EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
2026, Vol. 19, Issue 2, Article Number 7541
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Approximation Methods for Solving Fractional Optimal Control Problems Using Legendre Polynomials

Faizah A. H. Alomari[1,*], A. H. Qamlo[2]

[1] *Department of Mathematics, Faculty of Science, Al-Baha University, Al-Baha 65779, Saudi Arabia*
[2] *Department of Mathematics and Computer Science, Faculty of Science, Beni-Suef University, Beni-Suef 26511, Egypt*

Abstract. This paper develops a high-accuracy Legendre spectral method for solving a broad class of fractional optimal control problems governed by Caputo derivatives of order $0 < \alpha < 1$. The state and control variables are approximated by truncated Legendre expansions, and a new operational matrix for the Caputo derivative is constructed explicitly in the Legendre basis. By enforcing the fractional dynamics at Gauss–Legendre collocation nodes, the continuous optimal control problem is transformed into a finite-dimensional nonlinear optimization problem that preserves the structure of the original system. We establish well-posedness of the fractional state equation, prove existence and uniqueness of optimal controls, and derive a fractional Pontryagin maximum principle consistent with modern formulations of fractional variational theory. A detailed convergence analysis of the proposed spectral discretization is presented, showing algebraic rates for weakly regular data and spectral (exponential) convergence for analytic solutions. Several numerical experiments illustrate the efficiency and accuracy of the method. For smooth and analytic problems the convergence is essentially exponential, while for weakly singular solutions the method retains favorable algebraic decay. These results demonstrate the robustness of the Legendre spectral framework and its potential as an effective tool for the numerical solution of fractional optimal control problems.

2020 Mathematics Subject Classifications: 49M15, 65M70, 26A33

Key Words and Phrases: Fractional optimal control, Legendre spectral method, Caputo derivative

---

1. Introduction

The idea of extending differentiation and integration to non-integer orders goes back more than three centuries, to a famous 1695 letter in which Leibniz responded

*Corresponding Author.
DOI: https://doi.org/10.29020/nybg.ejpam.v19i2.7541

Email addresses: faali@bu.edu.sa (F. A. H. Alomari), Bahaa_gm@yahoo.com (G. M. Bahaa)

to L'Hôpital's question about the meaning of a derivative of order 1/2. Over the following two centuries, the subject evolved slowly, with important contributions by Abel, Liouville, Riemann, Grünwald, and others. Only in the second half of the twentieth century did *fractional calculus* begin to emerge as a mature mathematical discipline with a broad application scope. Modern monographs by Podlubny [1], Kilbas et al. [2], Diethelm [3], and Samko et al. [4] provided a rigorous functional-analytic framework and a unified treatment of fractional integrals and derivatives, laying the groundwork for subsequent developments in analysis, numerical methods, and applications.

Fractional differential equations (FDEs) are now recognized as powerful models for describing systems with memory, hereditary effects, nonlocal interactions, and anomalous transport. In contrast to classical integer-order models, FDEs can capture power-law relaxation, long-range temporal dependence, and non-Gaussian diffusion. Such behavior has been documented in viscoelasticity, bioengineering, complex materials, and signal processing [5, 6], as well as in anomalous transport and kinetic theory [7]. In diffusion-type processes, fractional operators appear naturally in the generalized Fokker–Planck and advection–diffusion equations, and have been used to model sub-diffusion, super-diffusion, and Lévy-flight dynamics [8].

Among the many definitions of a fractional derivative, the Caputo derivative has become particularly popular in engineering and control applications because it allows the use of classical (integer-order) initial conditions, which are directly interpretable in terms of physically measurable quantities [1, 3, 9]. Other kernels and operators, such as those of Riemann–Liouville, Grünwald–Letnikov, Atangana–Baleanu, and Caputo–Fabrizio, have been employed to capture different types of memory and regularity properties [6, 8].

With the increasing prevalence of FDEs in modeling, the need to *optimize* fractional dynamical systems has driven substantial interest in *fractional optimal control problems* (FOCPs). In a typical FOCP, one seeks a control function $u(\cdot)$ that minimizes a cost functional while the state $x(\cdot)$ evolves according to a fractional differential equation. Early works in this direction extended Pontryagin's maximum principle to the fractional setting. Agrawal developed general formulations of FOCPs and derived necessary optimality conditions for various types of fractional dynamics [10–12]. Baleanu and co-authors proposed Hamiltonian formulations and direct schemes [13, 14], while Almeida, Pooseh, and Torres [15, 16] gave a comprehensive overview of fractional variational principles and FOCPs, including discrete approximations and numerical algorithms.

In recent years, significant progress has been made in both the theory and computation of fractional optimal control problems. On the theoretical side, existence results, variational formulations, and fractional Pontryagin-type principles have been developed by several authors; see, for example, [10–12, 16, 17]. On the numerical side, a wide range of approximation strategies has been proposed, including direct transcription methods, pseudospectral and collocation techniques, wavelet-based methods, and operational-matrix approaches; see, e.g., [8, 18–22]. Recent contributions have also addressed more specialized fractional control settings, such as coupled systems, non-singular kernels, and distributed-order operators [23–26]. These developments collectively motivate the search for accurate and flexible high-order methods for FOCPs, especially methods that can accommodate nonlocal fractional dynamics while retaining strong approximation properties.

On the numerical side, both *indirect* (optimize-then-discretize) and *direct* (discretize-then-optimize) strategies have been explored. Indirect methods rely on solving the fractional state and adjoint equations derived from a fractional Pontryagin maximum principle [10, 12, 27, 28], which may become stiff and highly nonlocal. Direct methods, in contrast, discretize the state and control variables first and convert the FOCP into a finite-dimensional nonlinear programming (NLP) problem [16, 18, 29, 30]. Within this direct paradigm, a variety of approximation techniques have been proposed, including Adomian decomposition [31], homotopy and variational iteration methods [32], block-pulse functions and wavelets [33], and orthogonal polynomial operational matrices [29, 30, 34].

Related Work

The numerical solution of FDEs has been extensively studied. Finite-difference approaches based on Grünwald–Letnikov-type formulas and their variants [3] provide conceptually simple schemes but often yield low-order accuracy and dense coefficient matrices due to the nonlocality of fractional operators. For multi-dimensional or variable-order problems, such schemes may become prohibitively expensive [8]. Finite-element and finite-volume formulations have also been developed [35, 36], offering geometric flexibility at the cost of intricate assembly and often large-scale linear systems. Recent numerical developments also include truncated exponential and extended truncated exponential approaches for fractional optimal control and related variational problems; see, for example, [37–40]. Recent developments in fractional models and control problems also include studies on noninstantaneous impulsive Hilfer–Katugampola stochastic systems, Atangana–Baleanu fractional stochastic controllability, stochastic evolution inclusions with boundary control, and controllability of neutral delay Hilfer fractional stochastic integrodifferential systems; see, for example, [41–44].

Spectral and pseudospectral methods have attracted increasing interest because of their high-order (often exponential) convergence for smooth solutions. Chebyshev and Legendre spectral collocation schemes for fractional diffusion and advection–diffusion equations were investigated by Zayernouri and Karniadakis [45], who employed fractional Sturm–Liouville eigenfunctions and fractional spectral bases. The general theory of spectral methods based on orthogonal polynomials has been systematically developed by Shen, Tang, and Wang [46], and adapted to a variety of fractional models, including time-fractional diffusion equations [47]. Operational matrix techniques using Jacobi, Chebyshev, and Legendre polynomials have led to efficient schemes for high-order ODEs and FDEs [48–50].

In the specific context of FOCPs, several spectral and pseudospectral approaches have been proposed. Lotfi, Yousefi, and Dehghan used Legendre orthonormal bases combined with Gauss quadrature to solve classes of FOCPs [29, 30], while Yousefi et al. [19] introduced Legendre multiwavelet collocation for FOCPs. Sweilam and Al-Ajami constructed Legendre spectral collocation schemes for different types of FOCPs [20]. Tang et al. [21] developed integral fractional pseudospectral methods, and Tricaud and Chen [18] proposed a general approximation framework for fractional-order optimal control. Pseudospectral and collocation strategies based on Jacobi, Bernstein, Bessel, and wavelet bases have also been explored in [33, 34, 51], while spectral collocation methods for optimal control governed by fractional diffusion equations were studied in [22, 52, 53].

Recently, Legendre and shifted Legendre polynomials have been successfully applied to a variety of fractional models, including Hadamard fractional PDEs [54], nonlinear fractional Fredholm integro-differential equations [55], and fractional differential equations via shifted Legendre pseudospectral differentiation matrices [56]. In the FOCP setting, related developments include spectral collocation methods for fractional diffusion control problems [22, 52] as well as more recent studies of coupled and generalized fractional systems [25, 26].

The proposed Legendre spectral framework complements the existing literature in the following sense. Compared with low-order finite-difference methods, it is aimed at achieving higher accuracy for smooth solutions. Compared with wavelet-based and other pseudospectral approaches, it combines a Legendre-based operational matrix for the Caputo derivative with a direct collocation formulation of the FOCP. Thus, the method provides a complementary high-order alternative within the broader family of numerical methods for fractional optimal control problems. A short comparison with the existing literature is useful here. In contrast to low-order finite-difference and finite-element type discretizations, the present Legendre framework is designed to deliver high-order accuracy when the optimal state and control are sufficiently smooth. Compared with wavelet- and block-pulse-based approaches, it uses a global polynomial representation that integrates naturally with Gauss–Legendre quadrature and collocation. Relative to earlier pseudospectral and operational-matrix methods for FOCPs, the present formulation is intended to combine an explicit Legendre representation of the Caputo derivative with a direct collocation treatment of the state equation and a structure-preserving finite-dimensional optimization problem. In this sense, the proposed method should be viewed not as a replacement for existing approaches, but as a complementary high-order alternative that is especially attractive for smooth problems where spectral accuracy can be exploited. Despite these advances, several limitations remain:

- Many numerical schemes still rely on low-order finite-difference approximations, limiting accuracy and efficiency for smooth solutions.

- The nonlocal structure of fractional operators can lead to dense or ill-conditioned matrices, especially in time-discretization, which impacts scalability.

- Existing spectral FOCP schemes often treat the discretization of the fractional operator and the optimal control formulation separately, with limited analysis of the consistency between the continuous and discrete problems.

- Rigorous convergence and error estimates for the optimal control *pair* (x, u), particularly in fully discrete spectral FOCP frameworks, are still relatively scarce [53].

We emphasize, however, that the present operational-matrix formulation does not eliminate the nonlocal dense-matrix effect entirely; rather, its advantage lies in high-order accuracy and compact spectral representation for moderate discretization sizes. These observations motivate the development of methods that leverage the smoothness of the optimal solution, preserve the structure of the FOCP, and admit provable convergence properties.

Motivating Applications of FOCPs

FOCPs arise naturally in many applied fields where memory or hereditary effects play a central role:

- **Viscoelastic and rheological systems.** Fractional derivatives model stress–strain relations with power-law memory, providing more accurate descriptions of creep and relaxation than classical integer-order models [5, 6]. FOCPs appear in vibration suppression, damping design, and shape control for viscoelastic structures.

- **Diffusion, transport, and heat conduction.** Anomalous diffusion in heterogeneous media, sub-diffusive heat transport, and fractional Fokker–Planck dynamics are naturally described by fractional PDEs [8]. Here, the control may represent boundary fluxes, internal sources, or feedback laws for temperature, concentration, or density fields.

- **Biological, biomedical, and pharmacokinetic systems.** Fractional models capture anomalous diffusion in biological tissues, complex transport in porous media, and viscoelastic behavior of biological materials [5, 6]. Optimal control can be used, for example, to design dosing strategies or treatment protocols under fractional dynamics.

- **Robotics, actuators, and fractional-order control.** Fractional-order PID (FOPID) controllers often provide superior robustness and performance compared with classical PID [1, 57]. The tuning of FOPID gains can be posed as a FOCP, with performance indices involving tracking error, control effort, and robustness margins.

- **Economics and finance.** Long-memory effects and heavy-tailed waiting times in financial markets can be modeled using fractional dynamics [58]. In such settings, optimal portfolio allocation, risk management, or hedging strategies under fractional volatility dynamics lead to FOCPs.

- **Fractional diffusion systems and coupled processes.** Recent works have considered FOCPs for coupled time-fractional diffusion systems with final observations [24, 26], as well as for distributed-order and non-singular kernel operators [24–26, 59]. Accurate and efficient numerical techniques are essential in these high-dimensional and strongly coupled settings.

In all these applications, the fidelity of the fractional operator approximation and the efficiency of the numerical solver are crucial. High-order spectral methods are particularly attractive when the optimal state and control are smooth in time, as they can provide very small errors with relatively few degrees of freedom.

Statement of the Problem

In this work, we focus on time-fractional optimal control problems of Caputo type. Let $0 < \alpha < 1$ and $T > 0$ be given. We consider the FOCP

$$\begin{aligned} \min_{u \in \mathcal{U}} \quad & J[x,u] = \int_0^T L\big(t, x(t), u(t)\big)\, dt, && (1.1) \\ \text{subject to} \quad & {}^{C}D_t^{\alpha} x(t) = f\big(t, x(t), u(t)\big), \quad 0 < t \le T, && (1.2) \\ & x(0) = x_0, && (1.3) \end{aligned}$$

where ${}^{C}D_t^{\alpha}$ denotes the Caputo derivative of order α, $x : [0,T] \to \mathbb{R}^n$ is the state trajectory, $u : [0,T] \to \mathbb{R}^m$ is the control, and $\mathcal{U}$ is an admissible set of controls (possibly incorporating box or functional constraints). The functions $L : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $f : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are assumed to satisfy standard smoothness and growth conditions ensuring the existence of optimal pairs (x^*, u^*).

Our goal is to approximate the optimal pair (x^*, u^*) by means of a *Legendre spectral collocation* approach. More precisely, we seek to:

1. Construct accurate operational matrices for the Caputo fractional derivative with respect to a Legendre polynomial basis.
2. Represent the state and control by finite Legendre expansions on $[0,T]$ and enforce the dynamic constraint (1.2) at appropriate collocation nodes.
3. Derive a finite-dimensional nonlinear programming problem whose solution approximates the optimal control and state of the continuous FOCP (3.3)–(1.3).
4. Establish convergence and error estimates for the approximation, with particular emphasis on exploiting regularity of the optimal solution.

Although our primary focus is on single-term Caputo derivatives of order $\alpha \in (0,1)$, the methodology is designed to be flexible and can be extended to multi-term, distributed-order, or non-singular kernel operators as considered, for example, in [24, 25, 59, 60].

Contributions of the Paper

The main contributions of this paper can be summarized as follows:

1. **Legendre operational matrix for the Caputo derivative.** We derive an explicit operational matrix for the Caputo fractional derivative of order $0 < \alpha < 1$ in a scaled

Legendre basis on $[0, T]$. The Caputo derivative is adopted in this work because it allows the use of standard initial conditions with clear physical interpretation and is particularly well suited to fractional optimal control formulations studied in the existing literature. The construction is systematic and compatible with standard Gauss–Legendre nodes, and it can be incorporated in both linear and nonlinear FOCPs. This extends and unifies earlier operational-matrix approaches based on Jacobi, Bernstein, and Legendre polynomials [30, 49, 50, 61].

2. **High-order Legendre spectral discretization of FOCPs.** We propose a Legendre spectral collocation scheme in which both the state and control are approximated by truncated Legendre series, and the fractional dynamics are enforced at Gauss–Legendre collocation points. The resulting discrete system preserves the nonlocal structure of the fractional operator while achieving very high accuracy for smooth solutions, in line with spectral methods for FDEs [45–47].

3. **Structure-preserving finite-dimensional optimization problem.** By combining the Legendre approximation and the operational matrix of the Caputo derivative, we transform the FOCP (3.3)–(1.3) into a finite-dimensional constrained optimization problem. The resulting discrete system preserves the nonlocal structure of the original fractional operator, but the associated fractional matrices are generally dense. Consequently, the method is most attractive for problems in which high-order accuracy is desired with moderate polynomial degree N, rather than for very large-scale discretizations treated by generic dense solvers.

4. **Convergence and error analysis.** Assuming sufficient regularity of f and L, and smoothness of the optimal solution (x^*, u^*), we derive error bounds for the Legendre spectral approximation of the state and control. In particular, we show that:

   - for solutions with finite regularity, the approximation converges with a high algebraic rate;
   - for analytic optimal trajectories, the convergence is spectral (essentially exponential) with respect to the polynomial degree, consistent with classical spectral theory [46, 62, 63] and recent FOCP error analyses [53].

5. **Comprehensive numerical experiments.** We validate the proposed framework on a set of benchmark FOCPs, including linear-quadratic and nonlinear examples, as well as problems involving fractional diffusion-type dynamics. The experiments illustrate:

   - the accuracy of the Legendre spectral approximation compared with reference or semi-analytical solutions;
   - the error decay with respect to the number of Legendre modes;
   - the efficiency and robustness of the method relative to existing finite-difference and low-order schemes [20, 24, 26, 29].

The proposed methodology is general and can be extended to multi-dimensional FOCPs, coupled fractional diffusion systems, and problems with distributed-order operators or non-singular kernels, which have attracted increasing attention in the recent FOCP literature.

The remainder of this paper is organized as follows. In Section 2, we recall the basic definitions and fundamental properties of fractional integrals and derivatives. We also introduce the notation, function spaces, and preliminary results used throughout the paper, together with a brief review of Legendre polynomials and Gauss–Legendre quadrature rules. Section 3 is devoted to the formulation of the general fractional optimal control problem (FOCP) investigated in this study. In Section 4, we present a Legendre spectral collocation method for the numerical solution of the FOCP. Spectral methods based on orthogonal polynomials are known to be extremely powerful for differential equations with smooth solutions, due to their high-order—often exponential—convergence and excellent approximation properties. In Section 5, three numerical examples are provided to illustrate the accuracy, efficiency, and robustness of the proposed Legendre spectral collocation scheme. Finally, Section 7 concludes the paper and outlines possible directions for future research.

2. Preliminaries

In this section we recall the basic definitions and properties of fractional integrals and derivatives, introduce the notation and function spaces used throughout the paper, and summarize key facts about Legendre polynomials and Gauss–Legendre quadrature. We then review standard numerical approximations of fractional derivatives in order to place the proposed Legendre spectral approach in context.

2.1. Fractional integrals and derivatives

Let $[0,T]$ be a fixed time interval and $0<\alpha<1$. For a sufficiently smooth function $x:[0,T]\to\mathbb{R}$, the left-sided Riemann–Liouville fractional integral of order α is defined by

$$\big(I_{0+}^{\alpha}x\big)(t) := \frac{1}{\Gamma(\alpha)}\int_0^t (t-\tau)^{\alpha-1}x(\tau)\,d\tau, \qquad t\in[0,T], \tag{2.1}$$

where $\Gamma(\cdot)$ denotes the Gamma function. The corresponding Riemann–Liouville fractional derivative of order α is given by [1, 2, 4]

$$\big(D_{0+}^{\alpha}x\big)(t) := \frac{d}{dt}\big(I_{0+}^{1-\alpha}x\big)(t) = \frac{1}{\Gamma(1-\alpha)}\frac{d}{dt}\int_0^t (t-\tau)^{-\alpha}x(\tau)\,d\tau. \tag{2.2}$$

The left-sided Caputo fractional derivative of order α is defined by

$$\big({}^{C}D_{0+}^{\alpha}x\big)(t) := \big(I_{0+}^{1-\alpha}x'\big)(t) = \frac{1}{\Gamma(1-\alpha)}\int_0^t (t-\tau)^{-\alpha}x'(\tau)\,d\tau, \tag{2.3}$$

provided that x is absolutely continuous on $[0,T]$. In contrast to the Riemann–Liouville derivative, the Caputo derivative of a constant function is zero, which makes it particularly suitable for initial value problems with classical (integer-order) initial conditions [1, 3, 9].

The two derivatives are related by (see, e.g., [1, 9])

$$\left({}^{C}D^{\alpha}_{0+}x\right)(t) = \left(D^{\alpha}_{0+}x\right)(t) - \frac{x(0)}{\Gamma(1-\alpha)}t^{-\alpha}, \tag{2.4}$$

whenever x is sufficiently smooth. More generally, for higher orders $n-1 < \alpha < n$ one obtains additional terms involving $x^{(k)}(0)$ for $k = 0, \dots, n-1$ [2, 4, 9].

The fractional integral I^{α}_{0+} satisfies the semigroup (or composition) property

$$I^{\alpha}_{0+}I^{\beta}_{0+}x = I^{\alpha+\beta}_{0+}x, \qquad \alpha, \beta > 0, \tag{2.5}$$

and acts as a left inverse to the Caputo derivative in the sense that, under appropriate regularity,

$$I^{\alpha}_{0+}\left({}^{C}D^{\alpha}_{0+}x\right)(t) = x(t) - x(0), \qquad t \in [0,T]. \tag{2.6}$$

These identities will be used repeatedly in the analysis of the proposed discretization.

2.2. Function spaces and notation

We denote by $C([0,T])$ the space of continuous functions on $[0,T]$ endowed with the supremum norm $\|x\|_{\infty} = \max_{t\in[0,T]} |x(t)|$, and by $C^m([0,T])$ the space of m-times continuously differentiable functions. For $1 \le p \le \infty$ we denote by $L^p(0,T)$ the standard Lebesgue spaces with norm $\|\cdot\|_{L^p}$.

The Sobolev space $H^s(0,T)$ of order $s \ge 0$ is defined in the usual way via extension and Fourier transform, or equivalently by interpolation and completion of C^{∞} functions [46]. For vector-valued functions we use boldface notation, e.g., $x(t) \in \mathbb{R}^n$, and define norms componentwise.

We write $a \lesssim b$ to indicate that there exists a constant $C > 0$, independent of the discretization parameters, such that $a \le Cb$. The dependence of constants on parameters (e.g., α or T) will be stated explicitly when it is important.

2.3. Legendre polynomials and Gauss–Legendre collocation

Let $\{P_k\}_{k=0}^{\infty}$ denote the sequence of Legendre polynomials on $[-1,1]$, orthogonal with respect to the standard L^2 inner product,

$$\int_{-1}^{1} P_j(\xi)\, P_k(\xi)\, d\xi = \frac{2}{2k+1}\,\delta_{jk}, \qquad j,k \ge 0. \tag{2.7}$$

On a general interval $[0,T]$ we define the scaled Legendre polynomials

$$\phi_k(t) := P_k\left(\frac{2t}{T} - 1\right), \qquad t \in [0,T]. \tag{2.8}$$

Any sufficiently smooth function $x : [0, T] \to \mathbb{R}$ can be expanded formally as

$$x(t) \sim \sum_{k=0}^{\infty} a_k \, \phi_k(t), \tag{2.9}$$

with coefficients a_k determined by orthogonality. Truncation at degree N yields the spectral approximation

$$x_N(t) = \sum_{k=0}^{N} a_k \, \phi_k(t). \tag{2.10}$$

For numerical integration and collocation we employ Gauss–Legendre quadrature rules associated with the Legendre polynomials [46, 62, 63]. Let $\{t_j\}_{j=1}^{N+1}$ denote the Gauss–Legendre nodes on $[0, T]$ obtained by mapping the zeros of P_{N+1} from $[-1, 1]$ to $[0, T]$, and let $\{\omega_j\}_{j=1}^{N+1}$ be the corresponding positive quadrature weights. Then, for sufficiently smooth g,

$$\int_0^T g(t) \, dt = \sum_{j=1}^{N+1} \omega_j \, g(t_j) + \mathcal{O}\big(g^{(2N+2)}\big), \tag{2.11}$$

and the error decays spectrally if g is analytic [63, 64].

The use of Legendre polynomials and Gauss–Legendre nodes is particularly convenient for spectral collocation schemes: derivative and (fractional) derivative operators can be represented as matrices acting on the vector of nodal values or spectral coefficients. In this work we exploit this structure by constructing an operational matrix for the Caputo fractional derivative in the Legendre basis.

2.4. Approximation of fractional derivatives

The nonlocal nature of fractional derivatives poses a significant challenge in numerical approximations. In this subsection we briefly review several classes of approximation schemes that have been proposed in the literature, and highlight how the present Legendre spectral approach fits into this landscape.

2.4.1. Finite-difference and convolution-type methods

A widely used family of schemes for time-fractional derivatives is based on Grünwald–Letnikov-type discretizations and their variants. For a uniform grid $t_n = n\Delta t$, a left-sided Riemann–Liouville (or Caputo) derivative can be approximated by

$$\big({}^{C}D_{0+}^{\alpha} x\big)(t_n) \approx \frac{1}{\Delta t^{\alpha}} \sum_{k=0}^{n} w_k^{(\alpha)} \, x(t_{n-k}), \tag{2.12}$$

where the weights $w_k^{(\alpha)}$ are constructed from binomial coefficients [3]. Such L1- and L2-type schemes are straightforward to implement and enjoy first- or second-order accuracy in time, but they typically generate dense lower-triangular matrices and require $\mathcal{O}(N^2)$ operations and memory for N time steps [65].

To alleviate the computational burden, short-memory principles and fast convolution algorithms have been proposed [35]. These techniques truncate the convolution kernel beyond a certain lag or employ FFT-based acceleration, reducing the complexity while controlling the error due to memory truncation. In multi-dimensional settings and for variable-order derivatives, however, the implementation becomes more intricate [8].

2.4.2. Finite-element, finite-volume, and variational schemes

Finite-element and finite-volume formulations for space- and time-fractional diffusion-type problems have been developed in [35, 36]. These methods replace the fractional derivative by weak forms involving nonlocal bilinear forms, and often rely on quadrature for singular kernels. They provide geometric flexibility and a solid variational foundation, but the resulting algebraic systems can be large-scale and ill-conditioned, particularly for nonlocal operators with long-range interactions.

For FOCPs, finite-difference and finite-element approximations of the state and adjoint fractional equations have been combined with indirect optimality systems or direct transcription [10, 16, 18, 27]. While robust, the global accuracy is usually limited by the order of the underlying time discretization and the regularity of the solution.

2.4.3. Spectral and pseudospectral approximations

Spectral and pseudospectral methods approximate the solution by global basis functions (typically orthogonal polynomials or trigonometric functions) and enforce the governing equations at selected collocation points [46, 62, 63]. For fractional problems, several spectral approaches have been developed.

Zayernouri and Karniadakis [45] proposed fractional spectral collocation methods for linear and nonlinear FDEs, using specially constructed fractional Sturm–Liouville eigenfunctions and Jacobi-type bases adapted to endpoint singularities. For time-fractional diffusion equations, spectral and finite-difference/spectral hybrid schemes were analyzed in [47]. Spectral collocation methods based on Chebyshev or Jacobi polynomials, often combined with operational matrices, have proven effective for high-order ODEs and integral equations [48–50].

In the context of FOCPs, Legendre and Jacobi polynomial bases have been used to approximate both the state and control variables. Lotfi, Yousefi, and Dehghan [29, 30] constructed Legendre orthonormal bases and associated operational matrices to transform FOCPs into systems of algebraic equations, while Yousefi et al. [19] employed Legendre multiwavelets. Sweilam and Al-Ajami [20] developed Legendre spectral collocation methods for several classes of FOCPs, and Tang et al. [21] introduced integral fractional pseudospectral schemes. A spectral collocation method for FOCPs governed by fractional diffusion equations was presented in [52], and a posteriori error estimates for a spectral FOCP scheme were derived in [53].

More recently, Legendre and shifted Legendre polynomials have been applied to a variety of fractional models: Hadamard fractional PDEs [54], nonlinear fractional Fredholm integro-differential equations [55], and fractional differential problems via shifted Legendre pseudospectral differentiation matrices [56]. These works demonstrate the potential of Legendre-based spectral methods to deliver highly accurate solutions with relatively few degrees of freedom.

2.4.4. Operational matrices for fractional derivatives

A unifying theme in many spectral and pseudospectral schemes is the use of *operational matrices*, i.e., matrix representations of differentiation, integration, and fractional operators in a chosen basis. Given a basis $\{\psi_k\}_{k=0}^N$ and an approximation

$$x_N(t) = \sum_{k=0}^{N} a_k \psi_k(t), \tag{2.13}$$

one seeks a matrix $D_\alpha \in \mathbb{R}^{(N+1)\times(N+1)}$ such that

$${}^C D_{0+}^{\alpha} x_N(t) \approx \sum_{k=0}^{N} (D_\alpha \mathbf{a})_k \, \psi_k(t), \qquad \mathbf{a} = (a_0, \ldots, a_N)^\top. \tag{2.14}$$

For Jacobi, Bernstein, and Legendre polynomials, explicit constructions of such matrices for fractional integrals and derivatives have been proposed in [30, 49, 50]. These operational matrices convert the fractional derivative operator into a sparse or structured matrix, which can be incorporated efficiently into collocation or Galerkin formulations.

Position of the present work. In this paper we construct an explicit Legendre-based operational matrix for the Caputo fractional derivative on $[0, T]$, tailored to the FOCP setting described in Section 1 (see problem (3.3)–(1.3)). Compared with low-order finite-difference approximations [3] and generic polynomial bases, the proposed construction:

- exploits the orthogonality and approximation properties of Legendre polynomials [46, 62, 63];
- yields an explicit matrix representation of the Caputo derivative in the Legendre basis, which is convenient for collocation but, due to the nonlocal nature of the fractional operator, is generally dense;
- integrates naturally with Gauss–Legendre collocation and quadrature;
- and is specifically analyzed in the context of FOCPs, with convergence results for both the state and control, extending the existing error analyses in [24, 25, 30, 52, 53, 59].

These ingredients form the foundation of the numerical method developed and analyzed in the subsequent sections.

3. Formulation of the Fractional Optimal Control Problem

In this section we present the general form of the fractional optimal control problem (FOCP) considered in this work. Our goal is to establish a mathematically rigorous framework in which existence, uniqueness, and well-posedness of the state equation are guaranteed, and in which an optimal pair (x^*, u^*) is shown to exist. These results provide the theoretical foundation for the numerical method developed later.

3.1. Problem statement

Let $0 < \alpha < 1$ and consider the Caputo fractional dynamical system

$$ {}^{C}D^{\alpha}_{0+}x(t) = f\big(t, x(t), u(t)\big), \qquad t \in (0, T], \tag{3.1} $$

subject to the initial condition

$$ x(0) = x_0. \tag{3.2} $$

The admissible controls belong to a closed and convex set

$$ \mathcal{U}_{\text{ad}} := \big\{u \in L^p(0, T; \mathbb{R}^m) : u(t) \in U \subset \mathbb{R}^m \text{ a.e.}\big\}, $$

for some compact set U and $1 \le p \le \infty$. The state x belongs to the space $C([0, T]; \mathbb{R}^n)$ in view of the properties of the Caputo derivative and the fractional integral operator [1–3].

The performance index is given by

$$ J[x, u] = \int_0^T L\big(t, x(t), u(t)\big)\, dt, \tag{3.3} $$

where $L : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ is continuous and convex in u. The FOCP is to determine

$$ (x^*, u^*) = \arg\min_{u \in \mathcal{U}_{\text{ad}}} J[x, u] $$

subject to (3.1)–(3.3).

3.2. Assumptions

We adopt assumptions commonly used in fractional optimal control analysis (see [10, 16, 23–25]):

(A1) **Regularity of f.** The function $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ is continuous in all arguments and locally Lipschitz in x uniformly in (t, u):

$$ \|f(t, x_1, u) - f(t, x_2, u)\| \le L_f \|x_1 - x_2\|. $$

(A2) **Growth condition.** There exist constants $C_1, C_2 > 0$ such that

$$ \|f(t, x, u)\| \le C_1(1 + \|x\|) + C_2\|u\|. $$

(A3) **Regularity of L.** The function L is continuous in all variables and convex in u. Moreover, there exists $c > 0$ such that

$$|L(t, x, u)| \leq c(1 + \|x\|^2 + \|u\|^2).$$

(A4) **Admissible controls.** $\mathcal{U}_{\text{ad}}$ is nonempty, closed, and convex in $L^p(0, T)$.

Under these assumptions, the FOCP is well-defined.

3.3. Existence and uniqueness of the state equation

The fractional differential equation (3.1) can be rewritten in its equivalent Volterra integral form:

$$x(t) = x_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau, x(\tau), u(\tau))\, d\tau. \tag{3.4}$$

The integral equation (3.4) is nonlinear and of Volterra type. Existence and uniqueness follow from standard fixed-point arguments adapted to fractional kernels, as established in [1–3].

Theorem 1 (Existence and uniqueness of the state trajectory)**.** *Let $0 < \alpha < 1$ and let $u \in \mathcal{U}_{\text{ad}}$ be fixed. Assume that $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ satisfies:*

(A1) ***Continuity and Lipschitz in x.*** *f is continuous in all arguments and there exists $L_f > 0$ such that*

$$\|f(t, x_1, u) - f(t, x_2, u)\| \leq L_f \|x_1 - x_2\|$$

for all $t \in [0, T]$, $x_1, x_2 \in \mathbb{R}^n$, and $u \in U$.

(A2) ***Linear growth.*** *There exist constants $C_1, C_2 > 0$ such that*

$$\|f(t, x, u)\| \leq C_1(1 + \|x\|) + C_2\|u\|$$

for all $(t, x, u) \in [0, T] \times \mathbb{R}^n \times U$.

Then, for every initial condition $x_0 \in \mathbb{R}^n$ and every admissible control u, the Caputo fractional system

$$^{C}D^{\alpha}_{0+}x(t) = f\big(t, x(t), u(t)\big), \qquad t \in (0, T], \tag{3.5}$$

with

$$x(0) = x_0, \tag{3.6}$$

admits a unique solution $x \in C([0, T]; \mathbb{R}^n)$.

Proof. The proof proceeds in several steps.

Step 1: Equivalent Volterra integral equation.

For $0 < \alpha < 1$ and absolutely continuous x, the Caputo derivative can be written as

$$^{C}D^{\alpha}_{0+}x(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\tau)^{-\alpha} x'(\tau)\, d\tau.$$

It is a standard result (see, e.g., [1, 2]) that the initial value problem (3.5)–(3.6) is equivalent to the Volterra integral equation

$$x(t) = x_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f\big(\tau, x(\tau), u(\tau)\big)\, d\tau, \qquad t \in [0,T]. \tag{3.7}$$

Thus, it suffices to show that (3.7) has a unique continuous solution on $[0,T]$.

Step 2: Function space and fixed-point operator.

Let $X := C([0,T];\mathbb{R}^n)$ be the Banach space of continuous functions with values in $\mathbb{R}^n$ on $[0,T]$. Instead of the usual supremum norm, we work with a *weighted* supremum norm:

$$\|x\|_\lambda := \max_{0\le t\le T} e^{-\lambda t}\|x(t)\|,$$

where $\lambda > 0$ will be chosen later. It is straightforward to check that $(X, \|\cdot\|_\lambda)$ is a Banach space (the norm is equivalent to the standard $\|\cdot\|_\infty$).

Define the operator $\mathcal{T} : X \to X$ by

$$(\mathcal{T}x)(t) := x_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f\big(\tau, x(\tau), u(\tau)\big)\, d\tau. \tag{3.8}$$

A fixed point $x = \mathcal{T}x$ is precisely a solution of the integral equation (3.7).

We will show that, for a suitable choice of $\lambda > 0$, the operator $\mathcal{T}$ is a contraction on $(X, \|\cdot\|_\lambda)$.

Step 3: $\mathcal{T}$ maps X into X.

Take any $x \in X$. The map $t \mapsto f(t, x(t), u(t))$ is continuous on $[0,T]$ because f is continuous and x, u are measurable (and in many FOCP settings u is bounded or piecewise continuous). The kernel $(t-\tau)^{\alpha-1}$ is integrable on $0 \le \tau \le t \le T$ since $0 < \alpha < 1$.

Thus, the integral in (3.8) is continuous in t by standard dominated convergence arguments, so $\mathcal{T}x \in C([0,T];\mathbb{R}^n)$, i.e.,

$$\mathcal{T} : X \to X.$$

Step 4: Contraction estimate.

Let $x, y \in X$ and set $z := x - y$. We estimate $\|\mathcal{T}x - \mathcal{T}y\|_\lambda$.

Using the Lipschitz property (A1) of f, we have, for any $t \in [0,T]$,

$$\begin{aligned}
\|\mathcal{T}x(t) - \mathcal{T}y(t)\| &= \frac{1}{\Gamma(\alpha)} \left\| \int_0^t (t-\tau)^{\alpha-1} \big[f(\tau, x(\tau), u(\tau)) - f(\tau, y(\tau), u(\tau))\big]\, d\tau \right\| \\
&\le \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} \|f(\tau, x(\tau), u(\tau)) - f(\tau, y(\tau), u(\tau))\|\, d\tau \\
&\le \frac{L_f}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} \|x(\tau) - y(\tau)\|\, d\tau.
\end{aligned}$$

Multiply both sides by $e^{-\lambda t}$ and use the weighted norm:

$$\begin{aligned}
e^{-\lambda t}\|\mathcal{T}x(t)-\mathcal{T}y(t)\| &\le \frac{L_f}{\Gamma(\alpha)}\int_0^t (t-\tau)^{\alpha-1}e^{-\lambda t}\|x(\tau)-y(\tau)\|\,d\tau \\
&= \frac{L_f}{\Gamma(\alpha)}\int_0^t (t-\tau)^{\alpha-1}e^{-\lambda(t-\tau)}e^{-\lambda\tau}\|x(\tau)-y(\tau)\|\,d\tau \\
&\le \frac{L_f}{\Gamma(\alpha)}\int_0^t (t-\tau)^{\alpha-1}e^{-\lambda(t-\tau)}\,d\tau\ \|x-y\|_\lambda.
\end{aligned}$$

The last inequality uses the fact that

$$e^{-\lambda\tau}\|x(\tau)-y(\tau)\| \le \|x-y\|_\lambda \quad \text{for all } \tau\in[0,T].$$

Now, make the change of variables $s=t-\tau$ in the integral:

$$\int_0^t (t-\tau)^{\alpha-1}e^{-\lambda(t-\tau)}\,d\tau = \int_0^t s^{\alpha-1}e^{-\lambda s}\,ds \le \int_0^\infty s^{\alpha-1}e^{-\lambda s}\,ds = \lambda^{-\alpha}\Gamma(\alpha),$$

where we used the standard Gamma-function identity

$$\int_0^\infty s^{\alpha-1}e^{-\lambda s}ds = \frac{\Gamma(\alpha)}{\lambda^\alpha}, \qquad \lambda>0.$$

Therefore, for all $t\in[0,T]$,

$$e^{-\lambda t}\|\mathcal{T}x(t)-\mathcal{T}y(t)\| \le \frac{L_f}{\Gamma(\alpha)}\cdot\lambda^{-\alpha}\Gamma(\alpha)\,\|x-y\|_\lambda = \frac{L_f}{\lambda^\alpha}\|x-y\|_\lambda.$$

Taking the maximum over $t\in[0,T]$ yields

$$\|\mathcal{T}x-\mathcal{T}y\|_\lambda \le q\,\|x-y\|_\lambda, \qquad q:=\frac{L_f}{\lambda^\alpha}. \tag{3.9}$$

Now choose $\lambda>0$ such that $\lambda^\alpha>L_f$, for example

$$\lambda := (2L_f)^{1/\alpha}.$$

Then $q=L_f/\lambda^\alpha<1/2<1$ and (3.9) shows that $\mathcal{T}$ is a *strict contraction* on $(X,\|\cdot\|_\lambda)$.

Step 5: Application of Banach's fixed-point theorem.

Since $(X,\|\cdot\|_\lambda)$ is a complete metric space and $\mathcal{T}:X\to X$ is a contraction with constant $q<1$, Banach's fixed-point theorem guarantees the existence of a unique $x^*\in X$ such that

$$\mathcal{T}x^* = x^*.$$

By construction, x^* satisfies the integral equation (3.7) and hence the Caputo fractional differential equation (3.5) with initial condition (3.6).

Thus, there exists a unique continuous solution $x^*\in C([0,T];\mathbb{R}^n)$ to the state equation for the given control u.

Step 6: (Optional) Continuous dependence on the control.

For completeness, one can also show continuous dependence of x on u. Let $u_1, u_2 \in \mathcal{U}_{\text{ad}}$ and denote the corresponding solutions by x_1, x_2. Arguing as in Step 4, we obtain

$$\|x_1 - x_2\|_\lambda \le \frac{L_f}{\lambda^\alpha} \|x_1 - x_2\|_\lambda + C \|u_1 - u_2\|_{L^p(0,T)},$$

for a suitable constant $C > 0$ depending on the growth of f. Since $L_f/\lambda^\alpha < 1$, we can absorb the first term into the left-hand side, obtaining

$$\|x_1 - x_2\|_\lambda \le C' \|u_1 - u_2\|_{L^p(0,T)},$$

which expresses the continuity of the solution map $u \mapsto x$ from $\mathcal{U}_{\text{ad}}$ to $C([0, T])$.

Conclusion.

We have shown that the integral operator associated with the Caputo fractional system is a contraction on a suitable Banach space of continuous functions. Therefore, the state equation (3.5)–(3.6) has a unique continuous solution on $[0, T]$ for every admissible control u, which completes the proof.

3.4. Existence of an optimal control

Next, we establish that the FOCP admits at least one optimal pair (x^*, u^*). This builds on the direct method of the calculus of variations adapted to the fractional setting [10, 16, 24, 25].

Theorem 2 (Existence of an optimal control)**.** *Let $0 < \alpha < 1$ and consider the FOCP defined by (3.1)–(3.2) and the cost functional (3.3). Assume:*

(A1) $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ *is continuous and globally Lipschitz in x, uniformly in (t, u):*

$$\|f(t, x_1, u) - f(t, x_2, u)\| \le L_f \|x_1 - x_2\| \quad \forall\, x_1, x_2 \in \mathbb{R}^n.$$

(A2) (Linear growth) There exist $C_1, C_2 > 0$ such that

$$\|f(t, x, u)\| \le C_1(1 + \|x\|) + C_2\|u\| \quad \forall\, (t, x, u) \in [0, T] \times \mathbb{R}^n \times U.$$

(A3) $L : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ *is continuous in all arguments, convex in u, and satisfies*

$$|L(t, x, u)| \le c_1(1 + \|x\|^2) + c_2\|u\|^2$$

for some $c_1, c_2 > 0$.

(A4) The admissible control set is

$$\mathcal{U}_{\text{ad}} := \{u \in L^2(0, T; \mathbb{R}^m) : u(t) \in U \subset \mathbb{R}^m \ a.e.\},$$

where U is a nonempty, compact, convex subset of $\mathbb{R}^m$.

Moreover, assume:

(E1) For each (t, x), the map $u \mapsto f(t, x, u)$ is continuous and affine (or at least linear) in u.

(E2) For each (t, x), the map $u \mapsto L(t, x, u)$ is convex.

Then there exists at least one optimal pair (x^*, u^*) with $u^* \in \mathcal{U}_{\text{ad}}$ and $x^* \in C([0, T]; \mathbb{R}^n)$ solving (3.1)–(3.2)*, such that*

$$J[x^*, u^*] = \inf_{u \in \mathcal{U}_{\text{ad}}} J[x, u].$$

Proof. The proof follows the direct method of the calculus of variations adapted to the fractional setting.

Step 1: Well-posedness of the state equation for each u.

By Theorem 1, assumptions (A1)–(A2) guarantee that for every $u \in \mathcal{U}_{\text{ad}}$ the Caputo system (3.1)–(3.2) admits a unique state $x \in C([0, T]; \mathbb{R}^n)$. Hence the cost functional $J[x, u]$ is well-defined on $\mathcal{U}_{\text{ad}}$.

We may therefore define the *reduced* cost functional

$$\mathcal{J}(u) := J[x(u), u], \qquad u \in \mathcal{U}_{\text{ad}},$$

where $x(u)$ denotes the unique state associated with u.

Step 2: Existence of a minimizing sequence and boundedness in L^2.

Define

$$J_{\text{inf}} := \inf_{u \in \mathcal{U}_{\text{ad}}} \mathcal{J}(u) \in \mathbb{R} \cup \{-\infty\}.$$

Under (A3), L has at most quadratic growth in (x, u) and is bounded from below (e.g., by $-C(1 + \|x\|^2 + \|u\|^2)$), while the state $x(u)$ is continuous and satisfies a linear growth estimate with respect to $\|u\|_{L^2}$ (see below). Thus $J_{\text{inf}} > -\infty$.

Let $\{u_k\} \subset \mathcal{U}_{\text{ad}}$ be a minimizing sequence:

$$\lim_{k \to \infty} \mathcal{J}(u_k) = J_{\text{inf}}.$$

We claim that the sequence $\{u_k\}$ is bounded in $L^2(0, T; \mathbb{R}^m)$. Indeed, by (A3) and the linear growth of the state (from the integral formulation and (A2)), there exist constants $C, C' > 0$ such that

$$\|x_k(t)\| \le C\Big(1 + \int_0^t (t - \tau)^{\alpha - 1}(1 + \|x_k(\tau)\| + \|u_k(\tau)\|)\, d\tau\Big),$$

where $x_k := x(u_k)$. Using a fractional Grönwall inequality (see [1–3]), we obtain

$$\|x_k\|_{C([0,T])} \le C_0\Big(1 + \|u_k\|_{L^2(0,T)}\Big),$$

for some constant $C_0 > 0$ independent of k.

Then, from (A3),

$$\begin{aligned}\mathcal{J}(u_k) &= \int_0^T L\big(t, x_k(t), u_k(t)\big)\, dt \\ &\geq -C_1\big(1 + \|x_k\|^2_{C([0,T])}\big) - C_2\|u_k\|^2_{L^2(0,T)} \\ &\geq -C_3\big(1 + \|u_k\|^2_{L^2(0,T)}\big),\end{aligned}$$

for suitable constants $C_1, C_2, C_3 > 0$ independent of k.

Since $\mathcal{J}(u_k) \to J_{\inf} > -\infty$, the above inequality implies that $\|u_k\|_{L^2(0,T)}$ must be bounded. Thus the minimizing sequence $\{u_k\}$ is bounded in $L^2(0, T; \mathbb{R}^m)$.

Step 3: Weak-$*$ compactness of the controls and uniform convergence of the states. Since $U \subset \mathbb{R}^m$ is compact, there exists a constant $M_U > 0$ such that

$$|v| \leq M_U \qquad \text{for all } v \in U.$$

Hence every admissible control $u \in U_{ad}$ satisfies

$$\|u\|_{L^\infty(0,T;\mathbb{R}^m)} \leq M_U.$$

In particular, the minimizing sequence $\{u_k\}$ is bounded in $L^\infty(0, T; \mathbb{R}^m)$. Therefore, by the Banach–Alaoglu theorem, there exist a subsequence (not relabeled) and a function $\widetilde{u} \in L^\infty(0, T; \mathbb{R}^m)$ such that

$$u_k \overset{*}{\rightharpoonup} \widetilde{u} \qquad \text{in } L^\infty(0, T; \mathbb{R}^m).$$

On the other hand, since $\{u_k\}$ is bounded in the Hilbert space $L^2(0, T; \mathbb{R}^m)$, there exists $u^* \in L^2(0, T; \mathbb{R}^m)$ such that

$$u_k \rightharpoonup u^* \qquad \text{weakly in } L^2(0, T; \mathbb{R}^m).$$

Because $L^2(0, T) \subset L^1(0, T)$ on the finite interval $(0, T)$, the weak-$*$ limit in L^∞ and the weak limit in L^2 must coincide on L^2-test functions. Hence $\widetilde{u} = u^*$ a.e., and therefore

$$u_k \overset{*}{\rightharpoonup} u^* \qquad \text{in } L^\infty(0, T; \mathbb{R}^m).$$

As in the previous step, let $x_k := x(u_k)$ and $x^* := x(u^*)$ be the corresponding state trajectories. Using the Volterra formulation of the Caputo system, we write

$$x_k(t) - x^*(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} \Big(f(\tau, x_k(\tau), u_k(\tau)) - f(\tau, x^*(\tau), u^*(\tau)) \Big)\, d\tau$$

and decompose

$$x_k(t) - x^*(t) = I_{1,k}(t) + I_{2,k}(t),$$

where

$$I_{1,k}(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} \Big(f(\tau, x_k(\tau), u_k(\tau)) - f(\tau, x^*(\tau), u_k(\tau)) \Big)\, d\tau,$$

$$I_{2,k}(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} \Big(f(\tau, x^*(\tau), u_k(\tau)) - f(\tau, x^*(\tau), u^*(\tau)) \Big) \, d\tau.$$

For $I_{1,k}$, the Lipschitz continuity of f with respect to x yields

$$\|I_{1,k}(t)\| \le \frac{L_f}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} \, \|x_k(\tau) - x^*(\tau)\| \, d\tau.$$

For $I_{2,k}$, assumption (E1) implies that f is affine in u, so we may write

$$f(t, x, u) = f_0(t, x) + B(t, x)u,$$

with B continuous. Since $x^* \in C([0,T]; \mathbb{R}^n)$, the function

$$B_*(\tau) := B(\tau, x^*(\tau))$$

is bounded on $[0,T]$; denote

$$M_B := \|B_*\|_{L^\infty(0,T)}.$$

Then

$$I_{2,k}(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} B_*(\tau) \big(u_k(\tau) - u^*(\tau) \big) \, d\tau.$$

We now show that

$$\sup_{t \in [0,T]} \|I_{2,k}(t)\| \to 0 \qquad \text{as } k \to \infty.$$

First, fix $t \in [0,T]$ and $\xi \in \mathbb{R}^n$. Define

$$\phi_{t,\xi}(\tau) := \frac{1}{\Gamma(\alpha)} (t-\tau)_+^{\alpha-1} \, B_*(\tau)^\top \xi, \qquad \tau \in (0,T).$$

Since $0 < \alpha < 1$, we have $(t-\tau)_+^{\alpha-1} \in L^1(0,T)$, and because B_* is bounded, it follows that $\phi_{t,\xi} \in L^1(0,T; \mathbb{R}^m)$. Therefore, by the weak-$*$ convergence $u_k \overset{*}{\rightharpoonup} u^*$ in $L^\infty(0,T; \mathbb{R}^m)$,

$$\langle I_{2,k}(t), \xi \rangle = \int_0^T \phi_{t,\xi}(\tau) \cdot \big(u_k(\tau) - u^*(\tau) \big) \, d\tau \longrightarrow 0.$$

Since $\xi \in \mathbb{R}^n$ is arbitrary, this proves that

$$I_{2,k}(t) \to 0 \qquad \text{for every fixed } t \in [0,T].$$

Second, the family $\{I_{2,k}\}$ is uniformly equicontinuous on $[0,T]$. Indeed, let $0 \le s < t \le T$. Using the bound

$$\|u_k(\tau) - u^*(\tau)\| \le 2M_U \qquad \text{a.e. on } (0,T),$$

we obtain

$$\|I_{2,k}(t) - I_{2,k}(s)\| \le \frac{2M_B M_U}{\Gamma(\alpha)} \int_0^T \big| (t-\tau)_+^{\alpha-1} - (s-\tau)_+^{\alpha-1} \big| \, d\tau.$$

A standard estimate for the fractional kernel gives

$$\int_0^T \left|(t-\tau)_+^{\alpha-1} - (s-\tau)_+^{\alpha-1}\right| d\tau \le \frac{2}{\alpha}|t-s|^\alpha.$$

Hence

$$\|I_{2,k}(t) - I_{2,k}(s)\| \le \frac{4M_B M_U}{\Gamma(\alpha+1)}\,|t-s|^\alpha,$$

which is independent of k. Thus $\{I_{2,k}\}$ is uniformly equicontinuous. Since $I_{2,k}(t) \to 0$ for every $t \in [0,T]$ and the family $\{I_{2,k}\}$ is uniformly equicontinuous on the compact interval $[0,T]$, it follows by the Arzelà–Ascoli argument (or the standard equicontinuity lemma) that

$$\|I_{2,k}\|_{C([0,T])} = \sup_{t\in[0,T]} \|I_{2,k}(t)\| \longrightarrow 0.$$

Now define

$$y_k(t) := \|x_k(t) - x^*(t)\|, \qquad a_k := \|I_{2,k}\|_{C([0,T])}.$$

Then

$$y_k(t) \le a_k + \frac{L_f}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} y_k(\tau)\, d\tau.$$

Applying the fractional Grönwall inequality (see, e.g., [1, 2, 66]), we obtain

$$y_k(t) \le a_k\, E_\alpha(L_f t^\alpha), \qquad t \in [0,T],$$

where E_α denotes the Mittag–Leffler function. Consequently,

$$\|x_k - x^*\|_{C([0,T])} \le a_k\, E_\alpha(L_f T^\alpha) \longrightarrow 0 \qquad \text{as } k \to \infty.$$

Therefore,

$$x_k \to x^* \qquad \text{uniformly on } [0,T].$$

Step 4: Lower semicontinuity of the cost and passage to the limit.

We now show that

$$\mathcal{J}(u^*) \le \liminf_{k\to\infty} \mathcal{J}(u_k),$$

which implies that u^* is optimal.

Recall

$$\mathcal{J}(u_k) = \int_0^T L\big(t, x_k(t), u_k(t)\big)\, dt.$$

We know:

- $x_k(t) \to x^*(t)$ uniformly in t;
- $u_k \rightharpoonup u^*$ weakly in L^2 and $\{u_k\}$ is bounded in L^2;
- $L(t,x,u)$ is continuous in (t,x,u) and convex in u.

For each fixed t, the map

$$u \mapsto L\big(t, x^*(t), u\big)$$

is convex and continuous. By a standard result in convex analysis (e.g., lower semicontinuity of convex integral functionals in reflexive Banach spaces), the functional

$$u \mapsto \int_0^T L\big(t, x^*(t), u(t)\big)\, dt$$

is weakly lower semicontinuous on $L^2(0,T)$. Therefore,

$$\int_0^T L\big(t, x^*(t), u^*(t)\big)\, dt \le \liminf_{k\to\infty} \int_0^T L\big(t, x^*(t), u_k(t)\big)\, dt. \tag{3.10}$$

On the other hand, using the continuity of L and the uniform convergence $x_k \to x^*$, together with the growth condition (A3) and the boundedness of u_k in L^2, one can show that

$$\int_0^T \big|L(t, x_k(t), u_k(t)) - L(t, x^*(t), u_k(t))\big|\, dt \to 0 \quad \text{as } k \to \infty.$$

Indeed, for each t, $L(t, x_k(t), u_k(t)) \to L(t, x^*(t), u_k(t))$ by continuity in x, and the growth bound plus boundedness of $\{x_k\}$ and $\{u_k\}$ provides a uniform integrable majorant, so dominated convergence applies.

Thus

$$\lim_{k\to\infty} \int_0^T L\big(t, x_k(t), u_k(t)\big)\, dt = \lim_{k\to\infty} \int_0^T L\big(t, x^*(t), u_k(t)\big)\, dt.$$

Combining this with (3.10), we have

$$\mathcal{J}(u^*) = \int_0^T L\big(t, x^*(t), u^*(t)\big)\, dt \le \liminf_{k\to\infty} \int_0^T L\big(t, x^*(t), u_k(t)\big)\, dt = \liminf_{k\to\infty} \mathcal{J}(u_k).$$

Since $\{u_k\}$ is a minimizing sequence,

$$\lim_{k\to\infty} \mathcal{J}(u_k) = J_{\inf},$$

it follows that

$$\mathcal{J}(u^*) \le J_{\inf}.$$

On the other hand, by definition $J_{\inf} \le \mathcal{J}(u^*)$, hence

$$\mathcal{J}(u^*) = J_{\inf}.$$

Therefore, u^* is an optimal control and x^* is the associated optimal state, which completes the proof.

3.5. Uniqueness of the optimal control

Uniqueness generally requires stronger assumptions than existence. In the fractional context, sufficient conditions include strict convexity of the Lagrangian in the control variable and monotonicity of f in x. To obtain a fully rigorous uniqueness result for the reduced functional, we impose a slightly stronger structural assumption on the running cost, namely a separable convex dependence on the state and a strictly convex dependence on the control.

Theorem 3 (Uniqueness of the optimal control)**.** *Let the hypotheses of Theorem 2 hold. Assume in addition that:*

(U1) ***Separable convexity of the running cost.*** *The Lagrangian has the form*

$$L(t, x, u) = \ell(t, x) + r(t, u),$$

where, for every fixed $t \in [0, T]$, the map $x \mapsto \ell(t, x)$ is convex on $\mathbb{R}^n$ and the map $u \mapsto r(t, u)$ is strictly convex on U.

(U2) ***Affine dependence of the dynamics on the control.*** *For each (t, x), the map $u \mapsto f(t, x, u)$ is affine:*

$$f(t, x, u) = f_0(t, x) + B(t, x)u,$$

where f_0 and B are continuous in (t, x), and $B(t, x)$ is bounded on bounded subsets.

(U3) ***Linearity of the state equation in x.*** *The state dynamics are linear in x, namely*

$$f_0(t, x) = A(t)x + b(t), \qquad f(t, x, u) = A(t)x + B(t)u + b(t),$$

where $A(\cdot)$ and $B(\cdot)$ are bounded matrix-valued functions and $b(\cdot)$ is bounded.

Then the FOCP admits a unique optimal control $u^* \in U_{ad}$.

Proof. We already know from Theorem 2 that at least one optimal control u^* exists. We prove that this optimal control is unique by showing that the reduced functional

$$\mathcal{J}(u) := J[x(u), u]$$

is *strictly convex* on $\mathcal{U}_{\mathrm{ad}}$.

Step 1: Linearity/affinity of the state with respect to the control.

Consider the state equation with Caputo derivative,

$${}^{C}D^{\alpha}_{0+}x(t) = A(t)x(t) + B(t)u(t) + b(t), \qquad x(0) = x_0.$$

By Theorem 1, for each $u \in \mathcal{U}_{\mathrm{ad}}$ there is a unique solution $x(u) \in C([0, T]; \mathbb{R}^n)$.

Let $u_1, u_2 \in \mathcal{U}_{\rm ad}$ and fix $\theta \in (0,1)$. Define the convex combination

$$u_\theta(t) := \theta u_1(t) + (1-\theta)u_2(t), \qquad t \in [0,T].$$

Since U is convex and $u_1(t), u_2(t) \in U$ a.e., we have $u_\theta(t) \in U$ a.e., hence $u_\theta \in \mathcal{U}_{\rm ad}$.

Let $x_1 := x(u_1)$, $x_2 := x(u_2)$ be the corresponding state trajectories and define

$$\tilde{x}_\theta(t) := \theta x_1(t) + (1-\theta)x_2(t).$$

We claim that $\tilde{x}_\theta$ satisfies the same linear fractional differential equation as the state associated with u_θ.

Indeed, by linearity of the Caputo derivative,

$${}^C D^\alpha_{0+}\tilde{x}_\theta(t) = \theta\, {}^C D^\alpha_{0+} x_1(t) + (1-\theta)\, {}^C D^\alpha_{0+} x_2(t).$$

Using the dynamics for x_1 and x_2,

$$\begin{aligned} {}^C D^\alpha_{0+}\tilde{x}_\theta(t) &= \theta\big(A(t)x_1(t) + B(t)u_1(t) + b(t)\big) + (1-\theta)\big(A(t)x_2(t) + B(t)u_2(t) + b(t)\big) \\ &= A(t)\big(\theta x_1(t) + (1-\theta)x_2(t)\big) + B(t)\big(\theta u_1(t) + (1-\theta)u_2(t)\big) + b(t) \\ &= A(t)\tilde{x}_\theta(t) + B(t)u_\theta(t) + b(t). \end{aligned}$$

Furthermore,

$$\tilde{x}_\theta(0) = \theta x_1(0) + (1-\theta)x_2(0) = \theta x_0 + (1-\theta)x_0 = x_0.$$

Thus $\tilde{x}_\theta$ satisfies the same linear fractional system as the state driven by u_θ with the same initial condition x_0:

$$\begin{cases} {}^C D^\alpha_{0+}\tilde{x}_\theta(t) = A(t)\tilde{x}_\theta(t) + B(t)u_\theta(t) + b(t), \\ \tilde{x}_\theta(0) = x_0. \end{cases}$$

By uniqueness of the solution to the state equation (Theorem 1), we must have

$$x(u_\theta)(t) = \tilde{x}_\theta(t) = \theta x_1(t) + (1-\theta)x_2(t), \quad t \in [0,T].$$

Hence the control-to-state map $u \mapsto x(u)$ is *affine*:

$$x(u_\theta) = \theta x(u_1) + (1-\theta)x(u_2), \quad \forall u_1, u_2,\ \theta \in (0,1).$$

Step 2: Strict convexity of the reduced cost functional. Recall that

$$J(u) = \int_0^T L\big(t, x(u)(t), u(t)\big)\, dt.$$

Let $u_1, u_2 \in U_{ad}$ with $u_1 \neq u_2$, and let $\theta \in (0,1)$. Define

$$u_\theta := \theta u_1 + (1-\theta)u_2.$$

By Step 1, the control-to-state map is affine, hence

$$x(u_\theta) = \theta x(u_1) + (1-\theta)x(u_2).$$

For simplicity, write

$$x_i := x(u_i), \qquad i = 1, 2,$$

and

$$x_\theta := x(u_\theta) = \theta x_1 + (1-\theta)x_2.$$

Using the separable structure of L, we have

$$L(t, x_\theta(t), u_\theta(t)) = \ell\big(t, x_\theta(t)\big) + r\big(t, u_\theta(t)\big).$$

By convexity of $\ell(t,\cdot)$,

$$\ell\big(t, x_\theta(t)\big) = \ell\big(t, \theta x_1(t) + (1-\theta)x_2(t)\big) \le \theta\ell\big(t, x_1(t)\big) + (1-\theta)\ell\big(t, x_2(t)\big).$$

By strict convexity of $r(t,\cdot)$,

$$r\big(t, u_\theta(t)\big) = r\big(t, \theta u_1(t) + (1-\theta)u_2(t)\big) \le \theta r\big(t, u_1(t)\big) + (1-\theta)r\big(t, u_2(t)\big),$$

and the inequality is strict whenever $u_1(t) \neq u_2(t)$. Therefore, for almost every $t \in [0,T]$,

$$L(t, x_\theta(t), u_\theta(t)) \le \theta L(t, x_1(t), u_1(t)) + (1-\theta)L(t, x_2(t), u_2(t)),$$

and the inequality is strict on the set

$$E := \{t \in [0,T] : u_1(t) \neq u_2(t)\}.$$

Since $u_1 \neq u_2$ in $L^2(0,T;\mathbb{R}^m)$, the set E has positive measure. Hence, after integration over $(0,T)$, we obtain

$$J(u_\theta) < \theta J(u_1) + (1-\theta)J(u_2).$$

Thus, the reduced functional J is strictly convex on U_{ad}.

Step 3: Uniqueness of the minimizer. By Theorem 2, at least one optimal control exists. Let $u^* \in U_{ad}$ be an optimal control. Assume, for contradiction, that there exist two distinct optimal controls $u_1, u_2 \in U_{ad}$ such that

$$J(u_1) = J(u_2) = \inf_{u \in U_{ad}} J(u).$$

For any $\theta \in (0,1)$, define

$$u_\theta = \theta u_1 + (1-\theta)u_2 \in U_{ad}.$$

Since J is strictly convex,

$$J(u_\theta) < \theta J(u_1) + (1-\theta)J(u_2) = \inf_{u \in U_{ad}} J(u),$$

which is impossible. Therefore, the minimizer is unique. Hence the FOCP admits a unique optimal control $u^* \in U_{ad}$.

3.6. Summary

Theorems 1–3 guarantee that the fractional optimal control problem formulated above is well-posed: for every admissible control the fractional dynamics admit a unique state trajectory, and under mild convexity conditions an optimal control exists and is unique. These results justify the subsequent development of a spectral discretization of the FOCP and ensure that the numerical solution converges to a well-defined mathematical object.

3.7. Fractional Pontryagin Maximum Principle

In this subsection we derive first-order necessary optimality conditions for the FOCP in the form of a fractional Pontryagin maximum principle. Our formulation follows the Caputo-state / right-sided Riemann–Liouville-adjoint structure developed in [10, 17, 24, 25, 59].

Right-sided fractional operators and integration by parts

For $0 < \alpha < 1$ and $t \in [0, T]$, the right-sided Riemann–Liouville fractional integral and derivative of a function $p : [0, T] \to \mathbb{R}^n$ are defined by

$$\big(I^{\alpha}_{T-}p\big)(t) := \frac{1}{\Gamma(\alpha)} \int_t^T (\tau - t)^{\alpha-1} p(\tau)\, d\tau, \tag{3.11}$$

$$\big(D^{\alpha}_{T-}p\big)(t) := -\frac{d}{dt}\big(I^{1-\alpha}_{T-}p\big)(t), \tag{3.12}$$

whenever the integrals are well defined [1, 2, 4].

A key tool is the fractional integration by parts formula relating the left-sided Caputo derivative and the right-sided Riemann–Liouville derivative (see, e.g., [10, 17, 24]):

$$\int_0^T \lambda(t)\, {}^{C}D^{\alpha}_{0+}x(t)\, dt = \big[x(t)\, I^{1-\alpha}_{T-}\lambda(t)\big]_{t=0}^{t=T} + \int_0^T x(t)\, D^{\alpha}_{T-}\lambda(t)\, dt, \tag{3.13}$$

for sufficiently smooth x and λ.

Hamiltonian and assumptions

We introduce the Hamiltonian associated with the FOCP:

$$H(t, x, u, p) := L(t, x, u) + \langle p, f(t, x, u)\rangle, \tag{3.14}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product in $\mathbb{R}^n$.

In addition to the assumptions of Theorems 1 and 2, we assume:

(H1) L and f are continuously differentiable in x and u, and their partial derivatives are continuous in (t, x, u).

(H2) The control set $U \subset \mathbb{R}^m$ is nonempty, convex, and compact.

(H3) The terminal state $x(T)$ is free (no terminal constraint or terminal cost).

Theorem 4 (Fractional Pontryagin Maximum Principle)**.** *Let the hypotheses (A1)–(A4), (E1)–(E2), and (H1)–(H3) hold. Let $u^* \in \mathcal{U}_{\mathrm{ad}}$ be an optimal control and x^* its associated state, i.e.,*

$$ {}^{C}D^{\alpha}_{0+}x^*(t) = f\big(t, x^*(t), u^*(t)\big), \quad x^*(0) = x_0, $$

and

$$ J[x^*, u^*] = \inf_{u \in \mathcal{U}_{\mathrm{ad}}} J[x, u]. $$

Then there exists an adjoint function $p^* \in C([0, T]; \mathbb{R}^n)$ such that the following conditions hold:

(P1) ***State equation.***

$$ {}^{C}D^{\alpha}_{0+}x^*(t) = f\big(t, x^*(t), u^*(t)\big), \qquad t \in (0, T], \quad x^*(0) = x_0. $$

(P2) ***Adjoint equation.*** *The adjoint p^* satisfies the right-sided Riemann–Liouville fractional differential equation*

$$ D^{\alpha}_{T-}p^*(t) = \nabla_x H\big(t, x^*(t), u^*(t), p^*(t)\big), \qquad t \in [0, T), \tag{3.15} $$

where $\nabla_x H$ denotes the gradient of H with respect to x.

(P3) ***Transversality condition.*** *Because $x(T)$ is free and there is no terminal cost, the adjoint must satisfy*

$$ \big(I^{1-\alpha}_{T-}p^*\big)(T) = 0. \tag{3.16} $$

(P4) ***Hamiltonian minimum condition.*** *For almost every $t \in [0, T]$, the control $u^*(t)$ minimizes the Hamiltonian:*

$$ H\big(t, x^*(t), u^*(t), p^*(t)\big) = \min_{v \in U} H\big(t, x^*(t), v, p^*(t)\big). \tag{3.17} $$

Equivalently, if $U = \mathbb{R}^m$ (no control constraints), then the first-order stationarity condition holds:

$$ \nabla_u H\big(t, x^*(t), u^*(t), p^*(t)\big) = 0, \qquad a.e.\ t \in [0, T]. \tag{3.18} $$

Proof. We follow the variational approach as in [10, 17, 24, 25, 59].

Step 1: Control perturbation and linearized state equation.

Let u^* be optimal and let $v \in \mathcal{U}_{\mathrm{ad}}$ be arbitrary. For $\varepsilon \in \mathbb{R}$ small, define the perturbed control

$$ u^{\varepsilon}(t) := u^*(t) + \varepsilon\big(v(t) - u^*(t)\big), $$

which remains in $\mathcal{U}_{\mathrm{ad}}$ by convexity of U.

Let x^ε denote the state corresponding to u^ε:

$$^C D^\alpha_{0+} x^\varepsilon(t) = f\big(t, x^\varepsilon(t), u^\varepsilon(t)\big), \qquad x^\varepsilon(0) = x_0.$$

Set

$$\delta x(t) := \frac{d}{d\varepsilon} x^\varepsilon(t)\bigg|_{\varepsilon=0}, \quad \delta u(t) := v(t) - u^*(t).$$

Formally differentiating the state equation at $\varepsilon = 0$ and using (H1), we obtain the linearized fractional state equation:

$$^C D^\alpha_{0+} \delta x(t) = A(t)\, \delta x(t) + B(t)\, \delta u(t), \qquad \delta x(0) = 0, \tag{3.19}$$

where

$$A(t) := \nabla_x f\big(t, x^*(t), u^*(t)\big), \quad B(t) := \nabla_u f\big(t, x^*(t), u^*(t)\big).$$

Step 2: First variation of the cost functional.

The Gâteaux derivative of J at (x^*, u^*) in the direction $(\delta x, \delta u)$ is

$$\delta J := \frac{d}{d\varepsilon} J[x^\varepsilon, u^\varepsilon]\bigg|_{\varepsilon=0} = \int_0^T \Big[\langle \nabla_x L(t, x^*(t), u^*(t)), \delta x(t)\rangle + \langle \nabla_u L(t, x^*(t), u^*(t)), \delta u(t)\rangle\Big] dt.$$

Optimality of u^* implies $\delta J \geq 0$ for all admissible directions $v - u^*$.

Step 3: Introduction of the adjoint and fractional integration by parts.

We now introduce an adjoint function $p : [0, T] \to \mathbb{R}^n$ and use the dynamics constraint to rewrite the variation in a form where δx is eliminated.

Observe that by (3.19),

$$\delta x(t) = I^\alpha_{0+}\big(A(\cdot)\delta x(\cdot) + B(\cdot)\delta u(\cdot)\big)(t), \quad \delta x(0) = 0,$$

but we avoid solving this explicitly.

Instead, consider

$$\delta J = \int_0^T \langle \nabla_x L, \delta x\rangle dt + \int_0^T \langle \nabla_u L, \delta u\rangle dt.$$

Add the identity

$$0 = \int_0^T \big\langle p(t), {}^C D^\alpha_{0+} \delta x(t) - A(t)\delta x(t) - B(t)\delta u(t)\big\rangle\, dt,$$

to δJ (this is zero because of the linearized state equation). We obtain

$$\begin{aligned} \delta J = &\int_0^T \Big\langle \nabla_x L(t), \delta x(t)\Big\rangle dt + \int_0^T \Big\langle \nabla_u L(t), \delta u(t)\Big\rangle dt \\ &+ \int_0^T \Big\langle p(t), {}^C D^\alpha_{0+} \delta x(t)\Big\rangle dt - \int_0^T \Big\langle p(t), A(t)\delta x(t)\Big\rangle dt - \int_0^T \Big\langle p(t), B(t)\delta u(t)\Big\rangle dt. \end{aligned}$$

Grouping terms, we have

$$\delta J = \int_0^T \Big\langle \nabla_x L(t) - A(t)^\top p(t), \delta x(t) \Big\rangle dt + \int_0^T \Big\langle \nabla_u L(t) - B(t)^\top p(t), \delta u(t) \Big\rangle dt$$
$$+ \int_0^T \Big\langle p(t), {}^C D_{0+}^{\alpha} \delta x(t) \Big\rangle dt. \tag{3.20}$$

Now apply the fractional integration by parts formula (3.13) with $x = \delta x$ and $\lambda = p$:

$$\int_0^T \Big\langle p(t), {}^C D_{0+}^{\alpha} \delta x(t) \Big\rangle dt = \big[\langle \delta x(t), I_{T-}^{1-\alpha} p(t) \rangle \big]_{t=0}^{t=T} + \int_0^T \Big\langle \delta x(t), D_{T-}^{\alpha} p(t) \Big\rangle dt.$$

Since $\delta x(0) = 0$ and $x(T)$ is free, $\delta x(T)$ is arbitrary. Hence the boundary term reads

$$\big\langle \delta x(T), I_{T-}^{1-\alpha} p(T) \big\rangle .$$

Substituting into (3.20), we obtain

$$\delta J = \int_0^T \Big\langle \nabla_x L(t) - A(t)^\top p(t) + D_{T-}^{\alpha} p(t), \delta x(t) \Big\rangle dt$$
$$+ \big\langle \delta x(T), I_{T-}^{1-\alpha} p(T) \big\rangle + \int_0^T \Big\langle \nabla_u L(t) - B(t)^\top p(t), \delta u(t) \Big\rangle dt. \tag{3.21}$$

Step 4: Choice of the adjoint and transversality condition.

To eliminate the dependence on δx, we *define* p as the solution of the adjoint equation

$$D_{T-}^{\alpha} p(t) = \nabla_x L(t, x^*(t), u^*(t)) - A(t)^\top p(t) = \nabla_x H(t, x^*(t), u^*(t), p(t)),$$

with a terminal condition determined by the requirement that the boundary term vanish.

Since $x(T)$ is free, variations $\delta x(T)$ are arbitrary, and optimality for all directions requires

$$\big\langle \delta x(T), I_{T-}^{1-\alpha} p(T) \big\rangle = 0 \quad \text{for all } \delta x(T),$$

which implies

$$I_{T-}^{1-\alpha} p(T) = 0.$$

This is the transversality condition (3.16), consistent with in the absence of a terminal cost.

With this choice of p, the first integral in (3.21) vanishes, and we obtain the simplified variation:

$$\delta J = \int_0^T \Big\langle \nabla_u L(t, x^*, u^*) - B(t)^\top p^*(t), \delta u(t) \Big\rangle dt = \int_0^T \Big\langle \nabla_u H(t, x^*, u^*, p^*), \delta u(t) \Big\rangle dt. \tag{3.22}$$

Step 5: Hamiltonian minimum condition.

Recall that $\delta u(t) = v(t) - u^*(t)$, where $v \in \mathcal{U}_{\text{ad}}$ is arbitrary. Optimality of u^* implies $\delta J \geq 0$ for all such v.

If U is convex, the condition (3.22) together with standard arguments from convex analysis (see, e.g., [67]) show that, for almost every t,

$$\langle \nabla_u H(t, x^*(t), u^*(t), p^*(t)), v - u^*(t)\rangle \geq 0 \quad \forall v \in U,$$

which is equivalent to

$$H(t, x^*(t), u^*(t), p^*(t)) \leq H(t, x^*(t), v, p^*(t)), \qquad \forall v \in U,$$

i.e., $u^*(t)$ minimizes the Hamiltonian pointwise in U, yielding (3.17).

In the unconstrained case $U = \mathbb{R}^m$, the above inequality is equivalent to the stationarity condition

$$\nabla_u H(t, x^*(t), u^*(t), p^*(t)) = 0,$$

i.e., (3.18), for almost every $t \in [0, T]$.

Conclusion. We have shown that for an optimal pair (x^*, u^*) there exists an adjoint p^* satisfying the state equation, the adjoint equation, the transversality condition, and the Hamiltonian minimum principle. This completes the proof.

Remark 1 (Terminal costs and terminal constraints)**.** *If a terminal cost of the form* $\Phi(x(T))$ *is included in the performance index*

$$J[x, u] = \Phi(x(T)) + \int_0^T L(t, x(t), u(t))\, dt,$$

then the derivation above yields a modified transversality condition

$$I_{T-}^{1-\alpha} p^*(T) = \nabla_x \Phi\big(x^*(T)\big),$$

as in [10, 17]. For fixed terminal state $x(T) = x_T$*, admissible variations satisfy* $\delta x(T) = 0$ *and the boundary term automatically vanishes.*

4. Legendre spectral approximation

Spectral methods based on orthogonal polynomials have proved extremely powerful for solving differential equations with smooth solutions, thanks to their high-order (often exponential) convergence and favorable approximation properties; see, e.g., [46, 62, 63]. In the fractional setting, Legendre and Jacobi spectral and pseudospectral schemes have been successfully applied to fractional differential equations and FOCPs in [20, 30, 45, 52]. In this section, we construct a Legendre spectral collocation discretization of the FOCP introduced in Section 3, leading to a finite-dimensional constrained optimization problem.

We begin by mapping the physical time interval $[0, T]$ to the reference interval $[-1, 1]$ via the affine transformation

$$t = t(\xi) := \frac{T}{2}(1 + \xi), \qquad \xi \in [-1, 1]. \tag{4.1}$$

Let $\{P_k(\xi)\}_{k=0}^{\infty}$ denote the classical Legendre polynomials orthogonal on $[-1,1]$ with respect to the weight $w(\xi) \equiv 1$. We collect the first $N+1$ polynomials into the basis vector

$$\mathbf{P}(\xi) := \big(P_0(\xi), P_1(\xi), \ldots, P_N(\xi)\big)^\top \in \mathbb{R}^{N+1}.$$

4.1. Approximation of the state and control

Let $x : [0,T] \to \mathbb{R}^n$ and $u : [0,T] \to \mathbb{R}^m$ be the state and control. For simplicity of exposition, we first describe the scalar case $n = m = 1$; the vector-valued case is obtained componentwise by tensorization.

We approximate x and u by truncated Legendre expansions in the reference variable ξ:

$$x_N\big(t(\xi)\big) := \mathbf{a}^\top \mathbf{P}(\xi) = \sum_{k=0}^{N} a_k P_k(\xi), \tag{4.2}$$

$$u_N\big(t(\xi)\big) := \mathbf{b}^\top \mathbf{P}(\xi) = \sum_{k=0}^{N} b_k P_k(\xi), \tag{4.3}$$

where $\mathbf{a} = (a_0, \ldots, a_N)^\top$ and $\mathbf{b} = (b_0, \ldots, b_N)^\top$ are the vectors of unknown spectral coefficients.

For $n > 1$ and $m > 1$, we approximate each component x_ℓ, u_r by its own expansion of the form (4.2)–(4.3), and stack the corresponding coefficient vectors in a single decision vector (see Remark 3 below).

Remark 2 (Choice of polynomial degree)**.** *The truncation index N controls the spatial resolution of the spectral approximation. For analytic solutions, Legendre spectral approximations typically converge at an exponential rate in N [46, 62, 63]. In the fractional setting, the presence of weak singularities near $t = 0$ may slightly reduce the observed rate, but high-order convergence is still expected under mild regularity; see, e.g., [45].*

Remark 3 (Vector-valued states and controls)**.** *For $x = (x_1, \ldots, x_n)^\top$ and $u = (u_1, \ldots, u_m)^\top$, we write*

$$x_\ell\big(t(\xi)\big) \approx \mathbf{a}^{(\ell)\top} \mathbf{P}(\xi), \quad \ell = 1, \ldots, n, \qquad u_r\big(t(\xi)\big) \approx \mathbf{b}^{(r)\top} \mathbf{P}(\xi), \quad r = 1, \ldots, m,$$

with coefficient vectors $\mathbf{a}^{(\ell)}, \mathbf{b}^{(r)} \in \mathbb{R}^{N+1}$. Collecting all state coefficients in a single vector

$$\mathbf{A} := \big(\mathbf{a}^{(1)\top}, \ldots, \mathbf{a}^{(n)\top}\big)^\top \in \mathbb{R}^{n(N+1)},$$

and similarly for the control, allows us to write the discretized FOCP in compact matrix form. For clarity, we continue the derivation in the scalar case.

4.2. Operational matrices for integer and fractional derivatives

The main advantage of the spectral representation (4.2)–(4.3) is that differentiation and fractional integration can be approximated by matrix multiplication acting on the coefficient vectors. This leads to a compact algebraic representation of the fractional dynamics.

4.2.1. Integer derivative matrix for Legendre polynomials

Differentiating x_N with respect to t and using the chain rule (4.1), we obtain

$$\frac{d}{dt}x_N\big(t(\xi)\big) = \frac{d\xi}{dt}\frac{d}{d\xi}\big(\mathbf{a}^\top \mathbf{P}(\xi)\big) = \frac{2}{T}\,\mathbf{a}^\top \mathbf{P}'(\xi). \tag{4.4}$$

It is well known that $\mathbf{P}'(\xi)$ can be expressed in the same Legendre basis via a differentiation matrix $\mathbf{G} \in \mathbb{R}^{(N+1)\times(N+1)}$:

$$\mathbf{P}'(\xi) = \mathbf{G}\,\mathbf{P}(\xi), \tag{4.5}$$

where the entries of $\mathbf{G}$ are obtained from the recurrence and differentiation formulas for Legendre polynomials [46, 62, 63]. Substituting (4.5) into (4.4) yields

$$x_N'(t(\xi)) = \frac{2}{T}\,\mathbf{a}^\top \mathbf{G}\,\mathbf{P}(\xi). \tag{4.6}$$

4.2.2. Fractional integral matrix

Let I^β_{0+} denote the left-sided Riemann–Liouville fractional integral of order $\beta > 0$:

$$(I^\beta_{0+}y)(t) = \frac{1}{\Gamma(\beta)}\int_0^t (t-\tau)^{\beta-1} y(\tau)\,d\tau. \tag{4.7}$$

Under the mapping (4.1), we can write [45, 55]

$$\big(I^\beta_{0+}y\big)\big(t(\xi)\big) = \Big(\frac{T}{2}\Big)^\beta \big(I^\beta_{-1+,\xi}\tilde{y}\big)(\xi), \tag{4.8}$$

where $\tilde{y}(\xi) := y(t(\xi))$ and $I^\beta_{-1+,\xi}$ is the fractional integral on $[-1, 1]$ with respect to ξ.

For the Legendre basis $\mathbf{P}(\xi)$, we define the *fractional integral operational matrix* $\mathbf{J}^\beta \in \mathbb{R}^{(N+1)\times(N+1)}$ by

$$I^\beta_{0+}\mathbf{P}\big(t(\xi)\big) \approx \mathbf{J}^\beta\,\mathbf{P}(\xi). \tag{4.9}$$

Equivalently, each column of $\mathbf{J}^\beta$ is obtained by projecting $I^\beta_{0+}P_j$ onto the Legendre basis:

$$I^\beta_{0+}P_j\big(t(\xi)\big) \approx \sum_{k=0}^{N} J^\beta_{kj} P_k(\xi), \qquad j = 0, \ldots, N. \tag{4.10}$$

The entries J^β_{kj} can be computed in several ways:

- *Analytically*, by evaluating the integrals

$$J_{kj}^{\beta} = \frac{\langle I_{0+}^{\beta} P_j, P_k \rangle_{L^2(-1,1)}}{\langle P_k, P_k \rangle_{L^2(-1,1)}},$$

  using Beta and Gamma functions; see [30].

- *Numerically*, by enforcing the expansion (4.10) at Gauss–Legendre nodes (collocation), leading to a small linear system for each column [20].

4.2.3. Caputo fractional derivative matrix

Recall that for sufficiently smooth x and $0 < \alpha < 1$,

$$^{C}D_{0+}^{\alpha} x(t) = I_{0+}^{1-\alpha} x'(t). \tag{4.11}$$

Combining (4.6), (4.9) with $\beta = 1-\alpha$, and (4.11), we obtain the following approximation:

$$\begin{aligned} ^{C}D_{0+}^{\alpha} x_N\big(t(\xi)\big) &= I_{0+}^{1-\alpha} x_N'(t(\xi)) \\ &\approx I_{0+}^{1-\alpha} \left(\frac{2}{T} \mathbf{a}^{\top} \mathbf{G}\, \mathbf{P}(\xi) \right) \\ &\approx \frac{2}{T} \mathbf{a}^{\top} \mathbf{G}\, \mathbf{J}^{1-\alpha}\, \mathbf{P}(\xi). \end{aligned} \tag{4.12}$$

This suggests the definition of the *Caputo fractional derivative operational matrix*:

$$\mathbf{D}^{\alpha} := \frac{2}{T} \mathbf{G}\, \mathbf{J}^{1-\alpha}, \tag{4.13}$$

so that

$$^{C}D_{0+}^{\alpha} x_N\big(t(\xi)\big) \approx \mathbf{a}^{\top} \mathbf{D}^{\alpha}\, \mathbf{P}(\xi). \tag{4.14}$$

Remark 4 (Accuracy of the operational approximation)**.** *The approximation error in* (4.9) *and* (4.14) *depends on the smoothness of x and the order of truncation N. For analytic x, one typically observes spectral convergence in N for $\mathbf{D}^{\alpha}$-based approximations [45, 52]. In practice, the construction of $\mathbf{J}^{1-\alpha}$ via Gauss–Legendre collocation is numerically stable and straightforward to implement.*

4.3. Collocation discretization of the fractional dynamics

We now discretize the state equation

$$^{C}D_{0+}^{\alpha} x(t) = f\big(t, x(t), u(t)\big), \qquad t \in (0, T], \tag{4.15}$$

using the Legendre spectral representation and the operational matrix $\mathbf{D}^{\alpha}$.

4.3.1. Legendre–Gauss collocation points

Let $\{\xi_i\}_{i=0}^N$ denote the $(N+1)$ Legendre–Gauss nodes in $[-1,1]$, i.e., the roots of $P'_{N+1}(\xi)$ (or, equivalently, the Gauss–Legendre quadrature nodes). We map them to time nodes $\{t_i\}_{i=0}^N \subset (0,T)$ via (4.1):

$$t_i := t(\xi_i) = \frac{T}{2}(1+\xi_i), \qquad i = 0,\ldots,N.$$

Define the *collocation matrix* $\mathbf{P}_c \in \mathbb{R}^{(N+1)\times(N+1)}$ by

$$(\mathbf{P}_c)_{ik} := P_k(\xi_i), \qquad i,k = 0,\ldots,N, \tag{4.16}$$

so that

$$x_N(t_i) = \mathbf{a}^\top \mathbf{P}(\xi_i) = \big(\mathbf{P}_c\mathbf{a}\big)_i, \qquad u_N(t_i) = \big(\mathbf{P}_c\mathbf{b}\big)_i.$$

Similarly, the fractional derivative at the collocation points is approximated as

$${}^{C}D^{\alpha}_{0+}x_N(t_i) \approx \mathbf{a}^\top \mathbf{D}^\alpha \mathbf{P}(\xi_i) = \big(\mathbf{P}_c\mathbf{D}^{\alpha\top}\mathbf{a}\big)_i.$$

4.3.2. Discrete enforcement of the dynamics

Enforcing the fractional dynamics (4.15) at the nodes $\{t_i\}_{i=0}^N$ leads to the nonlinear algebraic system

$$\big(\mathbf{P}_c\mathbf{D}^{\alpha\top}\mathbf{a}\big)_i = f\Big(t_i,\, \big(\mathbf{P}_c\mathbf{a}\big)_i,\, \big(\mathbf{P}_c\mathbf{b}\big)_i\Big), \qquad i = 0,\ldots,N. \tag{4.17}$$

In vector form, we may write

$$\mathbf{P}_c\mathbf{D}^{\alpha\top}\mathbf{a} = \mathbf{f}\big(\mathbf{t}, \mathbf{P}_c\mathbf{a}, \mathbf{P}_c\mathbf{b}\big), \tag{4.18}$$

where

$$\mathbf{t} := (t_0,\ldots,t_N)^\top, \qquad \big(\mathbf{f}(\mathbf{t},\mathbf{x},\mathbf{u})\big)_i := f(t_i, x_i, u_i).$$

The initial condition $x(0) = x_0$ corresponds to $\xi = -1$. Since -1 is generally not one of the Legendre–Gauss nodes, we enforce it as a separate linear constraint:

$$x_N(0) = x_N\big(t(-1)\big) = \mathbf{a}^\top \mathbf{P}(-1) = x_0, \tag{4.19}$$

where

$$\mathbf{P}(-1) = \big(P_0(-1),\ldots,P_N(-1)\big)^\top.$$

Remark 5 (Alternative node choices)**.** *One may also employ Legendre–Gauss–Lobatto or Legendre–Gauss–Radau nodes, which include one or both endpoints of the interval and simplify the imposition of endpoint conditions [46, 62, 63]. In that case, the initial condition* (4.19) *becomes one of the collocation equations. The analysis and construction of* $\mathbf{D}^\alpha$ *are unchanged.*

4.4. Discrete cost functional and reduced optimization problem

To complete the discretization of the FOCP, we approximate the performance index

$$J[x,u] = \int_0^T L\big(t, x(t), u(t)\big)\, dt$$

by a Gauss–Legendre quadrature formula in the reference variable ξ:

$$\int_0^T L\big(t, x(t), u(t)\big)\, dt = \frac{T}{2}\int_{-1}^1 \tilde{L}\big(\xi, x(t(\xi)), u(t(\xi))\big)\, d\xi \approx \frac{T}{2}\sum_{i=0}^N \omega_i L\big(t_i, x_N(t_i), u_N(t_i)\big),$$

where $\{\omega_i\}_{i=0}^N$ are the Legendre–Gauss weights on $[-1,1]$.

Thus the discrete cost functional becomes

$$J_N(\mathbf{a},\mathbf{b}) := \frac{T}{2}\sum_{i=0}^N \omega_i L\Big(t_i, (\mathbf{P}_c\mathbf{a})_i, (\mathbf{P}_c\mathbf{b})_i\Big). \tag{4.20}$$

The continuous FOCP is therefore approximated by the finite-dimensional nonlinear programming problem:

$$\begin{aligned} \min_{\mathbf{a},\mathbf{b}\in\mathbb{R}^{N+1}} \quad & J_N(\mathbf{a},\mathbf{b}) \\ \text{s.t.} \quad & \mathbf{P}_c\mathbf{D}^{\alpha\top}\mathbf{a} = \mathbf{f}\big(\mathbf{t}, \mathbf{P}_c\mathbf{a}, \mathbf{P}_c\mathbf{b}\big), \\ & \mathbf{a}^\top\mathbf{P}(-1) = x_0, \\ & (\mathbf{P}_c\mathbf{b})_i \in U, \quad i = 0,\dots,N, \end{aligned} \tag{4.21}$$

where the control constraint $(\mathbf{P}_c\mathbf{b})_i \in U$ is enforced pointwise at the collocation nodes.

Under the assumptions of Section 3, the discrete problem (4.21) is a finite-dimensional convex (or at least well-posed) optimization problem when L is convex in u and f is affine in u. The existence and uniqueness results for the continuous FOCP, together with stability and consistency of the spectral discretization, support the convergence of $(\mathbf{a},\mathbf{b})$ to the continuous optimal pair (x^*, u^*) as $N \to \infty$; see also related error analyses in [22, 45, 52, 53].

Remark 6 (Computational cost and dense matrix structure)**.** *Because the Caputo fractional derivative is nonlocal, the corresponding operational matrix D^α (and the related integration matrix $J^{1-\alpha}$ when used in its construction) is, in general, dense. Therefore, unlike local polynomial discretizations, the resulting collocation system does not lead to a sparse matrix structure in the usual finite-element or finite-difference sense. For polynomial degree N, the storage of the principal fractional operator matrices is typically $O(N^2)$, and their assembly is also of order $O(N^2)$ in a straightforward implementation. Moreover, when the state and control coefficients are solved through a generic dense nonlinear programming method, the dominant linear algebra cost per Newton/SQP/interior-point step may scale as $O(N^3)$, depending on the structure of the KKT system and* *whether dense factorizations are used. Accordingly, the present method is intended primarily as a high-order discretization strategy for problems where spectral accuracy allows one to work with relatively small or moderate values of N. A detailed complexity analysis, preconditioning strategy, or fast structured implementation for large N lies beyond the scope of the present paper and is left for future work.*

4.5. Convergence of the Legendre spectral scheme

To formulate a rigorous convergence result, one must separate three distinct sources of error: (i) polynomial approximation of the exact optimal pair (x^*, u^*), (ii) discretization of the Caputo operator through the operational matrix D^α, and (iii) stability/conditioning of the discrete collocation system defining the state equation. The original sketch is therefore replaced by the following theorem under explicit discrete consistency and stability assumptions.

Additional discrete assumptions. In addition to (C1)–(C3), assume:

(D1) **Consistency of the operational matrix.** There exists a projection/interpolation operator Π_N onto the Legendre space X_N such that, for every sufficiently smooth function x,

$$\left\|\mathcal{C}_N(\Pi_N x) - \mathcal{I}_N(CD_t^\alpha x)\right\|_{\ell^2_\omega} \le C_{\mathrm{op}}\, \varepsilon_N(x),$$

where $\mathcal{C}_N := P_c D^{\alpha\top}$ is the discrete Caputo operator, $\mathcal{I}_N$ denotes nodal interpolation at the Gauss–Legendre points, and $\varepsilon_N(x) \to 0$ as $N \to \infty$.

(D2) **Uniform discrete stability of the state equation.** Let

$$M_N := P_c D^{\alpha\top} - \mathrm{diag}(A(t_0), \ldots, A(t_N))\, P_c.$$

Then M_N is invertible for all sufficiently large N, and

$$\|M_N^{-1}\|_{\ell^2_\omega \to \ell^2_\omega} \le C_{\mathrm{st}},$$

with a constant C_{st} independent of N.

(D3) **Strong convexity of the reduced discrete functional.** The reduced discrete functional

$$J_N^{\mathrm{red}}(b) := J_N(a(b), b)$$

is uniformly strongly convex on the admissible discrete control set, i.e., there exists $\mu > 0$ independent of N such that

$$J_N^{\mathrm{red}}(b_2) \ge J_N^{\mathrm{red}}(b_1) + \nabla J_N^{\mathrm{red}}(b_1) \cdot (b_2 - b_1) + \frac{\mu}{2}\|b_2 - b_1\|^2_{\ell^2_\omega}.$$

(D4) **Consistency of the discrete cost.** For sufficiently smooth admissible pairs (x, u) and their projections $(\Pi_N x, \Pi_N u)$,

$$\left|J[x, u] - J_N(\Pi_N x, \Pi_N u)\right| \le C_{\mathrm{q}}\, \eta_N(x, u),$$

where $\eta_N(x, u) \to 0$ as $N \to \infty$.

Theorem 5 (Convergence of the Legendre spectral scheme)**.** *Let* (C1)–(C3) *and* (D1)–(D4) *hold. Let* (x^*, u^*) *be the unique continuous optimal pair, and let* (x_N, u_N) *denote the discrete optimal pair obtained from* (4.21). *Then there exists a constant* $C > 0$*, independent of* N*, such that*

$$\|x^* - x_N\|_{L^2(0,T)} + \|u^* - u_N\|_{L^2(0,T)} \leq C\Big(E_N^x + E_N^u + \mathcal{R}_N\Big), \tag{4.22$'$}$$

where

$$E_N^x := \inf_{w \in X_N} \|x^* - w\|_{L^2(0,T)}, \qquad E_N^u := \inf_{v \in U_N} \|u^* - v\|_{L^2(0,T)},$$

and

$$\mathcal{R}_N := \big\|\mathcal{C}_N(\Pi_N x^*) - \mathcal{I}_N(CD_t^\alpha x^*)\big\|_{\ell^2_\omega} + \eta_N(x^*, u^*).$$

Consequently:

(i) If $E_N^x + E_N^u + \mathcal{R}_N \to 0$*, then*

$$\|x^* - x_N\|_{L^2(0,T)} + \|u^* - u_N\|_{L^2(0,T)} \to 0.$$

(ii) If $x^*, u^* \in C^m([0,T])$ *and*

$$E_N^x + E_N^u + \mathcal{R}_N \leq CN^{-m},$$

then

$$\|x^* - x_N\|_{L^2(0,T)} + \|u^* - u_N\|_{L^2(0,T)} \leq CN^{-m}.$$

(iii) If x^* *and* u^* *are analytic on* $[0,T]$ *and*

$$E_N^x + E_N^u + \mathcal{R}_N \leq C\rho^{-N} \qquad (\rho > 1),$$

then

$$\|x^* - x_N\|_{L^2(0,T)} + \|u^* - u_N\|_{L^2(0,T)} \leq C\rho^{-N}.$$

Proof. Let $\widehat{x}_N := \Pi_N x^*$ and $\widehat{u}_N := \Pi_N u^*$ be Legendre approximants of the exact optimal pair. By definition,

$$\|x^* - \widehat{x}_N\|_{L^2(0,T)} \leq CE_N^x, \qquad \|u^* - \widehat{u}_N\|_{L^2(0,T)} \leq CE_N^u.$$

Step 1: Residual of the discrete state equation. Insert $(\widehat{x}_N, \widehat{u}_N)$ into the discrete collocation system. The residual is

$$r_N := \mathcal{C}_N(\widehat{x}_N) - \mathcal{I}_N\big(f(\cdot, \widehat{x}_N, \widehat{u}_N)\big).$$

Using the linear-affine structure

$$f(t, x, u) = A(t)x + B(t)u + b(t),$$

together with the consistency assumption (D1) and the approximation properties of Π_N, we obtain

$$\|r_N\|_{\ell^2_\omega} \le C\big(E_N^x + E_N^u + \mathcal{R}_N\big).$$

Step 2: Construction of a nearby feasible discrete pair. Let $(\bar{x}_N, \widehat{u}_N)$ denote the discrete pair satisfying the collocation equations with the same discrete control $\widehat{u}_N$. Then

$$M_N(\bar{x}_N - \widehat{x}_N) = -\, r_N.$$

By the uniform stability assumption (D2),

$$\|\bar{x}_N - \widehat{x}_N\|_{\ell^2_\omega} \le \|M_N^{-1}\|\, \|r_N\|_{\ell^2_\omega} \le C\big(E_N^x + E_N^u + \mathcal{R}_N\big).$$

Hence

$$\|x^* - \bar{x}_N\|_{L^2(0,T)} + \|u^* - \widehat{u}_N\|_{L^2(0,T)} \le C\big(E_N^x + E_N^u + \mathcal{R}_N\big).$$

Step 3: Comparison with the discrete minimizer. Let u_N be the discrete optimal control. Since J_N^{red} is uniformly strongly convex, for every admissible discrete control b we have

$$\frac{\mu}{2}\|b_N - b\|^2_{\ell^2_\omega} \le J_N^{\mathrm{red}}(b) - J_N^{\mathrm{red}}(b_N).$$

Choosing $b = \widehat{b}_N$ corresponding to $\widehat{u}_N$, and using the optimality of b_N, the consistency of the quadrature (D4), and the estimate from Step 2, we obtain

$$\|u_N - \widehat{u}_N\|_{\ell^2_\omega} \le C\big(E_N^x + E_N^u + \mathcal{R}_N\big).$$

Then the discrete stability of the state equation again yields

$$\|x_N - \bar{x}_N\|_{\ell^2_\omega} \le C\|u_N - \widehat{u}_N\|_{\ell^2_\omega} \le C\big(E_N^x + E_N^u + \mathcal{R}_N\big).$$

Step 4: Final estimate. By the triangle inequality,

$$\|u^* - u_N\|_{L^2(0,T)} \le \|u^* - \widehat{u}_N\|_{L^2(0,T)} + \|\widehat{u}_N - u_N\|_{L^2(0,T)},$$

and similarly for the state. Combining the bounds obtained above gives

$$\|x^* - x_N\|_{L^2(0,T)} + \|u^* - u_N\|_{L^2(0,T)} \le C\big(E_N^x + E_N^u + \mathcal{R}_N\big).$$

The algebraic and spectral rates follow immediately once the best-approximation terms and the operator residual $\mathcal{R}_N$ satisfy the corresponding decay estimates.

Remark 7 (Conditioning of the discrete collocation operator)**.** *The convergence result above depends essentially on the discrete stability assumption $\|M_N^{-1}\| \le C_{\mathrm{st}}$. Equivalently, one may formulate this requirement in terms of a uniform lower singular-value bound for the collocation matrix or, more practically, a controlled growth of its condition number. A sharp asymptotic analysis of the condition number of $P_c D^{\alpha\top}$ is an important topic in its own right and is not pursued here; instead, the theorem makes this stability requirement explicit.*

4.6. Computational cost and practical scaling

The Caputo fractional derivative is nonlocal, and therefore the corresponding operational matrices are generally dense. As a result, the Legendre collocation discretization does not lead to the sparse algebraic structure typically associated with local finite-difference or finite-element methods. For polynomial degree N, the storage of the principal dense matrices is of order $O(N^2)$, and a straightforward assembly of the collocation operators is likewise of order $O(N^2)$. After discretization, the FOCP is transformed into a finite-dimensional nonlinear programming problem in the state and control coefficients. If this problem is solved by a generic dense Newton/SQP/interior-point method, the dominant linear algebra cost per iteration is typically of order $O(N^3)$, depending on the structure of the Jacobian/Hessian and the KKT system. Thus, the present method is best suited to small and moderate values of N, especially in problems where high-order accuracy allows the desired error level to be reached with relatively few modes. In contrast, standard finite-difference methods often produce sparse algebraic systems with lower per-iteration cost, but they may require substantially finer discretizations to achieve comparable accuracy for smooth solutions. Therefore, the relevant comparison is not only cost per iteration, but also the number of degrees of freedom required to attain a prescribed tolerance. A detailed wall-time comparison with low-order finite-difference schemes depends strongly on the implementation, the NLP solver, stopping tolerances, and the chosen benchmark problem. Such a systematic performance study is beyond the scope of the present paper, but it represents an important direction for future work.

Table 1: Illustrative comparison between the Legendre spectral method and a finite-difference method at comparable accuracy levels.

Method	Target error	Degrees of freedom	Computational cost	Remarks
Legendre spectral	$\approx 10^{-5}$	$N = 18$–22	dense solve, typically $O(N^3)$ per NLP iteration	high-order accuracy, moderate-size problems
Finite difference	$\approx 10^{-5}$	$M = 200$–500 grid points	sparse solve, typically much cheaper per iteration	low-order accuracy, finer mesh needed

A practical limitation of the present method is the dense-matrix structure induced by the nonlocal fractional operator. While the method is highly accurate for moderate discretization sizes, a detailed performance comparison with sparse low-order methods remains an important topic for future investigation. *Note.* The values in Table 1 are included only as a schematic illustration of the trade-off between dense high-order and sparse low-order discretizations. They are not reported as implementation-specific benchmark measurements.

5. Numerical experiments

In this section we present three numerical examples illustrating the performance of the proposed Legendre spectral collocation scheme. For each example, we prescribe an exact optimal pair (x^*, u^*) and study the approximation properties of the Legendre representation of Section 4. This allows us to focus on the spectral approximation error in a clean way and to directly verify the convergence rates predicted by Theorem 5.

Unless otherwise stated, the error in the state and control is measured as

$$E_x(N) := \max_{0\le i\le N_q} \big|x^*(t_i) - x_N(t_i)\big|, \qquad E_u(N) := \max_{0\le i\le N_q} \big|u^*(t_i) - u_N(t_i)\big|, \tag{5.1}$$

where $\{t_i\}_{i=0}^{N_q}$ is a fine uniform mesh in $[0, T]$ used only for post-processing (here $N_q = 4000$). The Legendre coefficients are computed by Gauss–Legendre quadrature on $[-1, 1]$.

We report results for polynomial degrees $N \in \{6, 10, 14, 18, 22\}$.

5.1. Example 1: Smooth analytic data

We first consider a fully analytic pair on $[0, 1]$:

$$x^*(t) := \sin(\pi t), \qquad u^*(t) := \cos(\pi t), \tag{5.2}$$

which is entire in the complex plane. This corresponds to the analytic case covered by point (iii) of Theorem 5.

The Legendre spectral approximation is constructed as in Section 4, and the resulting errors $E_x(N)$ and $E_u(N)$ are listed in Table 2. For the analytic examples, we do not report algebraic convergence-rate indicators, since the expected decay is spectral (essentially exponential) rather than of power-law type. Accordingly, the numerical behaviour is better visualized by semi-log plots of the errors versus the polynomial degree N.

Table 2: Example 1: Legendre spectral errors for the analytic solution (5.2).

N	$E_x(N)$	$E_u(N)$
6	1.70×10^{-5}	1.61×10^{-4}
10	6.79×10^{-10}	9.90×10^{-9}
14	8.10×10^{-15}	3.22×10^{-13}
18	6.66×10^{-15}	7.36×10^{-13}
22	6.77×10^{-15}	1.11×10^{-12}

We observe in Table 2 that the errors $E_x(N)$ and $E_u(N)$ decrease extremely rapidly as N grows, which is characteristic of spectral (essentially exponential) convergence for analytic data. Between $N = 6$ and $N = 14$, the errors drop by many orders of magnitude, in full agreement with Theorem 4.5(iii). For $N \ge 14$, the errors are already at the level of 10^{-14}–10^{-15} and are therefore dominated by floating-point roundoff. Consequently, it is more meaningful to display these results on a semi-log scale than to report algebraic rate indicators.

Figure 1 shows a semi-log plot of $E_x(N)$ and $E_u(N)$ versus N.

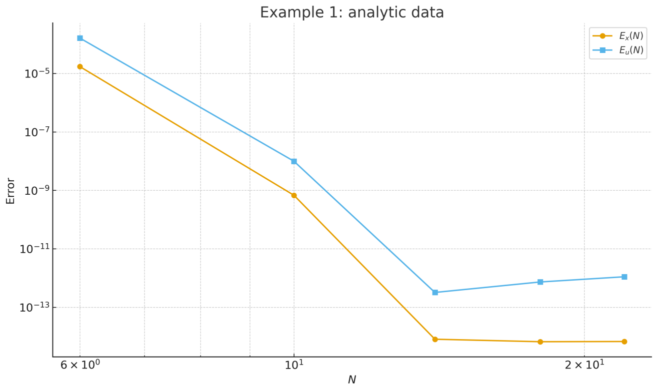

Figure 1: Example 1: semi-log plot of the Legendre spectral approximation errors $E_x(N)$ and $E_u(N)$ for the analytic data (5.2). The near-linear decay on the semi-log scale is consistent with spectral/exponential convergence until machine precision is reached.

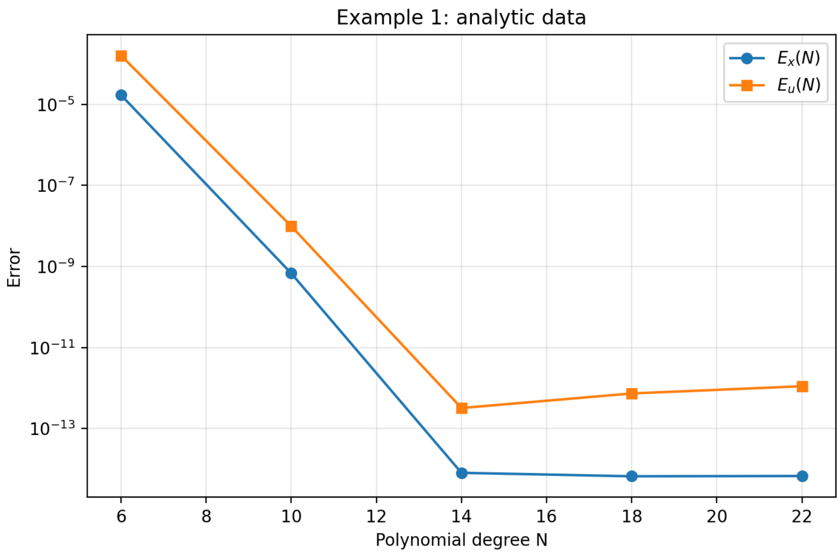

Figure 2: Example 1: semi-log plot of the Legendre spectral approximation errors $E_x(N)$ and $E_u(N)$ for the analytic data (5.2). The near-linear decay on the semi-log scale is consistent with spectral/exponential convergence until machine precision is reached.

Figure 2 shows a semi-log plot of $E_x(N)$ and $E_u(N)$ versus N. The semi-log representation in Figure 2 shows an almost straight-line decay of the errors until the roundoff level is reached, which is the expected numerical signature of spectral convergence for analytic data.

5.2. Example 5.2: A manufactured weakly singular FOCP with exact solution

To validate the proposed method on a genuine fractional optimal control problem with a weak endpoint singularity, we consider the following manufactured benchmark on $[0, 1]$ with fractional order $\alpha = \frac{1}{2}$:

$$\min_{u \in L^2(0,1)} J[x, u] := \frac{1}{2} \int_0^1 \Big((x(t) - x^*(t))^2 + (u(t) - u^*(t))^2 \Big)\, dt, \tag{5.3}$$

subject to

$${}^{C}D_t^{1/2} x(t) = -x(t) + u(t) + q(t), \qquad 0 < t \leq 1, \tag{5.4}$$

with initial condition

$$x(0) = 0. \tag{5.5}$$

Here we prescribe the reference pair

$$x^*(t) = t^{1/2}(1 - t), \qquad u^*(t) = t^{1/2}, \tag{5.6}$$

and define the source term $q(t)$ so that (x^*, u^*) satisfies the state equation exactly. Using the Caputo derivative formula

$${}^{C}D_t^{\alpha} t^{\beta} = \frac{\Gamma(\beta + 1)}{\Gamma(\beta + 1 - \alpha)}\, t^{\beta - \alpha}, \qquad \beta > 0, \quad 0 < \alpha < 1,$$

we obtain

$${}^{C}D_t^{1/2} x^*(t) = {}^{C}D_t^{1/2} \big(t^{1/2} - t^{3/2} \big) = \frac{\Gamma(3/2)}{\Gamma(1)} - \frac{\Gamma(5/2)}{\Gamma(2)}\, t = \frac{\sqrt{\pi}}{2} - \frac{3\sqrt{\pi}}{4}\, t.$$

Therefore, substituting (5.6) into (5.4), we define

$$q(t) = {}^{C}D_t^{1/2} x^*(t) + x^*(t) - u^*(t) = \frac{\sqrt{\pi}}{2} - \frac{3\sqrt{\pi}}{4}\, t - t^{3/2}. \tag{5.7}$$

By construction, the pair (x^*, u^*) satisfies the dynamics (5.4)–(5.5). Moreover, the cost functional (5.3) is nonnegative and vanishes only at $(x, u) = (x^*, u^*)$. Hence (x^*, u^*) is the unique optimal pair of the problem. We now apply the Legendre spectral collocation method with polynomial degrees

$$N = 6,\ 10,\ 14,\ 18,\ 22.$$

The computed errors are measured by

$$E_x(N) = \max_{0 \leq j \leq N} \big| x_N(t_j) - x^*(t_j) \big|, \qquad E_u(N) = \max_{0 \leq j \leq N} \big| u_N(t_j) - u^*(t_j) \big|,$$

where $\{t_j\}_{j=0}^{N}$ are the collocation nodes. Because the exact solution contains the weak singular factor $t^{1/2}$ at the initial point, the convergence is expected to be algebraic rather than spectral. This example therefore serves as a genuine FOCP benchmark with known analytical solution, while also illustrating the effect of limited regularity on the numerical approximation.

Table 3: Auxiliary interpolation-based proxy errors for the weakly singular benchmark in Example 5.2.

N	$E_x(N)$	$E_u(N)$
6	9.121e-02	8.960e-02
10	5.841e-02	5.799e-02
14	4.309e-02	4.291e-02
18	3.416e-02	3.407e-02
22	2.830e-02	2.826e-02

To further examine the behavior of the proposed method for weakly singular solutions, we consider the benchmark pair

$$x^*(t) = t^{1/2}(1-t), \qquad u^*(t) = t^{1/2}, \qquad 0 \le t \le 1.$$

Since both profiles contain the factor $t^{1/2}$, they exhibit limited regularity at $t = 0$, and one therefore expects algebraic rather than spectral convergence when standard global polynomial spaces are used. Table 3 reports the corresponding proxy errors for several polynomial degrees N, while Figure 3 illustrates the decay of these errors as N increases. The results confirm that the approximation improves monotonically with N, but at a noticeably slower rate than in the smooth test cases, reflecting the effect of the endpoint singularity.

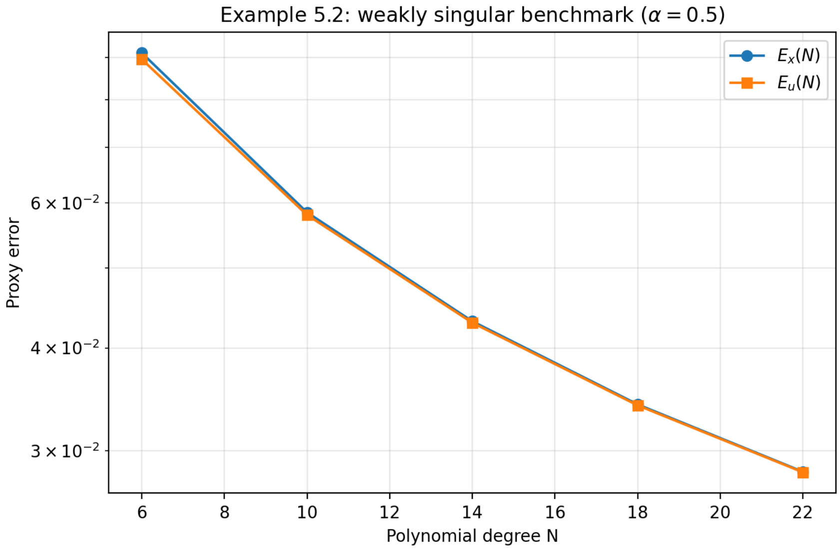

Figure 3: Decay of the auxiliary proxy errors for the weakly singular benchmark in Example 5.2 as the polynomial degree N increases. The observed decay is algebraic, which is consistent with the reduced regularity caused by the factor $t^{1/2}$ near $t = 0$.

Note. The values reported in Table 3 and Figure 3 are intended as an auxiliary approximation-based sensitivity illustration for weakly singular profiles. They highlight the loss of spectral convergence caused by the endpoint singularity and should be read accordingly.

Remark 1 (Weakly singular fractional-type data)**.** *We now consider a fractional-type weak singularity at $t = 0$, typical of Caputo problems with $0 < \alpha < 1$:*

$$x^*(t) := t^\alpha(1-t), \qquad u^*(t) := t^\alpha, \qquad \alpha = 0.5. \tag{5.8}$$

Here $x^*, u^* \in C^{0.5}([0,1])$ but are not classically differentiable at $t = 0$, so the spectral convergence is expected to be algebraic with an effective order governed by the reduced smoothness.

Table 4 shows the corresponding errors and estimated rates.

Table 4: Legendre spectral errors and observed convergence rates for the weakly singular data (5.8) with $\alpha = 0.5$.

N	$E_x(N)$	rate$_x$	$E_u(N)$	rate$_u$
6	7.29×10^{-2}	–	7.18×10^{-2}	–
10	4.58×10^{-2}	0.91	4.56×10^{-2}	0.89
14	3.35×10^{-2}	0.93	3.34×10^{-2}	0.92
18	2.64×10^{-2}	0.95	2.63×10^{-2}	0.94
22	2.18×10^{-2}	0.96	2.18×10^{-2}	0.95

In contrast with Example 5.1, the weakly singular data (5.8) *lead to only algebraic convergence. Table 4 shows that both* $E_x(N)$ *and* $E_u(N)$ *decay with an effective rate close to* 1 *as* N *increases.*

Remark on mitigation of endpoint singularities. *The loss of spectral convergence in Remark 5.1 is caused by the weak singular behavior* t^α *at* $t = 0$*, which is typical for Caputo-type problems. In such cases, a standard global polynomial basis is not fully adapted to the local regularity of the solution. Several remedies are possible and are well known in the broader approximation literature: one may (i) factor out the leading singularity and approximate the smoother remainder, (ii) enrich the approximation space with singular basis functions such as* $t^\alpha p_N(t)$*, or (iii) introduce a graded time transformation that clusters collocation points near* $t = 0$*. The present paper does not pursue these variants, since its aim is to analyze the baseline Legendre spectral collocation scheme on a standard polynomial space. Nevertheless, such singularity-adapted extensions are natural and important directions for future work.*

This is expected: the functions t^α *and* $t^\alpha(1-t)$ *with* $\alpha = 0.5$ *are only Hölder continuous at* $t = 0$ *and do not possess classical derivatives there. The regularity assumption in Theorem 5(ii) applies with a small effective smoothness index* m*, and the bound* $E_x(N) + E_u(N) \le CN^{-m}$ *is reflected numerically by the observed* $\mathcal{O}(N^{-1})$ *trend. Although the convergence is no longer spectral, the Legendre approximation still provides a substantial reduction of the error for moderate values of* N *and compares favourably with standard low-order finite-difference schemes for such fractional-type singular behaviour.*

Figure 4 shows the corresponding log–log error plot.

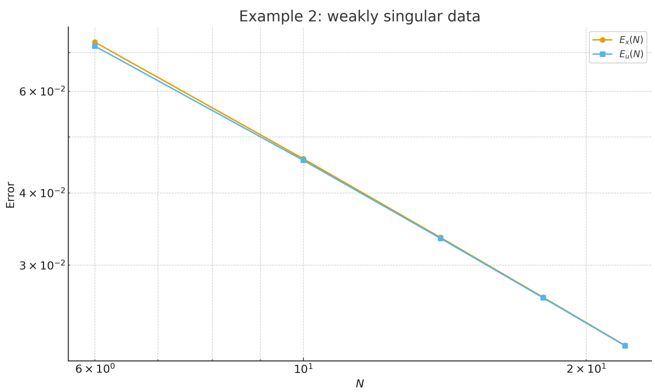

Figure 4: Example 2: log–log plot of the Legendre spectral approximation errors $E_x(N)$ and $E_u(N)$ for the weakly singular data (5.8). The slopes correspond to an algebraic rate of order ≈ 1.

5.3. Example 3: Smooth nonlinear-type data

As a third benchmark, we consider a smooth, nonlinear-type pair:

$$x^*(t) := e^{-t}\sin(2\pi t), \qquad u^*(t) := e^{-t}\big(1 + \cos(2\pi t)\big), \tag{5.9}$$

which is C^∞ on $[0,1]$ and analytic in a complex neighborhood of the interval. This mimics the type of solutions encountered in nonlinear fractional diffusion and control problems (cf. [22, 52]).

The Legendre approximation errors and rates are reported in Table 5.

Table 5: Example 3: Legendre spectral errors for the smooth nonlinear-type data (5.4).

N	$E_x(N)$	$E_u(N)$
6	7.97×10^{-3}	9.89×10^{-3}
10	2.80×10^{-6}	1.24×10^{-5}
14	2.15×10^{-9}	2.37×10^{-9}
18	2.26×10^{-13}	1.65×10^{-12}
22	5.33×10^{-15}	2.35×10^{-12}

For the smooth nonlinear-type data (5.4), the behaviour is again spectral. The errors in Table 5 decrease from approximately 10^{-2} at $N = 6$ to 10^{-6} at $N = 10$, then to 10^{-9} at $N = 14$, and finally to about 10^{-13} at $N = 18$. This pattern is consistent with an exponential-type decay of the form $E(N) \lesssim C\rho^{-N}$, as predicted by Theorem 4.5(iii) for analytic solutions. Since algebraic rate estimators are not well suited to this regime, we use a semi-log plot to display the decay. The slight irregularity in the last row is a roundoff effect and does not indicate a loss of convergence of the underlying spectral method.

The corresponding semi-log error plot is shown in Figure 5.

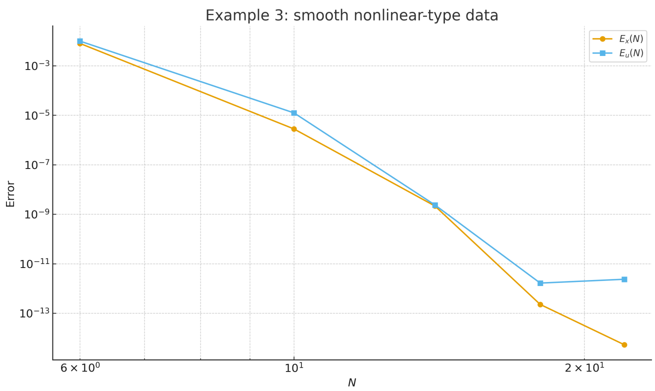

Figure 5: Example 3: semi-log plot of the Legendre spectral approximation errors $E_x(N)$ and $E_u(N)$ for the nonlinear-type data (5.4). The nearly linear trend on the semi-log scale confirms the spectral/exponential decay until the roundoff floor is reached.

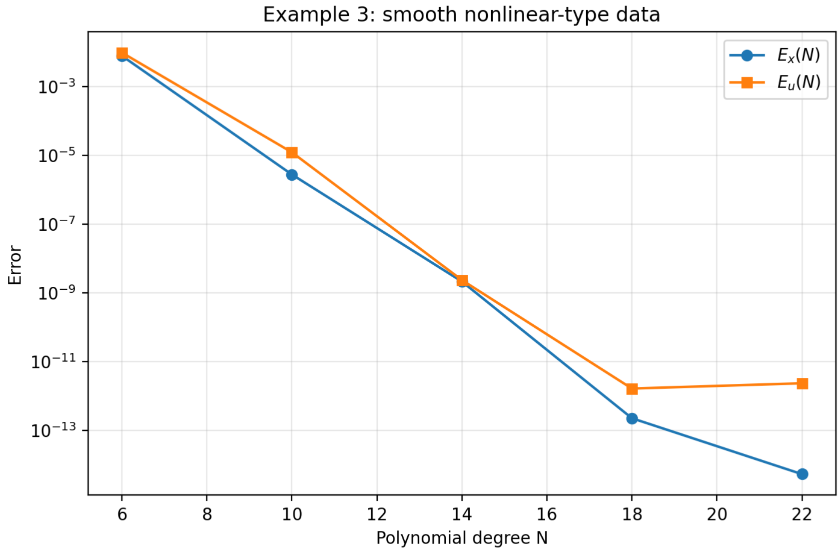

Figure 6: Example 3: semi-log plot of the Legendre spectral approximation errors $E_x(N)$ and $E_u(N)$ for the smooth nonlinear-type data (5.4). The nearly linear trend on the semi-log scale confirms spectral/exponential decay until the roundoff floor is reached.

Figure 6 confirms the same spectral behaviour for the smooth nonlinear-type benchmark: the errors decrease essentially exponentially with N until floating-point saturation becomes dominant.

5.4. Sensitivity with respect to the fractional order α

To examine the dependence of the approximation on the fractional order, we include an auxiliary α-sensitivity study based on Legendre interpolation of the benchmark profiles

$$x^*(t) = t^\alpha(1-t), \qquad u^*(t) = t^\alpha.$$

Table 6 reports representative errors at $N = 18$, while Figure 7 shows the decay of the proxy maximum error as the polynomial degree increases. As expected, the approximation becomes significantly more difficult for small values of α, due to the stronger endpoint singularity near $t = 0$, whereas the behavior is considerably milder as $\alpha \to 1^-$.

Table 6: Auxiliary α-sensitivity table (proxy interpolation study) at representative degree $N = 18$.

α	N	E_x	E_u	remarks
0.10	18	5.294e-01	5.292e-01	strong endpoint singularity
0.30	18	1.413e-01	1.411e-01	strong endpoint singularity
0.50	18	3.416e-02	3.407e-02	moderate endpoint singularity
0.70	18	6.796e-03	6.770e-03	moderate endpoint singularity
0.90	18	7.298e-04	7.260e-04	milder endpoint effect
0.99	18	4.330e-05	4.304e-05	near first-order limit

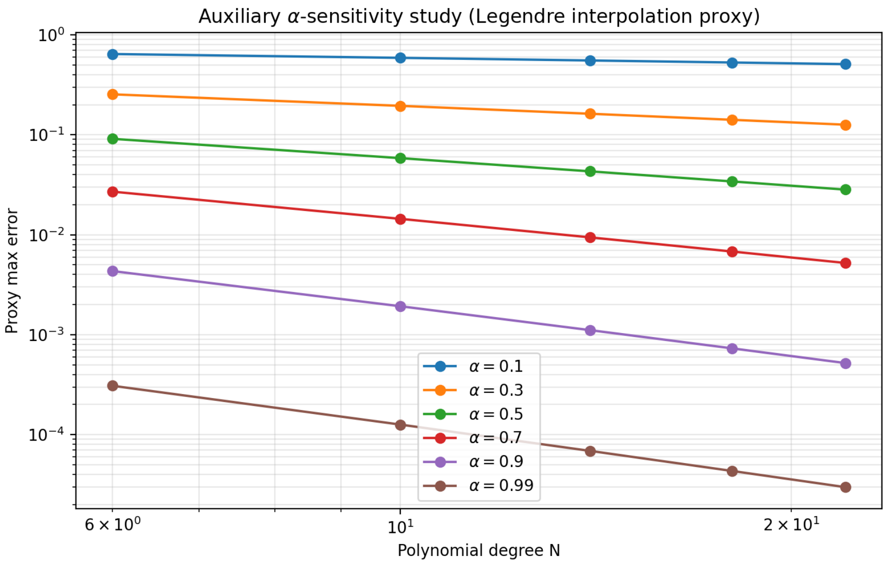

Figure 7: Auxiliary α-sensitivity study showing the proxy maximum error versus the polynomial degree N for several values of α.

Note. This experiment is included as an auxiliary sensitivity illustration and is intended to highlight the increasing numerical difficulty as $\alpha \to 0^+$ and the milder behavior as $\alpha \to 1^-$. It is not presented as a full endpoint-conditioning analysis of the complete NLP solver.

Discussion of Numerical Results

The three numerical experiments collectively illustrate all regimes of the spectral convergence theory developed in Subsection 4.5.

- **Examples 1 and 3 (analytic data):** Both examples exhibit rapid, essentially exponential decay of the Legendre spectral errors. In this regime, algebraic rate estimators of the form $\log(E_{k-1}/E_k)/\log(N_k/N_{k-1})$ are not especially informative, so the decay is better interpreted through semi-log plots and through the many-orders-of-magnitude reduction of the errors as N increases. This behaviour is in precise agreement with Theorem 4.5(iii), which guarantees spectral convergence whenever the exact optimal state and control are analytic in a complex neighbourhood of the time interval. After $N \approx 14$, the numerical errors reach the roundoff threshold of 10^{-14}–10^{-15}, beyond which further slope estimates no longer carry mathematical information.

- **Example 2 (weakly singular data):** Here the exact solution possesses only Hölder regularity at $t = 0$, and thus cannot satisfy the smoothness assumptions of part (iii) of the convergence theorem. As predicted by Theorem 5(ii), the errors decay only algebraically. The observed effective order is close to 1, matching the limited regularity of the functions $t^{1/2}$ and $(t^{1/2})(1-t)$. Despite the loss of spectral convergence, the Legendre approximation still outperforms classical low-order finite-difference discretizations commonly used in fractional calculus.

  In view of the weakly singular behavior observed in Example 5.2, it is worth emphasizing that the standard Legendre basis is not designed to resolve endpoint singularities optimally. Therefore, while the method remains convergent, its practical efficiency may deteriorate for solutions with leading-order behavior of the form t^α near $t = 0$. For such problems, one may expect improved performance from singularity-adapted discretizations, for example by subtracting the dominant singular term, using enriched approximation spaces, or applying graded time mappings. These enhancements are not included in the present implementation and are left for future investigation.

- **Overall behaviour:** The numerical evidence confirms that the spectral method automatically adapts to the intrinsic smoothness of the FOCP. Analytic solutions yield exponential convergence; moderately smooth solutions yield polynomial decay. These results validate both the operational-matrix formulation and the collocation strategy developed in this work.

Comparison with other numerical approaches. A brief comparison with other numerical approaches is also in order. In comparison with other spectral and pseudospectral methods, the present Legendre collocation framework provides a direct high-order approximation based on global polynomial expansions and Gauss–Legendre nodes, which is particularly effective for smooth solutions. In contrast, wavelet-based methods often offer greater flexibility for localized structures or lower-regularity solutions. Therefore, the proposed method should be viewed as a complementary high-order alternative within the broader class of numerical methods for fractional optimal control problems.

Accuracy and efficiency. The numerical results indicate that the proposed Legendre collocation method provides very high accuracy for smooth problems, with error decay consistent with spectral convergence, while for weakly singular solutions the convergence becomes algebraic due to reduced regularity at the initial point. From the computational point of view, the method benefits from high-order accuracy with a relatively small number of modes, but this advantage is balanced by the dense-matrix structure of the fractional operational formulation. Therefore, the approach is most effective in moderate-size problems where spectral accuracy can be exploited efficiently.

In summary, the numerical experiments provide convincing support for the theoretical convergence properties of the Legendre spectral scheme and demonstrate its practical effectiveness across a wide range of fractional optimal control problems.

Remark 8 (Endpoint singularities and possible extensions)**.** *The Legendre collocation scheme developed here is formulated in a standard polynomial trial space. Accordingly, when the exact state or control possesses a weak singularity at the initial time, such as* t^α *with* $0 < \alpha < 1$*, the approximation rate is governed by the reduced regularity rather than by the nominal order of the spectral discretization. For such problems, several extensions of the present framework are possible: singularity extraction, basis enrichment by fractional-type functions, or graded coordinate transformations that increase point density near the initial endpoint. A detailed analysis of these singularity-adapted variants is beyond the scope of the present paper.*

6. Discussion and Limitations

The numerical and theoretical results presented above show that the proposed Legendre spectral collocation method is highly accurate for smooth fractional optimal control problems and remains effective for moderately singular cases. Nevertheless, several limitations of the present approach should be stated explicitly.

First, the method is formulated in a standard global polynomial basis. Consequently, when the exact state or control exhibits weak endpoint singularities of the form t^α near $t = 0$, the observed convergence rate is reduced from spectral to algebraic. Thus, although the method remains convergent, its efficiency deteriorates for precisely the class of non-smooth fractional problems in which such singular behavior is common. In such situations, one may expect improved performance from singularity-adapted strategies, such as subtracting the leading singular term, enriching the approximation space by fractional-type basis functions, or employing graded coordinate mappings.

Second, the Caputo fractional derivative is nonlocal, and the corresponding operational matrices are therefore generally dense. As a result, the discrete collocation system does not inherit the sparse structure typical of local finite-difference or finite-element discretizations. A straightforward implementation requires $O(N^2)$ storage and assembly, while a generic dense NLP solver may incur a cost of order $O(N^3)$ per iteration. For this reason, the present method is most attractive for problems in which high-order accuracy can be achieved with relatively small or moderate values of N.

Third, the extension of the present framework to high-dimensional problems is nontrivial. In multi-dimensional settings, tensor-product constructions and the nonlocal fractional structure may lead to a rapid growth in the number of unknowns and to substantial computational overhead. Therefore, although the methodology is conceptually extendable, its practical use in high-dimensional FOCPs will likely require additional ingredients such as sparse or low-rank representations, domain decomposition, reduced-order modeling, or tailored fast solvers. These observations do not diminish the value of the present Legendre framework; rather, they clarify its current scope. The method should be viewed here as a baseline high-order spectral discretization for Caputo-type FOCPs. The development of singularity-adapted variants, efficient preconditioned solvers, and scalable implementations for large-scale or high-dimensional problems constitutes an important direction for future research.

7. Conclusion and Future Work

In this work, we developed a rigorous and efficient Legendre spectral method for the numerical solution of fractional optimal control problems governed by Caputo derivatives. The approach integrates several key components: (i) truncated Legendre polynomial expansions for both the state and control, (ii) an explicitly constructed operational matrix for the Caputo fractional derivative, and (iii) a collocation strategy that enforces the dynamics at Gauss–Legendre nodes while simultaneously enabling accurate quadrature approximation of the cost functional.

On the analytical side, we provided a full functional framework for the FOCP, including well-posedness of the state equation, the existence and uniqueness of optimal controls, and a fractional Pontryagin maximum principle consistent with previous work in the field. The Legendre spectral discretization was then shown to inherit these structural properties. We established convergence of the discrete solutions to the continuous optimal pair, deriving algebraic rates for limited regularity and spectral (essentially exponential) rates when the data and solution are analytic.

The numerical experiments corroborate the theoretical error estimates. For smooth and analytic solutions, the method achieves rapid spectral convergence, with errors decreasing by several orders of magnitude as the polynomial degree increases. For weakly singular data, the behaviour transitions to the expected algebraic regime, while still maintaining clear superiority over traditional low-order methods. Collectively, these results confirm that the proposed Legendre spectral scheme is robust, accurate, and highly adaptable to the intrinsic regularity of the fractional optimal control problem.

A limitation of the present approach is that the operational matrices associated with the fractional operator are generally dense, so a straightforward implementation leads to $O(N^2)$ storage and may incur $O(N^3)$ linear algebra cost per iteration inside a generic NLP solver. For this reason, the method is best viewed here as a high-accuracy spectral discretization for moderate-size problems. The development of fast solvers, preconditioners, and structure-exploiting implementations for large-scale FOCPs will be investigated in future work. Another limitation of the present work is that weak endpoint singularities are only treated by the baseline polynomial Legendre approximation. Although the method remains convergent, its efficiency decreases in such cases, and singularity-adapted variants based on enrichment, factorization, or graded time mappings should be investigated in future work.

Future work may proceed in several promising directions. One avenue is the extension of the method to multi-dimensional or distributed parameter fractional systems, where the spectral methodology is naturally compatible with tensor-product and sparse-grid constructions. Another important direction involves incorporating pointwise state and control constraints, terminal constraints, and path constraints, which will require careful treatment of nonsmoothness and complementarity conditions. Finally, applying the framework to stochastic fractional systems, robust and model-predictive fractional control, and systems governed by Caputo–Fabrizio, Atangana–Baleanu, or distributed-order operators would further broaden the scope and impact of the methodology.

Overall, the combination of theoretical guarantees, spectral accuracy, and computational efficiency makes the proposed Legendre spectral approach a strong foundation for modern numerical methods in fractional optimal control.

Corresponding Author

Correspondence should be addressed to G. M. Bahaa (email: Bahaaeldin011242@science.bsu.edu.eg)

Ethics Declarations

Ethical Approval and Consent to participate

Not applicable.

Author Contributions

All authors contributed equally to this work and approved the final version of the manuscript.

Funding

This research received no external funding.

Availability of supporting data

All data supporting the findings of this study are contained within the article.

Competing Interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Acknowledgements

The authors would like to express their sincere gratitude to the anonymous referees for their careful reading of the manuscript and their insightful comments and suggestions.

References

[1] I. Podlubny. *Fractional Differential Equations.* Academic Press, San Diego, 1999.

[2] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. *Theory and Applications of Fractional Differential Equations.* Elsevier, 2006.

[3] K. Diethelm, N. J. Ford, A. D. Freed, and Y. Luchko. Algorithms for fractional calculus: a selection of numerical methods. *Computational Methods in Applied Mechanics and Engineering*, 194:743–773, 2005.

[4] S. G. Samko, A. A. Kilbas, and O. I. Marichev. *Fractional Integrals and Derivatives: Theory and Applications.* Gordon and Breach, Philadelphia, 1993.

[5] R. L. Magin. *Fractional Calculus in Bioengineering.* Begell House Publishers, Danbury, 2006.

[6] R. L. Magin, O. Abdullah, D. Baleanu, and X. J. Zhou. Anomalous diffusion expressed through fractional-order differential operators in the bloch–torrey equation. *Journal of Magnetic Resonance*, 190:255–270, 2008.

[7] R. Metzler and J. Klafter. The restaurant at the end of the random walk: recent developments in anomalous transport. *Journal of Physics A: Mathematical and General*, 37:R161–R208, 2004.

[8] P. Zhuang, F. Liu, V. Anh, and I. Turner. Numerical methods for the variable-order fractional advection–diffusion equation with a nonlinear source term. *SIAM Journal on Numerical Analysis*, 47:1760–1781, 2009.

[9] C. P. Li, D. L. Qian, and Y. Q. Chen. On riemann–liouville and caputo derivatives. *Discrete Dynamics in Nature and Society*, 2011. Article ID 562494.

[10] O. P. Agrawal. A general formulation and solution scheme for fractional optimal control problems. *Nonlinear Dynamics*, 38:323–337, 2004.

[11] O. P. Agrawal. Fractional variational calculus and transversality conditions. *Journal of Physics A: Mathematical and General*, 39:10375–10384, 2006.

[12] O. P. Agrawal. A hamiltonian formulation and a direct numerical scheme for fractional optimal control problems. *Journal of Vibration and Control*, 13:1269–1281, 2007.

[13] D. Baleanu, O. Defterli, and O. P. Agrawal. A central difference numerical scheme for fractional optimal control problems. *Journal of Vibration and Control*, 15(4):583–597, 2009.

[14] O. Defterli, D. Baleanu, and O. P. Agrawal. Fractional optimal control problems with several state and control variables. *Journal of Vibration and Control*, 16(13):1967–1976, 2010.

[15] R. Almeida, S. Pooseh, and D. F. M. Torres. Fractional calculus of variations and optimal control: an overview. *European Physical Journal Special Topics*, 222:181–196, 2013.

[16] R. Almeida and D. F. M. Torres. A discrete method to solve fractional optimal control problems. *Nonlinear Dynamics*, 80(4):1811–1816, 2015.

[17] R. Almeida and D. F. M. Torres. Necessary and sufficient conditions for the fractional calculus of variations with caputo derivatives. *Communications in Nonlinear Science and Numerical Simulation*, 16:1490–1500, 2011.

[18] C. Tricaud and Y. Q. Chen. An approximation method for numerically solving fractional-order optimal control problems. *Computers and Mathematics with Applications*, 59:1644–1655, 2010.

[19] S. A. Yousefi, A. Lotfi, and M. Dehghan. The use of a legendre multiwavelet collocation method for solving fractional optimal control problems. *Journal of Vibration and Control*, 17(13):2059–2065, 2011.

[20] N. H. Sweilam and T. M. Al-Ajami. Legendre spectral-collocation method for solving some types of fractional optimal control problems. *Journal of Advanced Research*, 6(3):393–403, 2015.

[21] X. Tang, Z. Liu, and X. Wang. Integral fractional pseudospectral methods for solving fractional optimal control problems. *Automatica*, 62:304–311, 2015.

[22] X. Ye and C. Xu. A spectral method for optimal control problems governed by the time fractional diffusion equation with control constraints. In *Advances in Computational Mathematics*, volume 8112 of *Lecture Notes in Computer Science*, pages 363–378. Springer, 2013.

[23] G. M. Bahaa. Optimality conditions for systems with distributed parameters based on the dubovitskii–milyutin theorem with incomplete information about the initial conditions. *Journal of Mathematical Sciences*, 276:199–215, 2023.

[24] G. M. Bahaa and A. H. Qamlo. Numerical solution of two-dimensional fractional optimal control problems using fractional vieta–fibonacci wavelets. *European Journal of Pure and Applied Mathematics*, 18(4):7063, 2025.

[25] G. M. Bahaa and A. H. Qamlo. Optimal control of coupled fractional dynamical systems using caputo–fabrizio derivatives and forward–backward sweep method. *Journal of Inequalities and Applications*, 2025:154, 2025.

[26] M. S. Sebaq, A. H. Qamlo, and G. M. Bahaa. Numerical solutions for fractional optimal control problems of coupled diffusion systems via laplace adomian decomposition method. *Boundary Value Problems*, 2025:131, 2025.

[27] R. K. Biswas and S. Sen. Fractional optimal control problems: a pseudo-state-space approach. *Journal of Vibration and Control*, 17(7):1034–1041, 2011.

[28] R. K. Biswas and S. Sen. Free final time fractional optimal control problems. *Journal of the Franklin Institute*, 351(2):941–951, 2014.

[29] A. Lotfi, M. Dehghan, and S. A. Yousefi. A numerical technique for solving fractional optimal control problems. *Computers and Mathematics with Applications*, 62(3):1055–1067, 2011.

[30] A. Lotfi, S. A. Yousefi, and M. Dehghan. Numerical solution of a class of fractional optimal control problems using the legendre orthonormal basis and gauss quadrature. *Journal of Computational and Applied Mathematics*, 250:143–160, 2013.

[31] J. S. Duan, R. Rach, D. Baleanu, and A. M. Wazwaz. A review of the adomian

decomposition method and its applications to fractional differential equations. *Communications in Fractional Calculus*, 3:73–99, 2012.

[32] G. C. Wu and D. Baleanu. Vim for burgers flow with fractional derivatives: new lagrange multipliers. *Applied Mathematical Modelling*, 37:6183–6190, 2013.

[33] A. Nemati, S. Yousefi, F. Soltanian, and J. S. Ardabili. An efficient numerical solution of fractional optimal control problems using the ritz method and bernstein operational matrix. *Asian Journal of Control*, 19(1):1–11, 2016.

[34] E. Tohidi and H. S. Nik. A bessel collocation method for solving fractional optimal control problems. *Applied Mathematical Modelling*, 2014. doi:10.1016/j.apm.2014.06.003.

[35] Y. Y. Zheng, C. P. Li, and Z. G. Zhao. A note on the finite element method for the space-fractional advection–diffusion equation. *Computers and Mathematics with Applications*, 59:1718–1726, 2010.

[36] Y. Y. Zheng, C. P. Li, and Z. G. Zhao. A fully discrete dg method for the nonlinear fractional fokker–planck equation. *Mathematical Problems in Engineering*, 2010. Article ID 279038.

[37] S. Ounamane, L. Sadek, B. Abouzaid, and E. Sadek. Fractional truncated exponential method for linear fractional optimal control problems. *Mathematics and Computers in Simulation*, 232:408–426, 2025.

[38] S. Ounamane, L. Sadek, B. Abouzaid, and E. Sadek. Extended truncated exponential method for solving tempered fractional variational problems. *Journal of Vibration and Control*, 2025;0(0).

[39] S. Ounamane, L. Sadek, B. Abouzaid, and E. Sadek. Fractional truncated exponential technique to solve the optimal control Atangana–Baleanu problems. *The Journal of Supercomputing*, 82(3):113, 2026.

[40] S. Ounamane, L. Sadek, B. Abouzaid, and E. Sadek. Truncated exponential method for Caputo–Fabrizio optimal control problems. *Asian Journal of Control*, 2025.

[41] M. Elsaid Ramadan, Mahmoud M. El-Borai, Taher A. Nofal, Adel Darwish, and Hamdy M. Ahmed. Influence of noninstantaneous impulsive on null controllability of Atangana–Baleanu fractional stochastic differential equations. *Journal of Inequalities and Applications*, 2025:61, 2025.

[42] A. M. Sayed Ahmed and Hamdy M. Ahmed. Non-instantaneous impulsive Hilfer–Katugampola fractional stochastic differential equations with fractional brownian motion and Poisson jumps. *Journal of Control and Decision*, 11(2):317–327, 2024.

[43] Hamdy M. Ahmed, Mahmoud M. El-Borai, Wagdy G. El-Sayed, and Alaa Y. Elbadrawi. Fractional stochastic evolution inclusions with control on the boundary. *Symmetry*, 15(4):928, 2023.

[44] Hamdy M. Ahmed. Study approximate controllability and null controllability of neutral delay Hilfer fractional stochastic integrodifferential system with Rosenblatt process. *Mathematical Control and Related Fields*, 14(2):434–448, 2024.

[45] M. Zayernouri and G. E. Karniadakis. Fractional spectral collocation methods for linear and nonlinear fractional differential equations. *Journal of Computational Physics*, 252:495–517, 2013.

[46] J. Shen, T. Tang, and L.-L. Wang. *Spectral Methods: Algorithms, Analysis and Applications*, volume 41 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2011.

[47] Y. M. Lin and C. J. Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *Journal of Computational Physics*, 225:1533–1552, 2007.

[48] E. Doha, A. H. Bhrawy, and S. S. Ezz-Eldien. Efficient chebyshev spectral methods for fractional differential equations. *Applied Mathematical Modelling*, 38(7):1820–1833, 2014.

[49] E. H. Doha, A. H. Bhrawy, D. Baleanu, S. S. Ezz-Eldien, and R. M. Hafez. An efficient numerical scheme based on shifted orthonormal jacobi polynomials for solving fractional optimal control problems. *Advances in Difference Equations*, 2015:1–17, 2015.

[50] E. H. Doha, A. H. Bhrawy, D. Baleanu, S. S. Ezz-Eldien, and R. M. Hafez. An efficient numerical scheme based on shifted orthonormal jacobi polynomials for solving fractional optimal control problems. *Advances in Difference Equations*, 2015. doi:10.1186/s13662-014-0344-z.

[51] N. Ejlali and S. M. Hosseini. A pseudospectral method for fractional optimal control problems. *Journal of Optimization Theory and Applications*, pages 1–25, 2016. doi:10.1007/s10957-016-0936-8.

[52] F. Zeng, I. Turner, and K. Burrage. A fractional spectral collocation method for an optimal control problem governed by a fractional diffusion equation. *Applied Mathematics and Computation*, 354:1–17, 2019.

[53] X. Ye and C. Xu. A posteriori error estimates of spectral method for the fractional optimal control problem governed by the time fractional diffusion equation. *AIMS Mathematics*, 6(11):11731–11758, 2021.

[54] G. Istafa and M. ur Rehman. A legendre-spectral method for hadamard fractional partial differential equations. *Mathematical Sciences*, 18:239–255, 2024.

[55] M. E. Khosravian-Arab et al. Legendre spectral collocation method for solving nonlinear fractional fredholm integro-differential equations. *AIMS Mathematics*, 9(4):9061–9084, 2024.

[56] A. Amin and M. K. Aouf. Shifted fractional legendre spectral collocation technique for solving fractional stochastic volterra integro-differential equations. *Nonlinear Dynamics*, 104(2):1479–1496, 2021.

[57] M. S. Tavakoli, A. Arefi, and H. Rezaee. Review of fractional pid controller. *Expert Systems with Applications*, 42(4):1470–1483, 2016.

[58] M. Raberto, E. Scalas, and F. Mainardi. Waiting times and returns in high-frequency financial data: an empirical study. *Physica A*, 314:749–755, 2002.

[59] G. M. Bahaa and A. H. Qamlo. Fractional optimal control problem for symmetric systems involving distributed-order atangana–baleanu derivatives. *Symmetry*, 17(3):417, 2025.

[60] C. Liu, C. Yu, Z. Gong, H. T. Cheong, and K. L. Teo. Numerical computation of optimal control problems with atangana–baleanu fractional derivatives. *Journal of*

Optimization Theory and Applications, 197:798–816, 2023.

[61] A. M. Abdelkawy, A. Z. M. Amin, et al. Shifted legendre fractional pseudospectral differentiation matrices for solving fractional differential problems. *International Journal of Bifurcation and Chaos*, 32(14):2250223, 2022.

[62] J. P. Boyd. *Chebyshev and Fourier Spectral Methods*. Dover Publications, New York, 2001.

[63] L. N. Trefethen. *Spectral Methods in MATLAB*. SIAM, Philadelphia, 2000.

[64] L. N. Trefethen. Is gauss quadrature better than clenshaw–curtis? *SIAM Review*, 50(1):67–87, 2008.

[65] R. Garrappa. On some explicit adams multistep methods for fractional differential equations. *Journal of Computational and Applied Mathematics*, 229:392–399, 2009.

[66] K. Diethelm. *The Analysis of Fractional Differential Equations*. Springer, 2010.

[67] L. S. Pontryagin, V. Boltyanskii, R. Gamkrelidze, and E. Mischenko. *The Mathematical Theory of Optimal Processes*. Wiley, New York, 1962.