Sections 3 and 4 contain the results regarding feasibility and convergence of the discretized problem. In Section 5, the results are generalized to optimal control problems with a free final time. As an example, we apply the PS methods to a minimum time orbit transfer problem in Section 6.

2. THE PROBLEM AND ITS DISCRETIZATION

We consider the following mixed, state and control constrained nonlinear Bolza problem (Problem B) with feedback linearizable dynamics.

Problem B
Minimize

$$J[x(\cdot), u(\cdot)] = \int_{-1}^{1} F(x(t), u(t))\,\mathrm{d}t + E(x(-1), x(1)) \tag{1}$$

subject to the dynamics

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ &\vdots \\ \dot{x}_{r-1}(t) &= x_r(t) \\ \dot{x}_r(t) &= f(x(t)) + g(x(t))u(t) \end{aligned} \tag{2}$$

mixed path constraints

$$h(x(t), u(t)) \leqslant 0 \tag{3}$$

and endpoint conditions

$$e(x(-1), x(1)) = 0 \tag{4}$$

where $x \in \mathbb{R}^r$, $u \in \mathbb{R}$, and $F : \mathbb{R}^r \times \mathbb{R} \to \mathbb{R}$, $E : \mathbb{R}^r \times \mathbb{R}^r \to \mathbb{R}$, $f : \mathbb{R}^r \to \mathbb{R}$, $g : \mathbb{R}^r \to \mathbb{R}, e : \mathbb{R}^r \times \mathbb{R}^r \to \mathbb{R}^{N_e}$ and $h : \mathbb{R}^r \times \mathbb{R}^r \to \mathbb{R}^l$ are Lipschitz continuous (over the domain) with respect to their arguments. We assume $x(t)$ is absolutely continuous and $u(t)$ is L^1. For inversion reasons, we assume $g(x) \neq 0$ for all x. We also assume that an optimal solution, $(x^*(t), u^*(t))$, of Problem B exists. At several places in this paper, we use the norm $\|s(t)\|_\infty$ for a vector-valued function $s(t)$. If $s(t) \in \mathbb{R}^r$, then $\|s(t)\|_\infty$ is defined to be the maximum of $\|s_i(t)\|_\infty$ for $i = 1, 2, \ldots, r$.

Remark 2.1

Pseudospectral methods are not limited to dynamical systems in normal form; in fact, they are applicable to far more general nonlinear systems; see, for example, [24, 25, 29] and the references therein. What the normal form facilitates is the theoretical proof of the feasibility and convergence and the computational efficiency as illustrated in [20].

Remark 2.2

In Problem B, we assume the time interval to be fixed at $[-1, 1]$ in order to facilitate a simpler bookkeeping in using the Legendre PS method whose computational domain is $[-1, 1]$. If the physical time domain of the problem is not $[-1, 1]$, it can always be projected to the computational domain $[-1, 1]$ by a simple linear transformation [30].

Next, we apply the PS method to discretize the continuous-time optimal control Problem B. We focus on the Legendre PS method for the purpose of brevity; the extension to other PS methods is straight forward. The basic idea of Legendre PS method is to approximate $(x_1(t), \ldots, x_r(t))$ by Nth order polynomials $(x_1^N(t), \ldots, x_r^N(t))$ based on Lagrange interpolation of their values at the Legendre–Gauss–Lobatto (LGL) node points. Let $t_0 = -1 < t_1 < \cdots < t_N = 1$ be the LGL nodes defined as,

$$\begin{gathered} t_0 = -1, \quad t_N = 1, \text{ and} \\ \text{for } k = 1, 2, \ldots, N-1, t_k \text{ are the roots of } \dot{L}_N(t) \end{gathered}$$

where $\dot{L}_N(t)$ is the derivative of the Nth order Legendre polynomial $L_N(t)$. The distribution of the LGL nodes is illustrated in Figure 1. Note that the node distribution is not uniform. The high density of nodes near the end points is one of the key properties of PS discretizations in that it effectively prevents the Runge phenomenon. Computational advantages of such non-uniformly distributed quadrature nodes can be found in [18, 30, 31].

Let the pair, ($\bar{x}_k^N$ and $\bar{u}_k^N$), be an approximation of a feasible solution $(x(t), u(t))$ evaluated at the node t_k. Then, $x^N(t)$ is used to approximate $x(t)$ by

$$x(t) \approx x^N(t) = \sum_{k=0}^{N} \bar{x}_k^N \phi_k(t) \tag{5}$$

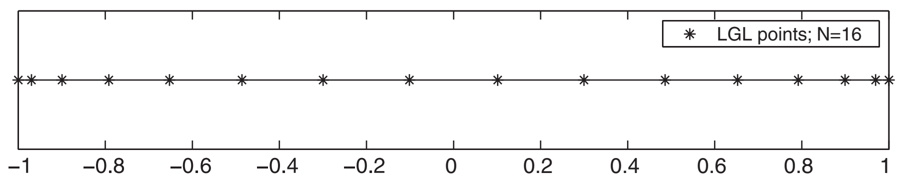

Figure 1. Distribution of LGL nodes.

where $\phi_k(t)$ is the Lagrange interpolating polynomial given by

$$\phi_k(t) = \frac{1}{N(N+1)L_N(t_k)} \frac{(t^2-1)\dot{L}_N(t)}{t-t_k} \tag{6}$$

From its definition (see [18]), $\phi_k(t_j)=1$, if $k=j$ and $\phi_k(t_j)=0$, if $k \neq j$. The precise nature of the approximation indicated in (5) is the main focus of this paper. From (2), the control that generates the approximate state is defined by

$$u^N(t) = \frac{\dot{x}_r^N(t) - f(x^N(t))}{g(x^N(t))} \tag{7}$$

Note that $u^N(t)$ is not necessarily a polynomial and hence differs from a standard PS approximation. It is a known result [18] that the derivative of $x_i^N(t)$ at the LGL node t_k satisfies

$$\dot{x}_i^N(t_k) = \sum_{j=0}^{N} D_{kj} x_i^N(t_j), \quad i = 1, 2, \ldots, r$$

where the $(N+1) \times (N+1)$ differentiation matrix D is defined by

$$D_{ik} = \begin{cases} \dfrac{L_N(t_i)}{L_N(t_k)} \dfrac{1}{t_i - t_k} & \text{if } i \neq k \\ -\dfrac{N(N+1)}{4} & \text{if } i = k = 0 \\ \dfrac{N(N+1)}{4} & \text{if } i = k = N \\ 0 & \text{otherwise} \end{cases}$$

Throughout the paper, we use the 'bar' notation to denote corresponding variables in the discrete space, and the superscript N to denote the number of nodes used in discretization. Thus, let

$$\bar{x}_0^N = \begin{pmatrix} \bar{x}_{10}^N \\ \vdots \\ \bar{x}_{r0}^N \end{pmatrix}, \ldots, \bar{x}_N^N = \begin{pmatrix} \bar{x}_{1N}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix}$$

Note that the subscript in $\bar{x}_k^N \in \mathbb{R}^r$ denotes an evaluation of the approximate state, $x^N(t) \in \mathbb{R}^r$, at the node t_k whereas $x_k(t)$ denotes the kth component of the exact state.

With these preliminaries, it is apparent that the approximate solutions must satisfy the following nonlinear algebraic equations:

$$D\begin{pmatrix} \bar{x}_{i0}^N \\ \vdots \\ \bar{x}_{iN}^N \end{pmatrix} = \begin{pmatrix} \bar{x}_{i+1,0}^N \\ \vdots \\ \bar{x}_{i+1,N}^N \end{pmatrix}, \quad i = 1, 2, \ldots, r-1 \tag{8}$$

$$D\begin{pmatrix} \bar{x}_{r0}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix} = \begin{pmatrix} f(\bar{x}_0^N) + g(\bar{x}_0^N)\bar{u}_0^N \\ \vdots \\ f(\bar{x}_N^N) + g(\bar{x}_N^N)\bar{u}_N^N \end{pmatrix}$$

for feasibility with respect to the dynamics. In a standard PS method, it is quite common [7, 19, 20, 23] to discretize the mixed state and control constraints as

$$h(\bar{x}_k^N, \bar{u}_k^N) \leqslant 0, \quad k = 0, 1, \ldots, N \tag{9}$$

Here, we propose the following relaxation:

$$h(\bar{x}_k^N, \bar{u}_k^N) \leqslant (N-r)^{-1/4} \cdot \mathbf{1}, \quad k = 0, 1, \ldots, N \tag{10}$$

where $\mathbf{1}$ denotes $[1, \ldots, 1]^{\mathrm{T}}$. When N tends to infinity, the difference between conditions (9) and (10) vanishes. Similarly, we relax the endpoint condition $e(x(-1), x(1)) = 0$, to an inequality, i.e.

$$\|e(\bar{x}_0^N, \bar{x}_N^N)\|_\infty \leqslant (N-r)^{-1/4} \tag{11}$$

The relaxation is necessary because it is impossible to numerically implement an equality constraint. In addition, examples can be found in which a discretization without relaxation is infeasible.

Finally, the cost functional $J[x(\cdot), u(\cdot)]$ is approximated by the Gauss–Lobatto integration rule,

$$J[x(\cdot), u(\cdot)] \approx \bar{J}^N(\bar{X}, \bar{U}) = \sum_{k=0}^{N} F(\bar{x}_k^N, \bar{u}_k^N) w_k + E(\bar{x}_0^N, \bar{x}_N^N)$$

where w_k are the LGL weights defined by

$$w_k = \frac{2}{N(N+1)} \frac{1}{[L_N(t_k)]^2}$$

and $\bar{X} = [\bar{x}_0^N, \ldots, \bar{x}_N^N]$, $\bar{U} = [\bar{u}_0^N, \ldots, \bar{u}_N^N]$. Hence, the optimal control Problem B can be approximated by a nonlinear programming problem with $\bar{J}^N$ as the objective function and (8), (10) and (11) as constraints; this is summarized below.

Problem B^N

Find $\bar{x}_k^N \in \mathbb{R}^r$ and $\bar{u}_k^N \in \mathbb{R}$, $k = 0, 1, \ldots, N$, that minimize

$$\bar{J}^N(\bar{X}, \bar{U}) = \sum_{k=0}^{N} F(\bar{x}_k^N, \bar{u}_k^N) w_k + E(\bar{x}_0^N, \bar{x}_N^N) \tag{12}$$

DOI: 10.1002/rnc

subject to

$$D\begin{pmatrix}\bar{x}_{i0}^N\\ \vdots\\ \bar{x}_{iN}^N\end{pmatrix}=\begin{pmatrix}\bar{x}_{i+1,0}^N\\ \vdots\\ \bar{x}_{i+1,N}^N\end{pmatrix},\quad i=1,2,\ldots,r-1 \tag{13}$$

$$D\begin{pmatrix}\bar{x}_{r0}^N\\ \vdots\\ \bar{x}_{rN}^N\end{pmatrix}=\begin{pmatrix}f(\bar{x}_0^N)+g(\bar{x}_0^N)\bar{u}_0^N\\ \vdots\\ f(\bar{x}_N^N)+g(\bar{x}_N^N)\bar{u}_N^N\end{pmatrix}$$

$$h(\bar{x}_k^N,\bar{u}_k^N)\leqslant(N-r)^{-1/4}\cdot\mathbf{1} \tag{14}$$

$$\|e(\bar{x}_0^N,\bar{x}_N^N)\|_\infty\leqslant(N-r)^{-1/4} \tag{15}$$

for all $0\leqslant k\leqslant N$.

The resulting nonlinear programming problem, i.e. Problem B^N can then be solved by an appropriate globally convergent algorithm [32], such as a sequential-quadratic programming method. This approach has been successfully used in solving an impressive array of problems (see, for example, [7, 14, 20, 23]).

If the constraints in Problem B^N result in a closed and bounded region for $\bar{x}_k^N$ and $\bar{u}_k^N$, $k=0,1,\ldots,N$, then Problem B^N has optimal solutions provided that feasible trajectories exist. Even when the region is unbounded, an artificial constraint can be added to Problem B^N so that the resulting region is bounded and large enough to contain the discretization of the true optimal solution. If the discrete optimal solutions converge to the continuous-time optimal solution, then the artificial bound finally becomes inactive for N that is large enough.

3. EXISTENCE OF FEASIBLE SOLUTIONS

For Problem B^N, a fundamental question that needs to be answered is the following: does a feasible solution satisfying the discretized constraints exist around a feasible solution of the continuous-time problem? In [26], the feasibility of Problem B^N is guaranteed under a critical assumption: the controller $u(t)$ is continuous. However, in many problems the optimal controller is discontinuous as in the case of a bang-bang controller. In this section, we extend the result in [26], and prove that Problem B^N is always feasible even when the optimal control of Problem B is discontinuous.

Definition 1

A function $\psi(t):[-1,1]\to\mathbb{R}^k$ is called piecewise C^1 if there exist finitely many points $\tau_0=-1<\tau_1<\cdots<\tau_{s+1}=1$ such that, on every subinterval (τ_i,τ_{i+1}), $i=0,\ldots,s$, $\psi(t)$ is continuously differentiable and both $\psi(t)$ and its derivative, $\dot{\psi}(t)$, are bounded.

Assumption 1

The optimal state, $x_r^*(t)$, is assumed to be continuous and piecewise C^1. The optimal control, $u^*(t)$, is assumed to be piecewise C^1.

DOI: 10.1002/rnc

Note that, according to Definition 1 and Assumption 1, $u^*(t)$ could have finitely many discontinuous points. In the following, a function $v'(t)$ is called the distributional derivative of a L^1 function $v(t)$ if

$$\int_{-1}^{1} v(t)\frac{\mathrm{d}\phi(t)}{\mathrm{d}t}\,\mathrm{d}t = -\int_{-1}^{1} v'(t)\phi(t)\,\mathrm{d}t$$

for all smooth functions $\phi(t)$ with compact support in $[-1,1]$ (see for instance [18]).

Assumption 2
The set $\{(x,u)|h(x,u)\leqslant 0\}$ is convex.

In the following, the results are proved for a subset of $[-1,1]$. The subset is defined as follows. Let $(x(t),u(t))$ be any feasible solution of Problem B, i.e. $(x(t),u(t))$ satisfying differential equation (2), constraint (3) and endpoint condition (4). Suppose Assumption 1 holds for $(x(t),u(t))$. Let $-1<\tau_1<\cdots<\tau_s<1$ represent the discontinuity points of $u(t)$, and define

$$I_\delta = [-1,1]\Big\backslash \bigcup_{j=1}^{s} (\tau_j-\delta,\tau_j+\delta) \tag{16}$$

where $\delta=(N-r)^{-1/2}$. In other words, I_δ represents the closed set in $[-1,1]$ by removing a δ neighbourhood around the discontinuous points of $u(t)$.

Assumption 2 and Lemma 1, to be proved in the next theorem, represent some major differences between the PS method for discontinuous control and the case of continuous control in [26]. For instance, the concept of convergence is different. It is impossible to prove the uniform convergence of the discrete solutions like in [26]. In this paper, the convergence is proved in I_δ, a subset of $[-1,1]$, in which an open neighbourhood around the discontinuities must be removed. Furthermore, we carefully select the rate at which the size of this open neighbourhood shrinks. Another difference from [26] is Assumption 2. It requires that the state-control constraint must be convex. This convexity property is not required for the continuous optimal control in [26]. Lemma 1 is fundamental in the proofs of the theorems. For optimal control with discontinuities, the discrete approximate solutions cannot be compared directly to the solutions of the original problem. The error must be estimated by comparing the discrete solution to the dummy solution developed in Lemma 1. Then the dummy solution is compared to the solution of the original problem. As a result, the proof of the existence of feasible solutions is much more involved than that in [26].

Theorem 1
Given any feasible solution $(x(t),u(t))$ of (2)–(4) in Problem B, suppose Assumptions 1 and 2 hold. Then there exists a positive integer N_1 such that, for any $N>N_1$, the constraints (13)–(15) of Problem B^N have a feasible solution $(\bar{x}_k^N,\bar{u}_k^N)$. Furthermore, the feasible solution satisfies

$$\|x(t_k)-\bar{x}_k^N\|_\infty \leqslant (N-r)^{-1/4}, \quad 0\leqslant k\leqslant N \tag{17}$$

$$|u(t_k)-\bar{u}_k^N|\leqslant (N-r)^{-1/4} \quad \forall t_k\in I_\delta \tag{18}$$

where I_δ is defined in (16).

Remark 3.1
The importance of Theorem 1 is self-evident. It guarantees that Problem B^N is well-posed with a non-empty set of feasible trajectories. If Problem B has infinitely many feasible trajectories satisfying Assumptions 1–2, then Problem B^N has infinitely many feasible trajectories provided a sufficient number of nodes are chosen. Furthermore, (17) and (18) imply the existence of a feasible discrete solution around any neighbourhood of the continuous trajectory.

Due to the discontinuity in the optimal control, the proof of this theorem calls for highly involved algebraic derivations and inequality estimations. We prove some of the key inequalities in the following lemma.

Lemma 1
Consider any feasible solution, $(x(t), u(t))$, of Problem B satisfying Assumptions 1–2. For any $N > 0$, there exist continuous and piecewise C^1 functions $(z_1(t), \ldots, z_r(t), v(t))$, such that $(z_1(t), \ldots, z_r(t), v(t))$ satisfy the differential equation (2) and the following conditions:

$$h(z(t), v(t)) \leqslant C_1 (N-r)^{-1/2} \cdot \mathbf{1} \tag{19}$$

$$\|e(z(-1), z(1))\|_\infty \leqslant C_2 (N-r)^{-1/2} \tag{20}$$

$$\|z(t) - x(t)\|_\infty \leqslant C_3 (N-r)^{-1/2} \tag{21}$$

$$|u(t) - v(t)| \leqslant C_4 (N-r)^{-1/2} \quad \forall t \in I_\delta, \quad \delta = (N-r)^{-1/2} \tag{22}$$

$$\sum_{i=1}^{2} \|z_r^{(i)}(t)\|_\infty \leqslant C_5 + C_6 (N-r)^{1/2} \tag{23}$$

where C_i, $1 \leqslant i \leqslant 6$, are positive constants independent of N and $z_r^{(i)}$ denotes the ith order distribution derivative of $z_r(t)$.

Proof
Define a continuous function $\hat{u}(t)$ as follows:

$$\hat{u}(t) = \begin{cases} (1-\alpha)u(\tau_i - \delta) + \alpha u(\tau_i + \delta) & \\ \quad \text{if } t \in [\tau_i - \delta, \tau_i + \delta], & 1 \leqslant i \leqslant s \\ u(t) & \text{otherwise} \end{cases} \tag{24}$$

where $\alpha = (1/2\delta)(t - \tau_i + \delta)$ and $\delta = (N-r)^{-1/2}$. So, $\hat{u}(t)$ agrees with $u(t)$ if t is not close to any point of discontinuity. If t is in a δ neighbourhood of discontinuity, $\hat{u}(t)$ interpolates the points $(\tau_i - \delta, u(\tau_i - \delta))$ and $(\tau_i + \delta, u(\tau_i + \delta))$ by a straight line. Let

$$q(t) = f(x(t)) + g(x(t))\hat{u}(t), \quad t \in [-1, 1] \tag{25}$$

Then both $\hat{u}(t)$ and $q(t)$ are bounded, continuous, and piecewise C^1. Next, define

$$\begin{aligned} z_r(t) &= \int_{-1}^{t} q(\tau)\,\mathrm{d}\tau + x_r(-1) \\ z_{r-1}(t) &= \int_{-1}^{t} z_r(\tau)\,\mathrm{d}\tau + x_{r-1}(-1) \\ &\vdots \\ z_1(t) &= \int_{-1}^{t} z_2(\tau)\,\mathrm{d}\tau + x_1(-1) \end{aligned} \tag{26}$$

and

$$v(t) = \frac{q(t) - f(z(t))}{g(z(t))} \tag{27}$$

Substituting the pair $(z(t), v(t))$ into (2), Equations (25)–(27) imply that $(z(t), v(t))$ satisfy the differential equation (2). Next, we will show that they also satisfy conditions (19)–(22).

Denote M_1 the upper bound of $|u(t)|$ for $t \in [-1, 1]$. From the definition of $\hat{u}(t)$, we have

$$\begin{aligned} \|u(t) - \hat{u}(t)\|_{L^1} &= \sum_{i=1}^{s} \int_{\tau_i-\delta}^{\tau_i+\delta} |(1-\alpha)(u(\tau_i - \delta) - u(t)) + \alpha(u(\tau_i + \delta) - u(t))|\,\mathrm{d}t \\ &\leqslant 4sM_1(N-r)^{-1/2} \end{aligned}$$

Therefore,

$$\begin{aligned} \|\dot{x}_r(t) - q(t)\|_{L^1} &= \|g(x(t))(u(t) - \hat{u}(t))\|_{L^1} \\ &\leqslant 4sM_1M_2(N-r)^{-1/2} \end{aligned} \tag{28}$$

where M_2 is an upper bound of $|g(x(t))|$ for $t \in [-1, 1]$. From (28), it is not difficult to show the following inequality:

$$|x_i(t) - z_i(t)| \leqslant 2^{r-i+2} sM_1M_2(N-r)^{-1/2} \quad \forall t \in [-1, 1] \tag{29}$$

where $i = 1, 2, \ldots, r$. Hence, (21) holds with $C_3 = 2^{r+1}sM_1M_2$. Next, for any t in $[-1, 1]$,

$$\begin{aligned} |v(t) - \hat{u}(t)| &= \left| \frac{q(t) - f(z(t))}{g(z(t))} - \frac{q(t) - f(x(t))}{g(x(t))} \right| \\ &\leqslant rK_1 \|z(t) - x(t)\|_\infty \\ &\leqslant 2^{r+1} srM_1M_2K_1(N-r)^{-1/2} \end{aligned} \tag{30}$$

where K_1 is determined by the upper bound of $q(t)$ and the Lipschitz constants of $1/g(x)$ and $f(x)/g(x)$. By definition, $u(t) = \hat{u}(t)$ for all $t \in I_\delta$; therefore, (22) is true with $C_4 = 2^{r+1}srM_1M_2K_1$.

For constraint (19), if $|t - \tau_i| > \delta$

$$h(x(t), \hat{u}(t)) = h(x(t), u(t)) \leqslant 0$$

If $|t-\tau_i| \leqslant \delta$, the convexity Assumption 2 implies

$$\begin{aligned} h(x(t), \hat{u}(t)) &= h((1-\alpha)x(\tau_i-\delta)+\alpha x(\tau_i+\delta), \hat{u}(t)) + h(x(t), \hat{u}(t)) \\ &\quad - h((1-\alpha)x(\tau_i-\delta)+\alpha x(\tau_i+\delta), \hat{u}(t)) \\ &\leqslant 0 + rK_2\|x(t) - ((1-\alpha)x(\tau_i-\delta)+\alpha x(\tau_i+\delta))\|_\infty \cdot \mathbf{1} \\ &\leqslant (rK_2(1-\alpha)\|x(t)-x(\tau_i-\delta)\|_\infty + rK_2\alpha\|x(t)-x(\tau_i+\delta)\|_\infty)\cdot \mathbf{1} \\ &\leqslant (2K_2(1-\alpha)rM_3\delta + 2K_2\alpha rM_3\delta)\cdot \mathbf{1} \\ &= 2rK_2M_3(N-r)^{-1/2}\cdot \mathbf{1} \end{aligned} \tag{31}$$

In the above derivation, K_2 represents a Lipschitz constant of $h(\cdot)$; M_3 is an upper bound of $|\dot{x}_i(t)|$, for $i=1,\dots,r$ and $t\in[-1,1]$.

From (29)–(31),

$$\begin{aligned} h(z(t), v(t)) &= h(x(t), \hat{u}(t)) + h(z(t), v(t)) - h(x(t), \hat{u}(t)) \\ &\leqslant 2rK_2M_3(N-r)^{-1/2}\cdot \mathbf{1} + K_2(r\|z(t)-x(t)\|_\infty + \|v(t)-\hat{u}(t)\|_\infty)\cdot \mathbf{1} \\ &\leqslant (2M_3 + (K_1+1)2^{r+1}sM_1M_2)rK_2(N-r)^{-1/2}\cdot \mathbf{1} \end{aligned}$$

Hence, constraint (19) holds with $C_1 = (2M_3 + (K_1+1)2^{r+1}sM_1M_2)rK_2$. Similarly,

$$\begin{aligned} \|e(z(-1), z(1))\|_\infty &\leqslant \|e(x(-1), x(1))\|_\infty + \|e(z(-1), z(1)) - e(x(-1), x(1))\|_\infty \\ &\leqslant rK_3(\|z(-1)-x(-1)\|_\infty + \|z(1)-x(1)\|_\infty) \\ &\leqslant 2^{r+2}rsM_1M_2K_3(N-r)^{-1/2} \end{aligned}$$

where K_3 represents a Lipschitz constant of $e(\cdot)$. Thus, (20) is verified.

Finally, because $\dot{z}_r(t) = q(t)$ and (25), we have

$$z_r^{(2)} = \frac{\mathrm{d}}{\mathrm{d}t}(f(x(t))) + \frac{\mathrm{d}}{\mathrm{d}t}(g(x(t)))\hat{u}(t) + g(x(t))\hat{u}^{(1)}(t)$$

From the definition of $\hat{u}(t)$,

$$|\hat{u}^{(1)}(t)| < \frac{M_1}{\delta} = M_1(N-r)^{1/2}$$

for sufficiently large N. In addition, the derivatives of $f(x(t))$ and $g(x(t))$ are bounded. Therefore, (23) holds. □

Proof of Theorem 1

From Lemma 1, there exists a continuous and piecewise C^1 function pair $(z(t), v(t))$ satisfying the differential equations (2) and inequalities (19)–(22). Let $p(t)$ be the $(N-r)$th order best approximation polynomial of $\dot{z}_r(t)$ in the norm of $L^\infty(-1,1)$. The following estimation has been proved in the literature of spectral methods [18]:

$$|\dot{z}_r(t) - p(t)| \leqslant C_0(N-r)^{-1} \sum_{i=1}^{2} \|z_r^{(i)}\|_{L^\infty(-1,1)} \tag{32}$$

$\forall t \in [-1, 1]$. Substituting (23) in (32) leads to

$$|\dot{z}_r(t) - p(t)| \leqslant C_0 C_5 (N-r)^{-1} + C_0 C_6 (N-r)^{-1/2} \tag{33}$$

Let us define

$$\begin{aligned}
\hat{x}_r(t) &= \int_{-1}^{t} p(\tau)\,\mathrm{d}\tau + x_r(-1)\\
\hat{x}_{r-1}(t) &= \int_{-1}^{t} \hat{x}_r(\tau)\,\mathrm{d}\tau + x_{r-1}(-1)\\
&\vdots\\
\hat{x}_1(t) &= \int_{-1}^{t} \hat{x}_2(\tau)\,\mathrm{d}\tau + x_1(-1)\\
\hat{v}(t) &= \frac{p(t) - f(\hat{x}_1(t), \ldots, \hat{x}_r(t))}{g(\hat{x}_1(t), \ldots, \hat{x}_r(t))}
\end{aligned}$$

From (33), it is easy to show

$$|z_i(t) - \hat{x}_i(t)| \leqslant 2^{r-i+1} C_0 [C_5 (N-r)^{-1} + C_6 (N-r)^{-1/2}] \quad \forall t \in [-1, 1] \tag{34}$$

and

$$\begin{aligned}
|v(t) - \hat{v}(t)| &= \left| \frac{\dot{z}_r(t) - f(z(t))}{g(z(t))} - \frac{p(t) - f(\hat{x}(t))}{g(\hat{x}(t))} \right|\\
&\leqslant K_1(|\dot{z}_r(t) - p(t)| + r\|z(t) - \hat{x}(t)\|_\infty)\\
&\leqslant C_0 K_1 (1 + r2^r)(C_5 (N-r)^{-1} + C_6 (N-r)^{-1/2})
\end{aligned} \tag{35}$$

Define

$$\bar{x}_k^N = \hat{x}(t_k), \quad \bar{u}_k^N = \hat{v}(t_k) \tag{36}$$

In the following, we prove that $(\bar{x}_k^N, \bar{u}_k^N)$ is a feasible solution of (13)–(15). Because $p(t)$ is an polynomial of degree less than or equal to $(N-r)$, the functions $\hat{x}_1(t), \ldots, \hat{x}_r(t)$ must be polynomials of degree less than or equal to N. Moreover, $(\hat{x}(t), \hat{v}(t))$ satisfies the differential equation (2) and has the same initial condition as $x(-1)$. Given any polynomial of degree less than or equal to N, it is known (see [18]) that its derivative at the nodes $t_0, \ldots, t_N$ are exactly equal to the value of the polynomial at the nodes multiplied by the differential matrix D. Thus,

DOI: 10.1002/rnc

we have

$$D\begin{pmatrix} \bar{x}_{i0}^N \\ \vdots \\ \bar{x}_{iN}^N \end{pmatrix} = D\begin{pmatrix} \hat{x}_i(t_0) \\ \vdots \\ \hat{x}_i(t_N) \end{pmatrix} = \begin{pmatrix} \dot{\hat{x}}_i(t_0) \\ \vdots \\ \dot{\hat{x}}_i(t_N) \end{pmatrix}$$

$$= \begin{pmatrix} \hat{x}_{i+1}(t_0) \\ \vdots \\ \hat{x}_{i+1}(t_N) \end{pmatrix} = \begin{pmatrix} \bar{x}_{i+1,0}^N \\ \vdots \\ \bar{x}_{i+1,N}^N \end{pmatrix}$$

where $i = 1, 2, \ldots, r-1$ and $\bar{x}_{ik}^N$ is the ith component of $\bar{x}_k^N$. At $i = r$, we have

$$D\begin{pmatrix} \bar{x}_{r0}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix} = D\begin{pmatrix} \hat{x}_r(t_0) \\ \vdots \\ \hat{x}_r(t_N) \end{pmatrix} = \begin{pmatrix} \dot{\hat{x}}_r(t_0) \\ \vdots \\ \dot{\hat{x}}_r(t_N) \end{pmatrix}$$

$$= \begin{pmatrix} f(\hat{x}(t_0)) + g(\hat{x}(t_0))\hat{v}(t_0) \\ \vdots \\ f(\hat{x}(t_N)) + g(\hat{x}(t_N))\hat{v}(t_N) \end{pmatrix}$$

Therefore, $(\bar{x}_k^N, \bar{u}_k^N)$, $k = 0, 1, \ldots, N$, satisfy the constraint equations in (13). Next, we prove that the mixed state-control constraint (14) is also satisfied. Because $h(\cdot)$ is Lipschitz continuous, the following estimation holds:

$$\begin{aligned} \|h(z(t), v(t)) - h(\hat{x}(t), \hat{v}(t))\|_\infty &\leqslant K_2(r\|z(t) - \hat{x}(t)\|_\infty + |v(t) - \hat{v}(t)|) \\ &\leqslant K_2C_0(r2^r + K_1 + r2^rK_1) \cdot [C_5(N-r)^{-1} + C_6(N-r)^{-1/2}] \end{aligned}$$

Hence, by (19),

$$h(\hat{x}(t), \hat{v}(t)) \leqslant (L_1(N-r)^{-1} + L_2(N-r)^{-1/2}) \cdot \mathbf{1}$$

where

$$\begin{aligned} L_1 &= K_2C_0C_5(r2^r + K_1 + r2^rK_1) \\ L_2 &= K_2C_0C_6(r2^r + K_1 + r2^rK_1) + C_1 \end{aligned}$$

Since constants L_1 and L_2 are independent of N, there exists a positive integer N_1 such that, for all $N > N_1$,

$$L_1(N-r)^{-1} + L_2(N-r)^{-1/2} \leqslant (N-r)^{-1/4}$$

Therefore, $\hat{x}_1(t_k), \ldots, \hat{x}_r(t_k), \hat{u}(t_k)$, $k = 0, 1, \ldots, N$, satisfy the mixed state and control constraints (14) for all $N > N_1$. The end-point conditions (15) can be proved in the same way. Thus, $(\bar{x}_k^N, \bar{u}_k^N)$ is a feasible discrete solution to Problem B^N.

As for (17)–(18), they can be easily deduced from (34)–(35) and (21)–(22) in Lemma 1. □

Remark 3.2
In the proof of Theorem 1 and Lemma 1, we actually established a stronger result than (17)–(18). That is

$$\|x(t) - \hat{x}(t)\|_\infty \leqslant (N - r)^{-1/4} \quad \forall t \in [-1, 1]$$

$$|u(t) - \hat{v}(t)| \leqslant (N - r)^{-1/4} \quad \forall t \in I_\delta$$

These properties will be used later in the proof of the convergence of Legendre PS method.

4. CONVERGENCE RESULTS

Once the feasibility of the discrete Problem B^N is established, one can apply nonlinear programming solver to compute the discrete optimal solution. Next, we focus on the challenging problem of proving the convergence of the discrete solutions of Problem B^N as an approximation of the original continuous-time optimal control problem. In this section, we will provide a sufficient condition under which the convergence of the Legendre PS method for the continuous-time optimal control problem can be guaranteed.

Let $(\bar{x}_k^N, \bar{u}_k^N)$, $k = 0, 1, \ldots, N$, be a feasible solution to Problem B^N, and $x^N(t) \in \mathbb{R}^r$ be the Nth order interpolating polynomials of $(\bar{x}_0^N, \ldots, \bar{x}_N^N)$, i.e.

$$x^N(t) = \sum_{k=0}^{N} \bar{x}_k^N \phi_k(t) \tag{37}$$

where $\phi_k(t)$ is defined by (6). Also denote

$$u^N(t) = \frac{\dot{x}_r^N(t) - f(x^N(t))}{g(x^N(t))}$$

Because

$$D\begin{pmatrix} \bar{x}_{r0}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix} = D\begin{pmatrix} \hat{x}_r(t_0) \\ \vdots \\ \hat{x}_r(t_N) \end{pmatrix} = \begin{pmatrix} \dot{\hat{x}}_r(t_0) \\ \vdots \\ \dot{\hat{x}}_r(t_N) \end{pmatrix}$$

The definition of $u^N(t)$ and (13) imply that $u^N(t_k) = \bar{u}_k^N$. Now consider a sequence of discrete feasible solution $\{(\bar{x}_k^N, \bar{u}_k^N), k = 0, \ldots, N\}_{N=N_1}^\infty$ and the corresponding interpolating polynomial sequence $\{x^N(t)\}_{N=N_1}^\infty$ and the non-polynomial sequence $\{u^N(t)\}_{N=N_1}^\infty$.

Assumption 3
(a) For all $1 \leqslant i \leqslant r$, the sequence $\{\bar{x}_{i0}^N\}_{N=N_1}^\infty$ converges as $N \to \infty$; (b) $\dot{x}_r^N(t)$ is uniformly bounded for $N \geqslant N_1$ and $t \in [-1, 1]$; (c) there exists a piecewise C^1 function $q(t)$ such that, for any fixed $\varepsilon > 0$, $\dot{x}_r^N(t)$ converges to $q(t)$ uniformly on the interval I_ε, where

$$I_\varepsilon = [-1, 1] \Big\backslash \bigcup_{j=1}^{s} (\tau_j - \varepsilon, \tau_j + \varepsilon) \tag{38}$$

and $-1 < \tau_1 < \cdots < \tau_s < 1$ represent the discontinuity points of $q(t)$.

In practical computations, Assumption 3 can be verified up to a large N, the number of nodes in discretization. Through the following theorem, the verification of this assumption provides the confidence on the optimality of the discrete solutions. Assumption 3 is made along the line of the consistent approximation theory [27] in which the discrete solutions are assumed to be 'epi-convergent.' Assumption 3 is more transparent in the sense that it requires the convergence of $\dot{x}_r^N(t)$ instead of the epigraph in the multiple dimensional state space. An important question that remains unanswered is: under what condition does an optimal control problem satisfy Assumption 3 for the PS methods? We have proved some results on this issue. It will be reported in a separate paper.

Theorem 2
Consider a sequence of feasible solutions $(\bar{x}_k^N, \bar{u}_k^N)$, $k = 0, 1, \ldots, N$, of (13)–(15) in Problem B^N. Suppose Assumption 3 holds. Then there exists a feasible solution, $(x^\infty(t), u^\infty(t))$, of (2)–(4) in the continuous-time optimal control Problem B such that the limit

$$\lim_{N\to\infty} (x^N(t) - x^\infty(t)) = 0 \tag{39}$$

converges uniformly on $[-1, 1]$, and the limit

$$\lim_{N\to\infty} (u^N(t) - u^\infty(t)) = 0 \tag{40}$$

converges uniformly on any closed set I_ε.

Proof
Let x_{i0} be the limit of $\{\bar{x}_{i0}^N\}_{N=N_1}^\infty$. Then, define the following functions:

$$\begin{aligned}
x_r^\infty(t) &= \textstyle\int_{-1}^t q(\tau)\,\mathrm{d}\tau + x_{r0}\\
x_{r-1}^\infty(t) &= \textstyle\int_{-1}^t x_r^\infty(\tau)\,\mathrm{d}\tau + x_{r-1,0}\\
&\ \vdots\\
x_1^\infty(t) &= \textstyle\int_{-1}^t x_2^\infty(\tau)\,\mathrm{d}\tau + x_{10}\\
u^\infty(t) &= \frac{q(t) - f(x_1^\infty(t), \ldots, x_r^\infty(t))}{g(x_1^\infty(t), \ldots, x_r^\infty(t))}
\end{aligned}$$

Obviously, $(x^\infty(t), u^\infty(t))$ satisfies the differential equation (2). Next, we prove (39)–(40) and the fact that $(x^\infty(t), u^\infty(t))$ satisfies both the mixed constraints in (3) and end-point condition (4).

Let $x_i^N(t)$ be the interpolating polynomial of $\bar{x}_{i0}^N, \ldots, \bar{x}_{iN}^N$. Because $(\bar{x}_k^N, \bar{u}_k^N)$ satisfies discrete state equation (13), it is easy to see

$$\begin{pmatrix} \dot{x}_i^N(t_0)\\ \vdots \\ \dot{x}_i^N(t_N)\end{pmatrix} = D\begin{pmatrix} \bar{x}_{i0}^N\\ \vdots \\ \bar{x}_{iN}^N\end{pmatrix} = \begin{pmatrix} \bar{x}_{i+1,0}^N\\ \vdots \\ \bar{x}_{i+1,N}^N\end{pmatrix} = \begin{pmatrix} x_{i+1}^N(t_0)\\ \vdots \\ x_{i+1}^N(t_N)\end{pmatrix}$$

for $i = 1, 2, \ldots, r-1$. Hence, the Nth order polynomial:

$$\dot{x}_i^N(t) - x_{i+1}^N(t)$$

has $N+1$ different roots: $t_0, \ldots, t_N$. Therefore, $\dot{x}_i^N(t) = x_{i+1}^N(t)$, $i = 1, \ldots, r-1$. Under Assumption 3, $\dot{x}_r^N(t)$ is a bounded sequence that converges to $q(t)$ almost everywhere, thus $\dot{x}_r^N(t)$ converges to $q(t)$ in L^1. Therefore,

$$\begin{aligned}\lim_{N\to\infty} |x_r^N(t) - x_r^\infty(t)| &= \lim_{N\to\infty} \left| \int_{-1}^{t} (\dot{x}_r^N(\tau) - q(t))\, \mathrm{d}\tau \right| \\ &\leqslant \lim_{N\to\infty} \int_{-1}^{1} |(\dot{x}_r^N(\tau) - q(t))|\, \mathrm{d}\tau \\ &= 0\end{aligned}$$

Moreover, the limit converge uniformly in t. Hence, the following limit converges uniformly:

$$\begin{aligned}\lim_{N\to\infty} x_{r-1}^N(t) &= \lim_{N\to\infty} \int_{-1}^{t} x_r^N(\tau)\, \mathrm{d}\tau + x_{r-1,0} \\ &= \int_{-1}^{t} x_r^\infty(\tau)\, \mathrm{d}\tau + x_{r-1,0} = x_{r-1}^\infty(t)\end{aligned}$$

Following the same procedure, we can prove

$$\lim_{N\to\infty} x_i^N(t) = x_i^\infty(t), \quad i = 1, 2, \ldots, r$$

uniformly in t. Thus, (39) is proved.

As for (40), it follows the following inequality:

$$\begin{aligned}|u^N(t) - u^\infty(t)| &= \left| \frac{\dot{x}_r^N(t) - f(x^N(t))}{g(x^N(t))} - \frac{q(t)}{g(x^N(t))} + \frac{q(t)}{g(x^N(t))} - \frac{q(t) - f(x^\infty(t))}{g(x^\infty(t))} \right| \\ &= \left| \frac{\dot{x}_r^N(t) - q(t)}{g(x^N(t))} + \frac{q(t) - f(x^N(t))}{g(x^N(t))} - \frac{q(t) - f(x^\infty(t))}{g(x^\infty(t))} \right| \\ &\leqslant K_1 |\dot{x}_r^N(t) - q(t)| + rK_1 \|x^N(t) - x^\infty(t)\|_\infty\end{aligned}$$

and the fact that both $\dot{x}_r^N(t) - q(t)$ and $x^N(t) - x^\infty(t)$ converge to zero uniformly on any closed set I_ε. In this inequality, K_1 is defined by (30).

The endpoint condition $e(x^\infty(-1), x^\infty(1)) = 0$ follows directly from the convergence property, since

$$\begin{aligned}e(x^\infty(-1), x^\infty(1)) &= \lim_{N\to\infty} e(x^N(-1), x^N(1)) \\ &= \lim_{N\to\infty} e(\bar{x}_0^N, \bar{x}_N^N) = 0\end{aligned}$$

Now, to show $(x^\infty(t), u^\infty(t))$ is a feasible solution of Problem B, it is enough to prove the mixed state-control constraint $h(x^\infty(t), u^\infty(t)) \leqslant 0$. Using contradiction argument, suppose at a time instance $\tau' \in (-1, 1)$ there exists a constraint $h_i(\cdot)$, $i \in \{1, 2, \ldots, l\}$, so that

$$h_i(x^\infty(\tau'), u^\infty(\tau')) > 0$$

Since $x^\infty(t)$ is continuous and $u^\infty(t)$ is piecewise C^1, without loss of generality, we can select τ' outside the set $\{\tau_1, \ldots, \tau_s\}$. By the fact that the nodes t_k are getting dense as N tends to infinity [33], there exists a sequence $\{j^N\}$ such that, $0 \leqslant j^N \leqslant N$, the LGL nodes $t_{j^N} \in I_\varepsilon$ and

$$\lim_{N\to\infty} t_{j^N} = \tau'$$

Then (39) and (40) imply

$$\lim_{N\to\infty} h_i(\bar{x}^N_{j^N}, \bar{u}^N_{j^N}) = h_i(x^\infty(\tau'), u^\infty(\tau')) > 0$$

It contradicts the mixed state-control constraint (14), in which the right side of the inequality approaches zero as N approaching infinity. □

Theorem 2 implies that for any convergent discrete solution sequence, the limit point of this sequence must be a feasible solution of the original continuous-time optimal control problem. Next, we study a special sequence of discrete feasible solutions. These are the optimal solutions of Problem B^N. Naturally, the question we must answer is: under what condition does the sequence converge to the optimal solution of the continuous-time problem, and the cost (12) converges to the optimal cost function defined by (1)? In the following we will show that, under Assumption 3, the convergence of the PS method can be guaranteed even if the optimal control of the Problem B is discontinuous. The notations in the next theorem, such as $(\bar{x}_k^{*N}, \bar{u}_k^{*N})$ and J^N, are defined in Section 2.

Theorem 3

Suppose Problem B satisfies Assumptions 1 and 2. Let $(\bar{x}_k^{*N}, \bar{u}_k^{*N})$, $k = 0, 1, \ldots, N$, be a sequence of discrete optimal solutions of Problem B^N. Assume the sequence satisfies Assumption 3. Then, there exists an optimal solution $(x^*(t), u^*(t))$ of the continuous-time optimal control Problem B such that the following limits converge uniformly:

$$\lim_{N\to\infty} (\bar{x}_k^{*N} - x^*(t_k)) = 0$$

$$\lim_{N\to\infty} (\bar{u}_k^{*N} - u^*(t_k)) = 0, \quad t_k \in I_\varepsilon$$

$$\lim_{N\to\infty} \bar{J}^N(\bar{X}^*, \bar{U}^*) = J(x^*(\cdot), u^*(\cdot))$$

for all $0 \leqslant k \leqslant N$ and any fixed $\varepsilon > 0$.

Before the proof of this convergence result, we need the following lemmas. The first two are known results in the literature (see [33] for the proof).

Lemma 2

Let t_k, $k = 0, 1, \ldots, N$, be the LGL nodes, and w_k be the LGL weights. Suppose $\xi(t)$ is Riemann integrable; then,

$$\int_{-1}^{1} \xi(t)\,\mathrm{d}t = \lim_{N\to\infty} \sum_{k=0}^{N} \xi(t_k) w_k$$

Lemma 3

Given any interval $[a, b]$ in $[-1, 1]$. Then

$$\lim_{N\to\infty} \sum_{\substack{k \\ t_k \in [a,b]}} \omega_k = b - a \tag{41}$$

where t_k are LGL nodes.

DOI: 10.1002/rnc

Lemma 4

Suppose $\{x^N(t)\}_{N\geqslant 1}$ is a sequence consisting of continuous functions. Suppose $\{u^N(t)\}_{N\geqslant 1}$ is a sequence of uniformly bounded piecewise C^1 functions. Moreover, assume there exists $x(t)$ so that

$$\lim_{N\to\infty} x^N(t) = x(t) \tag{42}$$

converges uniformly on $[-1, 1]$. Assume there exists a piecewise C^1 function $u(t)$ such that

$$\lim_{N\to\infty} u^N(t) = u(t) \tag{43}$$

converges uniformly on any I_ε, a closed set defined by ε and the discontinuous points of $u(t)$ (see (38) for the definition of I_ε). Then we have

$$\lim_{N\to\infty}\left\{\sum_{k=0}^{N} F(x^N(t_k), u^N(t_k))\omega_k + E(x^N(-1), x^N(1))\right\}$$
$$= \int_{-1}^{1} F(x(t), u(t))\,\mathrm{d}t + E(x(-1), x(1))$$

Proof

Let $-1<\tau_1<\cdots<\tau_s<1$ be the points of discontinuity of $u(t)$. From Lemma 3, given any $\varepsilon>0$, there exists $N_1>0$ so that

$$\sum_{\substack{k \\ t_k\in\bigcup_{j=1}^{s}[\tau_j-\varepsilon,\tau_j+\varepsilon]}} \omega_k < 3s\varepsilon \tag{44}$$

for all $N\geqslant N_1$. Furthermore, from (42) and (43), we can select N_1 large enough so that

$$|u(t_k) - u^N(t_k)| < \varepsilon \quad \forall t_k \in I_\varepsilon$$

$$\|x(t_k) - x^N(t_k)\|_\infty < \varepsilon, \quad 0\leqslant k\leqslant N$$

for all $N\geqslant N_1$. Thus

$$\left|\sum_{k=0}^{N}[F(x(t_k), u(t_k)) - F(x^N(t_k), u^N(t_k))]\omega_k\right|$$
$$\leqslant \left|\sum_{t_k\in I_\varepsilon}^{N}[F(x(t_k), u(t_k)) - F(x^N(t_k), u^N(t_k))]\omega_k\right|$$
$$+ \left|\sum_{t_k\notin I_\varepsilon}^{N}[F(x(t_k), u(t_k)) - F(x^N(t_k), u^N(t_k))]\omega_k\right|$$
$$\leqslant K\varepsilon + M\varepsilon \tag{45}$$

where K is determined by the Lipschitz constant of $F(x, u)$ and the fact

$$\sum_{k=0}^{N} \omega_k = 2 \tag{46}$$

M is determined by the upper bound of $F(x(t), u(t))$, $F(x^N(t), u^N(t))$ and inequality (44). Inequality (45) implies

$$\lim_{N\to\infty} \sum_{k=0}^{N} [F(x^N(t_k), u^N(t_k)) - F(x(t_k), u(t_k))]\omega_k = 0 \tag{47}$$

Therefore,

$$\begin{aligned}\int_{-1}^{1} F(x(t), u(t))\,\mathrm{d}t &= \lim_{N\to\infty} \sum_{k=0}^{N} F(x(t_k), u(t_k))\omega_k \\ &= \lim_{N\to\infty} \sum_{k=0}^{N} F(x^N(t_k), u^N(t_k))\omega_k \end{aligned} \tag{48}$$

Due to the convergence of $x^N(t)$, it is obvious that

$$\lim_{N\to\infty} E(x^N(-1), x^N(1)) = E(x(-1), x(1)) \tag{49}$$

Then, Lemma 4 follows from (48) and (49). □

Proof of Theorem 3

According to Theorem 2, the discrete optimal solutions uniformly converge to a feasible trajectory of the continuous-time problem. More specifically, there exists a feasible solution, $(x^\infty(t), u^\infty(t))$, of (2)–(4) in Problem B such that

$$\lim_{N\to\infty} (\bar{x}_k^* - x^\infty(t_k)) = 0$$

$$\lim_{N\to\infty} (\bar{u}_k^* - u^\infty(t_k)) = 0, \quad t_k \in I_\varepsilon$$

uniformly for $0 \leqslant k \leqslant N$ and any fixed $\varepsilon > 0$. In the next, we prove that $(x^\infty(t), u^\infty(t))$ is indeed an optimal solution of the continuous-time optimal control problem. To this end, denote $\bar{J}^N(\bar{X}^*, \bar{U}^*)$ and $J(x^*(\cdot), u^*(\cdot))$ the optimal cost of Problem B^N and Problem B, respectively, i.e.

$$\bar{J}^N(\bar{X}^*, \bar{U}^*) = E(\bar{x}_0^*, \bar{x}_N^*) + \sum_{k=0}^{N} F(\bar{x}_k^*, \bar{u}_k^*)w_k$$

$$J(x^*(\cdot), u^*(\cdot)) = E(x^*(-1), x^*(1)) + \int_{-1}^{1} F(x^*(t), u^*(t))\,\mathrm{d}t$$

where $(x^*(t), u^*(t))$ denotes any optimal solution of Problem B (the optimal solution may not be unique). According to Theorem 1, there exists a sequence of feasible solutions, $(\tilde{x}_k^N, \tilde{u}_k^N)$, of (13)–(15) that converges to $(x^*(t), u^*(t))$ in the way defined by (17)–(18). Now, from Lemma 4

DOI: 10.1002/rnc

and the optimality of $(x^*(t), u^*(t))$ and $(\bar{x}_k^{*N}, \bar{u}_k^{*N})$, we have

$$\begin{aligned} J(x^*(\cdot), u^*(\cdot)) &\leqslant J(x^{\infty}(\cdot), u^{\infty}(\cdot)) \\ &= \lim_{N\to\infty} \bar{J}^N(\bar{X}^*, \bar{U}^*) \\ &\leqslant \lim_{N\to\infty} \bar{J}^N(\tilde{X}, \tilde{U}) \\ &= J(x^*(\cdot), u^*(\cdot)) \end{aligned}$$

The last equation is deduced from Lemma 4 and Remark 3.2. Therefore, we proved $J(x^*(\cdot), u^*(\cdot)) = J(x^{\infty}(\cdot), u^{\infty}(\cdot))$. It is equivalent to say that $(x^{\infty}(t), u^{\infty}(t))$ is a feasible solution that achieves optimal cost. Hence, $(x^{\infty}(t), u^{\infty}(t))$ is an optimal solution to the continuous-time optimal control Problem B. □

5. OPTIMAL CONTROL WITH FREE FINAL TIME

The time interval in Problem B is $[-1, 1]$. If the original problem has a fixed time interval $[t_0, t_f]$, then it can be transformed into the interval of $[-1, 1]$ as follows:

$$\xi = \frac{2}{t_f - t_0} t - \frac{t_f + t_0}{t_f - t_0} \tag{50}$$

However, if the final time t_f is not fixed, then this transformation has a free parameter t_f. The resulting optimal control problem is different from the one defined by Problem B. Fortunately, we found that all the results proved in the previous sections can be extended to the case of free final time with some minor modifications. Consider the optimal control problem with free t_f

$$\min_{x(\cdot), u(\cdot), t_f} J[x(\cdot), u(\cdot), t_f]$$

$$J[x(\cdot), u(\cdot), t_f] = \int_{t_0}^{t_f} F(x(t), u(t))\, dt + E(x(t_0), x(t_f))$$

s.t.

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ &\vdots \\ \dot{x}_{r-1}(t) &= x_r(t) \\ \dot{x}_r(t) &= f(x(t)) + g(x(t))u(t) \\ t_f &> t_0 \\ h(x(t), u(t)) &\leqslant 0 \\ e(x(t_0), x(t_f)) &= 0 \end{aligned} \tag{51}$$

We use (50) to transform the problem into the interval of $[-1, 1]$. Then

$$\frac{dt}{d\xi} = \frac{t_f - t_0}{2}$$

DOI: 10.1002/rnc

For any differentiable function $h(\xi)$, the derivative $\mathrm{d}h/\mathrm{d}\xi$ is denoted by $h'(\xi)$. Under the new variable ξ, (51) is equivalent to the following optimal control problem:

$$\min_{x(\cdot),u(\cdot),t_f} J[x(\cdot),u(\cdot),t_f]$$

$$J[x(\cdot),u(\cdot),t_f] = \int_{-1}^{1} F(x(\xi),u(\xi))\frac{t_f-t_0}{2}\,\mathrm{d}\xi + E(x(-1),x(1))$$

s.t.

$$\begin{aligned}
x_1'(\xi) &= \frac{t_f-t_0}{2}x_2(\xi)\\
&\vdots\\
x_{r-1}'(\xi) &= \frac{t_f-t_0}{2}x_r(\xi)\\
x_r'(\xi) &= \frac{t_f-t_0}{2}(f(x(t))+g(x(t)u(t)))\\
t_f &> t_0\\
h(x(\xi),u(\xi)) &\leqslant 0\\
e(x(-1),x(1)) &= 0
\end{aligned} \tag{52}$$

Except for the free parameter t_f and the factor $(t_f-t_0)/2$, this problem is similar to Problem B. After Legendre PS discretization at LGL nodes, the resulting discrete optimization problem is similar to Problem B^N except for the variable t_f and the term $(t_f-t_0)/2$ on the right-hand side of the dynamics (13) and the cost function (12). Following the same idea in Section 3, we can prove that Theorem 1 holds true for the problem defined by (52). In other words, given any feasible solution of (52), there exists a feasible solution of the discretized problem such that the discrete-time solution satisfies (17)–(18) provided the number of nodes is large enough. The proof of the theorem is a copy of the proof in Section 3, except that a factor $(t_f-t_0)/2$ must be added to the integration terms in (26)–(27) and the definitions of $q(t)$ in (25).

Following the ideas in Section 4, results similar to Theorems 2 and 3 can be proved for the optimal control problem defined by (52). The discretization of (52) consists of unknown variables $(\bar{x}_k^N, \bar{u}_k^N, t_f^N)$. Suppose a sequence $(\bar{x}_k^N, \bar{u}_k^N)$ satisfies Assumption 3. In addition, assume that the sequence t_f^N converges as N approaches infinity. Then, it can be proved that there exists a feasible solution $(x^\infty(t), u^\infty(t), t_f)$ satisfying the constraints in (52) so that the sequence $\{(\bar{x}_k^N, \bar{u}_k^N)\}_{N=1}^{\infty}$ approaches $(x^\infty(t), u^\infty(t))$ in the way defined by (39)–(40) and t_f^N approaches t_f. If the sequence $(\bar{x}_k^N, \bar{u}_k^N)$ is the optimal discrete-time solution, then it converges to an optimal solution of the continuous-time optimal control problem defined in (52).

6. EXAMPLES

In this section we present an example to illustrate the main points of the PS method. The problem was programmed in MATLAB on a Pentium 4, 2.4 GHz PC with 256 MB of RAM.

The PS method was applied to this problem using the software package, DIDO [17]. Problems of continuous-thrust trajectory optimization have been serving as motivating problems for optimal control theory since its inception [34–36]. The classic problem posed by Moyer and Pinkham [36] is widely discussed in text books [34, 35] and research articles [8, 37]. When the continuity of thrust is removed from the problem formulation, the optimal control can be dramatically different and hence, the discontinuous thrusting problem remains an active research area of research in astronautical engineering, especially for low-thrust optimal trajectory design [38–40].

Consider the minimum time orbit transfer problem

$$\min \; J[\cdot] = t_{\rm f}$$

$$\text{s.t.} \quad \dot{r} = v_{\rm r}$$

$$\dot{\theta} = \frac{v_{\rm t}}{r}$$

$$\dot{v}_{\rm r} = \frac{v_{\rm t}^2}{r} - \frac{1}{r^2} + u_{\rm r}$$

$$\dot{v}_{\rm t} = -\frac{v_{\rm r} v_{\rm t}}{r} + u_{\rm t}$$

$$|u_{\rm r}| \leqslant 0.05; \; |u_{\rm t}| \leqslant 0.05$$

$$(r(0), v_{\rm r}(0), v_{\rm t}(0)) = (1, 0, 1)$$

$$(r(t_{\rm f}), v_{\rm r}(t_{\rm f}), v_{\rm t}(t_{\rm f})) = (4, 0, 0.5)$$

where r is the radial distance, θ is the true anomaly, $v_{\rm r}$ is the radial velocity, $v_{\rm t}$ is the transverse velocity, $u_{\rm r}$ is the radial thrust and $u_{\rm t}$ is transverse thrust. The problem has free final time. The system has multiple inputs. However, the dynamics are in the multi-input feedback linearizable normal form. It is equivalent to two subsystems in which each one has a single input.

Figure 2 shows the numerical optimal solution with $N = 100$. The optimal final time is 13.085. The first plot in Figure 2 shows the curves of the optimal thrusts $u_{\rm r}$ and $u_{\rm t}$, which appear to be bang-bang. In the second plot of Figure 2, we show the transfer trajectory as well as the direction of the thrust. It is interesting to note that, during the beginning of the transition, the thrust is pointing inwards because $u_{\rm r}$ is negative. This phenomenon is counter-intuitive, and raises suspicion with regards to the optimality of the solution.

Based on the convergence results presented in previous sections, we verify the optimality of the solution by increasing the number of nodes and check the convergence property of the discrete solution series. The simulation results are demonstrated in Figure 3. It can be observed from Figure 3 that the derivative of the interpolating polynomial sequences $\dot{v}_{\rm r}^N(t)$ and $\dot{v}_{\rm t}^N(t)$ converge very well except on small neighbourhoods around the discontinuous points. This is in concurrence with the theoretical results of this paper, and hence provides confidence on the optimality of the discrete solutions.

Next, we independently check the extremality of the discrete solution by verifying the necessary conditions. To this end, we construct the control Hamiltonian, H, and the Lagrangian

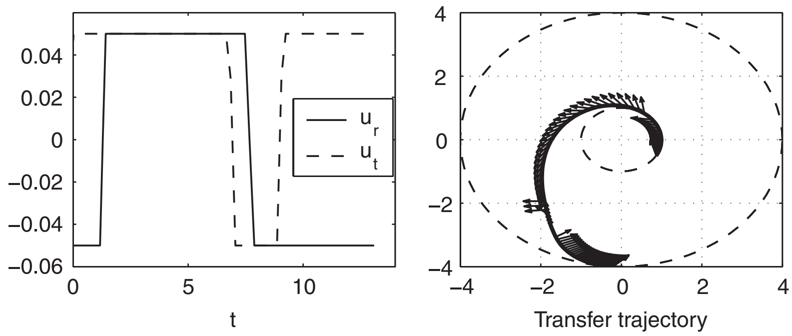

Figure 2. A benchmark minimum time low-thrust orbit transfer.

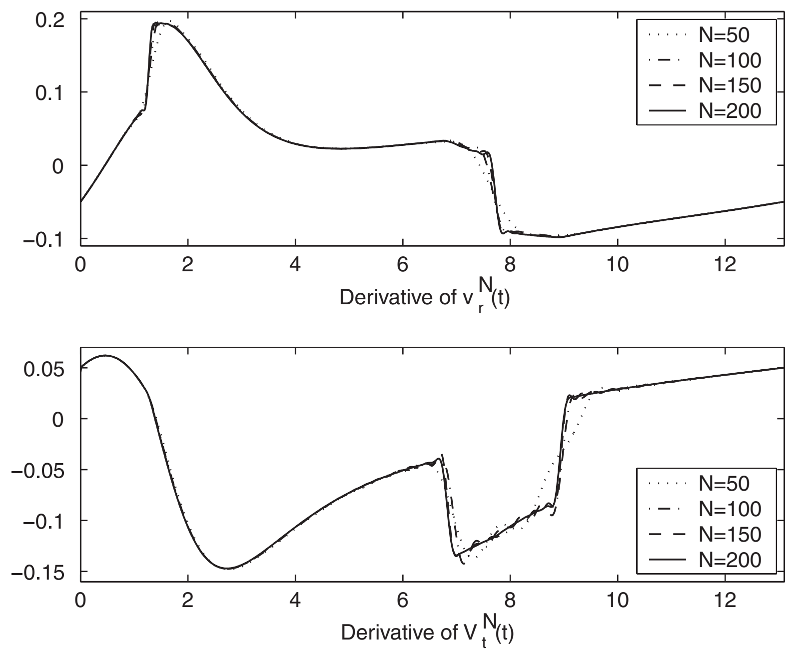

Figure 3. Convergence of discrete optimal solutions.

of the Hamiltonian, $\bar{H}$, as

$$H(x,u,\lambda) = \lambda_r v_r + \lambda_\theta \frac{v_t}{r} + \lambda_{v_r}\left(\frac{v_t^2}{r} - \frac{1}{r^2} + u_r\right) + \lambda_{v_t}\left(-\frac{v_r v_t}{r} + u_t\right)$$

$$\bar{H}(x,u,\lambda,\mu) = H(x,u,\lambda) + \mu_1 u_r + \mu_2 u_t$$

where $\lambda(t)$ is the costate and $\mu(t)$ is the instantaneous KKT multiplier associated with the Hamiltonian minimization condition. Based on the minimum principle, it is straightforward to draw the following conclusions:

$$\begin{aligned} &\lambda_\theta \equiv 0, \quad H \equiv -1 \\ &\mu_1 < 0 \quad \text{if } u_r = -0.05 \\ &\mu_1 > 0 \quad \text{if } u_r = 0.05 \\ &\mu_2 < 0 \quad \text{if } u_t = -0.05 \\ &\mu_2 > 0 \quad \text{if } u_t = 0.05 \end{aligned} \tag{53}$$

In Figure 4 we verify the aforementioned conditions. All the covectors are automatically computed within DIDO by an application of the covector mapping theorem [5, 24]. It can be seen from Figure 4 that the Hamiltonian, H, λ_θ and the covectors μ_1, μ_2 satisfy all the conditions in (53).

This paper is largely concerned about the convergence property of the PS methods but not the arguably more important problem of convergence rate. Like other optimal control algorithms, the convergence rate of 'standard' PS methods applied for discontinuous optimal

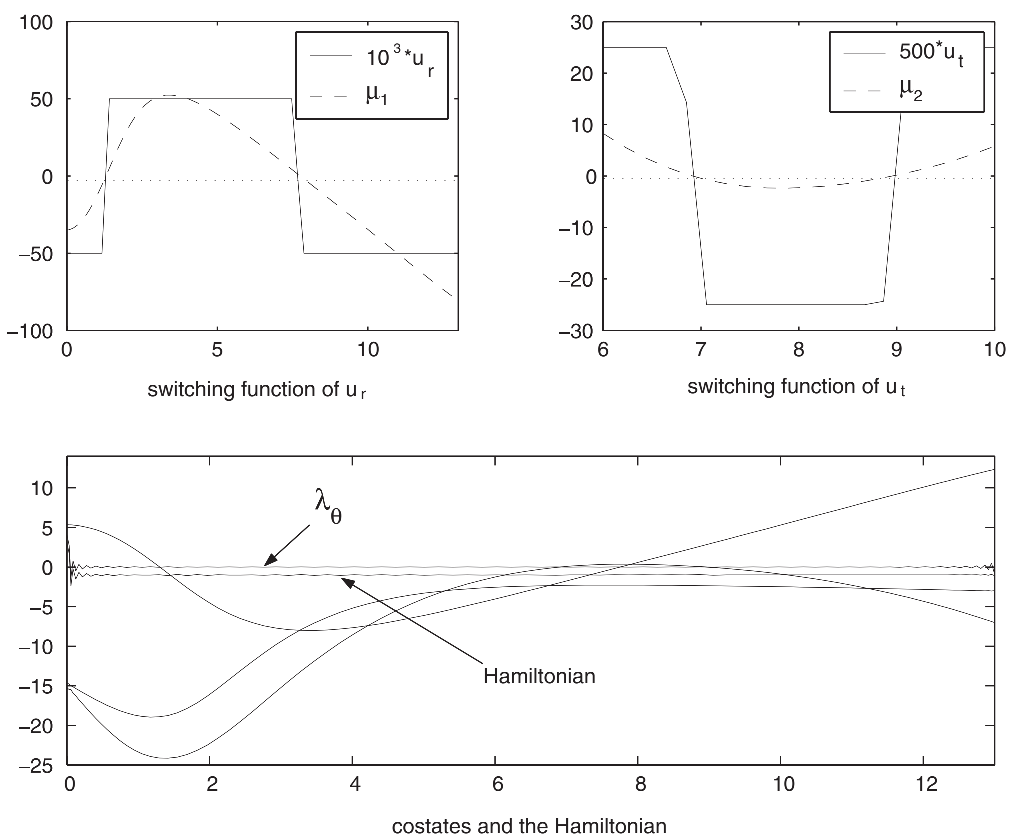

Figure 4. Verification of the necessary conditions.

control is not as impressive as the rate of solving problems with smooth solutions. Nevertheless, the convergence of the discrete solutions is extremely helpful because it provides key information about the point of discontinuity. A significant amount of work is ongoing to address the application and convergence rate of PS methods to both discontinuous solutions and discontinuous problems. All the ideas rely on the notion of PS knots [41, 42]. Based on the estimated location of the discontinuity, various mesh refinement techniques can be applied jointly with PS knotting methods to recover the fast convergence rate of smooth PS methods so that the accuracy of the approximate solution can be improved. In Figure 5 we plot out the control input obtained by using a PS knotting technique (with $N = 90$) together with a standard mesh refinement technique obtained by simply choosing a large number of nodes ($N = 200$). In the PS knotting technique, the smooth PS method is applied to each subinterval. The plot shows that the accuracy is improved by using PS knots with a much smaller number of nodes. How to analyse the rate of convergence of smooth and non-smooth PS methods for discontinuous control is an important issue that deserves further investigation, but is outside the scope of this paper.

The PS method is a robust approach for many optimal control problems. Interested readers are referred to [26] where it is shown by an example that the PS method converges for the problem while many other numerical optimal control methods fail to converge. The robustness

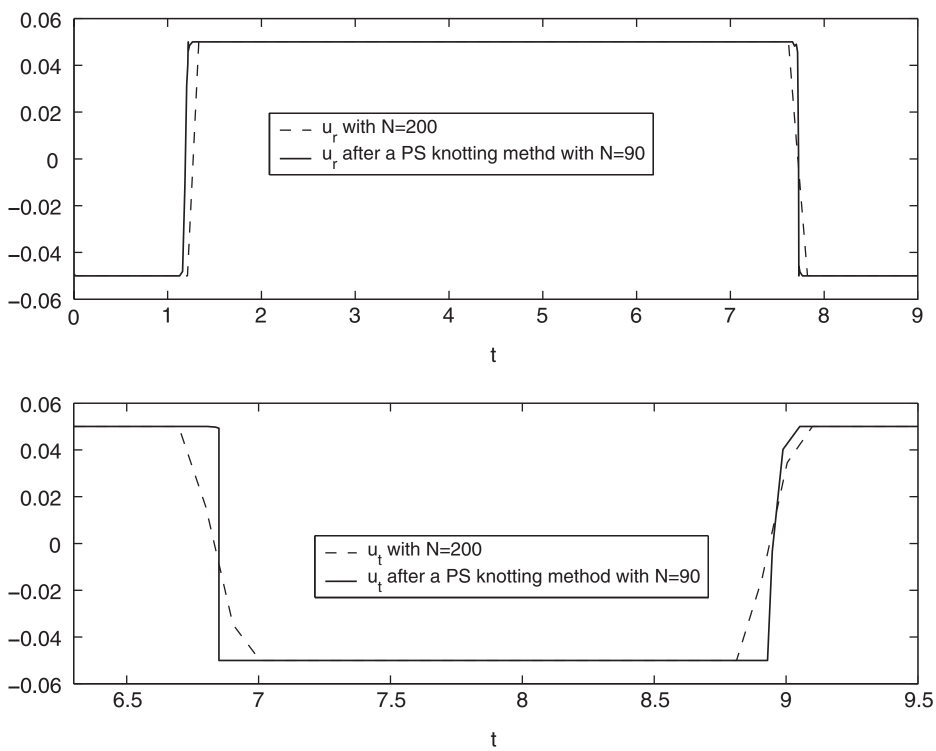

Figure 5. Optimal controls with and without the application of PS knots. Dashed lines are optimal solutions obtained by a 'brute force' mesh refinement technique.

of the PS approach continues to be independently verified numerically by many other studies [3, 4, 12] thus suggesting that a significant amount of theoretical analysis remains wide open for further study.

REFERENCES

1. Betts JT. *Practical Methods for Optimal Control Using Nonlinear Programming*. SIAM: Philadelphia, PA, 2001.
2. Betts JT. Survey of numerical methods for trajectory optimization. *Journal of Guidance, Control, and Dynamics* 1998; **21**(2):193–207.
3. Paris SW, Riehl JP, Sjauw WK. Enhanced procedures for direct trajectory optimization using nonlinear programming and implicit integration. *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Keystone, CO, 21–24 August 2006; *AIAA Paper No. 2006-6309*.
4. Riehl JP, Paris SW, Sjauw WK. Comparison of implicit integration methods for solving aerospace trajectory optimization problems. *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Keystone, CO, 21–24 August 2006; *AIAA Paper No. 2006-6033*.
5. Ross IM, Fahroo F. A perspective on methods for trajectory optimization. *Proceedings of the AIAA/AAS Astrodynamics Conference*, Monterey, CA, August 2002; *AIAA Paper No. 2002-4727*.
6. Elnagar G, Kazemi MA. Pseudospectral Chebyshev optimal control of constrained nonlinear dynamical systems. *Computational Optimization and Applications* 1998; **11**:195–217.
7. Elnagar G, Kazemi MA, Razzaghi M. The pseudospectral Legendre method for discretizing optimal control problems. *IEEE Transactions on Automatic Control* 1995; **40**:1793–1796.
8. Fahroo F, Ross IM. Costate estimation by a Legendre pseudospectral method. *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Boston, MA, 10–12 August 1998.
9. Fahroo F, Ross IM. Computational optimal control by spectral collocation with differential inclusion. *Proceedings of the 1999 Goddard Flight Mechanics Symposium*; *NASA/CP-1999-209235*, Greenbelt, MD, 185–200.
10. Fahroo F, Ross IM. Direct trajectory optimization by a Chebyshev pseudospectral method. *Journal of Guidance, Control and Dynamics* 2002; **25**(1):160–166.
11. Infeld SI, Murray W. Optimization of stationkeeping for a libration point mission. *AAS Spaceflight Mechanics Meeting*, Maui, HI, February 2004; AAS 04-150.
12. Lu P, Sun H, Tsai B. Closed-loop endoatmospheric ascent guidance. *Journal of Guidance, Control and Dynamics* 2003; **26**(2):283–294.
13. Rea J. Launch vehicle trajectory optimization using a Legendre pseudospectral method. *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Austin, TX, August 2003; *Paper No. AIAA 2003-5640*.
14. Stanton S, Proulx R, D'Souza C. Optimal orbit transfer using a Legendre pseudospectral method. *AAS/AIAA Astrodynamics Specialist Conference*, Big Sky, MT, 3–7 August 2003; AAS-03-574.
15. Williams P, Blanksby C, Trivailo P. Receding horizon control of tether system using quasilinearization and Chebyshev pseudospectral approximations. *AAS/AIAA Astrodynamics Specialist Conference*, Big Sky, MT, 3–7 August 2003; *Paper AAS 03-535*.
16. Paris SW, Hargraves CR. *OTIS 3.0 Manual*, Boeing Space and Defense Group, Seattle, WA, 1996.
17. Ross IM. User's manual for DIDO: a MATLAB application package for solving optimal control problems. *Technical Report 04-01.0*, Tomlab Optimization Inc, February 2004.
18. Canuto C, Hussaini MY, Quarteroni A, Zang TA. *Spectral Methods in Fluid Dynamics*. Springer: New York, 1988.
19. Ross IM, Fahroo F. Pseudospectral methods for optimal motion planning of differentially flat systems. *IEEE Transactions on Automatic Control* 2004; **49**(8):1410–1413.
20. Ross IM, Fahroo F. Issues in the real-time computation of optimal control. *Mathematical and Computer Modelling, An International Journal* 2006; **43**:1172–1188 (Special Issue: Optimization and Control for Military Applications).
21. Ross IM, Gong Q, Fahroo F, Kang W. Practical stabilization through real-time optimal control. *2006 American Control Conference*, Minneapolis, MN, June 2006.
22. Ross IM, Sekhavat P, Fleming A, Gong Q. Pseudospectral feedback control: foundations, examples and experimental results. *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Keystone, CO, August 2006; AIAA-2006-6354.
23. Fahroo F, Ross IM. Costate estimation by a Legendre pseudospectral method. *AIAA Journal of Guidance, Control and Dynamics* 2001; **24**(2):270–277.
24. Ross IM, Fahroo F. Legendre pseudospectral approximations of optimal control problems. *Lecture Notes in Control and Information Sciences*, vol. 295, Springer: New York, 2003; 327–342.
25. Gong Q, Ross IM, Kang W, Fahroo F. Convergence of pseudospectral methods for constrained nonlinear optimal control problems. *Proceedings of the IASTED International Conference on Intelligent Systems and Control*, Honolulu, HI, 2004; 209–214.

26. Gong Q, Kang W, Ross IM. A pseudospectral method for the optimal control of constrained feedback linearizable systems. *IEEE Transactions on Automatic Control* 2006; **51**(7):1115–1129.
27. Polak E. *Optimization*: *Algorithms and Consistent Approximations*. Springer: Heidelberg, 1997.
28. Vinter RB. *Optimal Control*. Birkhäuser: Boston, MA, 2000.
29. Ross IM, Fahroo F. A unified framework for real-time optimal control. *Proceedings of the IEEE Conference on Decision and Control*, Maui, December 2003.
30. Boyd JP. *Chebyshev and Fourier Spectral Methods* (2nd edn). Dover: New York, NY, 2001.
31. Trefethen LN. *Spectral Methods in MATLAB*. SIAM: Philadelphia, PA, 2000.
32. Boggs PT, Kearsley AJ, Tolle JW. A global convergence analysis of an algorithm for large-scale nonlinear optimization problems. *SIAM Journal of Optimization* 1999; **9**(4):833–862.
33. Freud G. *Orthogonal Polynomials*. Pergamon Press: Oxford, 1971.
34. Bryson AE, Ho YC. *Applied Optimal Control*. Hemisphere: New York, 1975.
35. Bryson AE. *Dynamic Optimization*. Addison-Wesley Longman, Inc., 1999.
36. Moyer HG, Pinkham G. Several trajectory optimization techniques. Part II: Applications. In *Computing Methods in Optimization Problems*, Balakrishnan AV, Neustadt LW (eds). Academic Press: New York, 1964; 91–105.
37. Wall B, Conway BA. Near-optimal low-thrust earth–mars trajectories via a genetic algorithm. *Journal of Guidance*, *Control and Dynamics* 2005; **28**(5):1027–1031.
38. Mengali G, Quarta AA. Fuel-optimal, power-limited rendezvous with variable thruster efficiency. *Journal of Guidance*, *Control and Dynamics* 2005; **28**(6):1194–1199.
39. Williams SN, Coverstone-Carroll V. Mars missions using solar electric propulsion. *Journal of Spacecraft and Rockets* 2000; **37**(5):71–77.
40. Ross IM, Gong Q, Sekhavat P. Low-thrust high-accuracy trajectory optimization. *Journal of Guidance*, *Control and Dynamics*, in press.
41. Ross IM, Fahroo F. Pseudospectral knotting methods for solving optimal control problems. *Journal of Guidance*, *Control and Dynamics* 2004; **27**(3):397–405.
42. Ross IM, Fahroo F. Discrete verification of necessary conditions for switched nonlinear optimal control systems. *Proceedings of the American Control Conference*, Boston, MA, June 2004.