STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **8**, December 2020, pp 858–875.
Published online in International Academic Press (www.IAPress.org)

A General Framework for Optimal Control of Fractional Nonlinear Delay Systems by Wavelets

Iman Malmir *

Department of Mechanical Engineering, Buali Sina University, Hamedan, Iran

Abstract An iterative procedure to find the optimal solutions of general fractional nonlinear delay systems with quadratic performance indices is introduced. The derivatives of state equations are understood in the Caputo sense. By presenting and applying a general framework, we use the Chebyshev wavelet method developed for fractional linear optimal control to convert fractional nonlinear optimal control problems as a sequence of quadratic programming ones. The concepts and computational procedure that are used for fractional linear optimal control are applied on fractional nonlinear optimal control. Different types of nonlinear optimal control problems with fractional or integer order can be solved. To see this, some numerical examples are solved. Another operational property of Chebyshev wavelets is presented.

Keywords Fractional constrained nonlinear delay system, Fractional time-delay optimal control, Wavelet method, Fractional time-varying delay system, Quasilinearization method, Isoperimetric constraint

AMS 2010 subject classifications 26A33, 65T60, 93C10, 90C20

DOI: 10.19139/soic-2310-5070-939

1. Introduction

In some processes like chemical, electronic, aerospace and mechanical processes, we have to deal with time-delay systems. The optimal control problem is to find the optimal solutions which cause a dynamical system to satisfy some conditions and at the same time minimize a performance index [1, 2, 3]. To carry out the computational procedures developed for linear optimal control systems to nonlinear optimal control delay-free systems, we may use the quasilinearization method [4] which is based on replacing a nonlinear optimal control problem by a sequence of linear optimal control problems. The iterative procedure has been studied for delay-free optimal control problems in some works like [5, 6, 7]. In this work, we are going to extend this iterative procedure to fractional time-delay optimal control problems. One of the most important features of a method for solving the optimal control problems is its generality. For handling fractional nonlinear delay optimization in the previous research [8, 9, 10], some methods for obtaining the solutions of fractional optimal control of nonlinear delays systems have been presented. The main idea in these works is converting a nonlinear optimal control problem to a system of algebraic equations. Regardless of how difficult this is, one has to calculate the value of the optimal cost by substituting the approximations of the state and control vectors in the cost function of the system under consideration and this causes at the cost of more computations. Also, some of the methods are only applicable to a specific class of nonlinear optimal control problems, for example see [10]. We proposed a simple method in [11] based on an exact Riemann–Liouville fractional integration operational matrix of Chebyshev wavelets to obtain the optimal control of fractional linear quadratic systems from solving quadratic optimization problems and without doing any work, the value of optimal cost is reached as a default output of the quadprog solver. Whereas some of

*Correspondence to: Iman Malmir (Email: iman.malmir@outlook.com).

ISSN 2310-5070 (online) ISSN 2311-004X (print)

real-world systems are modeled by nonlinear differential equations, for example a model of machine tool vibrations in the turning of metals [12], and differential equations of fractional orders are valuable tools in modeling of some phenomena in various fields of engineering [13, 14, 15], now the question is: Whether the method can be applied to fractional nonlinear systems? Of course, we implicitly implemented the method on a simple nonlinear example in [11]. In this research, we go into greater detail and illustrate how by a general useful wavelet-based method we can solve such problems. In addition, we implement Chebyshev wavelet method to some engineering systems. We see that the method can be applied to a wide variety of nonlinear systems. Also, it can be used for constrained and unconstrained systems having integer orders and/or fractional orders. The procedure proposed in this work, with some modifications, can be utilized in useful wavelet-based methods like Legendre wavelet [16] and Hermite wavelet [17] methods or other wavelet methods such as those based on Boubaker wavelets [18], Taylor wavelets [19], and etc.

2. Approximation process

2.1. Basic definitions of fractional operators

Definition 1
The Riemann–Liouville fractional integral operator of order α denoted by $\mathcal{I}^\alpha$ of a function $f(t)$ is defined by [11]

$$\mathcal{I}^\alpha f(t) = \begin{cases} \frac{1}{\Gamma(\alpha)} \int_0^t (t-\rho)^{\alpha-1} f(\rho) d\rho, & \alpha > 0 \\ f(t), & \alpha = 0, \end{cases}$$

where $\Gamma(\alpha)$ is the gamma function.

Definition 2
The Caputo fractional derivative of order α denoted by the symbol $\mathcal{D}^\alpha$ of a function $f(t)$ is defined by [20]

$$\mathcal{D}^\alpha f(t) = \begin{cases} \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-\rho)^{n-\alpha-1} f^{(n)}(\rho) d\rho, & n-1 < \alpha < n, n \in \mathbb{N} \\ \frac{d^n}{dt^n} f(t), & \alpha = n. \end{cases}$$

This operator arise in modeling of some physical systems. The following property will be needed

$$\mathcal{I}^\alpha \mathcal{D}^\alpha f(t) = f(t) - \sum_{\nu=0}^{n-1} f^{(\nu)}(0) \frac{t^\nu}{\nu!}.$$

2.2. An iterative procedure for optimal control of nonlinear systems

We select the general form of the performance index which includes Terminal Control, Minimum Control Energy, and Regulator problems as

$$J = \tfrac{1}{2} \mathbf{x}^\top(1) \mathbf{T} \mathbf{x}(1) + \tfrac{1}{2} \int_0^1 \left\{ \mathbf{x}^\top(t) \mathbf{Q}(t) \mathbf{x}(t) + \mathbf{u}^\top(t) \mathbf{R}(t) \mathbf{u}(t) \right\} dt, \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{R}^q$ and $\mathbf{u}(t) \in \mathbb{R}^r$ are the state and control vectors, $\mathbf{T} \in \mathbb{R}^{q \times q}$ and $\mathbf{Q}(t) \in \mathbb{R}^{q \times q}$ and $\mathbf{R}(t) \in \mathbb{R}^{r \times r}$ are matrices. It indicates that our objective is to keep the terminal state close to its desired value and the desired state vector close to its steady-state value with consuming minimum control energy [2]. The performance index (1) (quadratic form) leads to some very elegant results in optimal control systems [3].

Consider a fractional nonlinear time-delay system described by

$$\mathcal{D}^\alpha \mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{x}(t-h(t)), \mathbf{x}(t-h_x), \mathbf{u}(t-h_u), t), \tag{2}$$

$$\mathbf{x}(0) = \mathbf{x}_0, \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0, \begin{cases} \mathbf{x}(t) = \boldsymbol{\theta}(t), & t < 0 \\ \mathbf{u}(t) = \boldsymbol{\zeta}(t), & -h_u \le t \le 0, \end{cases} \tag{3}$$

where $0 < \alpha \leq 2$, $t \in [0,1]$, $\mathbf{f} \in \mathbb{R}^q$ is a nonlinear (differentiable) vector function, h_x and h_u are delays, $h(t)$ is a piecewise constant delay, and $\boldsymbol{\theta}(t) \in \mathbb{R}^q$ and $\boldsymbol{\zeta}(t) \in \mathbb{R}^r$ are specified initial functions. This fractional time-delay system is to be controlled to minimize the quadratic performance index given in (1).

Lemma 1
Consider the nonlinear delay optimal control problem which to minimize (1) subject to

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t),$$

where $\mathbf{x}(0) = \mathbf{x}_0$ is the initial condition, $\mathbf{f}$ is continuous in $(t, \mathbf{x}, \mathbf{u})$ and nonlinearity satisfies a local Lipschitz condition. The problem can be replaced by the following sequence of linear fractional delay problems which the sequence converges to a solution
for $i \geq 1$, minimize

$$J^{[i]} = \tfrac{1}{2}{\mathbf{x}^{[i]}}^{\top}(1)\mathbf{T}\mathbf{x}^{[i]}(1) + \tfrac{1}{2}\int_0^1 \left\{{\mathbf{x}^{[i]}}^{\top}(t)\mathbf{Q}(t)\mathbf{x}^{[i]}(t) + {\mathbf{u}^{[i]}}^{\top}(t)\mathbf{R}(t)\mathbf{u}^{[i]}(t)\right\} dt \tag{4}$$

subject to

$$\dot{\mathbf{x}}^{[i]}(t) = \mathbf{A}(\mathbf{x}^{[i-1]}(t))\mathbf{x}^{[i]}(t) + \mathbf{B}(\mathbf{x}^{[i-1]}(t), \mathbf{u}^{[i-1]}(t))\mathbf{u}^{[i]}(t) + \mathbf{d}^{[i-1]}(t), \tag{5}$$

where $[i]$ represents the iteration, $\mathbf{x}^{[0]}(t) = \mathbf{x}_0$, $\mathbf{u}^{[0]}(t) = \mathbf{0}$, $\mathbf{A} \in \mathbb{R}^{q\times q}$, $\mathbf{B} \in \mathbb{R}^{q\times r}$ and $\mathbf{d} \in \mathbb{R}^q$.

Proof
We prove this Lemma in like manner presented in [5]. To show that the iteration process converges to the solution of the problem, we must show that $||\mathbf{x}^{[i]} - \mathbf{x}^{[i-1]}|| \to 0$ as $i \to \infty$. Suppose that the optimal control of the problem (4), (5) is given by $\mathbf{u}^{[i]}(t) = -\mathbf{K}(\mathbf{x}^{[i]}(t), \mathbf{u}^{[i]}(t), \mathbf{d}^{[i-1]}(t))\mathbf{x}^{[i]}(t)$. By $\mathbf{K}^{[i-1]}(t) := \mathbf{K}(\mathbf{x}^{[i]}(t), \mathbf{u}^{[i]}(t), \mathbf{d}^{[i-1]}(t))$, from (5) it follows that $\dot{\mathbf{x}}^{[i]}(t) = (\mathbf{A}(\mathbf{x}^{[i-1]}(t)) - \mathbf{B}(\mathbf{x}^{[i-1]}(t), \mathbf{u}^{[i-1]}(t))\mathbf{K}^{[i-1]}(t))\mathbf{x}^{[i]}(t) + \mathbf{d}^{[i-1]}(t)$. By $\mathbf{G}(\mathbf{x}(t), \mathbf{u}(t)) := -\mathbf{B}(\mathbf{x}(t), \mathbf{u}(t))\mathbf{K}(t)$, we assume here that:

1. $\mathbb{T}^{[i-1]}(t, 0)$ is the transition matrices generated by $\mathbf{A}(\mathbf{x}^{[i-1]}(t))$.

2. $\mathbf{A}(\mathbf{x}(t))$ is δ-Lipschitz continuous, $||\mathbf{A}(\mathbf{x}(t)) - \mathbf{A}(\bar{\mathbf{x}}(t))|| \leq \delta||\mathbf{x}(t) - \bar{\mathbf{x}}(t)||$, where $\mathbf{x}(t), \bar{\mathbf{x}}(t) \in \mathbb{R}^q$.

3. $||\mathbf{G}(\mathbf{x}(t), \mathbf{u}(t))|| \leq \epsilon_1$ and $||\mathbf{G}(\mathbf{x}(t), \mathbf{u}(t)) - \mathbf{G}(\bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t))|| \leq \epsilon_2||\mathbf{x}(t) - \bar{\mathbf{x}}(t)|| + \epsilon_3||\mathbf{u}(t) - \bar{\mathbf{u}}(t)||$, where $\mathbf{u}(t), \bar{\mathbf{u}}(t) \in \mathbb{R}^r$ and ϵ_1, ϵ_2, $\epsilon_3 \in \mathbb{R}_{>0}$ are constants.

4. $||\mathbf{d}(\mathbf{x}(t), \mathbf{u}(t))|| \leq \epsilon_4$ and $||\mathbf{d}(\mathbf{x}(t), \mathbf{u}(t)) - \mathbf{d}(\bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t))|| \leq \epsilon_5||\mathbf{x}(t) - \bar{\mathbf{x}}(t)|| + \epsilon_6||\mathbf{u}(t) - \bar{\mathbf{u}}(t)||$, where ϵ_4, ϵ_5, $\epsilon_6 \in \mathbb{R}_{>0}$ are constants.

5. $\mu(\mathbf{A}(\mathbf{x}(t))) \leq \mu_0$, where $\mu(\mathbf{A})$ denotes the measure of the matrix $\mathbf{A}$ and is defined as its logarithmic norm.

In each iteration, we have

$$\dot{\mathbf{x}}^{[i]}(t) = \mathbf{A}(\mathbf{x}^{[i-1]}(t))\mathbf{x}^{[i]}(t) + \mathbf{G}(\mathbf{x}^{[i-1]}(t), \mathbf{u}^{[i-1]}(t))\mathbf{x}^{[i]}(t) + \mathbf{d}(\mathbf{x}^{[i-1]}(t), \mathbf{u}^{[i-1]}(t)).$$

So

$$\mathbf{x}^{[i]}(t) = \mathbb{T}^{[i-1]}(t, 0)\mathbf{x}_0 + \int_0^t \mathbb{T}^{[i-1]}(t, \varrho)\{\mathbf{G}(\mathbf{x}^{[i-1]}(\varrho), \mathbf{u}^{[i-1]}(\varrho))\mathbf{x}^{[i]}(\varrho) + \mathbf{d}(\mathbf{x}^{[i-1]}(\varrho), \mathbf{u}^{[i-1]}(\varrho))\}d\varrho. \tag{6}$$

Using this equation, for $i > 2$ we have

$$\begin{aligned}||\mathbf{x}^{[i]}(t) - \mathbf{x}^{[i-1]}(t)|| = &\Big|\Big| \mathbb{T}^{[i-1]}(t,0)\mathbf{x}_0 - \mathbb{T}^{[i-2]}(t,0)\mathbf{x}_0 \\ &+ \int_0^t \mathbb{T}^{[i-1]}(t,\varrho)\mathbf{G}(\mathbf{x}^{[i-1]}(\varrho),\mathbf{u}^{[i-1]}(\varrho))\{\mathbf{x}^{[i]}(\varrho) - \mathbf{x}^{[i-1]}(\varrho)\}d\varrho \\ &+ \int_0^t \mathbb{T}^{[i-1]}(t,\varrho)\{\mathbf{G}(\mathbf{x}^{[i-1]}(\varrho),\mathbf{u}^{[i-1]}(\varrho)) - \mathbf{G}(\mathbf{x}^{[i-2]}(\varrho),\mathbf{u}^{[i-2]}(\varrho))\}\mathbf{x}^{[i-1]}(\varrho)d\varrho \\ &+ \int_0^t \{\mathbb{T}^{[i-1]}(t,\varrho) - \mathbb{T}^{[i-2]}(t,\varrho)\}\mathbf{G}(\mathbf{x}^{[i-2]}(\varrho),\mathbf{u}^{[i-2]}(\varrho))\mathbf{x}^{[i-1]}(\varrho)d\varrho \\ &+ \int_0^t \mathbb{T}^{[i-1]}(t,\varrho)\{\mathbf{d}(\mathbf{x}^{[i-1]}(\varrho),\mathbf{u}^{[i-1]}(\varrho)) - \mathbf{d}(\mathbf{x}^{[i-2]}(\varrho),\mathbf{u}^{[i-2]}(\varrho))\}d\varrho \\ &+ \int_0^t \{\mathbb{T}^{[i-1]}(t,\varrho) - \mathbb{T}^{[i-2]}(t,\varrho)\}\mathbf{d}(\mathbf{x}^{[i-2]}(\varrho),\mathbf{u}^{[i-2]}(\varrho))d\varrho\Big|\Big|. \end{aligned} \tag{7}$$

By Brauer inequality [21],

$$||\mathbb{T}^{[i-1]}(t,0)|| \leq \mathrm{e}^{\int_0^t \mu(\mathbf{A}(\mathbf{x}^{[i-1]}(\varrho)))d\varrho}. \tag{8}$$

For $0 \leq \vartheta \leq 1$, we can write

$$||\mathbb{T}^{[i-1]}(t,\vartheta) - \mathbb{T}^{[i-2]}(t,\vartheta)|| \leq \delta \mathrm{e}^{\mu_0(t-\vartheta)}(t-\vartheta) \sup_{\varrho\in[0,t]} ||\mathbf{x}^{[i-1]}(\varrho) - \mathbf{x}^{[i-2]}(\varrho)||. \tag{9}$$

From (8), we see $||\mathbb{T}^{[i-1]}(t,\varrho)|| \leq \mathrm{e}^{\mu_0(t-\varrho)}$. Using the submultiplicativity property of the matrix norm, it follows from (6) that

$$||\mathbf{x}^{[i]}(t)|| \leq \mathrm{e}^{\mu_0 t}||\mathbf{x}_0|| + \int_0^t \mathrm{e}^{\mu_0(t-\varrho)}\left\{\epsilon_1||\mathbf{x}^{[i]}(\varrho)|| + \epsilon_4\right\}d\varrho.$$

Taking $\epsilon_4 = \epsilon_4'||\mathbf{x}^{[i]}(t)||$, we see that $\mathrm{e}^{-\mu_0 t}||\mathbf{x}^{[i]}(t)|| \leq ||\mathbf{x}_0|| + \int_0^t \mathrm{e}^{-\mu_0(\varrho)}\left\{\epsilon_1 + \epsilon_4'\right\}||\mathbf{x}^{[i]}(\varrho)||d\varrho$. Applying Gronwall–Bellman inequality [22] yields

$$||\mathbf{x}^{[i]}(t)|| \leq \mathrm{e}^{(\epsilon_1+\epsilon_4'+\mu_0)t}||\mathbf{x}_0||. \tag{10}$$

By

$$\beta^{[i]}(t) \coloneqq \sup_{\varrho\in[0,t]} ||\mathbf{x}^{[i]}(\varrho) - \mathbf{x}^{[i-1]}(\varrho)||, \gamma^{[i]}(t) \coloneqq \sup_{\varrho\in[0,t]} ||\mathbf{u}^{[i]}(\varrho) - \mathbf{u}^{[i-1]}(\varrho)|| \tag{11}$$

and using the bounds given in Assumptions 2–5 and (8)–(11) in (7), we write

$$\begin{aligned}\beta^{[i]}(t) \leq &\delta \mathrm{e}^{\mu_0 t}t\beta^{[i-1]}(t)||\mathbf{x}_0|| + \epsilon_1\beta^{[i]}(t)\left|\left|\int_0^t \mathrm{e}^{\mu_0(t-\varrho)}d\varrho\right|\right| \\ &+ \{\epsilon_2\beta^{[i-1]}(t) + \epsilon_3\gamma^{[i-1]}(t)\}||\mathbf{x}_0||\left|\left|\int_0^t \mathrm{e}^{(\epsilon_1+\epsilon_4'+\mu_0)\varrho}\mathrm{e}^{\mu_0(t-\varrho)}d\varrho\right|\right| \\ &+ \epsilon_1\delta\beta^{[i-1]}(t)||\mathbf{x}_0||\left|\left|\int_0^t \mathrm{e}^{(\epsilon_1+\epsilon_4'+\mu_0)\varrho}\mathrm{e}^{\mu_0(t-\varrho)}(t-\varrho)d\varrho\right|\right| \\ &+ \{\epsilon_5\beta^{[i-1]}(t) + \epsilon_6\gamma^{[i-1]}(t)\}\left|\left|\int_0^t \mathrm{e}^{\mu_0(t-\varrho)}d\varrho\right|\right| \\ &+ \epsilon_4\delta\beta^{[i-1]}(t)\left|\left|\int_0^t \mathrm{e}^{\mu_0(t-\varrho)}(t-\varrho)d\varrho\right|\right|. \end{aligned}$$

By considering similar assumptions and doing similar calculations for $||\mathbf{u}^{[i]}(t) - \mathbf{u}^{[i-1]}(t)||$, we see that

$$\beta^{[i]}(t) \leq l_1\beta^{[i-1]}(t) + l_2\gamma^{[i-1]}(t), \gamma^{[i]}(t) \leq l_3\beta^{[i-1]}(t) + l_4\gamma^{[i-1]}(t),$$

where l_1, l_2, l_3, and l_4 are constants. Under some assumptions leading to $\beta^{[i+1]} \leq \beta^{[i]}$ and $\gamma^{[i+1]} \leq \gamma^{[i]}$, we see the sequence converges to a solution. □

Lemma 2
A nonlinear time-delay system characterized by

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{x}(t - h(t)), \mathbf{x}(t - h_x), \mathbf{u}(t - h_u), t),$$

can be transformed to a nonlinear delay-free system in the form (the subscript 'mod' is a shorthand to writing modified)

$$\dot{\mathbf{x}}_{mod}(t) = \mathbf{A}(\mathbf{x}_{mod}(t))\mathbf{x}_{mod}(t) + \mathbf{B}(\mathbf{x}_{mod}(t), \mathbf{u}(t))\mathbf{u}(t) + \mathbf{d}(t).$$

Proof
We can use some techniques like the Taylor series approximation in [23] and Páde approximation in [24, 25] for systems characterized by $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{x}(t - h_x), t)$, and together with procedures in [26, 27, 28, 29] for systems in the form $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{x}(t - h(t)), \mathbf{x}(t - h_x), \mathbf{u}(t - h_u), t)$, by which nonlinear systems having delayed state and/or control functions are transformed into delay-free ones. □

Remark 1
It should be obvious that we just mention these methods like Páde approximation for using the proof of the iterative method for delay-free optimal control problems and we will not use any of them in our procedure.

2.2.1. General Framework In the discussion that follows, we present a general framework for optimal control of fractional nonlinear systems with delays which can be utilized by wavelet functions.

Theorem 1
The nonlinear delay optimal control problem described by (1)–(3), which nonlinearity satisfies a local Lipschitz condition, can be replaced by the following sequence of linear delay problems which the sequence converges to a solution
for $i \geq 1$, minimize

$$J^{[i]} = \tfrac{1}{2}{\mathbf{x}^{[i]}}^{\top}(1)\mathbf{T}\mathbf{x}^{[i]}(1) + \tfrac{1}{2}\int_0^1 \left\{{\mathbf{x}^{[i]}}^{\top}(t)\mathbf{Q}(t)\mathbf{x}^{[i]}(t) + {\mathbf{u}^{[i]}}^{\top}(t)\mathbf{R}(t)\mathbf{u}^{[i]}(t)\right\} dt \tag{12}$$

subject to

$$\begin{aligned}\mathcal{D}^{\alpha}\mathbf{x}^{[i]}(t) = \mathbf{A}^{[i-1]}(t)\mathbf{x}^{[i]}(t) + \mathbf{B}^{[i-1]}(t)\mathbf{u}^{[i]}(t) + \mathbf{C}^{[i-1]}(t)\mathbf{x}^{[i]}(t - h(t)) + \mathbf{E}^{[i-1]}(t)\mathbf{x}^{[i]}(t - h_x)\\ + \mathbf{F}^{[i-1]}(t)\mathbf{u}^{[i]}(t - h_u) + \mathbf{d}^{[i-1]}(t),\end{aligned} \tag{13}$$

$$\mathbf{x}^{[i]}(0) = \mathbf{x}_0, \begin{cases} \mathbf{x}^{[i]}(t) = \boldsymbol{\theta}(t), & t < 0 \\ \mathbf{u}^{[i]}(t) = \boldsymbol{\zeta}(t), & -h_u \leq t \leq 0, \end{cases} \tag{14}$$

where $\alpha = 1$, $\mathbf{A}^{[i-1]}, \mathbf{C}^{[i-1]}, \mathbf{E}^{[i-1]} \in \mathbb{R}^{q\times q}$, $\mathbf{B}^{[i-1]}, \mathbf{F}^{[i-1]} \in \mathbb{R}^{q\times r}$, and $\mathbf{d}^{[i-1]} \in \mathbb{R}^q$.

Proof
A proof can be obtained upon invoking Lemma 2 in the proof of Lemma 1. □

Corollary 1
For $\alpha \neq 1$, under assumptions of Theorem 1, the fractional nonlinear delay optimal control problem described by (1)–(3), can be replaced by the sequence of linear fractional delay problems given in (12)–(14) which this sequence converges to a solution. Also we can use this procedure by considering $\dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0$ in the case $1 < \alpha \leq 2$.

Proof
This can be proved in a straightforward manner using Theorem 1. □

Remark 2
We use $\mathbf{d}(t)$ as a qth vector to represent disturbances, or terms like $\cos(x_\varsigma(t))$ and $\mathrm{e}^{x_\varsigma(t)}$, $\varsigma = 1, 2, \ldots, q$, which cannot be decomposed, or a combination of both.

Remark 3
For $i = 1$, we take $\mathbf{O}^{[i-1]}(t) = \mathbf{O}(t = 0)$, where $\mathbf{O}$ indicates $\mathbf{A}^{[i-1]}(t), \mathbf{B}^{[i-1]}(t)$, and so on. For $i > 1$, we construct it from $\mathbf{O}^{[i-1]}(t) = \mathbf{O}(\mathbf{x}^{[i-1]}(t), \mathbf{x}^{[i-1]}(t - h(t)), \mathbf{x}^{[i-1]}(t - h_x), \mathbf{u}^{[i-1]}(t), \mathbf{u}^{[i-1]}(t - h_u))$. If $\mathbf{O}^{[i-1]}(t)$ is not a function of $\mathbf{x}^{[i-1]}$ and $\mathbf{u}^{[i-1]}$, it is replaced by $\mathbf{O}(t)$.

Remark 4
For a delayed term like $\mathbf{x}^{[i-1]}(t - h_x)$ as a state or input matrix, or whose entries, or as a disturbance or its entries, we expand it from $\mathbf{x}^{[i-1]}(t - h_x) = \begin{cases} \boldsymbol{\theta}(t - h_x), & 0 \le t \le h_x \\ \mathbf{x}^{[i-1]}(t - h_x), & h_x \le t \le 1 \end{cases}$. Also for a similar term having input delays, we do the same. For a term with a piecewise delay, we expand it according to values of delays in each subinterval.

Remark 5
To use this framework, we solve the obtained sequence of linear quadratic optimal control problems iteratively until for $i \ge 2$, $|J^{[i]} - J^{[i-1]}|$ or $|x_\varsigma^{[i]}(t) - x_\varsigma^{[i-1]}(t)|$ becomes negligible.

2.3. Chebyshev wavelet method

Here, we use Chebyshev wavelets with scaling [30] for their advantages over classical Chebyshev wavelets [31].

Definition 3
Chebyshev wavelets $\psi_{nm}^{\xi}(t)$ are defined on the interval $[0, 1]$ by

$$\psi_{nm}^{\xi}(t) = \begin{cases} \sqrt{2\xi^{k-1}}\mathrm{c}_m T_m(2\xi^{k-1}t - 2n + 1), & t \in \left[\frac{n-1}{\xi^{k-1}}, \frac{n}{\xi^{k-1}}\right] \\ 0, & \text{otherwise,} \end{cases} \tag{15}$$

where $\xi \in \mathbb{N}_{\ge 2}$, $k \in \mathbb{N}_{\ge 2}$, and are finite values, $n = 1, 2, \ldots, \xi^{k-1}$, $m = 0, 1, \ldots, M - 1$ is the degree of T_m, $\mathrm{c}_0 = 1/\sqrt{\pi}, \mathrm{c}_{m\neq 0} = \sqrt{2}/\sqrt{\pi}$, and t is the independent variable. A function $f(t)$ may be approximated in terms of Chebyshev wavelets over the interval $0 \le t \le 1$ by the Mth term in ξ^{k-1} subintervals as

$$f(t) \cong \mathbf{f}_{cw}\boldsymbol{\psi}(t), \tag{16}$$

where $\mathbf{f}_{cw}\boldsymbol{\psi}(t) = \sum_{n=1}^{\xi^{k-1}} \sum_{m=0}^{M-1} f_{nm}^{\xi}\psi_{nm}^{\xi}(t)$, $\mathbf{f}_{cw}$ is a row vector of constant coefficients $\{f_{nm}^{\xi}\}$, $\boldsymbol{\psi}(t)$ denotes the Chebyshev wavelets column vector consisting of Chebyshev wavelets $\{\psi_{nm}^{\xi}(t)\}$ and the subscript 'cw' refers to the Chebyshev wavelets expansion. The first summation denotes this expansion is piecewise-defined.

Theorem 2
Let $f(t)$ be a twice differentiable function on $[0, 1]$, then $f(t)$ can be expanded as $f(t) = \sum_{n=1}^{\xi^{k-1}} \sum_{m=0}^{\infty} f_{nm}^{\xi}\psi_{nm}^{\xi}(t)$, where this series converges uniformly to $f(t)$.

Proof
Since f is twice differentiable, f' is differentiable and bounded on $[0, 1]$ and f is bounded. By assuming that $|f(t)|, |f'(t)|, |f''(t)| \le l$, the idea of proof is the same that given in [30]. □

Lemma 3
For the Chebyshev wavelets vector $\boldsymbol{\psi}(t)$ we have: $\int_0^1 \boldsymbol{\psi}(t)\boldsymbol{\psi}^\top(t)\,dt = \boldsymbol{\Gamma}_{cw}$, $\mathbf{f}_{cw}\boldsymbol{\psi}(t)\boldsymbol{\psi}^\top(t) \cong \boldsymbol{\psi}^\top(t)\tilde{\mathbf{f}}_{cw}$, $\boldsymbol{\psi}(t - h_\iota) = \begin{cases} \mathbf{0}, & 0 \le t < h_\iota \\ \mathbf{D}_{\iota cw}\boldsymbol{\psi}(t), & h_\iota \le t \le 1 \end{cases}$, $\boldsymbol{\psi}(t - h(t)) = \begin{cases} \mathbf{0}, & 0 \le t < h(t) \\ \mathbf{D}_{cw}^{t}\boldsymbol{\psi}(t), & h(t) \le t \le 1 \end{cases}$ and $\mathcal{I}^{\alpha}\boldsymbol{\psi}(t) \cong \mathbf{P}_{cw}^{\alpha}\boldsymbol{\psi}(t)$, where h_ι is a time-delay, $h(t)$ is a piecewise delay and $\boldsymbol{\Gamma}_{cw}$, $\tilde{\mathbf{f}}_{cw}$, $\mathbf{D}_{\iota cw}$, $\mathbf{D}_{cw}^{t}$, and $\mathbf{P}_{cw}^{\alpha}$ are operational matrices of Chebyshev wavelets.

Proof
See [30] and [11]. □

Now, we use the concepts and computational procedures that have been developed for linear optimal control systems to nonlinear optimal control systems (in the Riemann–Liouville sense).

Theorem 3
A fractional nonlinear time-delay problem as minimizing the performance index (1) for the plant described by (2) and (3), by using (16) and the properties of Chebyshev wavelets can be solved by forming a sequence of quadratic programming (QP) problems as

$$\min_{\mathbf{z}^{[i]}} \ \tfrac{1}{2}{\mathbf{z}^{[i]}}^{\top}\boldsymbol{\Xi}\,\mathbf{z}^{[i]}$$
$$\text{subject to } \ \boldsymbol{\Lambda}^{[i-1]}(\alpha,\mathbf{f})\mathbf{z}^{[i]} = \mathbf{b}^{[i-1]}(\alpha,\mathbf{f}),$$

where $\boldsymbol{\Xi}$, $\boldsymbol{\Lambda}$, and $\mathbf{b}$ are known constant matrices and $\mathbf{z}$ as a combination of the state and control variables is the solution of the problem in terms of Chebyshev wavelets..

Proof
From Corollary 1, we model the problems given in (12)–(14) instead of the original problem. Using (16), we can write $x_{\varsigma}^{[i]}(t) = \boldsymbol{\psi}^{\top}(t){\mathbf{x}_{\varsigma cw}^{[i]}}^{\top}$, where $\varsigma = 1, 2, \ldots, q$ and ${\mathbf{x}_{\varsigma cw}^{[i]}}^{\top}$ is a vector of unknown parameters. For $\mathbf{x}^{[i]}(t) = [\{x_{\varsigma}^{[i]}(t)\}]^{\top}$, we find

$$\begin{aligned}
\mathbf{x}^{[i]}(t) &= [x_1^{[i]}(t), x_2^{[i]}(t), \ldots, x_q^{[i]}(t)]^{\top}\\
&= \left[\boldsymbol{\psi}^{\top}(t){\mathbf{x}_{1cw}^{[i]}}^{\top}, \boldsymbol{\psi}^{\top}(t){\mathbf{x}_{2cw}^{[i]}}^{\top}, \cdots, \boldsymbol{\psi}^{\top}(t){\mathbf{x}_{qcw}^{[i]}}^{\top}\right]^{\top}\\
&= [\psi_{10}^{\xi}(t){x_{10}^{\xi 1}}^{[i]} + \ldots + \psi_{\xi^{k-1}M-1}^{\xi}(t){x_{\xi^{k-1}M-1}^{\xi 1}}^{[i]}, \ldots, \psi_{10}^{\xi}(t){x_{10}^{\xi q}}^{[i]} + \ldots + \psi_{\xi^{k-1}M-1}^{\xi}(t){x_{\xi^{k-1}M-1}^{\xi q}}^{[i]}]^{\top}\\
&= \begin{bmatrix}\boldsymbol{\Phi}_1(t)\\ \boldsymbol{\Phi}_2(t)\\ \vdots\\ \boldsymbol{\Phi}_q(t)\end{bmatrix}\left[{x_{10}^{\xi 1}}^{[i]} \quad {x_{10}^{\xi 2}}^{[i]} \quad \ldots \quad {x_{10}^{\xi q}}^{[i]} \quad \ldots \quad {x_{\xi^{k-1}M-1}^{\xi 1}}^{[i]} \quad \ldots \quad {x_{\xi^{k-1}M-1}^{\xi q}}^{[i]}\right]^{\top},
\end{aligned}$$

where

$$\boldsymbol{\Phi}_{\varsigma}(t) = \boldsymbol{\psi}^{\top}(t) \otimes \boldsymbol{\varpi}_{\varsigma}, \ \boldsymbol{\varpi}_{\varsigma} = [\overbrace{0, 0, \ldots, 0}^{\varsigma-1}, 1, \overbrace{0, 0, \ldots, 0}^{q-\varsigma}].$$

Thus from the property of Kronecker product, we can write

$$\mathbf{x}^{[i]}(t) = (\boldsymbol{\psi}^{\top}(t) \otimes \mathbf{I}_q)\mathbf{x}_{cw}^{[i]},$$

where

$$\mathbf{x}_{cw}^{[i]} = \left[{x_{10}^{\xi 1}}^{[i]} \quad {x_{10}^{\xi 2}}^{[i]} \quad \ldots \quad {x_{10}^{\xi q}}^{[i]} \quad \ldots \quad {x_{\xi^{k-1}M-1}^{\xi 1}}^{[i]} \quad \ldots \quad {x_{\xi^{k-1}M-1}^{\xi q}}^{[i]}\right]^{\top}.$$

Similarly, for $\mathbf{u}^{[i]}(t) = [u_1^{[i]}(t), u_2^{[i]}(t), \ldots, u_r^{[i]}(t)]^{\top}$, we have

$$\mathbf{u}^{[i]}(t) = (\boldsymbol{\psi}^{\top}(t) \otimes \mathbf{I}_r)\mathbf{u}_{cw}^{[i]},$$

where

$$\mathbf{u}_{cw}^{[i]} = \left[{u_{10}^{\xi 1}}^{[i]} \quad {u_{10}^{\xi 2}}^{[i]} \quad \ldots \quad {u_{10}^{\xi r}}^{[i]} \quad \ldots \quad {u_{\xi^{k-1}M-1}^{\xi 1}}^{[i]} \quad \ldots \quad {u_{\xi^{k-1}M-1}^{\xi r}}^{[i]}\right]^{\top}.$$

Using these last two results, for $\mathbf{O}_s^{[i]}(t) \in \mathbb{R}^{q\times q}$, $\mathbf{O}_v^{[i]}(t) \in \mathbb{R}^{r\times r}$ and $\mathbf{O}_z^{[i]}(t) \in \mathbb{R}^{q\times r}$ in (12) and (13), we can write

$$\mathbf{O}_s^{[i]}(t) = \mathbf{O}_{scw}^{[i]}(\boldsymbol{\psi}(t) \otimes \mathbf{I}_q), \mathbf{O}_v^{[i]}(t) = \mathbf{O}_{vcw}^{[i]}(\boldsymbol{\psi}(t) \otimes \mathbf{I}_r), \mathbf{O}_z^{[i]}(t) = \mathbf{O}_{zcw}^{[i]}(\boldsymbol{\psi}(t) \otimes \mathbf{I}_r).$$

We express the initial condition in (14) as

$$\mathbf{x}_0 = (\boldsymbol{\psi}^\top(t) \otimes \mathbf{I}_q)\mathbf{x}_{cw}^0,$$

where

$$\mathbf{x}_{cw}^0 = \frac{\sqrt{\pi}}{\sqrt{2\xi^{k-1}}}[\boldsymbol{\chi}_1^0, \boldsymbol{\chi}_2^0, \ldots, \boldsymbol{\chi}_{\xi^{k-1}}^0]^\top,\ \boldsymbol{\chi}_n^0 = [\mathbf{x}_0^\top, \overbrace{0, 0, \ldots, 0}^{q(M-1)}].$$

We model each system in Theorem 1 for each iteration by proceeding exactly as we did in [11], and using the relations given in Lemma 3; by expanding the matrices terms of Chebyshev wavelets as $\mathbf{A}^{[i-1]}(t) = \mathbf{A}_{cw}^{[i-1]}(\boldsymbol{\psi}(t) \otimes \mathbf{I}_q)$, $\mathbf{B}^{[i-1]}(t) = \mathbf{B}_{cw}^{[i-1]}(\boldsymbol{\psi}(t) \otimes \mathbf{I}_r)$, $\mathbf{C}^{[i-1]}(t) = \mathbf{C}_{cw}^{[i-1]}(\boldsymbol{\psi}(t) \otimes \mathbf{I}_q)$, $\mathbf{E}^{[i-1]}(t) = \mathbf{E}_{cw}^{[i-1]}(\boldsymbol{\psi}(t) \otimes \mathbf{I}_q)$, $\mathbf{F}^{[i-1]}(t) = \mathbf{F}_{cw}^{[i-1]}(\boldsymbol{\psi}(t) \otimes \mathbf{I}_r)$, $\mathbf{d}^{[i-1]}(t) = (\boldsymbol{\psi}^\top(t) \otimes \mathbf{I}_q)\mathbf{d}_{cw}^{[i-1]}$, $\boldsymbol{\theta}(t - h(t)) = (\boldsymbol{\psi}^\top(t) \otimes \mathbf{I}_q)\boldsymbol{\theta}_{cw}^t$, $\boldsymbol{\theta}(t - h_x) = (\boldsymbol{\psi}^\top(t) \otimes \mathbf{I}_q)\boldsymbol{\theta}_{cw}$, $\boldsymbol{\zeta}(t - h_u) = (\boldsymbol{\psi}^\top(t) \otimes \mathbf{I}_r)\boldsymbol{\zeta}_{cw}$, and by applying the α-integral ($0 < \alpha \leq 1$) to both sides of (13), we have

$$\begin{aligned}(\boldsymbol{\psi}^\top(t) \otimes \mathbf{I}_q)\big(\mathbf{x}_{cw}^{[i]} - \mathbf{x}_{cw}^0\big) =& (\boldsymbol{\psi}^\top(t) \otimes \mathbf{I}_q)({\mathbf{P}_{cw}^\alpha}^\top \otimes \mathbf{I}_q)\big\{\tilde{\mathbf{A}}_{cw}^{[i-1]}\mathbf{x}_{cw}^{[i]} + \tilde{\mathbf{B}}_{cw}^{[i-1]}\mathbf{u}_{cw}^{[i]} \\ &+ \tilde{\mathbf{C}}_{cw}^{[i-1]}\boldsymbol{\theta}_{cw}^t + \tilde{\mathbf{C}}_{cw}^{[i-1]}({\mathbf{D}_{cw}^t}^\top \otimes \mathbf{I}_q)\mathbf{x}_{cw}^{[i]} \\ &+ \tilde{\mathbf{E}}_{cw}^{[i-1]}\boldsymbol{\theta}_{cw} + \tilde{\mathbf{E}}_{cw}^{[i-1]}(\mathbf{D}_{xcw}^\top \otimes \mathbf{I}_q)\mathbf{x}_{cw}^{[i]} \\ &+ \tilde{\mathbf{F}}_{cw}^{[i-1]}\boldsymbol{\zeta}_{cw} + \tilde{\mathbf{F}}_{cw}^{[i-1]}(\mathbf{D}_{ucw}^\top \otimes \mathbf{I}_r)\mathbf{u}_{cw}^{[i]} + \mathbf{d}_{cw}^{[i-1]}\big\},\end{aligned}$$

where $\mathbf{P}_{cw}^\alpha$, $\mathbf{D}_{cw}^t$, and $\mathbf{D}_{cw}$ are operational matrices, $\tilde{\mathbf{O}}_{cw}$ is the product operational matrix of the matrix $\mathbf{O}_{cw}$, the subscript 'cw' denotes that $\mathbf{O}_{cw}$ is the expansion of $\mathbf{O}(t)$ in terms of Chebyshev wavelets.

By expanding the matrices of the performance index (12) and using the concepts given in Lemma 3, we find

$$J = \tfrac{1}{2}{\mathbf{x}_{cw}^{[i]}}^\top(\boldsymbol{\psi}(1)\boldsymbol{\psi}^\top(1) \otimes \mathbf{T})\mathbf{x}_{cw}^{[i]} + \tfrac{1}{2}{\mathbf{x}_{cw}^{[i]}}^\top(\boldsymbol{\Gamma}_{cw} \otimes \mathbf{I}_q)\tilde{\mathbf{Q}}_{cw}\mathbf{x}_{cw}^{[i]} + \tfrac{1}{2}{\mathbf{u}_{cw}^{[i]}}^\top(\boldsymbol{\Gamma}_{cw} \otimes \mathbf{I}_r)\tilde{\mathbf{R}}_{cw}\mathbf{u}_{cw}^{[i]}.$$

Setting $I_n = [(n-1)/\xi^{k-1}, n/\xi^{k-1}]$, we see from (15) the fact that there exist some points as t_I, where $\{t_I\} = I_n \cap I_{n+1}$. At these points we must have $\mathbf{x}(t_I^-) = \mathbf{x}(t_I^+)$. Applying this compatibility constraint yields

$$(\boldsymbol{\psi}^\top(t_I^-) \otimes \mathbf{I}_q)\mathbf{x}_{cw}^{[i]} - (\boldsymbol{\psi}^\top(t_I^+) \otimes \mathbf{I}_q)\mathbf{x}_{cw}^{[i]} = \mathbf{0}.$$

This gives $q(\xi^{k-1} - 1)$ equations, written in matrix form,

$$\left[\begin{bmatrix} \boldsymbol{\psi}^\top(t_1^-) - \boldsymbol{\psi}^\top(t_1^+) \\ \boldsymbol{\psi}^\top(t_2^-) - \boldsymbol{\psi}^\top(t_2^+) \\ \vdots \\ \boldsymbol{\psi}^\top(t_{\xi^{k-1}-1}^-) - \boldsymbol{\psi}^\top(t_{\xi^{k-1}-1}^+) \end{bmatrix} \otimes \mathbf{I}_q\right]\mathbf{x}_{cw}^{[i]} = \mathbf{0}.$$

By denoting the first matrix in the bracket by $\boldsymbol{\Psi}_{cc}$,

$$(\boldsymbol{\Psi}_{cc} \otimes \mathbf{I}_q)\mathbf{x}_{cw}^{[i]} = \mathbf{0},$$

where $\boldsymbol{\Psi}_{cc}$ in a constant matrix obtained from the properties of Chebyshev wavelets.

Hence we can solve the following sequence of QP problems iteratively for $0 < \alpha \leq 1$ instead of the problem described by (12)–(14),

$$\text{for } i \geq 1, \min \quad \tfrac{1}{2}\begin{bmatrix} \mathbf{x}_{cw}^{[i]} \\ \hline \mathbf{u}_{cw}^{[i]} \end{bmatrix}^\top \left[\begin{array}{c|c} \boldsymbol{\Xi}_1 & \boldsymbol{\Xi}_2 \\ \hline \boldsymbol{\Xi}_3 & \boldsymbol{\Xi}_4 \end{array}\right] \begin{bmatrix} \mathbf{x}_{cw}^{[i]} \\ \hline \mathbf{u}_{cw}^{[i]} \end{bmatrix} \tag{17}$$

$$\text{subject to} \quad \left[\begin{array}{c|c} \boldsymbol{\Lambda}_1^{[i-1]} & \boldsymbol{\Lambda}_2^{[i-1]} \\ \hline \boldsymbol{\Lambda}_3 & \boldsymbol{\Lambda}_4 \end{array}\right] \begin{bmatrix} \mathbf{x}_{cw}^{[i]} \\ \hline \mathbf{u}_{cw}^{[i]} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1^{[i-1]} \\ \hline \mathbf{b}_2 \end{bmatrix}, \tag{18}$$

until the condition in Remark 5 is reached, where

$$\boldsymbol{\Xi}_1 = (\boldsymbol{\psi}(1)\boldsymbol{\psi}^\top(1) \otimes \mathbf{T}) + (\boldsymbol{\Gamma}_{cw} \otimes \mathbf{I}_q)\tilde{\mathbf{Q}}_{cw}, \tag{19}$$

$$\boldsymbol{\Xi}_2 = \mathbf{0}_{q\xi^{k-1}M \times r\xi^{k-1}M}, \tag{20}$$

$$\boldsymbol{\Xi}_3 = \mathbf{0}_{r\xi^{k-1}M \times q\xi^{k-1}M}, \tag{21}$$

$$\boldsymbol{\Xi}_4 = (\boldsymbol{\Gamma}_{cw} \otimes \mathbf{I}_r)\tilde{\mathbf{R}}_{cw}, \tag{22}$$

$$\boldsymbol{\Lambda}_1^{[i-1]} = ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\tilde{\mathbf{A}}_{cw}^{[i-1]} + ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\tilde{\mathbf{C}}_{cw}^{[i-1]}({\mathbf{D}_{cw}^{t}}^\top \otimes \mathbf{I}_q) + ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\tilde{\mathbf{E}}_{cw}^{[i-1]}(\mathbf{D}_{xcw}^\top \otimes \mathbf{I}_q) - \mathbf{I}_{q\xi^{k-1}M}, \tag{23}$$

$$\boldsymbol{\Lambda}_2^{[i-1]} = ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\tilde{\mathbf{B}}_{cw}^{[i-1]} + ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\tilde{\mathbf{F}}_{cw}^{[i-1]}(\mathbf{D}_{ucw}^\top \otimes \mathbf{I}_r), \tag{24}$$

$$\boldsymbol{\Lambda}_3 = \boldsymbol{\Psi}_C \otimes \mathbf{I}_q, \tag{25}$$

$$\boldsymbol{\Lambda}_4 = \mathbf{0}_{q\xi^{k-1} \times r\xi^{k-1}M}, \tag{26}$$

$$\mathbf{b}_1^{[i-1]} = -\mathbf{x}_{cw}^0 - ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\mathbf{d}_{cw}^{[i-1]} - ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\tilde{\mathbf{C}}_{cw}^{[i-1]}\boldsymbol{\theta}_{cw}^{t} - ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\tilde{\mathbf{E}}_{cw}^{[i-1]}\boldsymbol{\theta}_{cw} - ({\mathbf{P}_{cw}^{\alpha}}^\top \otimes \mathbf{I}_q)\tilde{\mathbf{F}}_{cw}^{[i-1]}\boldsymbol{\zeta}_{cw}, \tag{27}$$

$$\mathbf{b}_2 = \begin{bmatrix} \mathbf{0}_{q(\xi^{k-1}-1)\times 1} \\ \mathbf{x}_0 \end{bmatrix}. \tag{28}$$

Also $\boldsymbol{\Psi}_C \coloneqq \begin{bmatrix} \boldsymbol{\Psi}_{cc} \\ \boldsymbol{\psi}^\top(0) \end{bmatrix}$ and to avoid needless repetition we do not present the details and the process of constructing the operational matrices and other required matrices was described in [30] and [11].

For $1 < \alpha \leq 2$, we must modify $\mathbf{b}_1^{[i-1]}$ as we did for linear systems in the latter reference. □

Remark 6
For a matrix $\mathbf{O}^{[i-1]}(t)$ which is not a function of $\mathbf{x}^{[i-1]}$ and $\mathbf{u}^{[i-1]}$, $\tilde{\mathbf{O}}_{cw}^{[i-1]}$ is replaced by $\tilde{\mathbf{O}}_{cw}$.

Remark 7
In order to use this sequence of QP problems for another wavelets by applying α-integral, we must replace $\boldsymbol{\psi}(1)$ and the matrices denoted by the subscript cw with those which are defined for the new wavelet.

3. Some nonlinear systems

In this section, to see the generality of the method, different kinds of nonlinear time-delay systems are considered in which we construct (17)–(18) from (19)-(28) and then we use the powerful QP solver quadprog in MATLAB.

3.1. Example 1

Consider the minimization of the performance index [32]

$$J = \tfrac{1}{2}x^2(2) + \tfrac{1}{2}\int_0^2 \left\{x^2(t) + u^2(t)\right\} dt$$

for the nonlinear delay system described by the fractional delay differential equation

$$\mathcal{D}^\alpha x(t) = x(t)\sin x(t) + x(t-1) + u(t), \quad 0 \leq t \leq 2$$

in which the initial function is

$$\theta(t) = \begin{cases} 10(t+1), & -1 \leq t \leq -0.5 \\ -10t, & -0.5 \leq t \leq 0 \end{cases}.$$

Using the given procedure, considering Remark 3 and setting $t/2 \longrightarrow t$, this problem is replaced by a sequence of linear fractional delay optimization problems as follows:
for $i \geq 1$, minimize

$$J = \tfrac{1}{2}{x^{[i]}}^2(2) + \tfrac{1}{2}\int_0^1 \left\{2{x^{[i]}}^2(t) + 2{u^{[i]}}^2(t)\right\} dt \tag{29}$$

subject to

$$\mathcal{D}^\alpha x^{[i]}(t) = 2^\alpha[\sin x^{[i-1]}(t)]x^{[i]}(t) + 2^\alpha x^{[i]}(t-\tfrac{1}{2}) + 2^\alpha u^{[i]}(t), \quad 0 \leq t \leq 1 \tag{30}$$

$$x^{[i]}(0) = 0, \ \theta(t) = \begin{cases} 20t + 10, & -0.5 \leq t \leq -0.25 \\ -20t, & -0.25 \leq t \leq 0 \end{cases}. \tag{31}$$

Our task is to find the final solutions; we do this by substituting $x^{[0]}(t) = x(0)$ in (30) and obtaining $x^{[1]}(t)$ from solving the QP model of (29)–(31)-this is the first step-then by substituting $x^{[1]}(t) = \boldsymbol{\psi}^\top(t)\mathbf{x}_{cw}^{[1]}$, where $\mathbf{x}_{cw}^{[1]}$ was obtained from the first step, and obtaining $x^{[2]}(t)$-this is the second step-and so on.

First we select $\xi = 4$, $k = 2$, and $M = 8$. Using the procedure, after 15 iterations, we reach the solutions for $\alpha = 1$, where $|J^{[16]} - J^{[15]}| \leq 6.67E - 08$. Also, we solve the problem for different α and the results for the optimal cost J^* are given in Table 1 and a comparison is made between the optimal cost obtained by this work and those reported in [32, 33, 34]. The optimal controls and states for $\alpha = 1, 0.975, 0.95, 0.925$ are shown in Figure 1.

Table 1. J^* for some α, $0.8 \leq \alpha \leq 1$ in Example 1.

α	[32]	[33]	[34]	This work
1	2.4371	2.5290	2.5077202960	2.54807
0.99				2.55616
0.95				2.59081
0.91				2.62918
0.9				2.63937
0.8				2.75520

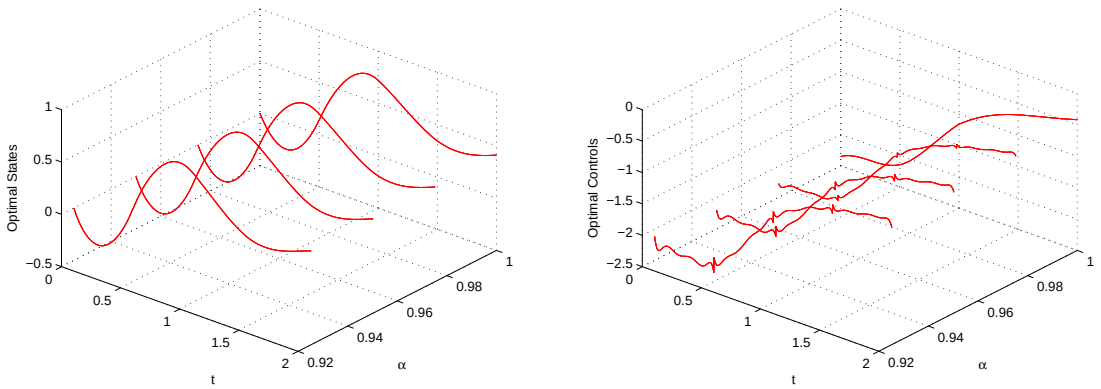


Figure 1. Optimal states and controls for Example 1.

3.2. Example 2

Consider a fractional nonlinear system a with time-varying delay described by

$$\mathcal{D}^\alpha x(t) = \frac{t-1}{t}x(t - h(t))x(t) + u(t), \quad 1 \leq t \leq 3$$

$$x(t) = 1, \quad t \leq 1.$$

We want to find the optimal state and control which minimize

$$J = \int_1^3 \left\{x^2(t) + u^2(t)\right\} dt,$$

where

Case 1: $h(t) = \ln(t) + 1, \quad 1 \leq t \leq 3$ (see [35, 36]).

Case 2: $h(t) = \ln(t) + 0.25, \quad 1 \leq t \leq 3.$

The reason for considering another case as Case 2 lies in the colored regions in Figure 2; we find that in Case 1 for $\forall t \in [1, 3]$, $t - h(t) < 1$ and the region is indicated in gray color, which means that in Form A given below, the initial function is expanded on full interval and $\mathbf{D}^t_{cw} = \mathbf{0}$ while in Case 2 for some $t \in [1, 3[$ which $t - h(t) < 1$, we should have a piecewise-defined initial function. The method for approximating and expanding this kind of the initial functions was introduced in [30] and [16]. To rescale the problem, first we set $t - 1 \longrightarrow t$, then $t/2 \longrightarrow t$. We apply our framework for linearization of the nonlinear system in two alternative forms:

A: $\mathcal{D}^\alpha x^{[i]}(t) = 2^\alpha[(1 - \frac{1}{2t+1})x^{[i-1]}(t)]x^{[i]}(t - h(t)) + 2^\alpha u^{[i]}(t), \quad 0 \leq t \leq 1.$

B: $\mathcal{D}^\alpha x^{[i]}(t) = 2^\alpha[(1 - \frac{1}{2t+1})x^{[i-1]}(t - h(t))]x^{[i]}(t) + 2^\alpha u^{[i]}(t), \quad 0 \leq t \leq 1.$

One should be a little bit careful in rescaling $(t - h(t))$. This argument in $[0, 2]$ should have same behavior as that in $[1, 3]$, not same values. We solve the problem with the use of the given forms by considering Remark 4 and the results are reported in Table 2.

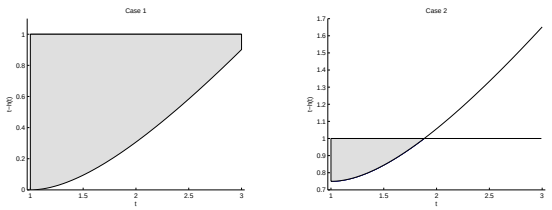


Figure 2. $t - h(t)$ for Cases 1 and 2, Example 2.

Table 2. Comparison of J^* for Example 2.

	Case 1				Case 2	
α	this work, A	this work, B	[35]	[36]	this work, A	this work, B
1	1.270217	1.243391	1.66793163716010	1.24339085679272	1.240470	1.222775
0.95	1.262642	1.232479			1.227841	1.208855
0.9	1.257003	1.223094			1.216226	1.195964

For Case 2, by using Remark 7 the solution curves (in the case $\alpha = 1$) for Legendre wavelets [16] are given in Figure 3, where we obtain $J^* = 1.222775$.

3.3. Example 3

In this example, we aim to minimize the criterion [37]

$$\int_0^{10} \{x_1^2(t) + x_2^2(t) + u^2(t)\} dt$$

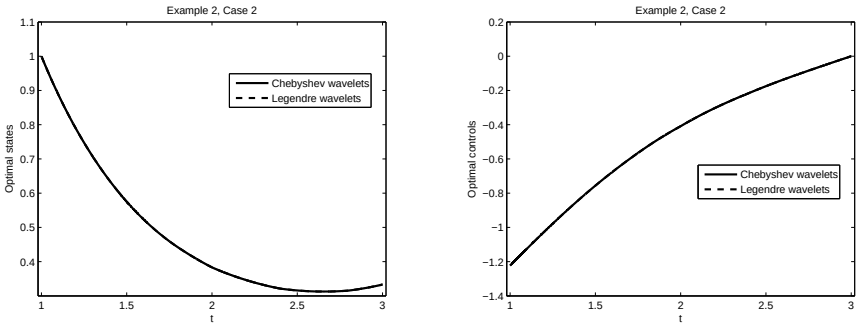


Figure 3. Solution curves for Case 2 in Example 2.

subject to

$$\mathcal{D}^{\alpha}x_1(t) = x_1(t)x_2(t) + 2x_1(t-2)x_2(t) - x_1(t-2) + u(t),$$
$$\mathcal{D}^{\alpha}x_2(t) = -x_1(t) + x_2(t-2)$$

and we have $\boldsymbol{\theta}(t) = [0.5, 0.5]^{\top}$.

Using the quasilinearization method, we solve the sequence of quadratic programming problems obtained for this problem. The values of J^* are given in Table 3. We see J^* decreases as α decreases. The optimal control and states for these values of α are shown in Figure 4. Also for showing the convergence of the method, the results for $\alpha = 0.9$ are given in Table 4. To overcome drawbacks of classical (conventional) Legendre wavelets, Ref. [37] was used the hybrid functions as a combinations of block pulse and Legendre functions. But we can use Legendre wavelets with scaling [16] by which those drawbacks are eliminated and they have advantages over these hybrid functions, for example $\boldsymbol{\Gamma}_{lw} = \mathbf{I}$.

Table 3. J^* for some α, Example 3.

α	[37]	This work
1	1.9955	2.022230
0.99		2.017244
0.98		2.012553
0.97		2.008170
0.96		2.004106
0.95		2.000377

3.4. Example 4

Consider a cascade chemical system with reactors 1 and 2, as shown in Figure 5 arises in the chemical industry, for details see [38]. The inputs of reactor 1 come from reactor 2 and the external disturbances; the inputs of reactor 2 are the delayed state of reactor 1, the control and the external disturbances. The plant is described by

$$\mathcal{D}^1x_1(t) = -k_1x_1(t) - \tfrac{1}{\theta_1}x_1(t) - \tfrac{1}{\theta_1}x_1(t-h_1) + \tfrac{1-R_2}{V_1}x_2(t) + \delta_1(t, x_1(t-h_1)),$$
$$\mathcal{D}^1x_2(t) = -k_2x_2(t) - \tfrac{1}{\theta_2}x_2^2(t) + \tfrac{R_1}{V_2}x_1(t-h_1) - \tfrac{1}{\theta_2}x_2(t) + \tfrac{R_2}{V_2}x_2(t-h_2) + \tfrac{F}{V_2}u(t) + \delta_2(t, x_2(t-h_1)),$$

where for $i = 1, 2$, R_i are the recycle flow rates, θ_i are the reactor residence times, k_i are the reaction constants, F is the feed rate, V_i are reactor volumes, and δ_i are nonlinear functions for describing the system uncertainties and disturbances. The performance index is

$$J = \int_0^{t_f} \{\mathbf{x}^{\top}(t)\mathbf{Q}\mathbf{x}(t) + Ru^2(t)\}\, dt.$$

Table 4. $|J^{[i]} - J^{[i-1]}|$ for $\alpha = 0.9$ in Example 3.

i	$J^{[i]}$	$\lvert J^{[i]} - J^{[i-1]}\rvert$
1	1.777053	
2	1.935883	0.158830
3	1.985908	0.050025
4	1.987188	0.001280
5	1.987283	0.000094
6	1.987306	0.000023
7	1.987316	0.000010
8	1.987319	0.000003
9	1.987321	0.000001
10	1.987321*	0.000000

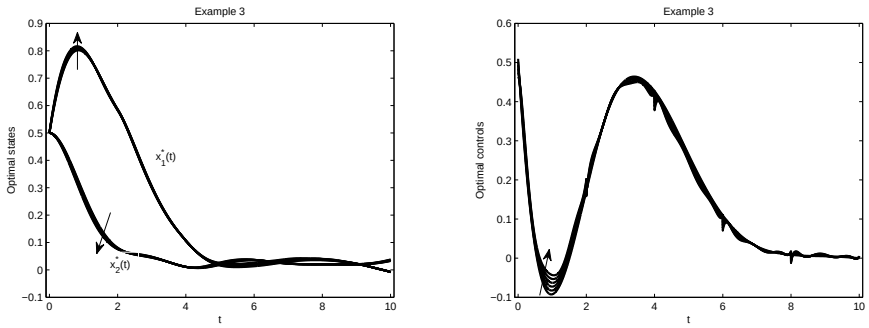


Figure 4. Solution curves of Example 3, α decreases in the direction of the arrow.

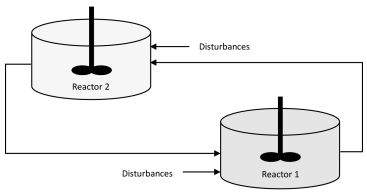


Figure 5. A schematic of cascade chemical reactor system, [38].

Also we have $R_i = 0.5$, $k_i = 0.5$, , $F = 0.5$, $V_i = 0.5$, $\delta_1(t, x_1(t-h_1)) = \theta_3 x_1(t-h_1)$ and

Case 1: $h = 0.5$, $t_f = 3$, $\delta_2(x_2(t-h)) = 0.5\theta_4 x_2^2(t-h)$, $\boldsymbol{\theta} = [1, -2]^\top$ and for $i = 1, \ldots, 4$, $\theta_i = 1$, $\mathbf{Q} = 100\mathbf{I}_2$, $R = 1$; see [37].

Case 2: $h = 0.25$, $t_f = 5$, , $\delta_2(x_2(t-h)) = 0.5\theta_4 x_2^2(t-h)\mathrm{e}^{0.01x_2(t-h)}$, $\boldsymbol{\theta} = [8, -8]^\top$, $\theta_1 = \theta_2 = 2$ and $\theta_3 = \theta_4 = 1$; see [38]. Also we take $\mathbf{Q} = \mathbf{I}_2$, $R = 0.01$.

We use the proposed method for solving both cases. Taking the optimal control law as a function of the state and delayed state vectors, the optimal curves for the two cases are shown in Figure 6. In addition, a comparison is made in Table 5.

We have considered a wide variety of constraints in the previous works. We again, to show the applicability of the method, we consider different kinds of constraints in this work. We impose another types of constraints to the same system of the Example 4, Case 2 as:

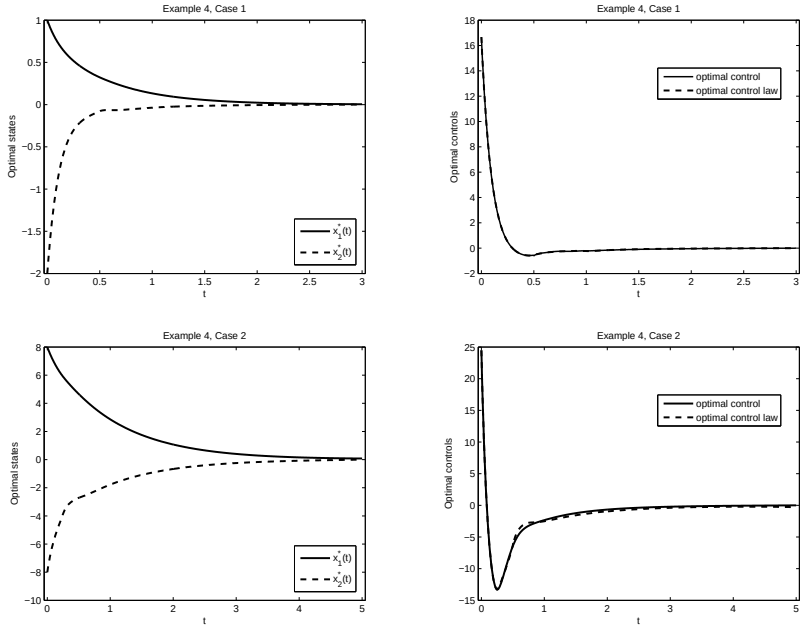


Figure 6. Solution curves of Example 4.

Table 5. Comparison of J^* for Example 4.

Case 1		Case 2
This Work	[37]	This Work
57.971696	57.9349	45.193082

Type 1: the pure state constraint $-1 \leq x_2(t) \leq 0$ is imposed for each $0.25 \leq t \leq 5$.

Type 2: the isoperimetric constraint $\int_{0.25}^{5} \{x_1(t) + x_2(t) - u(t)\}\, dt \leq 1.5$ is active.

Type 3: the quadratic state and control constraint $x_1^2(t) - x_2^2(t) + u(t) \leq 0$ is imposed for each $0.25 \leq t \leq 5$.

For adding the constraint in Type 1 to the QP model of the unconstrained system, we use the method given in [31, 14]. For the constraint in Type 2, in like manner we add it by using a new operational property as (A.1) which is presented in Appendix A. Finally for adding the constraint in Type 3, we use a similar linearization method by replacing it with a sequence of linear constraints. We find for Types 1, 2, and 3, $J^* = 49.710339$, $J^* = 46.614521$ and $J^* = 46.190119$, respectively. In the second type, we have $\int_{0.25}^{5} \{x_1^*(t) + x_2^*(t) - u^*(t)\} dt = 1.5$. The solution curves are presented in Figure 7. We see that the method can handle with complex nonlinear systems as real-world applications of time-delay optimal control theory. Figure 8 graphically illustrates the result of Type 3.

3.5. Example 5

Consider a continuous stirred tank reactor with the dynamics (see [39, 40])

$$\begin{aligned}
&\mathcal{D}^{\alpha} x_1(t) = -x_1(t) - r(t),\\
&\mathcal{D}^{\alpha} x_2(t) = -x_2(t) + 0.9u_2(t - h_u) + 0.1u_2(t),\\
&\mathcal{D}^{\alpha} x_3(t) = -2x_3(t) + 0.25r(t) - 1.05u_1(t)x_3(t - h_x),
\end{aligned}$$

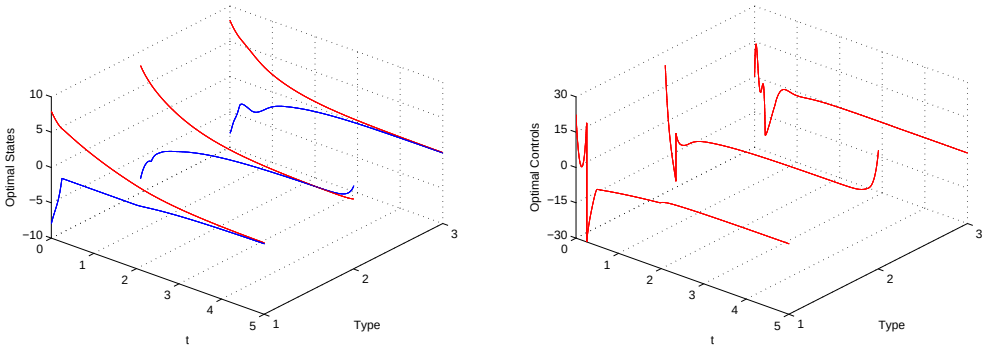


Figure 7. Solution curves of Example 4, Case 2 with constraints; $x_1^*(t)$ —, $x_2^*(t)$ —.

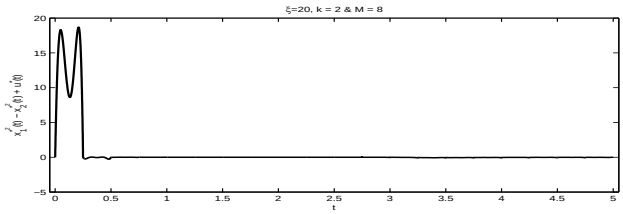


Figure 8. Results of Example 4, Case 2, Type 3.

where $r(t) = (1 + x_1(t))(1 + x_2(t)) \exp\left[\dfrac{25x_3(t)}{1 + x_3(t)}\right]$, $h_u = 0.02$ and we take $h_x \simeq 0.02$; also

$$\begin{aligned} x_3(t) &= -0.02, \quad h_x \le t \le 0 \\ u_2(t) &= 1, \quad h_u \le t \le 0 \\ \mathbf{x}(0) &= [0.49, -0.0002, -0.02]^\top. \end{aligned}$$

The cost function is

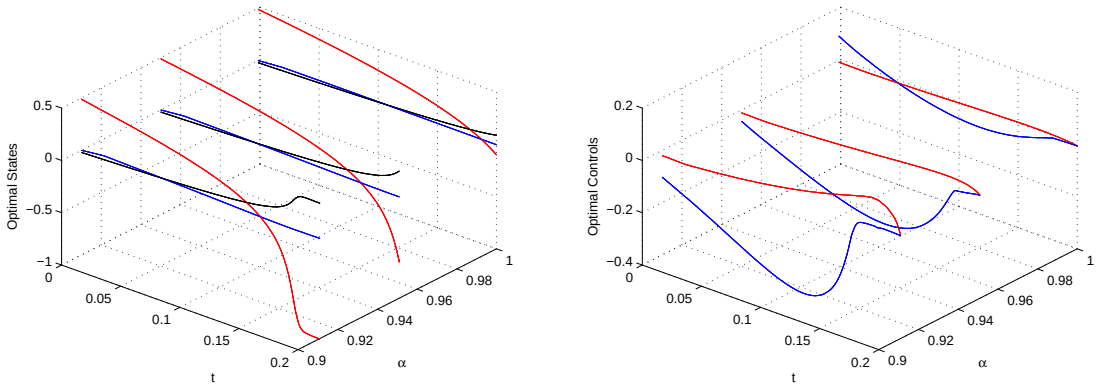
$$J = \int_0^{0.2} \left\{\mathbf{x}^\top(t)\mathbf{I}_3\mathbf{x}(t) + 0.01\mathbf{u}^\top(t)\mathbf{I}_2\mathbf{u}(t)\right\} dt.$$

By using Remark 2 for the term $\exp[25x_3(t)/(1 + x_3(t))]$, we model this system and our result are presented in Figure 9 and Table 6. Also for $\alpha = 0.91$, we impose the constraint $u_2(t) \ge -0.1$. By modeling the constraint and adding it to the QP formulation of the system, we obtain the solutions and J^* is given in Table 6.

Again as a practical application, we applied the method to a real-world nonlinear system.

Table 6. J^* for some α, Example 5.

α	J^*
0.999	0.021312
0.95	0.021614
0.91	0.041571
0.91, constrained system	0.043957



(a) $x_1^*(t)$ —, $x_2^*(t)$ —, $x_3^*(t)$ —

(b) $u_1^*(t)$ —, $u_2^*(t)$ —

Figure 9. Solution curves of Example 5.

4. Conclusion

In this study, by using the iterative method we have expanded the concepts and computational procedure presented for the optimal control of fractional linear quadratic systems to the optimal control of fractional nonlinear systems. By using the quasilinearization method and the powerful properties associated with Chebyshev wavelets, the optimal solutions of fractional nonlinear systems have been obtained. Since the method is based on converting the problem under consideration to the sequence of QP problems, it can be applied in the cases we have some constraints on the control (input) vector, and/or state (internal variable) vector, and/or on the energy of the system. After reaching the condition of the final (optimal) solutions, the value of optimal cost is available as a default output of the quadprog solver. Different classes of nonlinear systems have been studied to show the applicability and generality of the method. We have seen that the method can be applied to practical industrial systems.

Appendix

A. The integration operational property of the Chebyshev wavelets vector on $[0, 1]$

Lemma 4
The integration of the Chebyshev wavelets vector on the interval $[0, 1]$ is obtained as

$$\int_0^1 \boldsymbol{\psi}(t)\, dt = \boldsymbol{\gamma}_{cw}, \tag{A.1}$$

where $\boldsymbol{\gamma}_{cw}$ is the $\xi^{k-1}M \times 1$ integration operational vector of the Chebyshev wavelets vector over $[0, 1]$ and it consists of ξ^{k-1} replicated column vectors.

Proof
From the definition of Chebyshev wavelets, by $\boldsymbol{\varphi}_n(t) \coloneqq [\psi_{n0}^{\xi}(t), \psi_{n1}^{\xi}(t), \ldots, \psi_{nM-1}^{\xi}(t)]$, we have

$$\int_0^1 \boldsymbol{\psi}(t)\, dt = \int_0^{\frac{1}{\xi^{k-1}}} [\boldsymbol{\varphi}_1(t), \mathbf{0}, \mathbf{0}, \ldots, \mathbf{0}]^\top dt + \ldots + \int_{\frac{\xi^{k-1}-1}{\xi^{k-1}}}^1 [\mathbf{0}, \mathbf{0}, \ldots, \mathbf{0}, \boldsymbol{\varphi}_{\xi^{k-1}}(t)]^\top dt.$$

We conclude by considering three cases as: $m = 0$, $m = 1$, and $m \geq 2$ that

I. if $m = 0$,

$$\int_{\frac{n-1}{\xi^{k-1}}}^{\frac{n}{\xi^{k-1}}} \psi_{n0}^{\xi}(t)dt = \frac{\sqrt{2}}{\sqrt{\xi^{k-1}\pi}},$$

II. if $m = 1$, by substituting $\cos\theta = 2\xi^{k-1}t - 2n + 1$,

$$\int_{\frac{n-1}{\xi^{k-1}}}^{\frac{n}{\xi^{k-1}}} \psi_{n1}^{\xi}(t)dt = \frac{2\sqrt{\xi^{k-1}}}{\sqrt{\pi}} \frac{1}{2\xi^{k-1}} \int_0^{\pi} \cos\theta \, \sin\theta \, d\theta = 0,$$

III. if $m \geq 2$,

$$\int_{\frac{n-1}{\xi^{k-1}}}^{\frac{n}{\xi^{k-1}}} \psi_{nm}^{\xi}(t)dt = \frac{2\sqrt{\xi^{k-1}}}{\sqrt{\pi}} \frac{1}{2\xi^{k-1}} \int_0^{\pi} \cos m\theta \, \sin\theta \, d\theta = \frac{1}{\sqrt{\xi^{k-1}\pi}} \left(-\frac{1+(-1)^m}{m^2-1} \right).$$

It is apparent the values of the entries of $\boldsymbol{\gamma}_{cw}$ in the nth subinterval depend not on n but only on m, thus we find

$$\boldsymbol{\gamma}_{cw} = \frac{1}{\sqrt{\xi^{k-1}\pi}} \left\{ \begin{bmatrix} \bar{\boldsymbol{\gamma}} \\ \bar{\boldsymbol{\gamma}} \\ \vdots \\ \bar{\boldsymbol{\gamma}} \end{bmatrix} \right\} (\xi^{k-1} \text{ times})$$

as the $\xi^{k-1}M \times 1$ vector consisting of ξ^{k-1} replicated vectors, where the $M \times 1$ vector $\bar{\boldsymbol{\gamma}}$ is

$$\bar{\boldsymbol{\gamma}} = \left[\sqrt{2}, 0, -\tfrac{2}{3}, 0, -\tfrac{2}{15}, 0, \ldots, 0, -\tfrac{1+(-1)^m}{m^2-1}, 0, \ldots, -\tfrac{1+(-1)^{M-1}}{(M-1)^2-1} \right]^{\top}, m \geq 2.$$

□

REFERENCES

1. M. Malek-Zavarei, M. Jamshidi, *Time-Delay Systems: Analysis, Optimization and Applications*, Elsevier Science Inc., North-Holland, 1978.
2. D.E. Kirk, *Optimal Control Theory: An Introduction*, Courier Corporation, 2004.
3. D.S. Naidu, *Optimal Control Systems*, Idaho State University. Pocatello. Idaho. USA, CRC PRESS, 2003.
4. A.P. Sage, *Optimum Systems Control*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.
5. T. Çimen, S.P. Banks, *Global optimal feedback control for general nonlinear systems with nonquadratic performance criteria*, Systems & Control Letters, vol. 53, no. 5, pp. 327–346, 2004.
6. M. Tomas-Rodrigueza, S.P. Banks, *An Iterative Approach to Eigenvalue Assignment for Nonlinear Systems*, International Journal of Control, vol. 86, no. 5, pp. 883–892, 2013. https://doi.org/10.1080/00207179.2013.765037
7. H. Jaddu, A. Majdalawi, *An Iterative Technique for Solving a Class of Nonlinear Quadratic Optimal Control Problems Using Chebyshev Polynomials*, International Journal of Intelligent Systems and Applications, vol. 6, no. 6, pp. 53–57, 2014. DOI: 10.5815/ijisa.2014.06.06
8. E. Ziaei, M.H. Farahi, *The approximate solution of non-linear time-delay fractional optimal control problems by embedding process*, IMA Journal of Mathematical Control and Information, vol. 36, no. 3, pp. 713–727, 2019. DOI: 10.1093/imamci/dnx063
9. F. Kheyrinataj, A. Nazemi, *Fractional power series neural network for solving delay fractional optimal control problems*, Connection Science, vol. 32, no. 1, pp. 53–80, 2020. https://doi.org/10.1080/09540091.2019.1605498
10. H.R. Marzban, F. Malakoutikhah, *Solution of delay fractional optimal control problems using a hybrid of block-pulse functions and orthonormal Taylor polynomials*, Journal of the Franklin Institute, vol. 356, no. 15, pp. 8182–8215, 2019. doi: https://doi.org/10.1016/j.jfranklin.2019.07.010
11. I. Malmir, *A New Fractional Integration Operational Matrix of Chebyshev Wavelets in Fractional Delay Systems*, Fractal and Fractional, vol. 3, no. 3, 2019: 46. Doi:10.3390/fractalfract3030046
12. M.A. Johnson, F.C. Moon, *Experimental characterization of quasiperiodicity and chaos in a mechanical system with delay*, International journal of Bifurcation and Chaos, vol. 9, no. 01, pp. 49–65, 1999. http://dx.doi.org/10.1142/S0218127499000031
13. M. Caputo, *Linear models of dissipation whose Q is almost frequency independent*, Annals of Geophysics, vol. 19, no. 4, pp. 383–393, 1966.

14. I. Malmir, S.H. Sadati, *Transforming linear time-varying optimal control problems with quadratic criteria into quadratic programming ones via wavelets*, Journal of Applied Analysis, vol. 26, no. 1, pp. 131–152, 2020. https://doi.org/10.1515/jaa-2020-2011
15. M. Barrios, G. Reyero, *An Euler-Lagrange Equation only Depending on Derivatives of Caputo for Fractional Variational Problems with Classical Derivatives*, Statistics, Optimization & Information Computing, vol. 8, no. 2, pp. 590–601, 2020. https://doi.org/10.19139/soic-2310-5070-865
16. I. Malmir, *Legendre wavelets with scaling in time-delay systems*, Statistics, Optimization & Information Computing, vol. 7, no. 1, pp. 235–253, 2019. https://doi.org/10.19139/soic.v7i1.460
17. A. Kheirabadi, A. Mahmoudzadeh Vaziri, S. Effati, *Liner optimal control of time delay systems via Hermite wavelet*, Numerical Algebra, Control and Optimization, vol. 10, no. 2, pp. 143–156., 2020. doi:10.3934/naco.2019044
18. M.A. Sarhan, S. Shihab, M. Rasheed, *A New Boubaker Wavelets Operational Matrix of Integration*, Journal of Southwest Jiaotong University, vol. 55, no. 2, 2020.
19. E. Keshavarz, Y. Ordokhani, *A fast numerical algorithm based on the Taylor wavelets for solving the fractional integrodifferential equations with weakly singular kernels*, Mathematical Methods in the Applied Sciences, vol. 42, no. 13, pp. 4427–4443, 2019.
20. M. Caputo, F. Mainardi, *A new dissipation model based on memory mechanism*, Pure and Applied Geophysics, vol. 91, no. 1, pp. 134–147, 1971. doi: https://doi.org/10.1007/BF00879562
21. F. Brauer, *Perturbations of nonlinear systems of differential equations, II*, Journal of Mathematical Analysis and Applications, vol. 17, no. 3, pp. 418–434, 1967.
22. B.G. Pachpatte *A Note on Gronwali-Bellman Inequality*, Journal of Mathematical Analysis and Applications, vol. 44, no. 3, pp. 758–762, 1973.
23. R. Luus, X. Zhang, F. Hartig, F. J. Keil, *Use of piecewise linear continuous optimal control for time-delay systems*, Industrial & engineering chemistry research, vol. 34, no. 11, pp. 4136–4139, 1995.
24. A.Y. Lee, *Numerical solution of time-delayed optimal control problems with terminal inequality constraints*, Optimal Control Applications and Methods, vol. 14, no. 3, pp. 203–210, 1993.
25. A. Nazemi, E. Fayyazi, M. Mortezaee, *Solving optimal control problems of the time-delayed systems by a neural network framework* Connection Science, vol. 31, no. 4, pp. 342–372, 2019. DOI: 10.1080/09540091.2019.1604627
26. Y.H. Roh, J.H. Oh, *Sliding Mode Control for Robust Stabilization of Uncertain Input-Delay Systems*, The Institute of Control, Automation and Systems Engineers, vol. 2, no. 2, pp. 98–103, 2000.
27. D. Bresch-Pietri, N. Petit, M. Krstic, *Prediction-based control for nonlinear state- and input-delay systems with the aim of delay-robustness analysis*, 54th IEEE Conference on Decision and Control (CDC 2015), Osaka, Japan, pp. 6403–6409, 2015. DOI: 10.1109/CDC.2015.7403228
28. B. Zhou, *Truncated Predictor Feedback for Time-Delay Systems*, Springer-Verlag, 2014. DOI: 10.1007/978-3-642-54206-0-1
29. I. Estrada-Sánchez, M. Velasco-Villa, H. Rodríguez-Cortés, *Prediction-Based Control for Nonlinear Systems with Input Delay*, Mathematical Problems in Engineering Volume 2017, Article ID 7415418. https://doi.org/10.1155/2017/7415418
30. I. Malmir, *Novel Chebyshev wavelets algorithms for optimal control and analysis of general linear delay models*, Applied Mathematical Modelling, vol. 69, pp. 621–647, 2019. doi:10.1016/j.apm.2018.12.009.
31. I. Malmir, *Optimal control of linear time-varying systems with state and input delays by Chebyshev wavelets*, Statistics, Optimization & Information Computing, vol. 5, no. 4, pp. 302–324 2017. http://dx.doi.org/10.19139/soic.v5i4.341
32. H.T. Banks, *Approximation of Nonlinear Functional Differential Equation Control Systems*, Journal of Optimization Theory and Applications, vol. 29, no. 3, pp. 383–408, 1979.
33. K.H. Wong, D.J. Clements, K.L. Teo *Optimal Control Computation for Nonlinear Time-Lag Systems*, Journal of Optimization Theory and Applications, vol. 47, no. 1, pp. 91–107, 1985.
34. G.N. Elnagar, M.A. Kazemi, *Numerical solution of time-delayed functional differential equation control systems*, Journal of computational and applied mathematics, vol. 130, no. 1–2, pp. 75–90, 2001.
35. S.M. Hoseini, H.R. Marzban, *Costate Computation by an Adaptive Pseudospectral Method for Solving Optimal Control Problems with Piecewise Constant Time Lag*, Journal of Optimization Theory and Applications, vol. 170, no. 3, pp. 735–755, 2016. DOI: 10.1007/s10957-016-0957-3
36. X. Tang, H. Xu, *Multiple-Interval Pseudospectral Approximation for Nonlinear Optimal Control Problems with Time-Varying Delays*, Applied Mathematical Modelling, vol. 68, pp. 137–151, 2018. doi: https://doi.org/10.1016/j.apm.2018.09.039
37. M.K. Bouafoura, N.B. Braiek, *Hybrid Functions Direct Approach and State Feedback Optimal Solutions for a Class of Nonlinear Polynomial Time Delay Systems*, Complexity, 2019, Article ID 9596253. doi: http://dx.doi.org/10.1155/2019/9596253
38. C. Hua, P.X. Liu, X. Guan, *Backstepping control for nonlinear systems with time delays and applications to chemical reactor systems*, IEEE Transactions on Industrial Electronics, vol. 56, no. 9, pp. 3723–3732, 2009.
39. J.T. Betts, S. L. Campbell, K.C. Thompson, *Optimal Control Software for Constrained Nonlinear Systems with Delays*, CACSD, IEEE Multi Conference on Systems and Control, Denver, USA, 2011. 444–449. DOI: 10.1109/CACSD.2011.6044560
40. J.T. Betts, S. L. Campbell, K.C. Thompson, *Solving optimal control problems with control delays using direct transcription*, Applied Numerical Mathematics, vol. 108, pp. 185–203, 2016.