J Optim Theory Appl (2017) 174:83–107
DOI 10.1007/s10957-016-0936-8

A Pseudospectral Method for Fractional Optimal Control Problems

Nastaran Ejlali[1] · **Seyed Mohammad Hosseini**[1]

Received: 27 July 2015 / Accepted: 9 April 2016 / Published online: 26 April 2016

Abstract In this article, a direct pseudospectral method based on Lagrange interpolating functions with fractional power terms is used to solve the fractional optimal control problem. As most applied fractional problems have solutions in terms of the fractional power, using appropriate characteristic nodal-based functions with suitable power leads to a more accurate pseudospectral approximation of the solution. The Lagrange interpolating functions and their fractional derivatives belong to the Müntz space; such functions are employed to show that a relationship exists between the Karush–Kukn–Tucker conditions associated with nonlinear programming and the first optimal necessary conditions. Furthermore, the convergence of the method is investigated. The obtained numerical results are an indication of the behavior of the algorithm.

Keywords Müntz basis Lagrange nodal function · Pseudospectral method · Fractional optimal control problems · Fractional power Lagrange functions · KKT conditions

Mathematics Subject Classification 26A33 · 49K05

1 Introduction

Optimal control problems emerge in a variety of fields such as science, aerospace, economics and finance. Numerical methods for solving the optimal control problems are divided into direct and indirect ones. The pseudospectral method is one of the

✉ Nastaran Ejlali
n.ejlali@modares.ac.ir

1 Department of Mathematics, Tarbiat Modares University, P.O. Box 14115-175, Tehran, Iran

most popular direct methods which has been applied by many researchers; see [1–6]. In these methods, the optimal control problem is reduced to a nonlinear programming (NLP) problem by discretizing the state and control variables. When the solution of the problem is smooth, pseudospectral approximations converge exponentially.

Fractional-order differential equations can be used to model many physical systems with more accuracy in relation to the integer-order equations due to their non-local properties [7]; hence, they have been further discussed and applied in the last few decades. Accordingly, since finding an analytical solution has its own difficulties and seems impossible, there has been a noticeable rise in the study and use of numerical methods; see [8].

Fractional optimal control problems are generalizations of classical optimal control problems; more precisely, fractional derivatives or integrals can appear in the performance index or constraints. The state of the art on fractional optimal control problems is given in some texts [9–11]. A first work in this field was conducted by Riewe [12] on the formulation of the fractional variational mechanics problem, in 1996.

Euler–Lagrange equations of fractional optimal control problems, which depend on Riemann–Liouville fractional derivatives, have been derived in [13], and a numerical method based on Legendre polynomials has been presented. Almedia and Torres [14] applied a Leitmann's direct method to solve fractional optimal control problems. In [15], they discretized the problem using a finite difference method. Pooseh et al. [16] employed Grünwald–Letnikov definition to solve fractional variational problems. Furthermore, Baleanu et al. [17] used a direct method and a modified Grünwald–Letnikov approach to approximate the Riemann–Liouville fractional derivative at some points. Tricaud and Chen [18,19] considered a large class of fractional optimal control problems and solved them by approximating the fractional operator. Biswas and Sen [20] presented a shooting-like numerical scheme for solving free final time fractional optimal control problems. In [21], the Legendre multiwavelet basis is used to solve the problem by an indirect collocation method.

In this paper, we use a direct method to approximate the solutions of the fractional constrained optimal control problems, in which their dynamics depend on Riemann–Liouville fractional derivatives. The aim of this paper is to use the pseudospectral method with special fractional powers belonging to a suitable Müntz space. The use of these functions is important for two reasons. First, we can obtain accurate solutions with a few number of Lagrange functions. Second, we can obtain the relation between the Karush–Kukn–Tucker (KKT) conditions and the first optimal necessary conditions.

The structure of the paper is as follows. In Sect. 2, the definitions of fractional integrals and its derivatives and some associated properties are presented which can be applied in the upcoming chapters. Section 3 is devoted to Müntz polynomials. In Sect. 4, the fractional optimal control problem is presented. In Sect. 5, the Müntz basis Lagrange nodal functions are introduced, and the differentiation matrix is evaluated. In Sect. 6, the pseudospectral method is described and it is shown that the KKT conditions are connected with discretized form of the first-order optimality conditions. Furthermore, the convergence of the method is discussed. In order to show the efficiency of the method, some numerical examples are considered in Sect. 7, and at the end, the final conclusions are presented in Sect. 8.

2 Some Definitions in Fractional Calculus

In this section, we present some definitions in fractional calculus [22], which are required throughout the paper.

Let $\Omega = [a, b]$ be a finite interval on the real line and $y(t)$ be an integrable function on $[a, b]$. The left and right Riemann–Liouville fractional integrals of order $\alpha \in \mathbb{C}$, $\Re(\alpha) > 0$, are defined, respectively, as follows:

$$ {}_aI_x^\alpha y(x) := \frac{1}{\Gamma(\alpha)} \int_a^x \frac{y(t)\mathrm{d}t}{(x-t)^{1-\alpha}}, \qquad (x > a), \tag{1} $$

$$ {}_xI_b^\alpha y(x) := \frac{1}{\Gamma(\alpha)} \int_x^b \frac{y(t)\mathrm{d}t}{(t-x)^{1-\alpha}}, \qquad (x < b). \tag{2} $$

The left and right Riemann–Liouville fractional derivatives of order $\alpha \in \mathbb{C}, \Re(\alpha) \geq 0$, are defined by

$$ {}_aD_x^\alpha y(x) := \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{\mathrm{d}x}\right)^n \int_a^x \frac{y(t)\mathrm{d}t}{(x-t)^{\alpha-n+1}}, \qquad (n = [\Re(\alpha)] + 1;\, x > a), $$

$$ {}_xD_b^\alpha y(x) := \frac{1}{\Gamma(n-\alpha)} \left(-\frac{d}{\mathrm{d}x}\right)^n \int_x^b \frac{y(t)\mathrm{d}t}{(t-x)^{\alpha-n+1}}, \qquad (n = [\Re(\alpha)] + 1;\, x < b). $$

Let $\Re(\alpha) \geq 0$, $y(t) \in AC^n([a, b])$ (absolutely continuous functions), and $n = [\Re(\alpha)] + 1$. Then, the left and right Caputo fractional derivatives of $y(t)$ of order α are defined as follows

$$ {}_a^cD_x^\alpha y(x) := \frac{1}{\Gamma(n-\alpha)} \int_a^x \frac{y^{(n)}(t)\mathrm{d}t}{(x-t)^{\alpha-n+1}}, \tag{3} $$

$$ {}_x^cD_b^\alpha y(x) := \frac{(-1)^n}{\Gamma(n-\alpha)} \int_x^b \frac{y^{(n)}(t)\mathrm{d}t}{(t-x)^{\alpha-n+1}}. \tag{4} $$

The following relations show the connection between Caputo and Riemann–Liouville fractional derivatives

$$ {}_a^cD_x^\alpha y(x) = {}_aD_x^\alpha y(x) - \sum_{k=0}^{n-1} \frac{y^{(k)}(a)}{\Gamma(k+1-\alpha)} (x-a)^{k-\alpha}, \tag{5} $$

$$ {}_x^cD_b^\alpha y(x) = {}_xD_b^\alpha y(x) - \sum_{k=0}^{n-1} \frac{y^{(k)}(b)}{\Gamma(k+1-\alpha)} (b-x)^{k-\alpha}. \tag{6} $$

When $0 < \Re(\alpha) < 1$, we have

$$ {}_a^cD_x^\alpha y(x) = {}_aD_x^\alpha y(x) - \frac{y(a)}{\Gamma(1-\alpha)} (x-a)^{-\alpha}, \tag{7} $$

$$ {}_x^cD_b^\alpha y(x) = {}_xD_b^\alpha y(x) - \frac{y(b)}{\Gamma(1-\alpha)} (b-x)^{-\alpha}. \tag{8} $$

Lemma 2.1 *Suppose that* $f(t) \in AC^n([a,b])$, $\Re(\alpha) > 0$, $n = [\Re(\alpha)] + 1$ *for* $\alpha \notin \mathbb{N}_0$, *and* $n = \alpha$ *for* $\alpha \in \mathbb{N}_0$; *then,*

$$ {}_0I_t^{\alpha}\, {}_0^cD_t^{\alpha} f(t) = f(t) - \sum_{j=0}^{n-1} \frac{f^{(j)}(a)}{j!}(t-a)^j, \tag{9}$$

$$ {}_tI_b^{\alpha}\, {}_t^cD_b^{\alpha} f(t) = f(t) - \sum_{j=0}^{n-1} \frac{(-1)^j f^{(j)}(b)}{j!}(b-t)^j. \tag{10}$$

If $\Re(\alpha) \in]0,1]$, *then*

$$ {}_aI_t^{\alpha}\, {}_a^cD_t^{\alpha} f(t) = f(t) - f(a), \quad {}_tI_b^{\alpha}\, {}_t^cD_b^{\alpha} f(t) = f(t) - f(b). \tag{11}$$

Lemma 2.2 *Let* $\Re(\alpha) > 0$ *and* $f(t)$ *be a continuous function on* $[a,b]$. *If* $\Re(\alpha) \in \mathbb{N}$ *or* $\alpha \in \mathbb{N}$, *then*

$$ {}_a^cD_t^{\alpha}\, {}_aI_t^{\alpha} f(t) = f(t), \quad {}_t^cD_b^{\alpha}\, {}_tI_b^{\alpha} f(t) = f(t). \tag{12}$$

The Caputo fractional differentiation and Riemann–Liouville fractional integration operators of the power functions $(x-a)^{\beta-1}$ and $(b-x)^{\beta-1}$ yield power functions of the same form. If $\Re(\alpha) > 0$ and $\beta \in \mathbb{C}$ ($\Re(\beta) > 0$), then

$$\left({}_a^cD_x^{\alpha}(t-a)^{\beta-1}\right)(x) = \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)}(x-a)^{(\beta-\alpha-1)}, \quad (\Re(\beta) > n), \tag{13}$$

$$\left({}_x^cD_b^{\alpha}(b-t)^{\beta-1}\right)(x) = \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)}(b-x)^{(\beta-\alpha-1)}, \quad (\Re(\beta) > n), \tag{14}$$

$$\left({}_aI_x^{\alpha}(t-a)^{\beta-1}\right)(x) = \frac{\Gamma(\beta)}{\Gamma(\beta+\alpha)}(x-a)^{(\beta+\alpha-1)}, \tag{15}$$

$$\left({}_xI_b^{\alpha}(b-t)^{\beta-1}\right)(x) = \frac{\Gamma(\beta)}{\Gamma(\beta+\alpha)}(b-x)^{(\beta+\alpha-1)}, \tag{16}$$

$$\left({}_0^cD_x^{\alpha}(t-a)^k\right)(x) = 0, \quad \left({}_x^cD_b^{\alpha}(b-t)^k\right)(x) = 0, \qquad k = 0, 1, \ldots n-1. \tag{17}$$

Finally, we state the integration by parts formula for the Riemann–Liouville fractional derivative and the fractional Taylor expansion in the Caputo sense; see, for example, [23]. Let $\alpha \in \mathbb{R}^+$, $p, q \geq 1$, and $\frac{1}{p} + \frac{1}{q} \leq 1 + \alpha$ ($p \neq 1$ and $q \neq 1$ in case when $\frac{1}{p} + \frac{1}{q} = 1 + \alpha$). If $f(t) \in {}_tI_b^{\alpha}(L_p)$ and $g(t) \in {}_aI_t^{\alpha}(L_q)$, then

$$\int_a^b f(t)\, {}_aD_t^{\alpha} g(t)\mathrm{d}t = \int_a^b g(t)\, {}_tD_b^{\alpha} f(t)\mathrm{d}t, \tag{18}$$

where the spaces of functions ${}_aI_t^{\alpha}(L_p)$ and ${}_tI_b^{\alpha}(L_p)$ are defined, for $1 \leq p \leq \infty$, by

$$ {}_aI_t^{\alpha}(L_p) := \{f :\ f = {}_aI_t^{\alpha}\phi,\ \phi \in L_p(]a,b[)\}, \tag{19}$$

$$ {}_tI_b^{\alpha}(L_p) := \{f :\ f = {}_tI_b^{\alpha}\phi,\ \phi \in L_p(]a, b[)\}. \tag{20} $$

Theorem 2.1 (Generalized Taylor's formula) *Let* ${}_a^cD_t^{i\alpha} f(t) \in C(]a, b])$ *for* $i = 0, 1, \ldots, N+1$ *and* $0 < \alpha \leq 1$. *Then,*

$$ f(t) = \sum_{k=0}^{N} \frac{({}_a^cD_t^{k\alpha} f)(a)}{\Gamma(k\alpha+1)}(t-a)^{k\alpha} + R_{N,a}(t), \tag{21} $$

where $R_{N,a}(t) = \frac{({}_a^cD_t^{(N+1)\alpha} f)(\xi)}{\Gamma((N+1)\alpha+1)}(t-a)^{(N+1)\alpha}$, *with* $a \leq \xi \leq t$, $\forall\ t \in]a, b]$, *and* ${}_a^cD_t^{N\alpha} := {}_a^cD_t^{\alpha}\, {}_a^cD_t^{\alpha} \cdots {}_a^cD_t^{\alpha}$ *(N-times).*

3 Müntz Systems and Müntz–Legendre Polynomials

Let $\Lambda := \{\lambda_0, \lambda_1, \lambda_2, \ldots\}$ be an increasing sequence of positive real numbers. Any element belonging to the space, $Span\{x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n}\}$ is called a Müntz polynomial or a Λ-polynomial. The set of all such polynomials is denoted by $M_n(\Lambda)$, that is, $M_n(\Lambda) := Span\{x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n}\}$. It is proved in the Müntz–Szász theorem [24] that, if $\lambda_0 = 0$, then Müntz polynomials are dense in $C([0, 1])$ in the uniform norm iff $\sum_{k=1}^{\infty} \frac{1}{\lambda_k} = +\infty$. Let $\Lambda_n := \{\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_n\}$ be a complex sequence, which its elements satisfy $Re(\lambda_k) > -1/2$ and $\lambda_k \neq \lambda_j$ $(k \neq j)$; then, the nth Müntz–Legendre polynomial for every $n = 1, 2, 3, \ldots$, is defined in power form as

$$ L_n(x; \Lambda_n) := \sum_{k=0}^{n} c_{nk} x^{\lambda_k}, \qquad c_{nk} := \frac{\prod_{\nu=0}^{n-1}(\lambda_k + \overline{\lambda}_\nu + 1)}{\prod_{\substack{\nu=0\\ \nu\neq k}}^{n}(\lambda_k - \lambda_\nu)}. \tag{22} $$

The orthogonality relation of the Müntz–Legendre polynomials is proved as

$$ \int_0^1 L_n(x)L_m(x)\mathrm{d}x = \frac{\delta_{m,n}}{1 + \lambda_n + \overline{\lambda}_n}, \tag{23} $$

for every $m, n = 0, 1, 2, \ldots$ [25].

If we consider $0 < \alpha \leq 1$ and set $\lambda_k = k\alpha$, $k = 0, 1, 2, \ldots, n$, then we have the space of Müntz polynomials $M_{n,\alpha} := Span\{1, x^{\alpha}, \ldots, x^{n\alpha}\}$. In order to evaluate the Müntz–Legendre polynomials, where $\lambda_k = k\alpha$, $k = 0, 1, 2, \ldots, n$, we use the following theorem which is proved in [26].

Theorem 3.1 *Let* $\alpha > 0$ *be a real number and* $t \in [0, 1]$; *then, the Müntz–Legendre polynomials can be represented as* $L_n(t; \alpha) = J_n^{(0,-1+\frac{1}{\alpha})}(2t^{\alpha} - 1)$, *where* $J_n^{(0,-1+\frac{1}{\alpha})}(t)$ *is a Jacobi polynomial.*

4 Fractional Optimal Control Problems

This section is devoted to fractional optimal control problems, in which the left Riemann–Liouville fractional derivative is included in the dynamic system. Problems are first stated and then their appropriate necessary optimality conditions are derived.

4.1 The Problem

Our goal is to solve the following fractional optimal control problem:

$$\min \quad J(x,u) := \int_0^1 F(\tau, x(\tau), u(\tau))\mathrm{d}\tau, \tag{24}$$

$$\text{s.t.} \qquad {}_0D_\tau^\alpha x(\tau) = G(\tau, x(\tau), u(\tau)), \quad x(0) = 0, \tag{25}$$

where $0 < \alpha \leq 1$, x and u are vectors with n and m components, respectively. $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $G : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, are continuously differentiable with respect to their arguments, and their gradients are Lipschitz continuous over the domain. The problem is considered on [0, 1] with homogeneous initial condition. One can easily transform a problem with non-homogeneous condition into a problem with homogeneous condition. The first-order optimality conditions derived by Agrawal [13] are

$$ {}_0D_\tau^\alpha x(\tau) = G(\tau, x(\tau), u(\tau)), \tag{26}$$

$$ {}_\tau D_1^\alpha \lambda(\tau) = H_x(\tau, x(\tau), u(\tau), \lambda(\tau)), \tag{27}$$

$$ 0 = H_u(\tau, x(\tau), u(\tau), \lambda(\tau)), \tag{28}$$

$$ x(0) = 0, \quad \lambda(1) = 0, \tag{29}$$

where $H(\tau, x, u, \lambda) = F(\tau, x, u) + \lambda^T G(\tau, x, u)$ is the Hamiltonian of problem (24)–(25) and $\lambda(\tau)$ is a Lagrange multiplier vector. Then, to establish a relationship between the KKT conditions and the discretized form of the first optimality conditions, we will consider the necessary conditions (26)–(29).

4.2 A Restatement of the Necessary Conditions (26)–(29)

Since the condition $\lambda(1) = 0$ holds, we can conclude, from (8), that the right Riemann–Liouville fractional derivative of $\lambda(t)$ is equal to its right Caputo fractional derivative. Taking the right Riemann–Liouville fractional integration of both sides of the costate equation, we get

$$ {}_tI_1^\alpha \, {}_t^cD_1^\alpha \lambda(t) = {}_tI_1^\alpha \; H_x(t, x(t), u(t), \lambda(t)).$$

Then, from Lemma 2.1, and since $\lambda(1) = 0$, we obtain

$$\lambda(t) = {}_tI_1^\alpha \; H_x(t, x(t), u(t), \lambda(t)).$$

Change of variable $t = 1 - \tau$, generates

$$\lambda(1-\tau) = {}_{(1-\tau)}I_1^{\alpha}\ H_x(\tau, x(\tau), u(\tau), \lambda(\tau)). \tag{30}$$

Taking the left Caputo fractional derivative of both sides of (30) results in

$${}_0^cD_{(1-\tau)}^{\alpha}\lambda(1-\tau) = {}_0^cD_{(1-\tau)}^{\alpha}\ {}_{(1-\tau)}I_1^{\alpha}\ H_x(\tau, x(\tau), u(\tau), \lambda(\tau)),$$

from which by Lemma 2.2 and change of variable, we obtain

$${}_0^cD_{(1-\tau)}^{\alpha}\lambda(1-\tau) = H_x(\tau, x(\tau), u(\tau), \lambda(\tau)). \tag{31}$$

This shows that we can replace Eq. (27) with Eq. (31) and rewrite the first-order optimality conditions as follows

$${}_0D_{\tau}^{\alpha}x(\tau) = G(\tau, x(\tau), u(\tau)), \tag{32}$$

$${}_0^cD_{(1-\tau)}^{\alpha}\lambda(1-\tau) = H_x(\tau, x(\tau), u(\tau), \lambda(\tau)), \tag{33}$$

$$0 = H_u(\tau, x(\tau), u(\tau), \lambda(\tau)), \tag{34}$$

$$x(0) = 0, \quad \lambda(1) = 0. \tag{35}$$

Hence, in the rest of the paper the new obtained conditions (32)–(35) are considered and applied as necessary conditions.

5 New Formulation of Pseudospectral Method

In this section, we construct fractional power Lagrange functions. Then, we use these functions to discretize the fractional optimal control problem.

Consider the set of $\{t_j\}_{j=0}^{N-1}$ of zeros of Jacobi polynomials $\{J_N^{a,b}\}$ with parameters $a = 0$, $b = -1 + \frac{1}{\alpha}$ on $[0, 1]$. The jth Lagrange polynomial, $L_j(t)$, of order $N - 1$ is defined by

$$L_j(t) := \prod_{\substack{i=0 \\ i \neq j}}^{N-1} \frac{t - t_i}{t_j - t_i}, \qquad j = 0, 1, \ldots, N-1.$$

We make the change of variable $t = \tau^{\alpha}$ and obtain the jth fractional power Lagrange function defined by

$$L_j^{\alpha}(\tau) := \prod_{\substack{i=0 \\ i \neq j}}^{N-1} \frac{\tau^{\alpha} - t_i}{t_j - t_i}, \qquad j = 0, 1, \ldots, N-1,$$

based on the distinct interpolating points $t_i = {\tau_i}^{\alpha}$, $i = 0, 1, \ldots, N-1$. It is worth noting that these functions preserve the property of Lagrange polynomials at the nodes $\{\tau_j\}_{j=0}^{N-1}$ or $\{t_j^{\frac{1}{\alpha}}\}_{j=0}^{N-1}$, in other words, they satisfy

$$L_j^\alpha(\tau) = \begin{cases} 1, & \tau = \tau_j, \\ 0, & \tau \neq \tau_j. \end{cases}$$

In order to impose the approximated state to satisfy the homogeneous initial condition, we add an extra node (0, 0) to the nodes of fractional power Lagrange functions. Then, we approximate the state variables $x(\tau)$ by fractional power Lagrange functions of degree at most $N\alpha$ as follows

$$x(\tau) \approx x^{N\alpha}(\tau) := \sum_{j=0}^{N-1} \bar{x}_j L_j^\alpha(\tau), \qquad \bar{x}_j := x^{N\alpha}(\tau_j), \quad j = 0, 1, \dots N-1,$$

in which $L_j^\alpha(\tau)$ is reexpressed as

$$L_j^\alpha(\tau) := \frac{\tau^\alpha}{t_j} \prod_{\substack{i=0 \\ i\neq j}}^{N-1} \left(\frac{\tau^\alpha - t_i}{t_j - t_i} \right). \tag{36}$$

In the same manner, to guarantee the condition $\lambda(1) = 0$, we add the node $(0, 0)$ to the nodes $\{(1-\tau_j)^\alpha\}_{j=0}^{N-1}$ and approximate the costate variables as

$$\lambda(1-\tau) \approx \lambda^{N\alpha}(1-\tau) := \sum_{j=0}^{N-1} \bar{\lambda}_j \bar{L}_j^\alpha(\tau), \qquad \bar{\lambda}_j := \lambda^{N\alpha}(\tau_j), \quad j = 0, 1, \dots N-1,$$

in which $\bar{L}_j^\alpha(\tau)$ is expressed as

$$\bar{L}_j^\alpha(\tau) := \left(\frac{\tau^\alpha}{(1-\tau_j)^\alpha} \right) \prod_{\substack{i=0 \\ i\neq j}}^{N-1} \left(\frac{\tau^\alpha - (1-\tau_i)^\alpha}{(1-\tau_j)^\alpha - (1-\tau_i)^\alpha} \right). \tag{37}$$

Theorem 5.1 *The Jacobi Gauss quadrature formula*

$$\int_0^1 p(t) w^{a,b}(t) dt \approx \sum_{j=0}^{N-1} w_j p(t_j),$$

is exact for any $p \in P_{2N-1}$ with the Jacobi Gauss nodes $\{t_j\}_{j=0}^{N-1}$, the zeros of $J_N^{a,b}(t)$, with the corresponding weights given by

$$w_j := \frac{\tilde{G}_{N-1}^{a,b}}{\left(1 - t_j^2\right) \left[\left(J_N^{a,b} \right)' (t_j) \right]^2}, \tag{38}$$

where $\tilde{G}_{N-1}^{a,b} := \frac{2^{a+b+1} \Gamma(N+a+1) \Gamma(N+b+1)}{N! \Gamma(N+a+b+1)}$.

Proof See [27]. □

Lemma 5.1 *Let* $\{J_N^{a,b}\}$ *be a sequence of orthogonal Jacobi polynomials with parameters* $a = 0$, $b = -1 + \frac{1}{\alpha}$, *and* $\{t_j\}_{j=0}^{N-1}$ *be the set of zeros of* $J_N^{a,b}(t)$. *Then, we have*

$$\int_0^1 p(\tau)d\tau = \frac{1}{\alpha 2^{\frac{1}{\alpha}}} \sum_{j=0}^{N-1} w_j p(\tau_j), \tag{39}$$

where $\tau_j = t_j^{\frac{1}{\alpha}}$, $j = 0, 1, \ldots, N-1$, $\{w_j\}_{j=0}^{N-1}$ *are the corresponding Jacobi quadrature weights, and* $p(\tau)$ *is a Müntz polynomial of order at most* $(2N-1)\alpha$, *that is,* $p(\tau) \in M_{2N-1,\alpha}$.

Proof Let $\{t_j\}_{j=0}^{N-1}$ and $\{w_j\}_{j=0}^{N-1}$ be the set of zeros and weights related to Jacobi polynomial $J_N^{a,b}(t)$ with $a = 0$, $b = -1 + \frac{1}{\alpha}$. It is known that the Gauss quadrature integration formula is exact for polynomials $q(t)$ with degree at most $2N-1$, that is,

$$\int_{-1}^1 q(t)(1+t)^{-1+\frac{1}{\alpha}} \mathrm{d}t = \sum_{j=0}^{N-1} w_j q(t_j).$$

Then, by change of variable, this relation reduces to (39). □

5.1 Fractional Differentiation Matrix

Now, we assume that $0 < \alpha \leq 1$ and obtain the left Riemann–Liouville fractional differentiation matrix. Consider the approximate solution and take the fractional derivative as

$$\begin{aligned} {}_0D_\tau^\alpha x^{N\alpha}(\tau) &= \sum_{j=0}^{N-1} \bar{x}_j \, {}_0D_\tau^\alpha \left(\frac{\tau^\alpha}{t_j} \prod_{\substack{i=0 \\ i \neq j}}^{N-1} \left(\frac{\tau^\alpha - t_i}{t_j - t_i} \right) \right) \\ &= \sum_{j=0}^{N-1} \bar{x}_j \frac{1}{t_j} \, {}_0D_\tau^\alpha \left(\tau^\alpha \prod_{\substack{i=0 \\ i \neq j}}^{N-1} \left(\frac{\tau^\alpha - t_i}{t_j - t_i} \right) \right). \end{aligned}$$

By putting $h_j(\tau) := \prod_{\substack{i=0 \\ i\neq j}}^{N-1} (\frac{\tau^\alpha - t_i}{t_j - t_i})$, $j = 0, 1, \ldots, N-1$, it becomes apparent that they are Müntz polynomials of order $(N-1)\alpha$; thus, it is possible to express these polynomials explicitly in terms of the Müntz–Legendre polynomials as follows

$$h_j(\tau) = \sum_{k=0}^{N-1} c_{kj} L_k(\tau;\alpha),$$

where $c_{ij} = \frac{2\alpha i+1}{\alpha 2^{\frac{1}{\alpha}}} w_j L_i(\tau_j;\alpha)$. Hence, we have

$$_0D_\tau^\alpha x^{N\alpha}(\tau) = \sum_{j=0}^{N-1} \bar{x}_j \frac{1}{t_j}\, {}_0D_\tau^\alpha(\tau^\alpha h_j(\tau)) = \sum_{j=0}^{N-1} \bar{x}_j \frac{1}{t_j} \sum_{k=0}^{N-1} c_{kj}\, {}_0D_\tau^\alpha(\tau^\alpha L_k(\tau;\alpha)). \tag{40}$$

According to Theorem 3.1, and the Jacobi polynomials with parameters $a = 0, b = -1 + \frac{1}{\alpha}$, the Müntz–Legendre polynomials can be expressed as

$$L_n(t;\alpha) = \sum_{m=0}^{n} \frac{(-1)^{(n-m)}\Gamma(1+b+n)\Gamma(1+a+b+n+m)}{m!(n-m)!\Gamma(1+b+m)\Gamma(1+a+b+n)} t^{m\alpha}. \tag{41}$$

Then, if we substitute (41) in (40), we get

$$\begin{aligned} &{}_0D_\tau^\alpha x^{N\alpha}(\tau) \\ &\quad = \sum_{j=0}^{N-1} \bar{x}_j \frac{1}{t_j} \sum_{k=0}^{N-1} c_{kj} \sum_{m=0}^{k} \frac{(-1)^{(k-m)}\Gamma(1+b+k)\Gamma(1+a+b+k+m)}{m!(k-m)!\Gamma(1+b+m)\Gamma(1+a+b+k)} \\ &\qquad {}_0D_\tau^\alpha(\tau^{(m+1)\alpha}). \end{aligned}$$

Using (7) and (13), we then obtain

$$_0D_\tau^\alpha x^{N\alpha}(\tau) = \sum_{j=0}^{N-1} \bar{x}_j \frac{1}{t_j} \sum_{k=0}^{N-1} c_{kj} \sum_{m=0}^{k} d_{km}(\tau^{m\alpha}), \tag{42}$$

where $d_{km} = \frac{(-1)^{(k-m)}\Gamma(1+b+k)\Gamma(1+a+b+k+m)\Gamma(m\alpha+\alpha+1)}{m!(k-m)!\Gamma(1+b+m)\Gamma(1+a+b+k)\Gamma(m\alpha+1)}$. Now, we obtain the fractional differentiation matrix, by evaluating (42) at the collocation points $\{\tau_i\}_{i=0}^{N-1}$, as

$$_0D_\tau^\alpha x^{N\alpha}(\tau)\Big|_{\tau=\tau_i} = \sum_{j=0}^{N-1} \bar{x}_j \left[\frac{1}{t_j} \sum_{k=0}^{N-1} c_{kj} \sum_{m=0}^{k} d_{km}(\tau_i^{m\alpha})\right] =: \sum_{j=0}^{N-1} \bar{x}_j D_{ij}^\alpha,$$

where D_{ij}^α denotes the ijth entry of the fractional differentiation matrix.

6 Direct Application of the Method to the Problem

We use Lemma 5.1 and approximate the performance index (24) as

$$J(x,u) := \int_0^1 F(\tau, x(\tau), u(\tau))\mathrm{d}\tau \approx \frac{1}{\alpha 2^{\frac{1}{\alpha}}} \sum_{j=0}^{N-1} w_j F(\tau_j, \bar{x}_j, \bar{u}_j). \tag{43}$$

Then, the discretization of the continuous-time fractional optimal control problem leads to the following nonlinear programming problem

$$\min \quad \bar{J}^N(\bar{X},\bar{U}) := \frac{1}{\alpha 2^{\frac{1}{\alpha}}}\sum_{i=0}^{N-1} w_i F(\tau_i,\bar{x}_i,\bar{u}_i), \tag{44}$$

$$\text{s.t.} \quad G(\tau_i,\bar{x}_i,\bar{u}_i) - \sum_{k=0}^{N-1}\bar{x}_k D^{\alpha}_{ik} = 0, \qquad i=0,1,\ldots,N-1, \tag{45}$$

where $\bar{X}=[\bar{x}_0,\ldots,\bar{x}_{N-1}]$ and $\bar{U}=[\bar{u}_0,\ldots,\bar{u}_{N-1}]$.
Now, the KKT conditions for (44)–(45) can be obtained by differentiating the Lagrangian with respect to the variables, $\bar{x}_i$ and $\bar{u}_i$, $i=0,1,\ldots,N-1$. The Lagrangian for the nonlinear programming (44)–(45) is obtained as

$$Ł(\lambda_i,\bar{x}_i,\bar{u}_i) := \sum_{i=0}^{N-1}\left(\lambda_i^T\left(G(\tau_i,\bar{x}_i,\bar{u}_i) - \sum_{k=0}^{N-1}\bar{x}_k D^{\alpha}_{ik}\right) + \frac{1}{\alpha 2^{\frac{1}{\alpha}}} w_i F(\tau_i,\bar{x}_i,\bar{u}_i)\right),$$

where λ_i, $i=0,1,\ldots,N-1$, is an n-tuple vector of Lagrange multipliers related to the collocation points. By differentiating the Lagrangian with respect to the variables, we obtain the following optimality conditions

$$\sum_{k=0}^{N-1}\lambda_k D^{\alpha}_{ki} = G_x^T(\tau_i,\bar{x}_i,\bar{u}_i)\lambda_i + \frac{1}{\alpha 2^{\frac{1}{\alpha}}} w_i F_x(\tau_i,\bar{x}_i,\bar{u}_i), \tag{46}$$

$$0 = G_u^T(\tau_i,\bar{x}_i,\bar{u}_i)\lambda_i + \frac{1}{\alpha 2^{\frac{1}{\alpha}}} w_i F_u(\tau_i,\bar{x}_i,\bar{u}_i), \qquad i=0,1,\ldots,N-1. \tag{47}$$

Now, we demonstrate that there is a very close relation between the discretization form of the first-order optimality conditions (33)–(34) and the optimal conditions (46)–(47). To this end, we set

$$\bar{\lambda}_i := \alpha 2^{\frac{1}{\alpha}}\frac{\lambda_i}{w_i}, \tag{48}$$

$$D^{\alpha^\dagger}_{ij} := \frac{1}{w_i} D^{\alpha}_{ji} w_j. \tag{49}$$

Then, the optimality conditions (46)–(47) can be rewritten as

$$\sum_{i=0}^{N-1}\bar{\lambda}_i D^{\alpha^\dagger}_{ij} = H_x(\bar{x}_i,\bar{u}_i,\bar{\lambda}_i), \tag{50}$$

$$0 = H_u(\bar{x}_i,\bar{u}_i,\bar{\lambda}_i). \tag{51}$$

Theorem 6.1 *The matrix $\boldsymbol{D}^{\alpha^\dagger}$, with entries $D^{\alpha^\dagger}_{ij}$ defined in* (49)*, is a differentiation matrix associated with the left fractional derivative in* (33)*.*

Proof Let $p(t)$ and $q(1-t)$ with $p(0)=q(1)=0$, be Müntz polynomials of degree at most $N\alpha$, that is, $p(t), q(1-t) \in M_{N\alpha}$. Let $\mathbf{R}^\alpha$ be the $N \times N$ differentiation matrix associated with the left fractional derivative that satisfies

$$(\mathbf{R}^\alpha \mathbf{q})_i = {}_0D^\alpha_{(1-\tau)} q(1-\tau)|_{\tau=\tau_i}, \qquad i = 0, 1, \ldots, N-1, \tag{52}$$

where $\mathbf{q}$ is an N-tuple vector with components $q(\tau_i)$, $i = 0, 1 \ldots, N-1$. Using the integration by parts given in (18) and the relation, ${}_\tau D^\alpha_1 q(\tau) = {}_0D^\alpha_{1-\tau} q(1-\tau)$, we have

$$\int_0^1 {}_0D^\alpha_\tau p(\tau)\, q(\tau)\mathrm{d}\tau = \int_0^1 p(\tau)\, {}_\tau D^\alpha_1 q(\tau)\mathrm{d}\tau = \int_0^1 p(\tau)\, {}_0D^\alpha_{1-\tau} q(1-\tau)\mathrm{d}\tau. \tag{53}$$

Then, the Lemma 5.1 gives

$$\sum_{i=0}^{N-1} w_i\, {}_0D^\alpha_\tau p(\tau)|_{\tau=\tau_i}\, q(\tau_i) \approx \sum_{i=0}^{N-1} w_i\, p(\tau_i)\, ({}_0D^\alpha_{1-\tau} q(1-\tau))|_{\tau=\tau_i}. \tag{54}$$

This approximation trivially has no error for $\alpha = 1$; however, for $0 < \alpha < 1$, its error is proportional to the numerical error of Müntz–Jacobi integral (53). By rewriting (54) in matrix form, the following relation is then obtained

$$(\mathbf{W}\mathbf{D}^\alpha \mathbf{p})^T \mathbf{q} \approx (\mathbf{W}\mathbf{p})^T \mathbf{R}^\alpha \mathbf{q},$$

where $\mathbf{p}$ and $\mathbf{q}$ are vectors with the components $p(\tau_i)$ and $q(\tau_i)$, respectively. Hence, we get

$$R^\alpha_{ij} \approx \frac{w_j}{w_i} D^\alpha_{ji}. \tag{55}$$

Therefore, the approximation error of differentiation matrix $\mathbf{D}^{\alpha^\dagger}$, related to left Caputo fractional derivative, is controlled directly by the numerical Müntz–Jacobi integral error in (54).

Thus, the following theorem has been proved. □

Theorem 6.2 *The KKT conditions, that is, the relations given in* (50)–(51) *together with* (45)*, represent a pseudospectral discretization form of the continuous system of fractional differential equations* (32)–(35).

6.1 Convergence of the Method

In this subsection, we investigate the feasibility of the discretized problem. An analysis of the convergence of the primal and a related discussion on the dual variables is given. We note that some authors have discussed the convergence of the pseudospectral methods for integer-order optimal control problems; see, for example, Kang et al. [28], Qi Gong et al. [29], Ruths et al. [30], Hou et al. [31].

In order to make sure that the discretized problem is feasible, the following relaxations can be offered

$$\|G(\tau_i, \bar{x}_i, \bar{u}_i) - \sum_{k=0}^{N-1} \bar{x}_k D_{ik}^{\alpha}\|_\infty \le \frac{m_1}{\Gamma(N\alpha+1)}, \qquad i = 0, 1, \dots, N-1, \tag{56}$$

where m_1 is a positive constant.

6.1.1 Feasibility

To prove the feasibility of the discretized problem, we state and prove the following Lemma.

Lemma 6.1 *Let* ${}_0^cD_t^{i\alpha} f(t) \in C(]0,1])$ *be bounded on* $[0,1]$ *for* $i = 0, 1, \dots, N+1$ *and* $0 < \alpha \le 1$*. Then, there is a Müntz polynomial of degree at most* $N\alpha$*,* $P_{N\alpha}(t) \in Span\{1, t^\alpha, \dots, t^{N\alpha}\}$*, such that*

$$\|f(t) - P_{N\alpha}(t)\|_\infty \le \frac{c_1}{\Gamma((N+1)\alpha+1)}, \tag{57}$$

where $c_1 = \|{}_0^cD_t^{(N+1)\alpha} f(t)\|_\infty$*.*

Proof Let $P_{N\alpha}(t)$ be best uniform approximation out of the space, $Span\{1, t^\alpha, \dots, t^{N\alpha}\}$, to $f(t)$. Then, we have

$$\|f(t) - P_{N\alpha}(t)\|_\infty \le \|f(t) - h(t)\|_\infty, \qquad \forall\, h(t) \in Span\{1, t^\alpha, \dots, t^{N\alpha}\}. \tag{58}$$

If we consider $h(t) = \sum_{k=0}^{N} \frac{({}_0^cD_t^{k\alpha} f)(0)}{\Gamma(k\alpha+1)} t^{k\alpha}$, by the generalized Taylor's formula, we have

$$\begin{aligned}\|f(t) - P_{N\alpha}(t)\|_\infty \le \|f(t) - h(t)\|_\infty &= \max_{t\in]0,1]} \left| \frac{\left({}_0^cD_t^{(N+1)\alpha} f\right)(\xi)}{\Gamma((N+1)\alpha+1)} t^{(N+1)\alpha} \right| \\ &\le \frac{c_1}{\Gamma((N+1)\alpha+1)},\end{aligned}$$

where $0 \le \xi \le t, \forall t \in]0,1]$ and $c_1 = \max_{\xi\in]0,1]} \left|({}_0^cD_t^{(N+1)\alpha} f)(\xi)\right|$. □

Theorem 6.3 *Given the feasible solution* (x, u) *of the fractional optimal control problem* (24)–(25)*, where* ${}_0^cD_t^{k\alpha} x(t) \in C(]0,1])$ *and is bounded on* $[0,1]$*, for* $k = 0, 1, \dots, N+1$*, then, the discretized problem* (44)–(45) *has a feasible solution* $(\bar{x}_i, \bar{u}_i)$*, with* $\bar{u}_i := u(\tau_i)$ *and*

$$\|x(\tau_i) - \bar{x}_i\|_\infty \le \frac{c_2}{\Gamma(N\alpha+1)},$$

where τ_i, $i = 0, 1, \ldots, N-1$ *are the zeros of Müntz–Legendre polynomial, and* $(0, 0)$ *is also added to the nodes.* c_2 *is a positive constant.*

Proof Let $P_{(N-1)\alpha}(t)$ be best approximation of ${}_0^cD_t^\alpha x(t)$ out of the space, $Span\{1, t^\alpha, \ldots, t^{(N-1)\alpha}\}$. According to Lemma 6.1, there is a positive constant l_1, such that

$$\|{}_0^cD_t^\alpha x(t) - P_{(N-1)\alpha}(t)\|_\infty \le \frac{l_1}{\Gamma(N\alpha + 1)}.$$

Now, we define $x^{N\alpha}(t)$ as follows

$$x^{N\alpha}(t) := {}_0I_t^\alpha\ P_{(N-1)\alpha}(t), \quad x^{N\alpha}(0) = 0, \quad \bar{x}_i := x^{N\alpha}(\tau_i), \quad \bar{u}_i := u(\tau_i)$$

where $i = 0, 1, \ldots, N-1$. According to (15), $x^{N\alpha}(t)$ is a Müntz polynomial of degree at most $N\alpha$. Moreover, we have

$$\begin{aligned}
\left|x(t) - x^{N\alpha}(t)\right| &= \left|{}_0I_t^\alpha\left({}_0^cD_t^\alpha x(t) - P_{(N-1)\alpha}(t)\right)\right| \\
&= \left|\frac{1}{\Gamma(\alpha)}\int_0^t \frac{{}_0^cD_\tau^\alpha x(\tau) - P_{(N-1)\alpha}(\tau)}{(t-\tau)^{(1-\alpha)}}\mathrm{d}\tau\right| \\
&\le \|{}_0^cD_\tau^\alpha x(\tau) - P_{(N-1)\alpha}(\tau)\|_\infty \left|\frac{1}{\Gamma(\alpha)}\int_0^t \frac{1}{(t-\tau)^{(1-\alpha)}}\mathrm{d}\tau\right| \\
&\le \frac{c_2}{\Gamma(N\alpha+1)}, \qquad \forall t \in [0, 1],
\end{aligned} \tag{59}$$

where $c_2 = \frac{2l_1}{\Gamma(\alpha+1)}$. As mentioned above, $x^{N\alpha}(t)$ is a Müntz polynomial of degree at most $N\alpha$, and as a result, we have

$$\sum_{k=0}^{N-1} \bar{x}_k D_{ik}^\alpha = {}_0^cD_t^\alpha x^{N\alpha}(t)\big|_{t=\tau_i}, \tag{60}$$

and

$$\begin{aligned}
&\left\|\sum_{k=0}^{N-1} \bar{x}_k D_{ik}^\alpha - G(\tau_i, \bar{x}_i, \bar{u}_i)\right\|_\infty \le \|{}_0^cD_t^\alpha x^{N\alpha}(t)\big|_{t=\tau_i} - {}_0^cD_t^\alpha x(t)\big|_{t=\tau_i}\|_\infty \\
&\quad + \|{}_0^cD_t^\alpha x(t)\big|_{t=\tau_i} - G(\tau_i, \bar{x}_i, \bar{u}_i)\|_\infty = \|P_{(N-1)\alpha}(\tau_i) - {}_0^cD_t^\alpha x(t)\big|_{t=\tau_i}\|_\infty \\
&\quad + \|G(\tau_i, x(\tau_i), u(\tau_i)) - G(\tau_i, \bar{x}_i, \bar{u}_i)\|_\infty \le \frac{l_1}{\Gamma(N\alpha+1)} + k_1\|x(\tau_i) - \bar{x}_i\|_\infty \\
&\quad \le \frac{m_1}{\Gamma(N\alpha+1)},
\end{aligned}$$

where $m_1 = l_1 + k_1c_2$, and k_1 is the Lipschitz constant of G with respect to x (G is Lipschitz continuous since it is continuously differentiable on a compact set).

Therefore, a feasible solution $(\bar{x}_i, \bar{u}_i)$ has been constructed for which, from (59), we conclude

$$\|x(\tau_i) - \bar{x}_i\|_\infty \leq \frac{c_2}{\Gamma(N\alpha + 1)}.$$

□

6.1.2 Convergence of the Primal Variables

We state the following definition from Qi Gong et al. [29].

Definition 6.1 A continuous function, $t \mapsto f(t) \in \mathbb{R}^n$, $t \in [0, 1]$, is called a uniform accumulation point of a function sequence, $\{t \mapsto f^N(t)\}_{N=0}^{\infty}$, $t \in [0, 1]$, if there exists a subsequence of $\{t \mapsto f^N(t)\}_{N=0}^{\infty}$ that uniformly converges to $t \mapsto f(t)$. Similarly, a point $v \in \mathbb{R}^n$ is called an accumulation point of a sequence $\{v^N\}_{N=0}^{\infty}$, if there exists a subsequence of $\{v^N\}_{N=0}^{\infty}$ that converges to v.

Assumption 6.1 It is assumed that the sequence $\{{}_0^cD_\tau^\alpha x^{N\alpha}(\cdot), u^N(\cdot)\}_{N=1}^{\infty}$ with $x^{N\alpha}(0) = 0$, has a uniform accumulation point, $(q(\cdot), u^\infty(\cdot))$. $q(t)$ and $u(t)$ are continuous on $[0, 1]$.

Let $(\bar{x}_i^*, \bar{u}_i^*)$, $i = 0, \ldots, N-1$, be an optimal solution of the discretized problem (44)–(45). Let $x^{N\alpha}(t) \in \mathbb{R}^n$ with $x^{N\alpha}(0) = 0$, be the interpolating Müntz polynomial of $(0, \bar{x}_0^*, \ldots, \bar{x}_{N-1}^*)$ and $u^N(t) \in \mathbb{R}^m$ be an interpolating function, but not necessarily a Müntz polynomial of $(\bar{u}_0^*, \ldots, \bar{u}_{N-1}^*)$.

Theorem 6.4 *Let $\{(\bar{x}_i^*, \bar{u}_i^*), i = 0, \ldots, N-1\}_{N=1}^{\infty}$ be the sequence of optimal solutions to the discretized problem and $\{(x^{N\alpha}(t), u^N(t)\}_{N=1}^{\infty}$ be the interpolating function sequence of $(0, \bar{x}_0^*, \ldots, \bar{x}_{N-1}^*)$ and $(\bar{u}_0^*, \ldots, \bar{u}_{N-1}^*)$, which satisfies Assumption 6.1. Then, $u^\infty(t)$ is an optimal control to the continuous problem and $x^\infty(t) = \frac{1}{\Gamma(\alpha)} \int_0^t \frac{q(\tau)}{(t-\tau)^{1-\alpha}} d\tau$ is an optimal state.*

Proof From Assumption 6.1, there is a subsequence $N_i \in 1, 2, \ldots$, with $\lim_{i\to\infty} N_i = \infty$, such that

$$\lim_{i\to\infty} ({}_0^cD_t^\alpha x^{N_i\alpha}(t), u^{N_i}(t)) = (q(t), u^\infty(t)).$$

Accordingly, we have $\lim_{i\to\infty} x^{N_i\alpha}(t) = x^\infty(t)$, because $x^{N_i\alpha}(0) = 0$, and ${}_0^cD_t^\alpha\, x^{N_i\alpha}(t)$ converges uniformly to $q(t)$. Our aim is to show that $(x^\infty(t), u^\infty(t))$ is a feasible solution to the original optimal control problem, the cost function $\bar{J}^{N_i}(\bar{X}^*, \bar{U}^*)$ converges to the continuous cost function $J(x^\infty, u^\infty)$, and $(x^\infty(t), u^\infty(t))$ is an optimal solution of the original problem.

1. In order to prove the feasibility of $(x^\infty(t), u^\infty(t))$ to the original problem, we show that

$$ {}_0D_t^\alpha x^\infty(t) - G(t, x^\infty(t), u^\infty(t)) = 0, \tag{61}$$

for any $t \in [0, 1]$. From (56), we have

$$\lim_{i\to\infty} ({}_0D_t^{\alpha} x^{N_i\alpha}(t)|_{t=t_j} - G(t_j, x^{N_i\alpha}(t_j), u^{N_i}(t_j)) = \lim_{i\to\infty} \frac{m_1}{\Gamma(N_i\alpha + 1)} = 0,$$

where t_j, $j = 0, \dots, N_i - 1$, are zeros of Müntz–Legendre polynomial of degree $N_i\alpha$. These nodes are dense in $[0, 1]$ as N_i tends to infinity; therefore, the last relation implies that (61) is true for all $t \in [0, 1]$.

2. Now, we prove that

$$\lim_{i\to\infty} \frac{1}{\alpha 2^{\frac{1}{\alpha}}} \sum_{j=0}^{N_i-1} w_j F(\tau_j, \bar{x}_j^*, \bar{u}_j^*) = \int_0^1 F(\tau, x^{\infty}(\tau), u^{\infty}(\tau))\mathrm{d}\tau. \tag{62}$$

From Lemma 5.1, we have

$$\int_0^1 F(t, x^{\infty}(t), u^{\infty}(t))\mathrm{d}t = \lim_{i\to\infty} \frac{1}{\alpha 2^{\frac{1}{\alpha}}} \sum_{j=0}^{N_i-1} w_j F(\tau_j, x^{\infty}(\tau_j), u^{\infty}(\tau_j)). \tag{63}$$

As shown before and from Assumption 6.1, $x^{N_i\alpha}(t)$ and $u^{N_i}(t)$ converge uniformly on $[0, 1]$, to $x^{\infty}(t)$ and $u^{\infty}(t)$, respectively. Since $\sum_{j=0}^{N_i-1} w_j = \alpha 2^{\frac{1}{\alpha}}$, we have

$$\lim_{i\to\infty} \left| \frac{1}{\alpha 2^{\frac{1}{\alpha}}} \sum_{j=0}^{N_i-1} w_j \big(F(\tau_j, x^{\infty}(\tau_j), u^{\infty}(\tau_j)) - F(\tau_j, \bar{x}_j^*, \bar{u}_j^*) \big) \right|$$
$$\leq \lim_{i\to\infty} K \frac{1}{\alpha 2^{\frac{1}{\alpha}}} \sum_{j=0}^{N_i-1} w_j (\|x^{\infty}(\tau_j) - \bar{x}_j^*\|_{\infty} + \|u^{\infty}(\tau_j) - \bar{u}_j^*\|_{\infty}) = 0,$$

where $K > 0$ is the Lipschitz constant of $F(\tau, x, u)$. Thus, (63) leads to

$$\lim_{i\to\infty} \bar{J}^{N_i}(\bar{X}^*, \bar{U}^*) = J(x^{\infty}, u^{\infty}).$$

3. Suppose that $(\hat{x}(t), \hat{u}(t))$ is an optimal solution of the continuous optimal control problem such that ${}_0^cD_t^{j\alpha}\hat{x}(t) \in C(]0, 1])$, and is bounded on $[0, 1]$, for $j = 0, 1, \dots, N_i+1$. By Theorem 6.3, there exists a sequence of feasible solutions, $(\hat{x}_k^N, \hat{u}_k^N)_{N=1}^{\infty}$, of the discretized problem that converges uniformly to $(\hat{x}(t), \hat{u}(t))$. Analogous to step 2, it can be shown that

$$J(\hat{x}, \hat{u}) = \lim_{i\to\infty} \bar{J}^{N_i}(\hat{X}, \hat{U}). \tag{64}$$

From (64) and the fact that $(\bar{x}_k^*, \bar{u}_k^*)$ and $(\hat{x}(t), \hat{u}(t))$ are optimal solutions, we then have

$$J(\hat{x}, \hat{u}) \leq J(x^{\infty}, u^{\infty}) = \lim_{i\to\infty} \bar{J}^{N_i}(\bar{X}^*, \bar{U}^*) \leq \lim_{i\to\infty} \bar{J}^{N_i}(\hat{X}, \hat{U}) = J(\hat{x}, \hat{u}).$$

Therefore, $(x^{\infty}(t), u^{\infty}(t))$ is an optimal solution to the continuous problem. □

6.1.3 Discussion on the Dual Variables

Remark 6.1 From the relation (55), we have

$$\left\| \frac{w_i}{\alpha 2^{\frac{1}{\alpha}}} \left(\sum_{j=0}^{N-1} \bar{\lambda}_j R_{ij}^{\alpha} - G_x^T(\tau_i, \bar{x}_i, \bar{u}_i)\bar{\lambda}_i - F_x(\tau_i, \bar{x}_i, \bar{u}_i) \right) \right\|_{\infty} \le \sigma_N, \qquad (65)$$
$$i = 0, 1, \ldots, N-1.$$

where σ_N is proportional to the numerical error of Müntz–Jacobi integral (53). Through some numerical experiments, we have come up with this impression that the relaxation bound σ_N for $\alpha \ge \frac{2}{3}$ tends to zero with order $\mathcal{O}(N^{-3})$. This is just a guess that needs to be investigated carefully; however, we will not go on with this investigation in this article.

Remark 6.2 It should be mentioned that the results obtained can be extended to the fractional optimal control problems with inequality constraints, $C(\tau, x(\tau), u(\tau)) \le 0$. In this case, the inequality constraints, $C(\tau_i, \bar{x}_i, \bar{u}_i) \le 0$, are added to the discretized problem (44)–(45). Considering γ_i as the KKT multipliers associated with the inequality constraints, and setting $\bar{\gamma}_i := \alpha 2^{\frac{1}{\alpha}} \frac{\gamma_i}{w_i}$, in the same manner as we did in (48), one can proceed to obtain the rest of the results.

7 Numerical Examples

Before implementing the method, we express the following algorithm which displays the steps of the numerical approach:

1. Transform the original problem into the problem with homogeneous initial value (if required).
2. Let $\{t_j\}_{j=0}^{N-1}$ be the N Gauss–Jacobi points in [0, 1], zeros of Jacobi polynomial of degree N with parameters $a = 0, b = -1 + \frac{1}{\alpha}$.
3. Construct the state approximation, by using the Lagrange functions (37).
4. Evaluate the system dynamic at the collocation points $\{\tau_j\}_{j=0}^{N-1}$, where $\tau_j = t_j^{\frac{1}{\alpha}}$, $j = 0, \ldots, N-1$.
5. Approximate the integral in the performance index, using Lemma 5.1.
6. Solve the NLP problem (44)–(45) (using, e.g., Fmincon in MATLAB or SNOPT in Gams).
7. Evaluate $\bar{\lambda}_0, \ldots, \bar{\lambda}_{N-1}$ from (48) .

Example 7.1 The following time-invariant example is considered by Agrawal [13]. The problem is to find the control $u(t)$ which minimizes the quadratic performance index

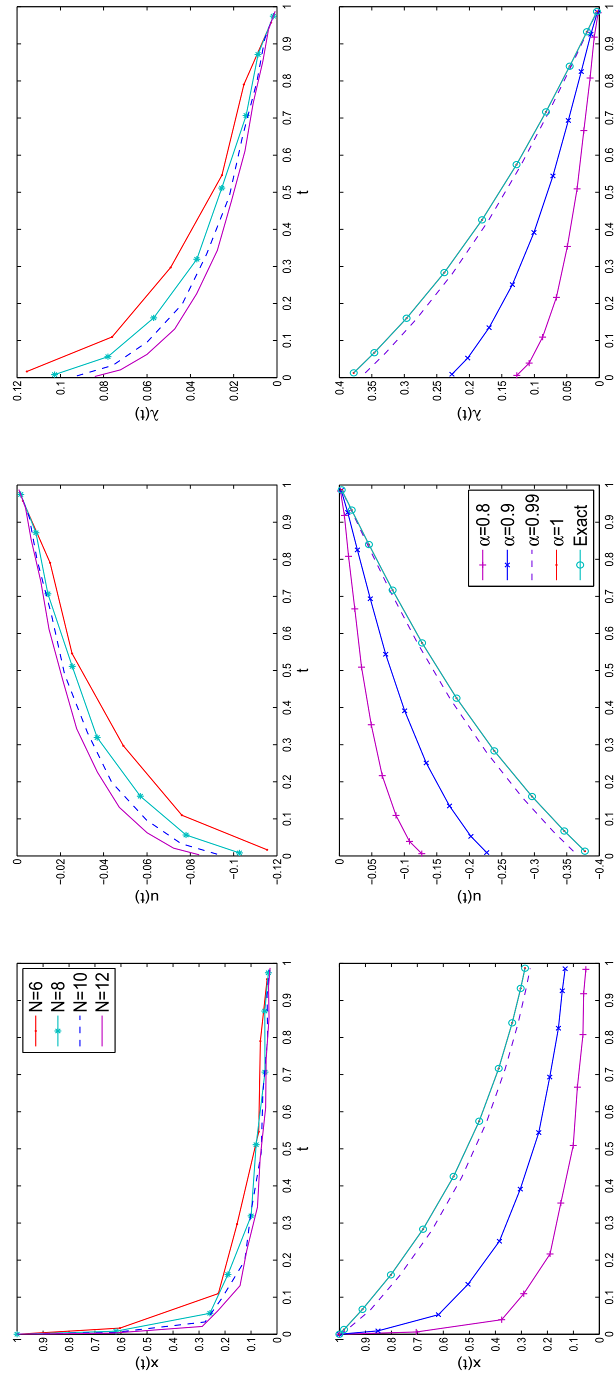

Fig. 1 Example 7.1: state, control and costate variables for $N = 6, 8, 10, 12$ with $\alpha = \frac{3}{4}$ (*upper-row*); state, control and costate variables for different α with $N = 10$ (*lower-row*)

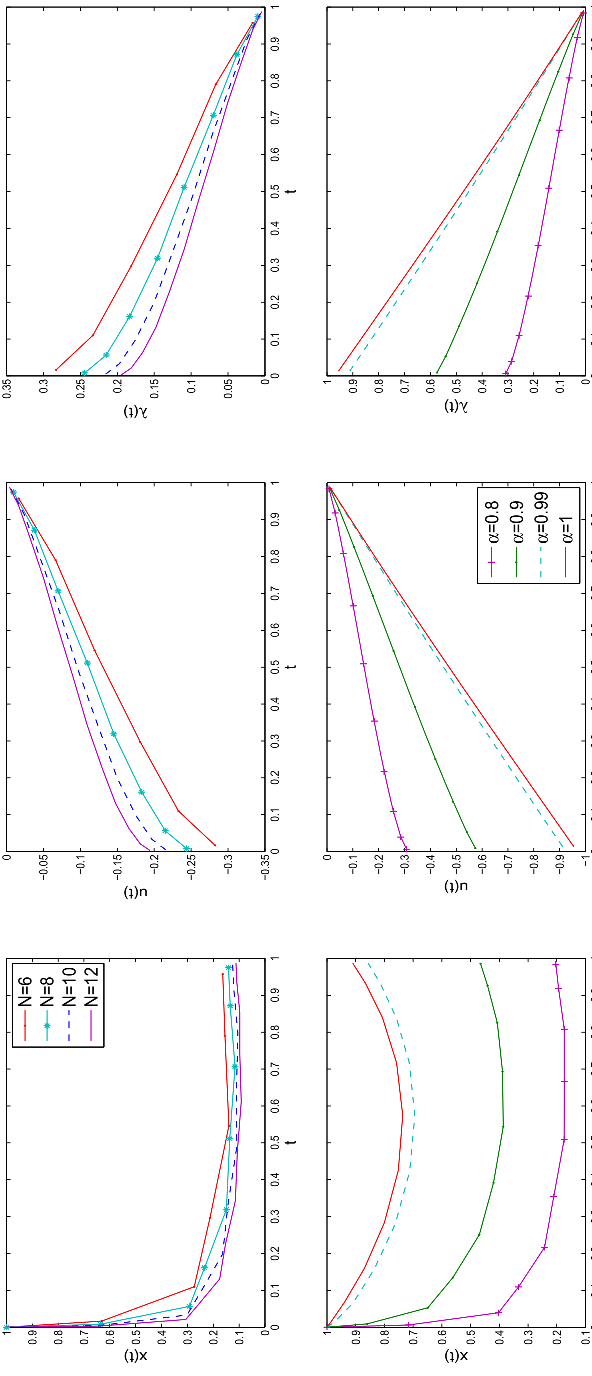

Fig. 2 Example 7.2: state, control and costate variables for $N = 6, 8, 10, 12$ with $\alpha = \frac{3}{4}$ (*upper-row*); state, control and costate variables for different α with $N = 10$ (*lower-row*)

Table 1 Error for the approximation of $x(t)$ for different values of α and N

N	$\alpha = \frac{1}{2}$	$\alpha = \frac{3}{4}$	$\alpha = \frac{7}{8}$	$\alpha = \frac{15}{16}$
2	0.01913570	0.00878672	0.00396563	0.00186817
4	9.476743×10^{-9}	$2.94217717 \times 10^{-4}$	2.0024896×10^{-4}	$1.09403399 \times 10^{-4}$
6	1.0505404×10^{-8}	$4.81442529 \times 10^{-5}$	3.8337106×10^{-5}	$2.22736042 \times 10^{-5}$
8	9.673289×10^{-9}	$1.34120798 \times 10^{-5}$	1.1846878×10^{-5}	$7.17323895 \times 10^{-6}$
10	7.627288×10^{-9}	$4.95808344 \times 10^{-6}$	4.7399761×10^{-6}	$2.96199192 \times 10^{-6}$

$$\begin{aligned} \min \ J(x,u) &:= \frac{1}{2}\int_0^1 [x^2(t) + u^2(t)]\mathrm{d}t, \\ \text{s.t.} \ \ {}_0D_t^{\alpha} x &= -x + u, \quad x(0) = 1. \end{aligned}$$

For $\alpha = 1$, the solution to this optimal control problem is $x(t) = \cosh(\sqrt{2}t) + \beta \sinh(\sqrt{2}t)$, and $u(t) = (1 + \sqrt{2}\beta)\cosh(\sqrt{2}t) + (\sqrt{2} + \beta)\sinh(\sqrt{2}t)$, where $\beta = -\frac{\cosh(\sqrt{2}) + \sqrt{2}\sinh(\sqrt{2})}{\sqrt{2}\cosh(\sqrt{2}) + \sinh(\sqrt{2})}$ [32]. Figure 1 (upper-row) shows the state, control and costate, for $N = 6, 8, 10, 12$ and $\alpha = \frac{3}{4}$. It is shown that the state, control and costate variables converge as N increases. Figure 1 (lower-row) shows the state, control and costate, for fixed $N = 10$ and different values of α. It can be seen that as α tends to 1, the numerical solutions of the problem approach to the solution of the case $\alpha = 1$.

Example 7.2 The second example is a linear time-variant which was again considered by Agrawal.

$$\begin{aligned} \min \ J(x,u) &:= \frac{1}{2}\int_0^1 [x^2(t) + u^2(t)]\mathrm{d}t, \\ \text{s.t.} \ \ {}_0D_t^{\alpha} x &= tx + u, \quad x(0) = 1. \end{aligned}$$

Similar to Example 7.1, Fig. 2 (upper-row) shows the state, control and costate, for $N = 6, 8, 10, 12$ and $\alpha = \frac{3}{4}$. We observe that the state, control and costate converge as N increases. Figure 2 (lower-row) shows that, for fixed $N = 10$ and as α tends to 1, the numerical solutions of the problem approach to the solution of the case $\alpha = 1$.

Example 7.3 The third example is considered as follows

$$\begin{aligned} &\min \ \ J(x,u) := \int_0^1 \left((u(t) - t)^2 + \left(x(t) - \frac{t^{\alpha+1}}{\Gamma(\alpha+2)} - \frac{t^{\alpha}}{\Gamma(\alpha+1)} \right)^2 \right) \mathrm{d}t, \\ &\text{s.t.} \ \ {}_0D_t^{\alpha} x(t) = u(t) + 1, \quad x(0) = 0. \end{aligned}$$

Exact solution to this problem is $x(t) = \frac{t^{\alpha+1}}{\Gamma(\alpha+2)} + \frac{t^{\alpha}}{\Gamma(\alpha+1)}$, and $u(t) = t$. Estimated solutions of the state variable for different values $N = 2, 4$ and $\alpha = \frac{1}{2}$ have been obtained as

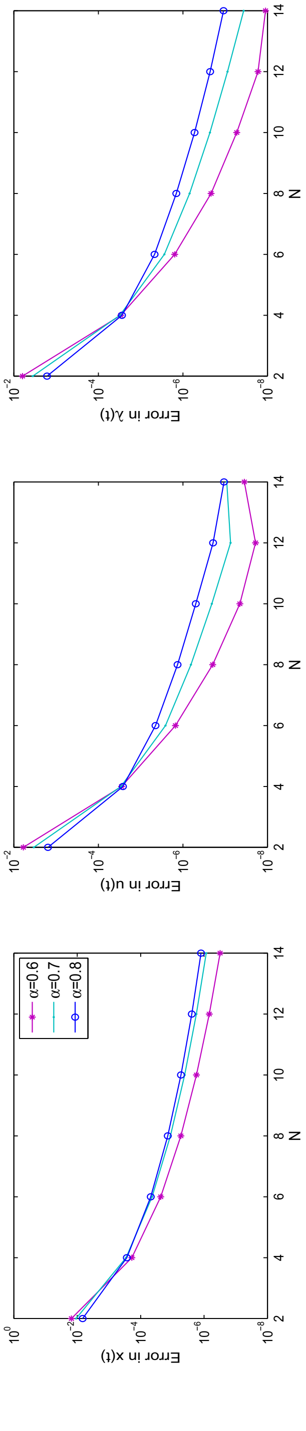

Fig. 3 Example 7.3: error in state, control and costate variables for $\alpha = 0.6, 0.7, 0.8$

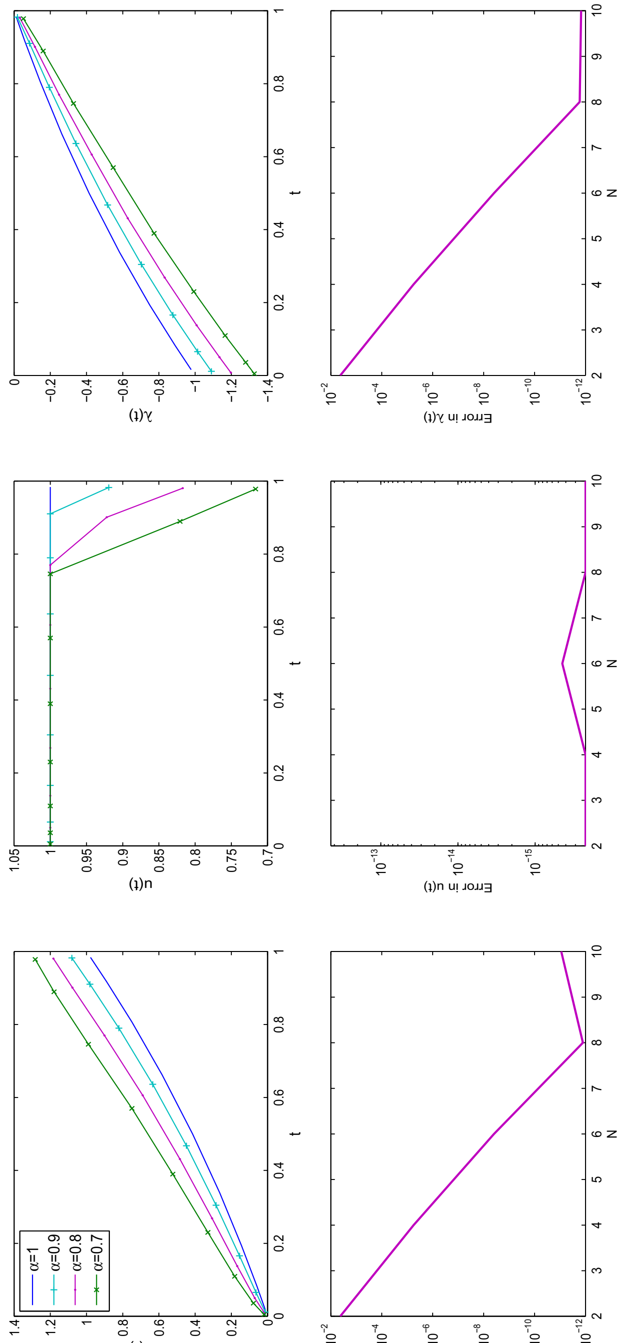

Fig. 4 Example 7.4: state, control and costate variables for different α with $N = 9$ (*upper-row*); error in state, control and costate variables for $\alpha = 1$ (*lower-row*)

$$x_{\text{est}}(t) = 0.79770769\sqrt{t} + 1.04662628t,$$
$$x_{\text{est}}(t) = 1.12837915\sqrt{t} + 6.69382928 \times 10^{-9}t + 0.75225277t^{\frac{3}{2}} + 1.05384331 \times 10^{-9}t^2,$$

respectively. On the other hand, the exact solution is $x_{\text{ex}}(t) = 1.12837916\sqrt{t} + 0.75225277t^{\frac{3}{2}}$. Now, clearly the obtained solutions for $N \geq 4$ are equal to the exact solution within nine digits of accuracy. However, it was expected because the presented method in this article is exact for Müntz polynomial solutions of degree $N\alpha$. In Table 1 and Fig. 3, the maximum error is shown for different values of α and N; it is shown that an accurate approximation has been obtained with $N = 10$.

Example 7.4 Consider the following example with inequality constraints

$$\begin{aligned}
min \quad & J(x,u) := \int_0^1 -ln(2)x(t)\mathrm{d}t, \\
s.t. \quad & {}_0D_t^{\alpha}x(t) = ln(2)(x(t) + u(t)), \\
& |u(t)| \leq 1, \quad x(t) + u(t) \leq 2, \quad x(0) = 0.
\end{aligned}$$

This example for $\alpha = 1$ is adopted from Elnager [33]. The exact solution of the problem for $\alpha = 1$ is $x(t) = \exp((\ln 2)t) - 1$, $u(t) = 1$, and $\lambda(t) = -2\exp((-\ln 2)t) + 1$. Figure 4 (upper-row) shows the state, control and costate, for $\alpha = 1, 0.9, 0.8, 0.7$ and $N = 9$. The numerical solutions of the problem approach to their corresponding solutions for $\alpha = 1$ as α approaches 1. Figure 4 (lower-row) shows the maximum error for different values of N and $\alpha = 1$.

8 Conclusions

A new direct pseudospectral method has been presented to solve fractional constrained optimal control problems. The Lagrange polynomials are transformed into the Lagrange functions, and zeros of fractional Jacobi polynomials are used as collocation points. It is shown that the KKT conditions of the NLP are connected with the continuous costate dynamics. This property confirms that the solution of the NLP and the solution of the first-order optimality conditions are asymptotically equivalent. This means that the direct and indirect solutions are approximately the same; hence, the presented method here essentially inherits the advantages of both direct and indirect methods.

Acknowledgements The authors would like to thank the Editor and anonymous reviewers for their helpful comments.

References

1. Elnagar, G., Kazemi, M., Razzaghi, M.: The pseudospectral Legendre method for discretizing optimal control problems. IEEE Trans. Autom. Control **40**, 1793–1796 (1995)

2. Fahroo, F., Ross, I.M.: Costate estimation by a Legendre pseudospectral method. J. Guid. Control Dyn. **24**, 270–277 (2001)
3. Benson, D.A., Huntington, G.T., Thorvaldsen, T.P., Rao, A.V.: Direct trajectory optimization and costate estimation via an orthogonal collocation method. J. Guid. Control Dyn. **29**, 1435–1440 (2006)
4. Garg, D., Patterson, M.A., Darby, C.L., Francolin, C., 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. **49**, 335–358 (2011)
5. Garg, D., Patterson, M.A., Hager, W.W., Rao, A.V., Benson, D.A., Huntington, G.T.: A unified framework for the numerical solution of optimal control problems using pseudospectral methods. Automatica **49**, 1843–1851 (2010)
6. Garg, D., Hager, W.W., Rao, A.V.: Pseudospectral methods for infinite-horizon optimal control problems. Automatica **47**, 829–837 (2011)
7. Podlubny, I.: Fractional Differential Equations. Academic Press, San Diego (1999)
8. Zayernouri, M., Karniadakis, G.: Fractional spectral collocation method. SIAM J. Sci. Comput. **36**, A40–A62 (2014)
9. Malinowska, A.B., Torres, D.F.M.: Introduction to the Fractional Calculus of Variations. Imperial College Press, London (2012)
10. Almeida, R., Pooseh, S., Torres, D.F.M.: Computational Methods in the Fractional Calculus of Variations. Imperial College Press, London (2015)
11. Malinowska, A.B., Odzijewicz, T., Torres, D.F.M.: Advanced Methods in the Fractional Calculus of Variations. Springer Briefs in Applied Sciences and Technology. Springer, New York (2015)
12. Riewe, F.: Nonconservative Lagrangian and Hamiltonian mechanics. Phys. Rev. E **53**, 1890–1899 (1996)
13. Agrawal, O.M.P.: A general formulation and solution scheme for fractional optimal control problems. Nonlinear Dyn. **38**, 323–337 (2004)
14. Almedia, R., Torres, D.F.M.: Leitmann's direct method for fractional optimization problems. Appl. Math. Comput. **217**, 956–962 (2010)
15. Almedia, R., Torres, D.F.M.: A discrete method to solve fractional optimal control problems. Nonlinear Dyn. **80**, 1811–1816 (2015)
16. Pooseh, S., Almeida, R., Torres, D.F.M.: A discrete time method to the first variation of fractional order variational functionals. Cent. Eur. J. Phys. **11**, 1262–1267 (2013)
17. Baleanu, D., Defterli, O., Agrawal, O.M.P.: A central difference numerical scheme for fractional optimal control problems. J. Vib. Control **15**, 583–597 (2009)
18. Tricaud, C., Chen, Y.Q.: An approximate method for numerically solving fractional order optimal control problems of general form. Comput. Math. Appl. **59**, 1644–1655 (2010)
19. Tricaud, C., Chen, Y.Q.: Solving fractional order optimal control problems in riots_95 a general purpose optimal control problem solver. In: Proceedings of the 3rd IFAC Workshop on Fractional Differentiation and its Applications, Ankara, Turkey (2008)
20. Biswas, R.K., Sen, S.: Free final time fractional optimal control problems. J. Frankl. Inst. **351**, 941–951 (2014)
21. Yousefi, S.A., Lotfi, A., Dehghan, M.: The use of a Legendre multiwavelet collocation method for solving the fractional optimal control problems. J. Vib. Control **17**, 1–7 (2011)
22. Kilbas, A.A., Srivastava, H.M., Trujillo, J.J.: Theory and Applications of Fractional Differential Equations. North-Holland Mathematics Studies, Amsterdam (2006)
23. Odibat, Z.M., Shawagfeh, N.T.: Generalized Taylor's formula. Appl. Math. Comput. **186**, 286–293 (2007)
24. Borwein, P., Erdélyi, T.: Polynomials and Polynomial Inequalities. Springer, New York (1995)
25. Borwein, P., Erdélyi, T., Zhang, J.: Müntz systems and orthogonal Müntz-Legendre polynomials. Trans. Am. Math. Soc. **342**, 523–542 (1994)
26. Esmaeili, S., Shamsi, M., Luchko, Y.: Numerical solution of fractional differential equation with a collocation method based on Müntz polynomials. Comput. Math. Appl. **62**, 918–929 (2011)
27. Shen, J., Tang, T., Wang, L.L.: Spectral Methods: Algorithms, Analysis and Applications. Springer, Berlin (2011)
28. Kang, W., Gong, Q., Ross, I.M., Fahroo, F.: On the convergence of nonlinear optimal control using pseudospectral methods for feedback linearizable systems. Int. J. Robust. Nonlinear Control **17**, 1251–1277 (2007)

29. Qi Gong, I., Ross, M., Kang, W.: Connection between the covector mapping theorem and convergence of pseudospectral methods for optimal control. Comput. Optim. Appl. **41**, 307–335 (2008)
30. Ruths, J., Zlotnik, A., Li Jr S.: Convergence of a Pseudospectral Method for Optimal Control of Complex Dynamical Systems. In: 50th IEEE Conference on Decision and Control. Orlando, FL, December, pp. 5553–5558 (2011)
31. Hou, H., Hager, W. W., Rao, A. V.: Convergence of a Gauss pseudospectral method for optimal control. In: AIAA Guidance, Navigation, and Control Conference and Exhibit. American Institute of Aeronautics and Astronautics, Minnesota, vol. 8, pp. 1–9 (2012)
32. Agrawal, O.P.: On a general formulation for the numerical solution of optimal control problems. Int. J. Control **50**, 627–638 (1989)
33. Elnager, G. N.: Legendre and pseudospectral Legendre approaches for solving optimal control problems. Ph.D. Thesis, Mississippi State University (1993)