CAPUTO FRACTIONAL DERivative OPERATIONAL MATRICES OF LEGENDRE AND CHEBYSHEV WAVELETS IN FRACTIONAL DELAY OPTIMAL CONTROL

IMAN MALMIR

Department of Mechanical Engineering
Buali Sina University, Hamedan, Iran

(Communicated by Chongyang Liu)

ABSTRACT. Caputo derivative operational matrices of the arbitrary scaled Legendre and Chebyshev wavelets are introduced by deriving directly from these wavelets. The Caputo derivative operational matrices are used in quadratic optimization of systems having fractional or integer orders differential equations. Using these operational matrices, a new quadratic programming wavelet-based method without doing any integration operation for finding solutions of quadratic optimal control of traditional linear/nonlinear fractional time-delay constrained/unconstrained systems is introduced. General strategies for handling different types of the optimal control problems are proposed.

1. Introduction. Quadratic programming (QP) is a powerful aid in finding the optimal solutions of systems with quadratic performance indices. There are many works which studied the algorithms and the methods of solutions. [19,20,37,41]. Also a wide variety of software packages have been designed for solving QP problems like MATLAB. Some control systems are characterized by delay-free state equations [18,36] and some of them are characterized by time-delay state equations in which we have some delayed terms in the systems [7,17,24]. Moreover, some systems are characterize by fractional order state equations [22,35,38]. The fractional order differential operator can always characterize complex phenomena involving memory better than integer order differential operator, [22,42]. Considering these facts, we need a method which can handle all the mentioned types of optimal control systems. In [25–31], we transformed the optimization of different types of control systems into QP problems. But in all of them, we have used the integration operations of the state equations by using the integration operational matrices, that is, first we have integrated both sides of state equations and then we have replaced the approximated terms; especially in fractional systems, we have utilized the fractional integral operator. We can transform a time-delay system to a delay-free one by using some approximations methods, which enables us to use a method developed for such systems. For example in [23], the authors presented a novel second order numerical integration technique for the fractional dynamical system using a set of quadrature

2010 Mathematics Subject Classification. Primary: 26A33, 49M99; Secondary: 90C20.

Key words and phrases. Caputo derivative operational matrices of Legendre wavelets and Chebyshev wavelets, Direct optimization by wavelets, Fractional time-delay optimal control, Multifractional delay systems, Derivative operational matrix, Nonlinear optimal control.
points. Then a discretization scheme for solving a class of nonlinear optimal problem has been proposed. In this work, we are going to present a general method for fractional/integer order linear/nonlinear delay/free-delay optimal control problems. Several works for optimal control of fractional delayed systems were used the fractional integral operator where some of them have been cited in [29,31]. In some of them, one can see there are certain issues of the accuracy of the final solutions, for example, the initial condition $x(0)$ were not satisfied, or their results for $\alpha = 1$ (in which we have an integer order optimal control problem) did not match with the previous results presented in many literature. In our method, the initial condition is satisfied exactly, the accurate results are obtained for $\alpha = 1$ and the value of the optimal performance index without doing any work is reached as a default output of the QP solver. Some other works are as: in [39] by using Bernstein polynomials, the fractional linear optimal control problem has been transformed into a system of algebraic equations; [13] by using the Păde approximation, the authors transformed the delayed systems into free-delay systems and then by using the fractional operational matrix of Müntz polynomials, the optimization problem is converted to a nonlinear programming one; in [6], the authors applied a collocation method to convert the time-delay fractional optimal control problem to a nonlinear programming one; in [33], the author used the hybrid of block-pulse and Legendre polynomials and by approximating the operational matrix of fractional integration, the resulting equations form a system of nonlinear equations. Most of these method are based on transforming the original problem to a set of nonlinear problems, regardless of whether we can apply them to constrained complex systems, they are defined in the case we have a system described by a single fractional order state equation, the fractional order is less than 1 and some can be applied only to one class of control systems.

In this study, we apply a direct method for finding QP formulations of different systems by using fractional derivative operational matrices. We obtain these matrices for Legendre and Chebyshev wavelets with scaling based on the Caputo sense. In [15], the authors presented a formulation for classic Chebyshev wavelets which has been used by the shifted Chebyshev polynomials. Also in [3] a similar formulation was presented which have been applied by using the shifted Chebyshev polynomials. In [27], we have emphasized that in the definition of Chebyshev wavelets, we must have $k \in \mathbb{N}_{\geq 2}$ and the value $k = 1$ (or $k = 0$ depending on formulations of Chebyshev wavelets which were used) must be avoided because by using this value, we revert to the shifted Chebyshev polynomials. To illustrate, let us show that by using the concepts of them. The classic Chebyshev wavelets are defined as

$$
\psi_{nm} = \begin{cases} 
\alpha_m \sqrt{2^k} T_m(2^k t - 2n + 1), & \frac{n-1}{2^k-1} \leq t \leq \frac{n}{2^k-1}, \\
0, & \text{otherwise},
\end{cases}
$$

where $\alpha_m$ is for orthonormality, $n = 1, 2, \ldots, 2^k-1$, and $T_m$ is the Chebyshev polynomial of degree $m$. In some texts $k$ is replaced with $k + 1$. Choosing $k = 1$, we see that $n = 1$, $T_m(2^k t - 2n + 1) = T_m(2t - 1)$ and $[\frac{n-1}{2^k-1}, \frac{n}{2^k-1}] = [0, 1]$ which means that we have the shifted Chebyshev polynomials defined on $[0, 1]$, see [12], and by multiplying $\alpha_m$ with that one obtained from the formula for finding constant coefficients of the expansion, we return to the shifted Chebyshev polynomial expansion that has already been mentioned in many texts. By selecting this $k$, the dilation property of wavelets has been removed [11]. There is a similar statement for derivative operational matrices of Legendre wavelets; in [5,40], the operational matrices are constructed for shifted Legendre polynomials by choosing $k = 0$. An obvious
indication is that the solutions were presented in one interval while the wavelets solutions are piecewise-defined functions that are presented in two (at least) or more subintervals and because of this, we consider the compatibility constraint in our QP formulations. Also regardless of this discussion, the mentioned formulations are presented for classic Chebyshev and Legendre wavelets and we cannot use those formulations for Chebyshev or Legendre wavelets with arbitrary scaled parameters in which the scaling parameter for a problem may vary with that one for another problem.

In this work, we obtain formulas for Caputo fractional derivative operational matrices of Chebyshev and Legendre wavelets directly from their definitions without using any approximated function and these formulations will be applied on some fractional delay optimal control problems.

After presenting some definitions in Section 2, we introduce general strategies for the optimal control of different types in Section 3. In the meantime we are going to introduce the formulations of the Caputo derivative operational matrices of both Legendre and Chebyshev wavelets. Also we shall give additional formulations for the fractional orders of higher degree. The proposed wavelet-based methods will be applied on several examples in Section 4.

2. Basic definitions and concepts.

**Definition 2.1.** The Caputo fractional derivative of order \( \alpha \) denoted by \( D^\alpha \) arises in modeling of some physical systems \([4,8,9,21]\) and it is defined for a function \( f(t) \) as

\[
D^\alpha f(t) = \begin{cases} 
\frac{1}{\Gamma(m-\alpha)} \int_0^t (t-\rho)^{m-\alpha-1} f^{(m)}(\rho) \, d\rho, & m-1 < \alpha < m \\
\frac{d^m}{dt^m} f(t), & \alpha = m,
\end{cases}
\]

where \( \Gamma(\cdot) \) is the gamma function and \( m \in \mathbb{N} \).

\( D^\alpha \rho^n \), where \( \rho \) is the independent variable, possesses the property

\[
D^\alpha \rho^n = \begin{cases} 
0, & \alpha > n \\
\frac{\Gamma(n+1)}{\Gamma(n-\alpha+1)} \rho^{n-\alpha}, & \alpha \leq n.
\end{cases}
\]

If \( f(t) \) is a constant function, we see that \( D^\alpha f(t) = 0 \). Also for the Caputo derivative we have \([16,38]\)

\[
D^\alpha (D^n f(t)) = D^{\alpha+n} f(t), \text{ where } n = 0, 1, 2, \ldots.
\]

**Definition 2.2.** The Riemann–Liouville fractional integral operator of order \( \alpha \) denoted by \( I^\alpha \) for a function \( f(t) \) as is defined as \([10]\)

\[
I^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\rho)^{\alpha-1} f(\rho) \, d\rho.
\]

For the Riemann–Liouville fractional integral and Caputo fractional derivative, we have

\[
I^\alpha(D^\alpha f(t)) = f(t) - \sum_{\nu=0}^{m-1} f^{(\nu)}(0) \frac{t^\nu}{\nu!}.
\]

**Definition 2.3.** Wavelets: Legendre wavelets \( \phi_{nm}^\xi \) are defined on \([0, 1]\) as \([28]\)

\[
\phi_{nm}^\xi(t) = \begin{cases} 
\sqrt{\xi} c_{m, 2\xi^{k-1}} P_m(2\xi^{k-1} t - 2n + 1), & t \in \left[ \frac{n-1}{\xi^{k-1}}, \frac{n}{\xi^{k-1}} \right] \\
0, & \text{otherwise},
\end{cases}
\]

\( c_{m, 2\xi^{k-1}} \) is the Legendre polynomial of degree \( m \) at \( 2\xi^{k-1} \) and \( P_m(\cdot) \) is the Legendre polynomial of degree \( m \) at \( \cdot \).
where \( P_m \) are the well-known Legendre polynomials, \( \xi \in \mathbb{N}_{\geq 2} \) is an arbitrarily selected scaling parameter, \( k \in \mathbb{N}_{\geq 2} \) together with \( \xi \) specify the number of subintervals and both are finite values, \( n = 1, 2, \ldots, \xi^{k-1} \) refers to the number of subinterval and specifies the location of the subinterval, \( m = 0, 1, \ldots \) is the degree of \( P_m \), \( c_m = \sqrt{2m+1} \), and \( t \in [0, 1] \) is as an independent variable.

Chebyshev wavelets \( \psi_{nm}^\xi \) are defined on \([0, 1]\) as [27]

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

where \( T_m \) are the well-known Chebyshev polynomials, \( \xi, k, n, \) and \( m \) are the same as before and \( c_0 = 1/\sqrt{\pi}, \ c_m \neq 0 = \sqrt{2/\sqrt{\pi}}. \)

These definitions of Legendre and Chebyshev wavelets have some advantages over classic definitions of them (when \( \xi = 2 \)) specially in time-delay systems, see [27]. We usually take \( k = 2 \), hence in both definition we can set \( N = \xi^{k-1} \).

Legendre wavelets form an orthogonal basis with respect to the weight functions \( \psi_{n0}(t) = 1 \) and Chebyshev wavelets form an orthogonal basis with respect to \( \psi_{0w}^\xi(t) = 1/\sqrt{1 - (2\xi^{k-1}t - 2n + 1)^2} \).

**Definition 2.4.** Wavelet series expansions: We can expand a function \( f(t) \) in a series of Legendre or Chebyshev wavelets denoted by \( \{w_{nm}^\xi(t)\} \) (the \( M \)th term in \( \xi^{k-1} \) subintervals) as

\[
f(t) \approx \sum_{n=1}^{\xi^{k-1}} \sum_{m=0}^{M-1} f_{nm} w_{nm}^\xi(t) = \mathbf{f}_w \mathbf{w}(t),
\]

where \( \mathbf{f}_w \) is a \( 1 \times \xi^{k-1}M \) vector consists of constants\(^1\), \( \mathbf{w}(t) \) as a vector consisting of any of these wavelets is a \( \xi^{k-1}M \times 1 \) vector and

\[
\mathbf{f}_w = [f_{10}, f_{11}, \ldots, f_{1M-1}, f_{20}, \ldots, f_{2M-1}, \ldots, f_{\xi^{k-1}0}, \ldots, f_{\xi^{k-1}M-1}],
\]

\[
\mathbf{w}(t) = [w_{10}^\xi(t), w_{11}^\xi(t), \ldots, w_{1M-1}^\xi(t), w_{20}^\xi(t), \ldots, w_{2M-1}^\xi(t), \ldots, w_{\xi^{k-1}M-1}^\xi(t)]^\top.
\]

In (7), the application of the first summation differs from that of the second one, see [30]. The constant coefficients of the scaling functions \( \{w_{nm}^\xi(t)\} \) can be obtained from

\[
f_{nm} = \int_{\frac{n-1}{\xi^{k-1}}}^{\frac{n}{\xi^{k-1}}} f(t) w_{nm}^\xi(t) \psi_{nm}^\xi dt.
\]

3. Strategies for optimal control problems and formulations of Caputo operational matrices.

3.1. **Linear-quadratic fractional time-delay optimal control, Type I.** We consider the general linear problem of minimizing the real-valued performance index

\[
J = \frac{1}{2} \mathbf{x}^\top(t_f) \mathbf{T} \mathbf{x}(t_f) + \frac{1}{2} \int_0^{t_f} \{ \mathbf{x}^\top(t) \mathbf{Q}(t) \mathbf{x}(t) + \mathbf{u}^\top(t) \mathbf{R}(t) \mathbf{u}(t) \} dt,
\]

\(^1\text{Note that "lw" and "cw" refer to Legendre and Chebyshev wavelets and "w" refers to both wavelets.}\)
subject to
\[ D^{\alpha_1} x(t) + C(t) D^{\alpha_2} x(t) = A(t) x(t) + B(t) u(t) + \sum_{\mu=1}^{a} E_\mu(t) x(t - h_\mu) \]
\[ + \sum_{\nu=1}^{b} F_\nu(t) u(t - h_\nu) + d(t), \]
where \( t \in [0, t_f] \), \( t_f \) as the final time is fixed, \( T, Q(t) \in \mathbb{R}^{q \times q} \) are symmetric, positive semi-definite matrices, \( R(t) \in \mathbb{R}^{r \times r} \) a is symmetric, positive definite matrix, \( D^{\alpha_1} \) and \( D^{\alpha_2} \) are the Caputo fractional derivatives of orders \( \alpha_1 \) and \( \alpha_2 \), and \( \alpha_1, \alpha_2 \in (0, 1] \), \( x: [-h_x, t_f] \rightarrow \mathbb{R}^q \) and \( u: [-h_u, t_f] \rightarrow \mathbb{R}^r \) are the state and control vectors, \( C(t), A(t), \{ E_\mu(t) \} \in \mathbb{R}^{q \times q} \) and \( B(t), \{ F_\nu(t) \} \in \mathbb{R}^{q \times r} \) are matrices, \( d: [0, t_f] \rightarrow \mathbb{R}^q \) represents the disturbances, \( \{ h_\mu \} \) and \( \{ h_\nu \} \) are delays, \( x_0 \) is the initial condition, \( \Theta: [-h_x, 0] \rightarrow \mathbb{R}^q \) and \( \zeta: [-h_u, 0] \rightarrow \mathbb{R}^r \) are the initial state and control vector functions, \( h_x \) and \( h_u \) are the maximum values of \( h_\mu \) and \( h_\nu \), respectively. All matrices can be either time-varying or time-invariant.

By setting \( t/t_f \rightarrow t \) and using (7), we parameterize the state and the control vectors as follows
\[ x(t) = y^\top(t) x_w, \]
\[ u(t) = y^\top(t) u_w, \]
where for a matrix \( O, \hat{O} := O \otimes I_q, O := O \otimes I_r \). Similarly, the matrices \( A(t), B(t), \{ E_\mu(t) \} \) and \( \{ F_\nu(t) \} \) and the disturbances can be expressed as
\[ A(t) = A_w \hat{w}(t), B(t) = B_w \hat{w}(t), E_\mu(t) = E_{\mu w} \hat{w}(t), F_\nu(t) = F_{\nu w} \hat{w}(t), d(t) = \hat{w}^\top(t) d_w. \]

We start with the right-hand side of (10). For the delayed terms, we use the following property
\[ y(t - h) = \begin{cases} 0, & 0 \leq t < h_i, \\ D_i y(t), & h_i \leq t \leq 1, \end{cases} \]
where for the time-delay \( h_i, D_i \) is its delay operational matrix. For the both wavelets, \( D_i \) has the same structure which was explained in detail in \([27, 28]\). When \( \{ 0 \leq t < h_\mu \} \), \( \{ -h_\mu \leq t - h_\mu \leq 0 \} \), \( \{ 0 \leq t < h_\nu \} \), \( \{ -h_\nu \leq t - h_\nu \leq 0 \} \); hence \( \{ x(t - h_\mu) = \Theta(t - h_\mu), 0 \leq t < h_\mu \} \), \( \{ u(t - h_\mu) = \zeta(t - h_\mu), 0 \leq t < h_\mu \} \). Thus by putting this fact together with (12), (13), and (15) and by using the results and (14) to the right-hand side of (10), we obtain
\[ A(t) x(t) + B(t) u(t) + \sum_{\mu=1}^{a} E_\mu(t) x(t - h_\mu) + \sum_{\nu=1}^{b} F_\nu(t) u(t - h_\nu) + d(t) = \]
\[ A_w \hat{w}(t) \hat{w}^\top(t) x_w + B_w \hat{w}(t) \hat{w}^\top(t) u_w + \sum_{\mu=1}^{a} (E_{\mu w} \hat{w}(t) \Theta(t - h_\mu) + E_{\mu w} \hat{w}(t) \times \]
\[ \hat{w}^\top(t) D_\mu x_w) + \sum_{\nu=1}^{b} (F_{\nu w} \hat{w}(t) \zeta(t - h_\nu) + F_{\nu w} \hat{w}(t) \hat{w}^\top(t) D_\nu u_w) + \hat{w}^\top(t) d_w. \]
Now, we expand \( \{ \Theta(t - h_\mu) \} \) and \( \{ \zeta(t - h_\nu) \} \) as
\[ \Theta(t - h_\mu) = \hat{w}^\top(t) \Theta_{\mu w}, \]
\[ \zeta(t - h_\mu) = \hat{w}^\top(t) \zeta_{\mu w}, \]
and then we use another properties of these wavelets as

\[ f_w(t)w^T(t) \cong w^T(t)\hat{f}_w. \]  

Hence from (17)–(19), (16) becomes

\[
A(t)x(t) + B(t)u(t) + \sum_{\mu=1}^{a} E_{\mu}(t)x(t - h_{\mu}) + \sum_{\nu=1}^{b} F_{\nu}(t)u(t - h_{\nu}) + d(t) = w^T(t)\{ \\
\hat{A}_w x_w + \hat{B}_w u_w + \sum_{\mu=1}^{a} (\hat{E}_{\mu \nu} \theta_{\mu \nu} + \hat{E}_{\mu \nu} \hat{D}_w^T x_w) + \sum_{\nu=1}^{b} (\hat{F}_{\nu \nu} \zeta_{\nu \nu} + \hat{F}_{\nu \nu} \hat{D}_w^T u_w) + d_w \}.
\]  

For the left-hand side of (10), we will use a property which is introduced in the following theorem.

**Theorem 3.1.** The Caputo fractional derivative of Legendre and Chebyshev wavelets vectors can be obtained directly from the desired wavelets as

\[ D^\alpha w(t) \cong \mathbf{D}_w^\alpha w(t), \]  

where \( \mathbf{D}_w^\alpha \) is the \( \xi^{-1} M \times \xi^{-1} M \) Caputo fractional derivative operational matrix of the desired wavelets.

**Proof.** We use the procedure introduced in [29]. For Legendre and Chebyshev polynomials, we have

\[ P_m(t) = \sum_{j=0}^{m} \frac{(m + j)!}{2^j (m - j)! (j)!^2} (t - 1)^j, \]  

\[ T_m(t) = (-1)^m \sum_{j=0}^{m} \frac{(-2)^j (a_m + j - 1)!}{(m - j)! (2j)!} (t + 1)^j, \]

whence

\[ a_m = \begin{cases} 1, & m = 0, \\ m, & m \neq 0. \end{cases} \]  

First, we prove this theorem for Chebyshev wavelets as

\[ D^\alpha \psi(t) \cong \mathbf{D}_{cw}^\alpha \psi(t), \]

where \( \mathbf{D}_{cw}^\alpha \) is the \( \xi^{-1} M \times \xi^{-1} M \) Caputo fractional derivative operational matrix of the Chebyshev wavelets vector \( \psi(t) \) and

\[ \psi(t) = [\psi_{10}^\xi(t), \psi_{11}^\xi(t), \ldots, \psi_{1M-1}^\xi(t), \psi_{20}^\xi(t), \ldots, \psi_{2M-1}^\xi(t), \ldots, \psi_{nM-1}^\xi(t)]^T. \]

From (6), we have

\[ D^\alpha \psi_{nm}^\xi(t) = \mathbf{D} \psi_{nm}^\xi(t), \]

\[ \psi_{nm}^\xi(t), \quad t \in \left[ \frac{n-1}{\xi^{-1} M}, \frac{n}{\xi^{-1} M} \right). \]

When \( t < n/\xi^{-1} \) or \( t \notin \left[ \frac{n-1}{\xi^{-1} M}, \frac{n}{\xi^{-1} M} \right) \), we can write

\[ D^\alpha \psi_{1m}^\xi(t) = z s_1^m s_1^{om} s_2^{om} \ldots s_{1M-1}^{om} \psi_{10}^\xi(t), \psi_{11}^\xi(t), \psi_{12}^\xi(t), \ldots, \psi_{1M-1}^\xi(t))^T, \]  

(24)
where \( z_s \) is a constant and for \( i = 0, 1, 2, \ldots, M - 1 \), \( \{ s_{1i}^{\alpha m} \} \) are constants. Assume that \( 0 < \alpha \leq 1 \). From (6) and (2) it follows that \( \mathcal{D}^\alpha \psi_{10}^\xi(t) = 0 \), so we have
\[
\mathcal{D}^\alpha \psi_{10}^\xi(t) = z_s [0, 0, 0, \ldots, 0] [\psi_{10}^\xi(t), \psi_{11}^\xi(t), \psi_{12}^\xi(t), \ldots, \psi_{1M-1}^\xi(t)]^T. \tag{25}
\]
By (23), for \( m \geq 1 \), we can write
\[
\mathcal{D}^\alpha \psi_{1m}^\xi(t) = \mathcal{D}^\alpha \left\{ \sqrt{2\xi^{k-1} c_m} T_m(2\xi^{k-1} t - 1) \right\}
= (-1)^m \sqrt{2\xi^{k-1} c_m} \mathcal{D}^\alpha \sum_{j=1}^{m} \frac{(-2)^j (m+j-1)!}{(m-j)!(2j)!} (2\xi^{k-1})^j v^j.
\]
Now from (2), we have
\[
\mathcal{D}^\alpha \psi_{1m}^\xi(t) = (-1)^m \sqrt{2\xi^{k-1} c_m} \mathcal{D}^\alpha \sum_{j=1}^{m} \frac{(-2)^j (m+j-1)!}{(m-j)!(2j)!} (2\xi^{k-1})^j \Gamma(j+1) \Gamma(j-\alpha+1) t^j \alpha.
\]
For finding \( \{ s_{1m}^{\alpha m} \} \) in (24), we use (8) by setting \( \cos \theta = 2\xi^{k-1} t - 1 \); we see that
\[
(-1)^m \sqrt{2\xi^{k-1} c_m} \int_0^\pi \sum_{j=1}^{m} \frac{(-2)^j (m+j-1)!}{(m-j)!(2j)!} (2\xi^{k-1})^j \Gamma(j+1) \Gamma(j-\alpha+1) \cos^j \theta \theta d\theta.
\]
Taking the latter equations and (25) together, by
\[
\Theta_n(t) := [w_{10}^\xi(t), w_{11}^\xi(t), w_{12}^\xi(t), \ldots, w_{nM-1}^\xi(t)]^T,
\]
in which \( w_{nm}^\xi(t) = \psi_{nm}^\xi(t) \), and using the given information yield \( \mathcal{D}^\alpha \Psi(t) = \mathcal{D}^\alpha [\Theta_1(t); 0; 0; \cdots; 0] \), where semicolons are used to separate the rows. Hence, when \( 0 \leq t < 1/\xi^{k-1} \)
\[
\mathcal{D}^\alpha \Psi(t) = (2\xi^{k-1})^\alpha \begin{bmatrix}
0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0
\end{bmatrix} \begin{bmatrix}
\psi(t),
\end{bmatrix}, \tag{26}
\]
where
\[
s_{11}^{\alpha m} = [s_{10}^{\alpha m}, s_{11}^{\alpha m}, s_{12}^{\alpha m}, \ldots, s_{1M-1}^{\alpha m}]_{1 \times M}
\]
and
\[
s_{11}^{\alpha m} = (-1)^m mc_m c_i \int_0^\pi \sum_{j=1}^{m} \frac{(-2)^j (m+j-1)!}{(m-j)!(2j)!} (\cos \theta + 1)^j \alpha \cos(j \theta) d\theta.
\]
When \( t = 1/\xi^{k-1} \), we have \( \mathcal{D}^\alpha \psi_{1m}^\xi(t) = z_v \sum_{i=2}^{\xi^{k-1}} [v_{10}^{\xi m}, v_{11}^{\xi m}, v_{12}^{\xi m}, \ldots, v_{1M-1}^{\xi m}] \Theta_1(t), \) where \( z_v \) and \( \{ v_{1i}^{\xi m} \} \) are constants. From (1), for \( 0 < \alpha < 1 \) we have
\[
\mathcal{D}^\alpha \psi_{1m}^\xi(t) = \frac{1}{\Gamma(-\alpha+1)} \int_0^t (t - \rho)^{-\alpha} \sqrt{2\xi^{k-1} c_m} d\rho \frac{d}{d\rho} T_m(2\xi^{k-1} \rho - 1) d\rho.
\]
For $m = 0$, we get $\mathcal{D}_\alpha \psi_{10}^\xi(t) = \frac{1}{\Gamma(-\alpha + 1)} \int_0^{\frac{1}{\xi^\alpha}} (t - \rho)^{-\alpha} \frac{d}{d\rho} \sqrt{2\xi^{k-1}} c_0 \rho = 0$. For $m = 1$, we find
\[
\mathcal{D}_\alpha \psi_{11}^\xi(t) = \frac{1}{\Gamma(-\alpha + 1)} \int_0^{\frac{1}{\xi^\alpha}} (t - \rho)^{-\alpha} \sqrt{2\xi^{k-1}} c_1 \frac{d}{d\rho} (2\xi^{k-1} \rho - 1) d\rho
\]
\[
= \frac{\sqrt{2\xi^{k-1}} c_1}{\Gamma(-\alpha + 1)} \int_0^{\frac{1}{\xi^\alpha}} (t - \rho)^{-\alpha} (2\xi^{k-1}) d\rho
\]
\[
= \frac{\sqrt{2\xi^{k-1}} c_1}{\Gamma(-\alpha + 1)} \left( \frac{2\xi^{k-1} t^{1-\alpha}}{\Gamma(2-\alpha)} - \frac{2\xi^{k-1} (t - \frac{1}{\xi^\alpha})^{1-\alpha}}{\Gamma(2-\alpha)} \right).
\]

Let $m = 2$, then we have
\[
\mathcal{D}_\alpha \psi_{12}^\xi(t) = \frac{1}{\Gamma(-\alpha + 1)} \int_0^{\frac{1}{\xi^\alpha}} (t - \rho)^{-\alpha} c_2 \frac{d}{d\rho} \left(2((2\xi^{k-1} \rho - 1)^2 - 1) \right) d\rho
\]
\[
= \frac{1}{\Gamma(-\alpha + 1)} \sqrt{2\xi^{k-1}} c_2 \left[ \frac{16(2\xi^{k-1})^2 t^{2-\alpha}}{(2-\alpha)(1-\alpha)} - \frac{16(2\xi^{k-1})^2 (t - \frac{1}{\xi^\alpha})^{1-\alpha}}{(2-\alpha)(1-\alpha)} - \frac{8\xi^{k-1} t^{1-\alpha}}{(1-\alpha)} + \frac{8\xi^{k-1} (t - \frac{1}{\xi^\alpha})^{1-\alpha}}{(1-\alpha)} \right]
\]
\[
= \frac{\sqrt{2\xi^{k-1}} c_2}{\Gamma(-\alpha + 1)} \left[ \frac{4(2\xi^{k-1})^2 t^{2-\alpha}}{\Gamma(3-\alpha)} - \frac{4(2\xi^{k-1})^2 (t - \frac{1}{\xi^\alpha})^{2-\alpha}}{\Gamma(3-\alpha)} - \frac{8\xi^{k-1} t^{1-\alpha}}{\Gamma(2-\alpha)} \right]
\]
\[
= \frac{\sqrt{2\xi^{k-1}} c_2}{\Gamma(-\alpha + 1)} \left[ - \frac{8\xi^{k-1} (t - \frac{1}{\xi^\alpha})^{1-\alpha}}{\Gamma(2-\alpha)} \right].
\]

By using induction, we obtain the following extension for $m > 2$,
\[
\mathcal{D}_\alpha \psi_{1m}^\xi(t) = \frac{1}{\Gamma(-\alpha + 1)} \int_0^{\frac{1}{\xi^\alpha}} (t - \rho)^{-\alpha} \sqrt{2\xi^{k-1}} c_m \frac{d}{d\rho} T_m(2\xi^{k-1} \rho - 1) d\rho
\]
\[
= \frac{\sqrt{2\xi^{k-1}} c_m}{\Gamma(-\alpha + 1)} \int_0^{\frac{1}{\xi^\alpha}} (t - \rho)^{-\alpha} (-1)^m m \frac{d}{d\rho} \sum_{j=0}^{m} (-2)^j (m+j-1)! \frac{d}{d\rho} \left(2\xi^{k-1}\right)^j \rho^j d\rho
\]
\[
= \frac{\sqrt{2\xi^{k-1}} c_m}{\Gamma(-\alpha + 1)} \int_0^{\frac{1}{\xi^\alpha}} (t - \rho)^{-\alpha} (-1)^m \sum_{j=1}^{m} j \left( (-2)^j (m+j-1)! \frac{d}{d\rho} \left(2\xi^{k-1}\right)^j \rho^j \right) d\rho
\]
\[
= (-1)^m \frac{\sqrt{2\xi^{k-1}} c_m}{\Gamma(-\alpha + 1)} \sum_{j=1}^{m} (-2)^j (m+j-1)! \frac{d}{d\rho} \left(2\xi^{k-1}\right)^j \rho^j \left( \frac{\Gamma(j+1)}{\Gamma(j-\alpha+1)} \right)^{\alpha-j} \rho^{-\alpha}
\]
\[
= (-1)^m \frac{\sqrt{2\xi^{k-1}} c_m}{\Gamma(-\alpha + 1)} \sum_{j=1}^{m} (-2)^j (m+j-1)! \frac{d}{d\rho} \left(2\xi^{k-1}\right)^j \rho^j \left( \frac{\Gamma(j+1)}{\Gamma(j-\alpha+1)} \right)^{\alpha-j} \rho^{-\alpha}.
\]

Expanding in terms of $[\mathcal{O}_n(t)]$, we find $\mathcal{D}_\alpha \psi_{1m}^\xi(t) = (2\xi^{k-1})^\alpha \sum_{\eta=2}^{\xi^{k-1}} \mathbf{v}_{\eta m}^\alpha \mathcal{O}_n(t)$, where $\mathbf{v}_{\eta m}^\alpha = [\mathbf{v}_{\eta 0}^m, \mathbf{v}_{\eta 1}^m, \mathbf{v}_{\eta 2}^m, \ldots, \mathbf{v}_{\eta M-1}^m]_{1 \times M}$. In the second interval $\eta = 2$ and it is defined over $[1/\xi^{k-1}, 2/\xi^{k-1}]$. So, by $C_{mi} = (-1)^m m c_m c_i$ and setting $t = (\cos \theta + 3)/(2\xi^{k-1})$, we have
\[
\mathbf{v}_{2m}^\alpha \mathbf{c}_i = C_{mi} \int_0^{\pi} \sum_{j=1}^{m} \frac{(-2)^j (m+j-1)! \Gamma(j+1)}{\Gamma(j-\alpha+1)} \left( \frac{\cos \theta + 3}{2\xi^{k-1}} \right)^{-\alpha} \cos^{i\theta} \sin \theta d\theta.
\]
By letting \( t = \frac{n}{\xi^{k-1}} \), we find that

\[
 v_{\xi_{k-1}}^{\alpha, m} = C_m \int_0^\pi \sum_{j=1}^m \frac{(-2)^j (m + j - 1)! \Gamma(j + 1)}{(m - j)! (2j)! \Gamma(j - \alpha + 1)} \{ (\cos \theta + 2 \xi^{k-1} - 1)^j - \alpha \}
\]

\[- (1)^{n-j} (\cos \theta + 2 \xi^{k-1} - 3)^j - \alpha \} \cos(i\theta) \, d\theta.
\]

So, when \( t = 1/\xi^{k-1} \) we have

\[
 D^\alpha \psi(t) = (2\xi^{k-1})^\alpha \begin{bmatrix}
 0 & 0 & 0 & \cdots & 0 \\
 0 & v_{21}^\alpha & v_{31}^\alpha & \cdots & v_{\xi_{k-1}}^\alpha \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 & v_{2M-1}^\alpha & v_{3M-1}^\alpha & \cdots & v_{\xi_{k-1} M-1}^\alpha \\
 0 & 0 & 0 & \cdots & 0
\end{bmatrix} \psi(t),
\] (27)

where

\[
 v_{\xi_{k-1}}^{\alpha, m} = C_m \int_0^\pi \sum_{j=1}^m \frac{(-2)^j (m + j - 1)! \Gamma(j + 1)}{(m - j)! (2j)! \Gamma(j - \alpha + 1)} \{ (\cos \theta + 2 \eta^{k-1} - 1)^j - \alpha \}
\]

\[- (1)^{n-j} (\cos \theta + 2 \eta^{k-1} - 3)^j - \alpha \} \cos(i\theta) \, d\theta.
\]

Extending to a specific interval \([(n-1)/\xi^{k-1}, n/\xi^{k-1}]\)

\[
 D^\alpha \psi_{nm}^\xi(t(n)) = \frac{1}{\Gamma(-\alpha+1)} \int_{\frac{n}{\xi^{k-1}}}^{\frac{n+1}{\xi^{k-1}}} (t(n) - \rho)^{-\alpha} \sqrt{2\xi^{k-1}} c_m \frac{d}{d\rho} T_m(2\xi^{k-1} \rho - 2n + 1) \, d\rho.
\]

Substituting \( \varepsilon = \rho - (n-1)/\xi^{k-1} \) and integrating, yields

\[
 D^\alpha \psi_{nm}^\xi(t(n)) = (-1)^n \sqrt{2\xi^{k-1}} c_m \sum_{j=1}^m \frac{(-2)^j (m + j - 1)! (2\xi^{k-1})^j \Gamma(j + 1)}{(m - j)! (2j)! \Gamma(j - \alpha + 1)} \times (t(n) - \frac{n-1}{\xi^{k-1}})^{-\alpha}.
\]

By letting \( t(n) = (\cos \theta + 2n - 1)/(2\xi^{k-1}) \) to expand the resulting function in terms of wavelets of the current subinterval, we find that \( D^\alpha \psi_{nm}^\xi(t(n)) \) has the same coefficients as \( D^\alpha \psi_{1m}^\xi(t) \) in terms of its current wavelets vector. Hence, when

\[
 (n-1)/\xi^{k-1} \leq t < n/\xi^{k-1}
\]

\[
 D^\alpha \psi(t) = (2\xi^{k-1})^\alpha \begin{bmatrix}
 0 & 0 & 0 & \cdots & 0 \\
 0 & 0 & 0 & \cdots & 0 \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & 0 & \cdots & 0
\end{bmatrix} \psi(t),
\] (28)

Again, when \( t(n) = n/\xi^{k-1} \), \( D^\alpha \psi_{nm}^\xi(t(n)) = 0 \). For \( m \geq 1 \), we obtain

\[
 D^\alpha \psi_{nm}^\xi(t(n)) = (-1)^m \sqrt{2\xi^{k-1}} c_m \sum_{j=1}^m \frac{(-2)^j (m + j - 1)! (2\xi^{k-1})^j \Gamma(j + 1)}{(m - j)! (2j)! \Gamma(j - \alpha + 1)}
\]
Thus we find that \( \eta \) is the same as \( \varphi \). In its next subsequent interval by setting \( t \) subintervals by considering (26)–(29) and by setting \( D \alpha = \alpha \in (\cdot \cdot \cdot , \cdot \cdot \cdot) \), we see that

\[
D \alpha(t) = (2^{\xi k-1})^\alpha \begin{bmatrix}
0 \\
0 \\
0 \\
\vdots \\
0 \cdot \cdot \cdot \\
0 \\
0 \\
\vdots \\
0 \\
\end{bmatrix} \psi(t). \tag{29}
\]

In its next subsequent interval by setting \( t(n) = \frac{(\cos \theta + 2\eta - 1)/(2^{\xi k-1})} \), we get

\[
\epsilon \eta = \epsilon_{\eta_{M}}^m = \int_{0}^{\pi} \sum_{j=1}^{m} \frac{(-2)^j(m + j - 1)! \Gamma(j + 1)}{(m - j)!2^j! \Gamma(j - \alpha + 1)} \left\{ (\cos \theta + 3)^j - \alpha \right\} - (-1)^{m-j}(\cos \theta + 1)^j - \alpha \} \cos(i\theta) d\theta.
\]

Thus we find that \( \varphi \) is calculated from

\[
C_{\alpha M} \int_{0}^{\pi} \sum_{j=1}^{m} \frac{(-2)^j(m + j - 1)! \Gamma(j + 1)}{(m - j)!2^j! \Gamma(j - \alpha + 1)} \left\{ (\cos \theta + 2\eta - 2 - 1)^j - \alpha \right\} - (-1)^{m-j}(\cos \theta + 2\eta - 2 - 1)^j - \alpha \} \cos(i\theta) d\theta.
\]

For \( \eta = n + 1 \), we see that \( \varphi \) is the same as \( \varphi_{2m} \) for \( n = 1 \); similarly \( \varphi \) is the same as \( \varphi_{3m} \), and finally \( \varphi \) is equal to \( \varphi_{n+1m} \). Now for all subintervals by considering (26)–(29) and by setting

\[
S^\alpha := \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
\end{bmatrix}; \quad S^\alpha_{\eta-1} := \begin{bmatrix}
0 \\
\varphi_{\eta_{1M-1}} \\
\vdots \\
\varphi_{\eta_{M-1}} \\
\end{bmatrix},
\]

we have \( D^\alpha(t) \cong D^\alpha_{cw} \psi(t) \), where

\[
D^\alpha_{cw} = (2^{\xi k-1})^\alpha \begin{bmatrix}
\hat{D}^\alpha_1 \\
\hat{D}^\alpha_2 \\
\vdots \\
\hat{D}^\alpha_{\xi k-1} \\
\end{bmatrix}, \quad D^\alpha_n = \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
\end{bmatrix} \times (n-1) \times S^\alpha \times V^\alpha_1 \times V^\alpha_2 \times \cdots \times V^\alpha_{\xi k-1},
\]

in which \( 0 \) is a zero matrix of size \( M \times M \) and the proof is complete. For the Legendre wavelets vector

\[
\Phi(t) = [\phi_{10}(t), \phi_{11}(t), \ldots, \phi_{1M-1}(t), \phi_{10}(t), \ldots, \phi_{1M-1}(t), \ldots, \phi_{1M-1}(t), \ldots, \phi_{1M-1}(t)]^T,
\]

we use a similar process as we did. When \( 0 \leq t \leq 1/\xi k^{-1} \), by (5) and (22) (after setting \( -t \rightarrow t \) and using \( P_m(-t) = (-1)^m P_m(t) \), for \( m \geq 1 \), we can write

\[
D^\alpha \phi_{1m}(t) = D^\alpha \left\{ \sqrt{\xi k^{-1} c_m} P_m(2^{\xi k^{-1} t - 1}) \right\}
\]
Thus when \( \alpha \leq \alpha < \phi \),

\[
(2\xi^{k-1})^{j} \Gamma(j+1) \Gamma(j-\alpha+1)^{j-\alpha}.
\]

When \( t = 1/\xi^{k-1} \), then

\[
\mathcal{D}^{\alpha} \phi_{m}^{\xi}(t) = \frac{1}{\Gamma(-\alpha+1)} \int_{0}^{\xi^{k-1}} (t - \rho)^{\alpha} \phi_{m}^{\xi} d\rho
\]

\[
= \frac{\sqrt{\xi^{k-1} c_{m}}}{\Gamma(-\alpha+1)} \int_{0}^{\xi^{k-1}} (t - \rho)^{\alpha} \frac{d}{d\rho} \phi_{m}^{\xi} P_{m}(2\xi^{k-1} \rho - 1) d\rho
\]

\[
= \frac{\sqrt{\xi^{k-1} c_{m}}}{\Gamma(-\alpha+1)} \int_{0}^{\xi^{k-1}} (t - \rho)^{\alpha} \frac{d}{d\rho} \phi_{m}^{\xi} (2\xi^{k-1})^{j} \rho^{j-1} d\rho
\]

\[
= (-1)^{m} \frac{\sqrt{\xi^{k-1} c_{m}}}{\Gamma(-\alpha+1)} \int_{0}^{\xi^{k-1}} (t - \rho)^{\alpha} \frac{d}{d\rho} \phi_{m}^{\xi} (2\xi^{k-1})^{j} \rho^{j-1} d\rho
\]

\[
= (-1)^{m-j} (t - \frac{1}{\xi^{k-1}})^{-\alpha}.
\]

Thus when \( 0 \leq t \leq 1/\xi^{k-1} \)

\[
\mathcal{D}^{\alpha} \phi(t) = (2\xi^{k-1})^{\alpha} \begin{bmatrix}
0 \\
\text{...} \\
0 \\
\text{...} \\
0 \\
\text{...} \\
0 \\
\text{...} \\
0 \\
\text{...} \\
0
\end{bmatrix}
\]

where by \( \dot{\phi} := 2\xi^{k-1} t - 2n + 1, \)

\[
s_{11}^{\alpha m} = (-1)^{m} \frac{c_{m} \xi}{2} \int_{-1}^{1} \sum_{j=1}^{m} \frac{(m+j)! \Gamma(j+1)}{(-2)^{j} (m-j)! (j)!^{2} \Gamma(j - \alpha + 1)} (\dot{\phi})^{j-\alpha} P_{i}(\dot{\phi}) d\phi,
\]

\[
v_{\eta}^{\alpha m} = (-1)^{m} \frac{c_{m} \xi}{2} \int_{-1}^{1} \sum_{j=1}^{m} \frac{(m+j)! \Gamma(j+1)}{(-2)^{j} (m-j)! (j)!^{2} \Gamma(j - \alpha + 1)} ((\dot{\phi} + 2\eta - 1)^{j-\alpha}
\]

\[
- (-1)^{m-j} (\dot{\phi} + 2\eta - 3)^{j-\alpha}) P_{i}(\dot{\phi}) d\phi.
\]

When \( n - 1/\xi^{k-1} \leq t(n) \leq n/\xi^{k-1} \),

\[
\mathcal{D}^{\alpha} \phi(t) = (2\xi^{k-1})^{\alpha} \begin{bmatrix}
0 \\
\text{...} \\
0 \\
\text{...} \\
0 \\
\text{...} \\
0 \\
\text{...} \\
0 \\
\text{...} \\
0
\end{bmatrix}
\]
where
\[ v_{n}^{m} = (-1)^{m} \frac{c_{n}}{2} \int_{-1}^{1} \sum_{j=1}^{m} \frac{(m+j)! \Gamma(j+1)}{(-2)^{(m-j)}(2)^{(j-1)} \Gamma(j+1)\Gamma(j-\alpha+1)} ((\rho+2\eta-2n+1)^{j-\alpha} - (-1)^{m-j}((\rho+2\eta-2n-1)^{j-\alpha}) P_{j}(\rho) \, d\rho. \]

From (30), (31), we see for 0 ≤ \( t \leq 1 \) that
\[ \mathcal{D}^{\alpha} \phi(t) \cong \mathcal{D}^{\alpha}_{w} \phi(t), \]
where by
\[ S^\alpha := \begin{bmatrix} 0 \\ s_{11}^\alpha \\ \vdots \\ s_{1M-1}^\alpha \end{bmatrix}, \]
\[ V^\alpha := \begin{bmatrix} 0 \\ v_{q1}^\alpha \\ \vdots \\ v_{qM-1}^\alpha \end{bmatrix}, \]
we have
\[ \mathcal{D}^{\alpha}_{w} = (2^{\xi k-1})^\alpha \begin{bmatrix} \mathcal{D}^{\alpha}_{1} \\ \vdots \\ \mathcal{D}^{\alpha}_{n} \\ \vdots \\ \mathcal{D}^{\alpha}_{k^{k-1}-1} \end{bmatrix}, \]
\[ \mathcal{D}^{\alpha}_{n} = \underbrace{0 \ \cdots \ 0}_{(n-1) \text{ times}} S^{\alpha} V_{1}^{\alpha} V_{2}^{\alpha} \cdots V_{k^{k-1}-1}^{\alpha}. \]

We go back to the left-hand side of (10). Using
\[ C(t) = C_{w} \hat{w}(t), \]
the wavelets expansion (12), the product property (19) and the fractional derivative property (21), yields
\[ \mathcal{D}^{\alpha_{1}} \phi(t) + C(t) \mathcal{D}^{\alpha_{2}} \phi(t) = \hat{w}^\top(t) \{ \mathcal{D}^{\alpha_{1}}_{w} \phi_{w} + \hat{C}_{w} \mathcal{D}^{\alpha_{2}}_{w} \phi_{w} \}. \]
From (20) and (32) and the relation of Caputo derivative for the rescaled time \((t/t_{f} \rightarrow t)\), we have
\[ t_{f}^{-\alpha_{1}} \mathcal{D}^{\alpha_{1}}_{w} \phi_{w} + t_{f}^{-\alpha_{2}} \hat{C}_{w} \mathcal{D}^{\alpha_{2}}_{w} \phi_{w} = \hat{A}_{w} \phi_{w} + \hat{B}_{w} \phi_{w} \]
\[ + \sum_{\mu=1}^{a} (\hat{E}_{\mu} \phi_{w} + \hat{E}_{\mu w} \hat{D}_{w}^{\top} \phi_{w}) + \sum_{\nu=1}^{b} (\hat{F}_{\nu w} \phi_{w} + \hat{F}_{\nu w} \hat{D}_{w}^{\top} \phi_{w}) + d_{w}. \]

Now, we impose the compatibility constraint as
\[ x(t_{i}^{-}) = x(t_{i}^{+}), \]
where \( i = 1, 2, \ldots, \xi^{k-1}-1 \), and \( t_{i} = i/\xi^{k-1} \); this constraint can be written
\[ (W_{ce} \otimes I_{q}) x_{w} = 0_{(\xi^{k-1}-1)q \times 1}. \]

where by
\[ \rho_{\xi} = \begin{bmatrix} 0_{1 \times M} \\ \vdots \\ 0_{1 \times M} \\ \omega_{n} \begin{bmatrix} 1/\xi^{k-1} \\ 0_{1 \times M} \end{bmatrix} - \omega_{n} \begin{bmatrix} 0_{1 \times M} \\ \cdots \\ 0_{1 \times M} \end{bmatrix} \end{bmatrix}, \]
\[ W_{ce} = [\rho_{1}; \rho_{2}; \cdots; \rho_{\xi^{k-1}-1}]. \]
From the initial condition, we have
\[ (w^\top(0) \otimes I_{q}) x_{w} = x_{0}. \]
(34) and (35), written in matrix form, yields
\[
\begin{align*}
W_C x_w &= \begin{bmatrix} 0_{q(\xi^{k-1}) \times 1} \end{bmatrix}, \\
W_C &= \begin{bmatrix} W_{cc} \end{bmatrix}.
\end{align*}
\]  
(36)

(33) together with (36) gives
\[
\begin{align*}
\begin{bmatrix} \tilde{A}_w + \sum_{\mu=1}^{a} \tilde{E}_{\mu w} \tilde{D}_{\mu}^\top - t_f^{-\alpha_1} \tilde{D}_w^\top - t_f^{-\alpha_2} \tilde{C}_w \tilde{D}_w^\top - \tilde{B}_w + \sum_{\nu=1}^{b} \tilde{F}_{\nu w} \tilde{D}_{\nu}^\top \end{bmatrix} \\
\times \begin{bmatrix} x_w \end{bmatrix} &= \begin{bmatrix} -d_w - \sum_{\mu=1}^{a} \tilde{E}_{\mu w} \theta_{\mu w} - \sum_{\nu=1}^{b} \tilde{F}_{\nu w} \xi_{\nu w} \end{bmatrix}.
\end{align*}
\]  
(37)

Setting
\[
Q(t) = Q_w \tilde{w}(t), R(t) = R_w \tilde{w}(t)
\]  
in (9), and using (19) and the property
\[
\int_0^1 w(t) w^\top (t) \, dt = \Gamma_w,
\]
we obtain
\[
J = \frac{1}{2} \begin{bmatrix} x_w \end{bmatrix}^\top \begin{bmatrix} \Psi_1 & \Psi_2 \\ \Psi_3 & \Psi_4 \end{bmatrix} \begin{bmatrix} x_w \end{bmatrix},
\]  
(38)

where the product operational matrices of $Q_w$ and $R_w$ are, in turn, $\tilde{Q}_w$ and $\tilde{R}_w$.

Considering (37), (38), the QP model of the fractional time-delay optimal control problem of Type I is
\[
\min \frac{1}{2} \begin{bmatrix} x_w \end{bmatrix}^\top \begin{bmatrix} \Psi_1 & \Psi_2 \\ \Psi_3 & \Psi_4 \end{bmatrix} \begin{bmatrix} x_w \end{bmatrix},
\]  
(39)

subject to
\[
\begin{align*}
\begin{bmatrix} \Lambda_1 & \Lambda_2 \\ \Lambda_3 & \Lambda_4 \end{bmatrix} & \begin{bmatrix} x_w \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},
\end{align*}
\]  
(40)

where
\[
\begin{align*}
\Psi_1 &= (w(1) w^\top (1) \otimes T) + t_f \tilde{\Gamma}_w \tilde{Q}_w, \\
\Psi_2 &= 0_{q(\xi^{k-1}) \times q(\xi^{k-1})}, \\
\Psi_3 &= 0_{q(\xi^{k-1}) \times q(\xi^{k-1})}, \\
\Psi_4 &= t_f \tilde{\Gamma}_w \tilde{R}_w, \\
\Lambda_1 &= \tilde{A}_w + \sum_{\mu=1}^{a} \tilde{E}_{\mu w} \tilde{D}_{\mu}^\top - t_f^{-\alpha_1} \tilde{D}_w^\top - t_f^{-\alpha_2} \tilde{C}_w \tilde{D}_w^\top, \\
\Lambda_2 &= \tilde{B}_w + \sum_{\nu=1}^{b} \tilde{F}_{\nu w} \tilde{D}_{\nu}^\top, \\
\Lambda_3 &= W_C, \\
\Lambda_4 &= 0_{q(\xi^{k-1}) \times q(\xi^{k-1})}, \\
b_1 &= -d_w - \sum_{\mu=1}^{a} \tilde{E}_{\mu w} \theta_{\mu w} - \sum_{\nu=1}^{b} \tilde{F}_{\nu w} \xi_{\nu w}, \\
b_2 &= \begin{bmatrix} 0_{q(\xi^{k-1}) \times 1} \end{bmatrix}. 
\end{align*}
\]  
(41)

(42)

So, fractional linear-quadratic delay optimal control problems by using the fractional derivative operational matrices are converted to QP problems. We see that there are some significant differences between this QP with those were obtained by using the integration operation in the previous works.

Remark 1. If there is a time-invariant matrix denoted by $O$ (or matrices) in the state equations, the related product operational matrix $\tilde{O}_w$ in (42), (43) must be replaced by $I_{\xi^{k-1} M} \otimes O$. For dealing with such matrices in (9), see [27], [28].
3.2. The Caputo fractional derivative operational matrices of Chebyshev and Legendre wavelets.

Corollary 1 (Corollary to Theorem 3.1). Setting $\eta - 1 \rightarrow \eta$ which yields $\eta = 1, 2, \ldots, \zeta^{k-1} - 1$, the Caputo fractional derivative operational matrix of Chebyshev wavelets $\mathbf{D}_c^\alpha$ for $0 < \alpha \leq 1$ is

$$\mathbf{D}_c^\alpha = (2\zeta^{k-1})^\alpha \begin{bmatrix} 0 & V_1^\alpha & V_2^\alpha & \cdots & V_{\zeta^{k-1}-1}^\alpha \\ 0 & S^\alpha & V_1^\alpha & V_2^\alpha & \cdots & V_{\zeta^{k-1}-2}^\alpha \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & S^\alpha \end{bmatrix},$$

where

$$S^\alpha = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ s_0^\alpha & s_1^\alpha & s_2^\alpha & \cdots & s_{M-1}^\alpha \\ s_0^\alpha & s_1^\alpha & s_2^\alpha & \cdots & s_{M-1}^\alpha \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_0^{M-1} & s_1^{M-1} & s_2^{M-1} & \cdots & s_{M-1}^{M-1} \end{bmatrix}, \quad V_1^\alpha = \begin{bmatrix} 0 \\ v_0^\alpha \\ v_1^\alpha \\ \vdots \\ v_{\zeta^{k-1}-1}^\alpha \end{bmatrix},$$

$$s_i^{\alpha m} = (-1)^m c_c \int_0^\pi \frac{(-2)^j (m + j - 1)! \Gamma(j + 1)}{(m - j)! (2j)! \Gamma(j - \alpha + 1)} ((\cos \theta + 1)^{j-\alpha} \cos(i \theta)) d\theta,$$

$$v_\eta^{\alpha m} = (-1)^m c_c \int_0^{2\eta} \frac{(-2)^j (m + j - 1)! \Gamma(j + 1)}{(m - j)! (2j)! \Gamma(j - \alpha + 1)} \{(\cos \theta + 2\eta + 1)^{j-\alpha} - (\cos \theta + 2\eta - 1)^{j-\alpha}\} \cos(i \theta) d\theta.$$

Corollary 2 (Corollary to Theorem 3.1). Setting $\eta - 1 \rightarrow \eta$, the Caputo fractional derivative operational matrix of Legendre wavelets $\mathbf{D}_l^\alpha$ for $0 < \alpha \leq 1$ is obtained from

$$\mathbf{D}_l^\alpha = (2\zeta^{k-1})^\alpha \begin{bmatrix} 0 & V_1^\alpha & V_2^\alpha & \cdots & V_{\zeta^{k-1}-1}^\alpha \\ 0 & S^\alpha & V_1^\alpha & V_2^\alpha & \cdots & V_{\zeta^{k-1}-2}^\alpha \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & S^\alpha \end{bmatrix},$$

where

$$S^\alpha = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ s_0^\alpha & s_1^\alpha & s_2^\alpha & \cdots & s_{M-1}^\alpha \\ s_0^\alpha & s_1^\alpha & s_2^\alpha & \cdots & s_{M-1}^\alpha \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_0^{M-1} & s_1^{M-1} & s_2^{M-1} & \cdots & s_{M-1}^{M-1} \end{bmatrix}, \quad V_1^\alpha = \begin{bmatrix} 0 \\ v_0^\alpha \\ v_1^\alpha \\ \vdots \\ v_{\zeta^{k-1}-1}^\alpha \end{bmatrix},$$

$$s_i^{\alpha m} = (-1)^m c_c \int_{\zeta^{k-1}-1}^1 \frac{(m + j)! \Gamma(j + 1)}{2 \sum_{j=1}^m (-2)^j (m - j)! (j)! \Gamma(j - \alpha + 1)} (\theta + 1)^{j-\alpha} P_i(\theta) d\theta,$$

$$v_\eta^{\alpha m} = (-1)^m c_c \int_0^{2\eta} \frac{(m + j)! \Gamma(j + 1)}{2 \sum_{j=1}^m (-2)^j (m - j)! (j)! \Gamma(j - \alpha + 1)} \{(\theta + 2\eta + 1)^{j-\alpha} - (\theta + 2\eta - 1)^{j-\alpha}\} P_i(\theta) d\theta.$$
Theorem 3.2. Assume that a function \( f(t) \) is a differentiable function defined on \([0, 1]\), with bounded derivatives as \( |D^{2+\alpha}f(t)| \leq \delta \). Then the following norm in \( L_2 \) with respect to the weight function \( \varsigma_w(t) \)
\[
\|D^\alpha f(t) - f_w D^\alpha_w w(t)\|_{\varsigma_w}^2
\]
tends to zero as \( M \) increases, where \( D^\alpha_w \) is Caputo derivative operational matrix of the desired wavelets.

Proof. By taking
\[
\frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\rho)^{-\alpha} f^{(1)}(\rho) d\rho = \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} g_{nm} w_{nm}(t),
\]
considering the point of application of the first summation given in Definition 2.4 and denoting the arrays of the Caputo derivative operational matrices by \( d_{nm} \), we can write
\[
\|D^\alpha f(t) - D^\alpha \{f_w w(t)\}\|_{\varsigma_w}^2 \leq \left\| \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} g_{nm} w_{nm}(t) - \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} f_{nm} w_{nm}(t) \right\|_{\varsigma_w}^2
\]

\[
\leq \left\| \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} g_{nm} w_{nm}(t) - \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} f_{nm} w_{nm}(t) \right\|_{\varsigma_w}^2
\]

\[
\leq \left\| \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} g_{nm} w_{nm}(t) - \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} f_{nm} w_{nm}(t) \right\|_{\varsigma_w}^2
\]

\[
\leq (\xi^{k-1} - (n-1)) h_w(\delta, \xi, k) l_m + \left\| \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} g_{nm} w_{nm}(t) \right\|_{\varsigma_w}^2
\]

\[
\leq (\xi^{k-1} - (n-2)) h_w(\delta, \xi, k) l_m,
\]
where \( h_w(\delta, \xi, k) \) and \( l_m \) are constants and obtained according to the relative wavelets; for example for Chebyshev wavelets [27] are \( h_w(\delta, \xi, k) l_m = \frac{\pi(\delta + ||f^{(0)}||)^2}{4w^2 - (\sum_{m=1}^{M-1} h_m)^2} \). Also, since
\[
\frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\rho)^{-\alpha} f^{(1)}(\rho) d\rho = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\rho)^{-\alpha} \{f_w w(\rho)\}^{(1)}(\rho) d\rho,
\]
we used the following assertions:

\[
\begin{align*}
    g_{nm}w_{nm}^\xi(t) &= f_{nm} \sum_{m'=-n}^n \sum_{m'=0}^{\infty} d_{n',m'}w_{n',m'}^\xi(t), \quad n-\frac{1}{\xi} < t < \frac{n-1}{\xi}, \\
    g_{nm}w_{nm}^\xi(t) &= f_{nm} \sum_{m'=-n+1}^{\infty} \sum_{m'=0}^{\xi-1} d_{n',m'}w_{n',m'}^\xi(t), \quad t = \frac{n-1}{\xi}.
\end{align*}
\]

From the values of \(l_m\), we see immediately that \(\|D^\alpha f(t) - D^\alpha \{f_{w,t}(t)\}\|_{\psi_{ch,w}} \to 0\) as \(M \to \infty\).

### 3.3. Extensions of the Caputo derivative operational matrices

Now assume that \(1 < \alpha \leq 2\). We can use two options for Chebyshev and Legendre wavelets (CWs and LWs).

For CWs, we can use one of the following choices.

1. In a similar way we obtain

\[
D^\alpha_{cw} = (2\xi^{k-1})^\alpha \begin{bmatrix}
S^\alpha & V_1^\alpha & V_2^\alpha & \cdots & V_{\xi^{k-2}}^\alpha \\
0 & S^\alpha & V_1^\alpha & \cdots & V_{\xi^{k-2}}^\alpha \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & S^\alpha
\end{bmatrix},
\]

where

\[
S^\alpha = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
g_{01}^\alpha & g_{11}^\alpha & g_{12}^\alpha & \cdots & g_{1M-1}^\alpha \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
g_{m1}^\alpha & g_{m1}^\alpha & g_{m2}^\alpha & \cdots & g_{mM-1}^\alpha \\
g_{mM}^\alpha & g_{mM}^\alpha & g_{mM-1}^\alpha & \cdots & g_{mM-1}^\alpha
\end{bmatrix}, \quad V_\alpha = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
v_{\alpha0}^\alpha & v_{\alpha1}^\alpha & v_{\alpha2}^\alpha & \cdots & v_{\alphaM-1}^\alpha
\end{bmatrix},
\]

\[
s_{ij}^\alpha = (-1)^{m(m-1)}c_i c_j \int_0^\pi \sum_{j=2}^m \frac{(-1)^{j-1}(r + j-1)!\Gamma(j + 1)}{(r - j)!\Gamma(j + 1)(\cos \theta + 1)^{j-\alpha} \cos(i\theta)} d\theta,
\]

\[
v_{\alpha i}^\alpha = (-1)^{m(m-1)}c_i c_j \int_0^\pi \sum_{j=2}^m \frac{(-1)^{j-1}(r + j-1)!\Gamma(j + 1)}{(r - j)!\Gamma(j + 1)(\cos \theta + 1)^{j-\alpha} \cos(i\theta)} d\theta.
\]

2. From the property (3), we can use the derivative operational matrix of Chebyshev wavelets

\[
D^\alpha\psi(t) = D^{\alpha-1}D^1\psi(t) = D_{cw}D^{\alpha-1}_{cw}\psi(t).
\]

\(D_{cw}\) is the \(\xi^{k-1-M} \times \xi^{k-1-M}\) derivative operational matrix of the Chebyshev wavelet vector on \([0,t]\) and

\[
D^1\psi(t) = \frac{d\psi(t)}{dt} = D_{cw}\psi(t).
\]

\(D^{\alpha-1}_{cw}\) is obtained from Corollary 1. We must find \(D_{cw}\). According to the definition of the Chebyshev polynomials [34], for \(m \geq 1\), we can write

\[
\frac{dT_m(x)}{dx} = m (T_0(x) + 2T_2(x) + 2T_4(x) + \ldots + 2T_{m-1}(x)), \quad m \text{ odd},
\]

\[
\frac{dT_m(x)}{dx} = m (2T_1(x) + 2T_3(x) + 2T_5(x) + \ldots + 2T_{m-1}(x)), \quad m \text{ even}.
\]
where the range of the variable \( x \) is \([-1, 1]\). So for \( m = 1, 3, 5, \ldots \), we have

\[
\frac{d\psi_{\xi_n}^\xi(t)}{dt} = \sqrt{2\xi^{k-1}} \frac{dT_m(2\xi^{k-1}t - 2n + 1)}{dt}
\]
\[
= 2m\xi^{k-1} \sqrt{2\xi^{k-1}} \frac{d}{dt}(T_0(2\xi^{k-1}t - 2n + 1) + 2T_2(2\xi^{k-1}t - 2n + 1) + \ldots + 2T_{m-1}(2\xi^{k-1}t - 2n + 1))
\]
\[
= 2m\xi^{k-1} \left( \frac{\xi}{\xi_0} \psi_{\xi_{n0}}(t) + 2\psi_{\xi_{n2}}(t) + \ldots + 2\psi_{\xi_{nm}}(t) \right)
\]
\[
= 2\xi^{k-1} \left[ \sqrt{2m, 0, 2m, 0, 2m, \ldots, 0, 0, \ldots, 0} \right] \left[ \psi_{\xi_{n0}}(t), \psi_{\xi_{n1}}(t), \ldots, \psi_{\xi_{nm}}(t) \right]^T.
\]

for \( m = 2, 4, 6, \ldots \), we get

\[
\frac{d\psi_{\xi_n}^\xi(t)}{dt} = \sqrt{2\xi^{k-1}} \frac{dT_m(2\xi^{k-1}t - 2n + 1)}{dt}
\]
\[
= 2m\xi^{k-1} \sqrt{2\xi^{k-1}} \frac{d}{dt}(T_1(2\xi^{k-1}t - 2n + 1) + 2T_3(2\xi^{k-1}t - 2n + 1) + \ldots + 2T_{m-1}(2\xi^{k-1}t - 2n + 1))
\]
\[
= 2m\xi^{k-1} \left( \psi_{\xi_{n1}}(t) + 2\psi_{\xi_{n3}}(t) + \ldots + 2\psi_{\xi_{nm}}(t) \right)
\]
\[
= 2\xi^{k-1} \left[ 0, 2m, 0, 2m, 0, 2m, \ldots, 0, 0, \ldots, 0 \right] \left[ \psi_{\xi_{n0}}(t), \psi_{\xi_{n1}}(t), \ldots, \psi_{\xi_{nm}}(t) \right]^T.
\]

As a result

\[
\mathcal{D}_{cw} = 2\xi^{k-1} \text{blkdiag} \left( \begin{array}{c} \mathcal{D}_{cw} \end{array} \right)
\] (57)

where when \( M \) is even, \( \mathcal{D}_{cw} \) is the \( M \times M \) matrix of the form

\[
\mathcal{D}_{cw} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\sqrt{2} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 4 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
3\sqrt{2} & 0 & 6 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 8 & 0 & 8 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\sqrt{3}(M-1) & 0 & 2(M-1) & 0 & 2(M-1) & 0 & \cdots & 2(M-1) & 0
\end{bmatrix}
\] (58)

and, if \( M \) is odd then

\[
\mathcal{D}_{cw} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\sqrt{2} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 4 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
3\sqrt{2} & 0 & 6 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 8 & 0 & 8 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 2(M-1) & 0 & 2(M-1) & 0 & 2(M-1) & \cdots & 2(M-1) & 0
\end{bmatrix}
\] (59)

For LWs, in the case \( 1 < \alpha \leq 2 \) we also use similar methods given in the following.

1. We obtain

\[
\mathcal{D}_{lw} = (2\xi^{k-1})^\alpha \begin{bmatrix}
S^\alpha & V_1^\alpha & V_2^\alpha & V_3^\alpha & \cdots & V_{\xi^{k-1}}^\alpha \\
0 & S^\alpha & V_1^\alpha & V_2^\alpha & \cdots & V_{\xi^{k-1}}^\alpha \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & S^\alpha
\end{bmatrix}
\] (60)
where

\[
S^\alpha = \begin{bmatrix}
0 & 0 & 0 & 0 & \cdots & 0 \\
0 & s_0^2 & s_1^2 & s_2^2 & \cdots & s_{M-1}^2 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & s_0^{2M}
\end{bmatrix}, \quad V^\alpha = \begin{bmatrix}
0 & 0 & 0 & 0 & \cdots & 0 \\
0 & v_{n0} & v_{n1} & v_{n2} & \cdots & v_{nM-1} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
v_{n0} & v_{n1} & v_{n2} & \cdots & v_{nM-1}
\end{bmatrix},
\]

(61)

\[
s_i^{am} = (-1)^m \frac{c_{m\alpha}}{2} \int_{-1}^{1} \left(2\right)^{j} (m-j)!! \Gamma(j+1) \frac{(q+1)^{j-\alpha}}{(q+1)^{j-\alpha}} P_i(q) \, dq, \quad (62)
\]

\[
v_{ni}^{am} = (-1)^m \frac{c_{m\alpha}}{2} \int_{-1}^{1} \left(2\right)^{j} (m-j)!! \Gamma(j+1) \frac{(q+2\eta+1)^{j-\alpha}}{(q+2\eta+1)^{j-\alpha}} P_i(q) \, dq.
\]

2. Similarly, by (3), we can use the derivative operational matrix of Legendre wavelets

\[
\mathcal{D}^\alpha \Phi(t) = \mathcal{D}^\alpha \mathcal{D}^1 \Phi(t) = \mathcal{D}_{lw} \mathcal{D}^\alpha_{lw} \Phi(t).
\]

(64)

\[\mathcal{D}^\alpha_{lw}\] is given in Corollary 2. The derivative operational matrix of the Legendre wavelet vector denoted by \(\mathcal{D}_{lw}\) is a matrix of order \(\xi^{k-1}M\),

\[
\frac{d\Phi(t)}{dt} = \mathcal{D}_{lw} \Phi(t),
\]

and it is obtained by using the recurrence relation of Legendre polynomials [1].

From the property \(P_{m+1}'(x) = P_m(x) + (2m+1)P_m(x)\), we can write

\[
\frac{dP_m(x)}{dx} = P_0(x) + 5P_2(x) + 9P_4(x) + \ldots + (2m-1)P_{m-1}(x), \quad m \text{ odd}
\]

\[
\frac{dP_m(x)}{dx} = 3P_1(x) + 7P_3(x) + 11P_5(x) + \ldots + (2m-1)P_{m-1}(x), \quad m \text{ even}.
\]

For \(m = 1, 3, 5, \ldots\), we have

\[
\frac{d\phi_{m0}(t)}{dt} = \sqrt{\xi^{k-1}c_m} \frac{dP_m(2\xi^{k-1}t - 2n + 1)}{dt}
\]

\[
= 2\xi^{k-1} \sqrt{\xi^{k-1}c_m} \left( P_0(2\xi^{k-1}t - 2n + 1) + 5P_2(2\xi^{k-1}t - 2n + 1) + \ldots + (2m-1)P_{m-1}(2\xi^{k-1}t - 2n + 1) \right)
\]

\[
= 2\xi^{k-1} \left( c_m \phi_{m0}^{\xi}(t) + c_m c_2 \phi_{m2}^{\xi}(t) + \ldots + c_m c_{m-1} \phi_{m-1}^{\xi}(t) \right)
\]

\[
= 2\xi^{k-1} \left[ \phi_{m0}^{\xi}(t), \ldots, \phi_{m-1}^{\xi}(t) \right]^T;
\]

for \(m = 2, 4, 6, \ldots\), we get

\[
\frac{d\phi_{m0}(t)}{dt} = \sqrt{\xi^{k-1}c_m} \frac{dP_m(2\xi^{k-1}t - 2n + 1)}{dt}
\]

\[
= 2\xi^{k-1} \sqrt{2\xi^{k-1}c_m} \left( 3P_1(2\xi^{k-1}t - 2n + 1) + 7P_3(2\xi^{k-1}t - 2n + 1) + \ldots + (2m-1)P_{m-1}(2\xi^{k-1}t - 2n + 1) \right)
\]

\[
= 2\xi^{k-1} \left( c_m c_1 \phi_{m1}^{\xi}(t) + c_m c_3 \phi_{m3}^{\xi}(t) + \ldots + c_m c_{m-1} \phi_{m-1}^{\xi}(t) \right)
\]

\[
= 2\xi^{k-1} \left[ 0, \phi_{m1}^{\xi}(t), 0, c_m c_3, 0, c_m c_5, \ldots, 0, 0, \ldots, 0 \right]^T.
\]
As a result
\[
\mathbf{D}_{lw} = 2\xi^{k-1} \cdot \text{blkdiag}\left(\mathbf{D}_{lw}, \mathbf{D}_{lw}, \ldots, \mathbf{D}_{lw}\right).
\]
where when \( M \) is even, \( \mathbf{D}_{lw} \) is the \( M \times M \) matrix of the form
\[
\mathbf{D}_{lw} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\sqrt{3} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & \sqrt{15} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & \sqrt{7} & \sqrt{35} & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \sqrt{(2M-1)} & 0 & \sqrt{5(2M-1)} & 0 & \sqrt{9(2M-1)} & \cdots & \sqrt{(2M-3)(2M-1)} & 0
\end{bmatrix}.
\]
(65)
and, if \( M \) is odd then
\[
\mathbf{D}_{lw} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\sqrt{3} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & \sqrt{15} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & \sqrt{7} & \sqrt{35} & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \sqrt{3(2M-1)} & 0 & \sqrt{7(2M-1)} & 0 & \sqrt{11(2M-1)} & \cdots & \sqrt{(2M-3)(2M-1)} & 0
\end{bmatrix}.
\]
(66)
Hence in the case of \( 1 < \alpha \leq 2 \), we see that for obtaining the Caputo operational matrices of these wavelets, we can use (52)–(55) and (60)–(63) directly or we can use (56), (64) in which \( \mathbf{D}_{lw}^{a-1} \) are obtained for each wavelets from (44)–(47) and (48)–(51) and \( \mathbf{D}_{lw} \) from (57)–(59) and (65)–(67). In a similar manner, we can extend the results for \( \alpha \) of higher degree.

3.4. Linear-quadratic fractional time-delay optimal control, Type II. Consider another type of linear-quadratic fractional time-delay optimal control in which (9)–(11) are the same, but \( 1 < \alpha_1 \leq 2 \). Also we have
\[
\dot{x}(t_f) = x_f,
\]
(68)
in which the required initial condition \( \dot{x}(0) = \dot{x}_0 \) is replaced with the final condition (68).

Since in the method of using derivative operation, we see in (32) that unlike the integral operation there is no arising term about the initial condition, so we must add the given constraint to the QP model of the new system (39), (40).

To handle such systems, we use one of the two options given for constructing the Caputo derivative operational matrix in the case \( 1 < \alpha_1 \leq 2 \), then we make the following changes
\[
\mathbf{W}_C = \begin{bmatrix}
\mathbf{W}_{cc} \\
\mathbf{w}_c^T(0) \\
\mathbf{w}_c^T(1)
\end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix}
0_{q(\xi^{k-1}-1) \times 1} \\
x_0 \\
x_f
\end{bmatrix}, \mathbf{A}_4 = 0_{q(\xi^{k-1}+1) \times r \xi^{k-1} M}.
\]
The other matrices are the same as before and just by these changes we can solve the problem.

3.5. Linear-quadratic fractional time-delay optimal control, Type III. Consider another type of linear-quadratic fractional time-delay optimal control in which (9)–(11) are the same as Type I, but \( 1 < \alpha_1 \leq 2 \) and
\[
\dot{x}(0) = \dot{x}_0.
\]
(69)
To handle such systems, we cannot add the condition (69) directly to the QP model of system. The reason lies in the fractional nature of the state solution.
Assume that \( x_i(t) = a_i^0 + a_i^1 t + \ldots + b_i^1 t^{\beta_1} + b_i^2 t^{\beta_2} + \ldots \), where \( i = 1, 2, \ldots, q \), \( a_i, b_i \in \mathbb{R}, \beta \), are fractional and \( \beta_1 < 1 \). The desired value in using the Caputo derivative is \( \hat{x}_i(0) = a_i^1 \), but the wavelet expansions compute \( \hat{x}_i(0) \) for the equivalent function including the fractional terms like \( b_i^1 t^{\beta_1} \). So we present the following two methods.

### 3.5.1. Method 1

Assume \( C(t) \) is constant. We return to the method based on Riemann–Liouville fractional integration operational matrix. Applying the \((\alpha_1 - 1)\)-integral (the Riemann–Liouville fractional integral) to the left side of (10), by using (4), recalling the operational properties, and considering (56), (64) we can write

\[
\mathcal{D}^{\alpha_1 - 1}\{t_f^{\alpha_1} \mathcal{D}^{\alpha_1} x(t) + t_f^{\alpha_2} C \mathcal{D}^{\alpha_2} x(t)\} = t_f^{\alpha_1} \mathcal{D}^{\alpha_1 - 1} \mathcal{D}^{\alpha_1} x(t) + t_f^{\alpha_2} \mathcal{D}^{\alpha_1 - 1} \mathcal{D}^{\alpha_2} x(t)
\]

where \( \mathcal{D}^{\alpha_1 - 1} \) is the fractional integration operational matrix of the desired wavelets in the Riemann–Liouville sense, \( \mathcal{D}^{\alpha_1 - 1} w(t) \equiv \mathcal{P}^{\alpha_1 - 1} w(t) \), and

\[
\mathcal{I}_w = \left[ \mathcal{I}_w, \mathcal{I}_w, \ldots, \mathcal{I}_w \right]^T, \quad \mathcal{I}_w = \left[ \frac{1}{w_0}, 0, 0, \ldots, 0 \right].
\]  

Hence by making the following changes in (42), (43), we construct another QP for this type

\[
\begin{align*}
\Lambda_1 &= \mathcal{P}^{\alpha_1 - 1}_w \mathcal{D}_w \mathcal{A}_w + \sum_{\mu=1}^a \mathcal{P}^{\alpha_1 - 1}_w \mathcal{E}_{\mu w} D_{\mu}^T - t_f^{\alpha_1} \mathcal{D}_w^T - t_f^{\alpha_2} (\mathcal{P}^{\alpha_1 - 1}_w \otimes C) \mathcal{D}_w^T, \\
\Lambda_2 &= \mathcal{P}^{\alpha_1 - 1}_w \mathcal{D}_w \mathcal{B}_w + \sum_{\nu=1}^b \mathcal{P}^{\alpha_1 - 1}_w \mathcal{F}_{\nu w} D_{\nu},
\end{align*}
\]

(70)

where \( \mathcal{P}^{\alpha_1 - 1}_w \) is the fractional integration operational matrix of the desired wavelets in the Riemann–Liouville sense, \( \mathcal{D}^{\alpha_1 - 1} w(t) \equiv \mathcal{P}^{\alpha_1 - 1} w(t) \), and

\[
\mathcal{I}_w = \left[ \mathcal{I}_w, \mathcal{I}_w, \ldots, \mathcal{I}_w \right]^T, \quad \mathcal{I}_w = \left[ \frac{1}{w_0}, 0, 0, \ldots, 0 \right].
\]

Hence (70) becomes

\[
\begin{align*}
\Lambda_1 &= \mathcal{P}^{\alpha_1 - 1}_w \mathcal{D}_w \mathcal{A}_w + \sum_{\mu=1}^a \mathcal{P}^{\alpha_1 - 1}_w \mathcal{E}_{\mu w} D_{\mu}^T - t_f^{\alpha_1} \mathcal{D}_w^T - t_f^{\alpha_2} (\mathcal{P}^{\alpha_1 - 1}_w \otimes C) \mathcal{D}_w^T, \\
\Lambda_2 &= \mathcal{P}^{\alpha_1 - 1}_w \mathcal{D}_w \mathcal{B}_w + \sum_{\nu=1}^b \mathcal{P}^{\alpha_1 - 1}_w \mathcal{F}_{\nu w} D_{\nu}.
\end{align*}
\]

(71)

Note, too, that \( \mathcal{D}_w(\tau = 0) = t_f \mathcal{D}_w(0) \), for \( \tau := t_f t \).

### 3.5.2. Method 2

First we define

\[
x_N(t) := \mathcal{D}^{1} x(t).
\]

(73)

Hence (10) becomes

\[
\mathcal{D}^{\alpha_1 - 1} x_N(t) + C(t) \mathcal{D}^{\alpha_2} x(t)
\]

\[
= \mathcal{A}(t)x(t) + \mathcal{B}(t)\mathbf{u}(t) + \sum_{\mu=1}^a \mathcal{E}_{\mu w} D_{\mu} x(t - h_{\mu}) + \sum_{\nu=1}^b \mathcal{F}_{\nu w} D_{\nu} \mathcal{u}(t - h_{\nu}) + d(t).
\]

(74)

Now by

\[
\mathbf{z}(t) := \begin{bmatrix} x(t) \\ x_N(t) \end{bmatrix},
\]
in which the first vector $x(t)$ is the state of the original system, we can write (73) and (74) in the form
\[
\begin{bmatrix}
I & 0 \\
0 & 0
\end{bmatrix}
\dot{z}(t) + \begin{bmatrix}
0 & 0 \\
C(t) & 0
\end{bmatrix} \mathcal{D}^{\alpha_1}z(t) + \begin{bmatrix}
0 & 0 \\
A(t) & 0
\end{bmatrix} \mathcal{D}^{\alpha_2}z(t) = \begin{bmatrix}
0 & I \\
0 & 0
\end{bmatrix} z(t)
\]
\[+ \begin{bmatrix}
0 & 0 \\
B(t) & 0
\end{bmatrix} \mathcal{D}^{\alpha_1}u(t) + \sum_{\mu=1}^{n} \begin{bmatrix}
0 & 0 \\
E_i(t) & 0
\end{bmatrix} z(t - h_{\mu}) + \sum_{\nu=1}^{b} \begin{bmatrix}
0 & 0 \\
F_{\nu}(t) & 0
\end{bmatrix} u(t - h_{\nu}) + \begin{bmatrix}
0 & 0 \\
d(t)
\end{bmatrix}.
\] (75)

By setting $t/t_f \rightarrow t$ and defining
\[
\begin{align*}
C_1 & := t_f^{-1} \begin{bmatrix}
I & 0 \\
0 & 0
\end{bmatrix}, \\
A_i(t) & := \begin{bmatrix}
0 & 0 \\
A(t) & 0
\end{bmatrix}, \\
C_N(t) & := t_f^{-\alpha_1} \begin{bmatrix}
0 & 0 \\
0 & C(t)
\end{bmatrix}, \\
A_N(t) & := \begin{bmatrix}
0 & 1 \\
B(t) & 0
\end{bmatrix}, \\
E_N(t) & := \begin{bmatrix}
0 & 0 \\
E_i(t) & 0
\end{bmatrix}, \\
F_{\nu}(t) & := \begin{bmatrix}
0 & 0 \\
F_{\nu}(t) & 0
\end{bmatrix}, \\
d_N(t) & := \begin{bmatrix}
0 & 0 \\
d(t)
\end{bmatrix}, \\
T_N & := \begin{bmatrix}
T & 0 \\
0 & 0
\end{bmatrix}, \\
Q_N(t) & := \begin{bmatrix}
Q(t) & 0 \\
0 & 0
\end{bmatrix}
\end{align*}
\] (76)

in (75) and then using
\[
\begin{align*}
C_N(t) & = C_{Nw}\hat{w}(t), \\
A_N(t) & = A_{Nw}\hat{w}(t), \\
B_N(t) & = B_{Nw}\hat{w}(t), \\
E_{N\mu}(t) & = E_{N\mu w}\hat{w}(t), \\
F_{N\nu}(t) & = F_{N\nu w}\hat{w}(t), \\
d_N(t) & = \hat{w}(t)d_w, \\
Q_N(t) & = Q_{Nw}\hat{w}(t),
\end{align*}
\]
we can model the problem as the QP one with the following elements
\[
\begin{align*}
\Xi_1 & = \mathcal{D}^{\mathcal{T}}(1) \otimes T_N + t_f \hat{F}_w Q_{Nw}, \\
\Xi_2 & = 0_{2q\mathcal{L}^{k-1}M \times 2q\mathcal{L}^{k-1}M}, \\
\Xi_3 & = 0_{\mathcal{L}^{k-1}M \times 2q\mathcal{L}^{k-1}M}, \\
\Xi_4 & = t_f \hat{F}_w R_w,
\end{align*}
\] (77)
\[
\begin{align*}
A_1 & = \hat{A}_{Nw} + \sum_{\mu=1}^{a} \hat{E}_{N\mu w} D_{\mu}^{\mathcal{T}} - (\mathcal{D}_w^{\mathcal{T}} \otimes C_1) - (\mathcal{D}_w^{\alpha_1 - 1 \mathcal{T}} \otimes C_2) - \hat{C}_{Nw} \hat{D}_w^{\mathcal{T}}, \\
A_2 & = B_w + \sum_{\nu=1}^{b} F_{N\nu w} D_{\nu}^{\mathcal{T}}, \\
A_3 & = \hat{w} C, \\
A_4 & = 0_{2q\mathcal{L}^{k-1}M \times 2q\mathcal{L}^{k-1}M},
\end{align*}
\] (78)
\[
\begin{align*}
b_1 & = -d_{Nw} - \sum_{\mu=1}^{a} \hat{E}_{N\mu w} (\theta_{\mu w} \otimes c_0) - \sum_{\nu=1}^{b} \hat{F}_{N\nu w} z_{\nu w}, \\
b_2 & = \begin{bmatrix}
0_{2q,(\mathcal{L}^{k-1} - 1) \times 1} \\
z_0
\end{bmatrix},
\end{align*}
\] (79)
where for a matrix $O, \hat{O} := O \otimes I_{2q}, z(t - h_{\mu}) = \begin{bmatrix} x(t - h_{\mu}) \\ 0 \end{bmatrix}$ and
\[
c_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, z_0 = \begin{bmatrix} x(0) \\ 0 \end{bmatrix}.
\] (80)

In order to solve the system of Type III with this method, first we construct the new components defined in (76), then we form the QP model from (77)–(80).

**Remark 2.** In the QP model given by (77)–(80), we first define a new state vector, that is, $x_N(t) := \mathcal{D}^1 x(t), 0 \leq t \leq t_f$; then we put it in the state equation, form a new state equation, and finally we rescale the time interval. If we first rescale the time interval, we must set $\dot{x}(\tau = 0) = t_f \hat{x}(t = 0)$ and modify any extended component in the QP model.
3.6. Fractional nonlinear time-delay optimal control. Consider a fractional nonlinear time-delay system described by
\[ D^\alpha x(t) = f(x(t), u(t), x(t - h(t)), x(t - h_x), u(t - h_u), t), \]  
(81)
\[ x(0) = x_0, \begin{cases} \dot{x}(t) = \theta(t) & t < 0, \\ u(t) = \zeta(t) & -h_u \leq t < 0, \end{cases} \]  
(82)
where \( f \) is continuous in \((t, x, u)\) and nonlinearity satisfies a local Lipschitz condition, \( h(t) \) is a piecewise constant delay, \( h_x \) and \( h_u \) are delays, \( x(0) = x_0 \) is the initial condition, and \( \theta(t) \in \mathbb{R}^q \) and \( \zeta(t) \in \mathbb{R}^r \) are specified functions. This fractional time-delay system is to be controlled to minimize the quadratic performance index given in (9).

We will use the technique of constructing a sequence of QP problems to handle this type of fractional optimal control which is presented in the following theorem.

**Theorem 3.3.** Consider the nonlinear delay optimal control problem which to minimize (9) subject to (81), (82). 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
\[ J[i] = \frac{1}{2} x[i]^\top(t_1) T x[i](1) + \frac{1}{2} \int_0^{t_f} \left\{ x[i]^\top(t) Q(t) x[i](t) + u[i]^\top(t) R(t) u[i](t) \right\} dt \]  
(83)
subject to
\[ D^\alpha x[i](t) = A[i-1](t)x[i](t) + B[i-1](t)u[i](t) + E[i-1](t)x[i](t - h_x) + F[i-1](t)u[i](t - h_u) + G[i-1](t)x[i](t - h(t)) + d[i-1](t), \]  
(84)
\[ x[i](0) = x_0, \begin{cases} x[i](t) = \theta(t) & t < 0, \\ u[i](t) = \zeta(t) & -h_u \leq t < 0, \end{cases} \]  
(85)
where \( x[0](t) = x_0, u[0](t) = 0, A, E, G \in \mathbb{R}^{q \times q}, B, F \in \mathbb{R}^{q \times r} \) and \( d \in \mathbb{R}^q \).

**Proof.** The proof can be found in [31].

From Theorem 3.3 after rescaling by \( t/t_f \rightarrow t \), by expanding \( A[i-1](t) = A_w[i-1] \bar{w}(t), B[i-1](t) = B_w[i-1] \bar{w}(t), E[i-1](t) = E_w[i-1] \bar{w}(t), F[i-1](t) = F_w[i-1] \bar{w}(t), G[i-1](t) = G_w[i-1] \bar{w}(t), d[i-1](t) = \bar{w}^\top(t) d_w[i-1], \theta(t - h(t)) = \bar{w}^\top(t) \theta_w, \zeta(t - h_u) = \bar{w}^\top(t) \zeta_u, \) using another useful property of the wavelets [27, 28] as
\[ w(t - h(t)) = \begin{cases} 0, & 0 \leq t < h(t) \\ D^t w(t), & h(t) \leq t \leq 1, \end{cases} \]
and the method developed for fractional linear optimal control, we must solve the following sequence of QP problems iteratively instead of the problem described by (83)–(85),
\[ \text{for } i \geq 1, \text{minimize } \frac{1}{2} \begin{bmatrix} x_w[i] \\ u_w[i] \end{bmatrix}^\top \begin{bmatrix} E_1 & E_2 \\ E_3 & E_4 \end{bmatrix} \begin{bmatrix} x_w[i] \\ u_w[i] \end{bmatrix}, \]  
(86)
subject to \[ \begin{bmatrix} A_1[i-1] & B_w[i-1] \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} x_w[i] \\ u_w[i] \end{bmatrix} = \begin{bmatrix} b_1[i-1] \\ b_2 \end{bmatrix}, \]  
(87)
until the conditions given in [31] is reached, where
\[
\begin{align*}
\Xi_1 &= (w(1)w^T(1) \otimes T) + tf_\xi Q_w, \\
\Xi_2 &= 0_{\xi^k \times 1}, \\
\Xi_3 &= 0_{\xi^k \times 1}, \\
\Xi_4 &= tf_\xi R_w, \\
\end{align*}
\]
(88)

\[
\begin{align*}
\Lambda_1^{[i-1]} &= \hat{A}_w^{[i-1]} + \hat{E}_w^{[i-1]} \hat{D}_w^{[i-1]} + \hat{G}_w^{[i-1]} \hat{D}_w^{[i-1]} - tf_\xi^\alpha \hat{D}_w^{[i-1]}, \\
\Lambda_2^{[i-1]} &= \hat{B}_w^{[i-1]} + \hat{F}_w^{[i-1]} \hat{D}_w^{[i-1]}, \\
\Lambda_3 &= \hat{W}_C, \\
\Lambda_4 &= 0_{\xi^k \times 1}, \\
\end{align*}
\]
(89)

\[
\begin{align*}
b_1^{[i-1]} &= -d_w^{[i-1]} - \hat{E}_w^{[i-1]} \bar{\theta}_{wx} - \hat{F}_w^{[i-1]} \bar{\zeta}_{wx} - \hat{G}_w^{[i-1]} \theta_w, \\
b_2 &= \begin{bmatrix} 0_{\xi^k \times 1} \\ x_0 \end{bmatrix}. \\
\end{align*}
\]
(90)

3.7. Fractional optimal control of a system having a performance index with delayed terms. Some texts study the optimal control problem consists of a quadratic performance index with delayed terms, for example see [32, 43]. Assume that for all types of the given systems, we have

\[
J = \frac{1}{2}x^T(t_f)T x(t) + \frac{1}{2} \int_0^{t_f} \{ x^T(t)Q(t)x(t) + u^T(t)R(t)u(t) \\
+ x^T(t-h_1)Q_h(t)x(t-h_1) + u^T(t-h_2)R_h(t)u(t-h_2) \} dt,
\]
(91)

where \( h_1 \in \{ h_w \}, h_2 \in \{ h_w \}, Q_h(t) \in \mathbb{R}^{n \times n} \) is a symmetric, positive semi-definite matrix, \( R_h(t) \in \mathbb{R}^{n \times n} \) is a symmetric, positive definite matrix.

To model this type of performance index, by rescaling, using the given expansions and
\[
Q_h(t) = Q_{hw}(\hat{w}(t)), R_h(t) = R_{hw}(\hat{w}(t)),
\]
we can write

\[
J = \frac{1}{2}x_w^T(w(1)w^T(1) \otimes T) + \hat{\Gamma}_w Q_w x_w + u_w^T R_w u_w \\
+ \frac{1}{2} tf_0^{h_1} \theta^T(t-h_1)Q_h(t)\theta(t-h_1) dt + \frac{1}{2} tf_0^{h_2} x_w^T \hat{D}_1 \hat{w}(t)Q_{hw}(\hat{w}(t)) x_w \hat{D}_1^T x_w dt \\
+ \frac{1}{2} tf_0^{h_2} \zeta^T(t-h_2)R_h(t)\zeta(t-h_2) dt + \frac{1}{2} tf_0^{h_2} u_w^T \hat{D}_2 \hat{w}(t)R_{hw}(\hat{w}(t))^T \hat{D}_2^T u_w dt \\
= c + \frac{1}{2}x_w^T((w(1)w^T(1) \otimes T) + tf_\xi Q_w + t f_\xi \hat{D}_1^{h_1-1} \hat{Q}_{hw} \hat{D}_1^T)x_w \\
+ \frac{1}{2} tf_0^{h_1} \theta^T(t-h_1)Q_h(t)\theta(t-h_1) dt + \frac{1}{2} tf_0^{h_2} \zeta^T(t-h_2)R_h(t)\zeta(t-h_2) dt
\]
(92)

where
\[
c := \frac{1}{2} tf_0^{h_1} \theta^T(t-h_1)Q_h(t)\theta(t-h_1) dt + \frac{1}{2} tf_0^{h_2} \zeta^T(t-h_2)R_h(t)\zeta(t-h_2) dt
\]
and
\[
\Gamma_w^{h_1} := \int_{t_1}^{t_2} w(t)w^T(t) dt.
\]

From the property of the entries of delay operational matrices, we see that
\[
\hat{D}_1 \Gamma_w^{h_1} = \hat{D}_1 \Gamma_w, \hat{D}_2 \Gamma_w^{h_2-1} = \hat{D}_2 \Gamma_w.
\]
Hence we model (91) as

\[ J = c + \frac{1}{2} \begin{bmatrix} x_w \\ u_w \end{bmatrix}^\top \]

this matrix must be used in the quadprog solver

\[
\begin{bmatrix}
(w(1)w^\top(1) \otimes T) + t_f \Gamma_w Q_w + t_f D_1 \Gamma_w Q_{hw} D_1^\top & 0_{q^k-1M \times q^k-1M} \\
0_{q^k-1M \times q^k-1M} & t_f \Gamma_w R_w + t_f D_2 \Gamma_w R_{hw} D_2^\top
\end{bmatrix}
\begin{bmatrix}
x_w \\
u_w
\end{bmatrix},
\]

(93)

To handle such problems, our task is to construct the matrix indicated in (93) instead of the previous one in (41), (88) or with some modifications in (77). The value of \( c \) can be obtained by integration from (92) and for polynomial functions (knowing \( \theta_1^\top \Gamma_w \theta_1 = \theta_{1w}^\top \Gamma_w \theta_{1w} \) and \( \zeta_2^\top \Gamma_w \zeta_2 = \zeta_{2w}^\top \Gamma_w \zeta_{2w} \)) is obtained exactly from

\[ c = \frac{t_f}{2} [\theta_{1w}^\top \Gamma_w \theta_{1w} + \zeta_{2w}^\top \Gamma_w \zeta_{2w}]. \]

The optimal cost for a system with this kind of the performance index is a cost as an output of the quadprog solver plus the value of \( c \).

4. Illustrative Examples. In this section we use the Caputo derivative operational matrices of Legendre and Chebyshev wavelets to find solutions of some fractional/integer order linear and nonlinear systems. By doing so we verify the accuracy and applicability of the method.

4.1. Example 1. The problem is finding the optimal control for the system described by the equation with retarded restoring force [2]

\[ \dddot{x}(t) + \dddot{x}(t) + x(t - 1) = u(t), \quad 0 \leq t \leq 2 \]

\[ x(t) = 1, \quad -1 \leq t \leq 0, \]

which minimizes the performance index

\[ J = \frac{1}{2} \{ x^2(2) + \dot{x}^2(2) \} + \frac{1}{2} \int_0^2 u^2(t) \, dt, \]

where \( x(0) = 1 \) and \( \dot{x}(0) = 0 \). The exact optimal control for this problem is given in [2].

We can handle this problem as Type III and Type I. For the both types with the use of minimum values of \( k \) and \( \xi \) in CW method, we solve the problem and the results are reported in Table 1. We increase the accuracy of the method by increasing \( \xi \) or \( k \). We do this for LW method and the results are given in this table.

Also for some \( \alpha \) in the new state equation as

\[ D^\alpha x(t) + \dot{x}(t) + x(t - 1) = u(t), \quad 1.9 \leq \alpha \leq 2 \]

(94)

the optimal state and control are shown in Figure 1. In addition, the values of the optimal cost \( J^* \) are given in Table 2. In applying Method 1, we have \( \Xi_1 = w(1)w^\top(1) + t_f^\alpha (D_w w(1))(D_w w(1))^\top \).
Table 1. $u^*$ for some $t$ in Example 1, $k = 2$.

| $t$  | Exact | CW, Type III, $M = 7$ | CW, Type III, $M = 7$ | CW, Type I, $M = 7$ | LW, Type I, $M = 7$ |
|------|-------|----------------------|----------------------|----------------------|----------------------|
| 0    | $-0.0870988$ | $-0.08655811$ | $-0.0870999$ | $-0.0870999$ | $-0.0870982$ |
| 0.2  | $-0.036738$  | $-0.0335485$  | $-0.0336745$  | $-0.0336745$  | $-0.0336737$  |
| 0.4  | 0.0218134  | 0.0219177  | 0.0218149  | 0.0218149  | 0.0218134  |
| 0.6  | 0.0774030  | 0.0773811  | 0.0774012  | 0.0774012  | 0.0774030  |
| 0.8  | 0.1301664  | 0.1298957  | 0.1301676  | 0.1301676  | 0.1301664  |
| 1    | 0.1758728  | 0.1757065  | 0.1758696  | 0.1758696  | 0.1758731  |
| 1.2  | 0.2205516  | 0.2079875  | 0.2205631  | 0.2205631  | 0.2205505  |
| 1.4  | 0.2751223  | 0.2785798  | 0.2751004  | 0.2751004  | 0.2751218  |
| 1.6  | 0.3417751  | 0.3464972  | 0.3417959  | 0.3417959  | 0.3417753  |
| 1.8  | 0.4231851  | 0.4093533  | 0.4231753  | 0.4231753  | 0.4231855  |
| 2    | 0.5226194  | 0.5872923  | 0.5226835  | 0.5226835  | 0.5226206  |

Figure 1. Numerical solutions for Example 1.

Table 2. $J^*$ for some $\alpha$ in (94), Example 1, $k = 2$.

| $\alpha$ | CW, Type I, $M = 7$ | LW, Type I, $M = 7$ | LW, Type I, $M = 8$ |
|----------|----------------------|----------------------|----------------------|
| 2        | 0.1974785            | 0.1974785            | 0.1974785            |
| 1.99     | 0.1934399            | 0.1934481            | 0.1934459            |
| 1.98     | 0.1894355            | 0.1894379            | 0.1894288            |
| 1.97     | 0.1854667            | 0.1854495            | 0.1854282            |
| 1.96     | 0.1815345            | 0.1814848            | 0.1814454            |
| 1.95     | 0.1776403            | 0.1775456            | 0.1774818            |
| 1.94     | 0.1737852            | 0.1736341            | 0.1735389            |
| 1.93     | 0.1699705            | 0.1697526            | 0.1696184            |
| 1.92     | 0.1661976            | 0.1659034            | 0.1657223            |
| 1.91     | 0.1624677            | 0.1620891            | 0.1618525            |
| 1.9      | 0.1587825            | 0.1583125            | 0.1580112            |

4.2. Example 2. The problem is minimizing

$$J = \int_0^2 \{ x^2(t) + u^2(t) \} \, dt,$$
subject to
\[ D^\alpha x(t) = tx(t) + x(t - 1) + u(t) + a(0.2 - 0.15 \cos t), \quad 0 \leq t \leq 2 \]
\[ \theta(t) = 1, \quad -1 \leq t < 0, \]
where \( 0.9 \leq \alpha \leq 1.3 \), \( x(0) = 1 \), \( \dot{x}(0) = 0.05 \) (as an initial condition only needed for the case \( 1 < \alpha \leq 1.3 \)), and \( a = 0.1 \).

For \( 0.9 \leq \alpha \leq 1 \), we use the method given for the optimization problem of Type I. The results of the optimal cost \( J^* \) are given in Table 3. For \( 1 < \alpha \leq 1.3 \), we use the two method given for Type III and the results are given in this table. In Method 1, we construct a new QP model by using (70)–(72), where \( q = 1 \) and \( \dot{x}(0) = 0.1 \). But in Method 2, we form another new QP problem from (76)–(80), where \( q = 2 \) and \( \dot{x}(0) = 0.05 \). We see for this value of \( \alpha \), the results of Method 1 are similar to those obtained by using the Riemann–Liouville integral operational matrix in the previous work and this is probably expected because we have used the fractional integral operator. For all values of \( \alpha \), we have \( x^*(0) = 1 \). In addition, one can see that the results for the case \( \alpha = 1 \) by which we have an integer order optimal control problem, are in excellent agreement with those obtained by the previous methods whose accuracies had been verified.

**Table 3. Comparison of \( J^* \) for Example 2, \( k = 2 \).**

| \( a \) | \( \alpha \) | \( \xi = 4, M = 8 \) | \( \xi = 2, M = 7 \) | \( [29], \xi = 2, M = 7 \) | \( [27], \xi = 2, M = 7 \) |
|-------|-------|----------------|----------------|----------------|----------------|
| 0     | 0.999 | 4.79679791916 | 4.79679791916 | 4.79679791913 | 4.79679791913 |
| 0     | 0.99  | 4.79728438115 | 4.79853220259 | 4.8137758646 | 4.8137758646 |
| 0     | 0.9   | 4.79931928713 | 4.80544625813 | 5.26128568218 | 5.26128568218 |
| 1     | 0.9   | 5.26128568218 | 5.24909761641 | 5.25573149155 | 5.25573149155 |
| 1     | 0.99  | 5.23983310060 | 5.24118253395 | 5.23778132890 | 5.23778132890 |
| 0     | 1     | 5.23778132890 | 5.23778132890 | 5.23778132890 | 5.23778132890 |
| 1     | 0.999 | 5.23755370619 | 5.23755370619 | 5.23755370619 | 5.23755370619 |
| 1     | 1.001 | 5.53898632105 | 5.49646031081 | 5.49646031081 | 5.49646031081 |
| 1     | 1.01  | 5.53527054351 | 5.49452438269 | 5.49452438269 | 5.49452438269 |
| 1     | 1.1   | 5.49587484403 | 5.46836520817 | 5.46836520817 | 5.46836520817 |
| 1     | 1.2   | 5.43737670256 | 5.42892272878 | 5.42892272878 | 5.42892272878 |
| 1     | 1.3   | 5.39396724356 | 5.38170076246 | 5.38170076246 | 5.38170076246 |

4.3. Example 3. Minimize
\[ J = \frac{1}{2} \int_0^1 \left\{ x^T(t) \begin{bmatrix} 1 & t \\ t & t^2 \end{bmatrix} x(t) + (t^2 + 1)u^2(t) \right\} dt, \]
subject to
\[ D^\alpha x(t) - a \begin{bmatrix} -2.7 & 3.1 \\ 0 & 0.08 \end{bmatrix} D^\alpha x(t) = \begin{bmatrix} t^2 + 1 & 1 \\ 0 & 2 \end{bmatrix} x(t - \frac{1}{2}) + \begin{bmatrix} 1 \\ t + 1 \end{bmatrix} u(t) \]
\[ ^2 \text{Type III, Method 2.} \]
\[ ^3 \text{For } \alpha \leq 1 \text{ by Type I and for } \alpha > 1 \text{ by Type III, Method 1.} \]
\[ \theta(t) = [1, 1]^T, \quad -\frac{1}{4} < t < 0 \]
\[ \zeta(t) = 1, \quad -\frac{1}{4} < t < 0 \]

where \( \alpha_1 = 1, \alpha_2 = 0 \) and \( 0.8 < \alpha_2 < 1 \).

By solving this problem with the use of the two methods of Caputo operational matrices, the results are given in Table 4. Although the QP models of this study have some differences with that obtained in [29], we see that the results are in good agreement with those were presented previously by using the Riemann–Liouville fractional integral operational matrix of Chebyshev wavelets.

Table 4. \( J^* \) for Example 3.

| \( a \) | \( \alpha_1 \) | \( \alpha_2 \) | This work, LW | This work, CW | [29] |
|---|---|---|---|---|---|
| 0 | 1 | 1.56224137366 | 1.562241373755 | 1.56224137355 |
| 1 | 1 | 0.999 | 1.41013313297 | 1.41013507255 | 1.41013747158 |
| 1 | 1 | 0.99 | 1.41080207527 | 1.41094921928 | 1.41106062244 |
| 1 | 1 | 0.95 | 1.41374036510 | 1.41394156982 | 1.41378641071 |
| 1 | 1 | 0.91 | 1.41655889207 | 1.41695140257 | 1.41668866306 |
| 1 | 1 | 0.9 | 1.41729512016 | 1.41766964017 | 1.41740062973 |
| 1 | 1 | 0.8 | 1.42440244581 | 1.42467128854 | 1.42442691113 |

4.4. Example 4. Minimize

\[ J = \frac{1}{2}x^2(2) + \frac{1}{2} \int_0^2 \{ x^2(t) + u^2(t) \} \, dt \]

subject to

\[ D^\alpha x(t) = x(t) \sin(x(t)) + x(t-1) + u(t), \quad 0 \leq t \leq 2 \]
\[ \theta(t) = \begin{cases} 
10(t+1), & -1 < t < -0.5 \\
-10t, & -0.5 < t \leq 0 
\end{cases} \]

Also assume that

Case 1: the control and state are unconstrained.
Case 2: the path constraint is \( x(t) + u(t) \geq 1, \quad 0 \leq t \leq 2 \).

Case 1 of this problem was studied in [31] by using the Riemann–Liouville integral operational matrix which the main problem was replaced with a sequence of linear problems. We see the proposed general framework by using the Caputo derivative operational matrices of the both wavelets, converges again to a solution and the results of \( J^* \) are reported in Table 5. Also, for Case 2, we show the optimal solutions obtained by the method in Figure 2.

4.5. Example 5. This example was investigated in [32] in the case \( \alpha = 1 \). The nonlinear time-varying optimal control problem with multiple delays described by

\[ D^\alpha x(t) = tx^2(t-\frac{1}{3}) + x^3(t-\frac{2}{3}) + u(t) + u(t-\frac{1}{3}) + u(t-\frac{2}{3}), \quad 0 \leq t \leq 1 \]
\[ x(t) = 1, \quad u(t) = 0, \quad -\frac{2}{3} \leq t \leq 0 \]

is to be controlled to minimize

\[ J = \frac{1}{2} \int_0^1 \{ u^2(t) + x^2(t-\frac{1}{3}) + u^2(t-\frac{2}{3}) \} \, dt. \]
Table 5. $J^*$ for Case 1 of Example 4.

| $\alpha$ | This work, LW | This work, CW | [31]       |
|----------|---------------|---------------|------------|
| 1        | 2.54807       | 2.54807       | 2.54807    |
| 0.999    | 2.54887       | 2.54884       | 2.54887    |
| 0.99     | 2.55613       | 2.55586       | 2.55616    |
| 0.95     | 2.59051       | 2.58935       | 2.59081    |
| 0.91     | 2.62858       | 2.62681       | 2.62918    |
| 0.9      | 2.63869       | 2.63682       | 2.63937    |
| 0.8      | 2.75391       | 2.75189       | 2.75520    |

Figure 2. $x^*(t)$ and $u^*(t)$ for Case 2 of Example 4.

It should be noted the performance index consists of constant term and the term which meets the specifications of the system. As mentioned in Section 3.7, we have $J^* = c + J^*_{QP}$. We replace the problem with the sequence of QP problems given in (86), (87), in which $\Xi_1$ and $\Xi_2$ are constructed from (93) and other components are the same as those given in (88)–(90). The computed results for values of the optimal costs are reported in Table 6. Also for $\alpha = 1$, in order to make a comparison with the solutions given in [32], we present the optimal solutions with grid lines in Figure 3 in which $i_{max}$ denotes the maximum number of iterations for both of our methods. In both methods we have $|J^{[i_{max}]} - J^{[i_{max} - 1]}| < 2.38E-10$. We used 0.000001 as the step size of the $t$-axis to plot the solutions, but since the data points are very close, there is no distinction between the two curves.

Figure 3. $x^*(t)$ and $u^*(t)$ for Example 5, $\alpha = 1$. 
The performance index is nonlinear functions for describing the system uncertainties and disturbances. The application of time-delay optimal control theory as a cascade chemical system with two reactors arises in the industry, see [14]. The plant is described by Example 5; for both wavelets, we set $k = 2, \xi = 3, M = 8.$

| $\alpha$ | $J^*_{QP}$ | $J^*$ | $J^*_{QP}$ | $J^*$ | 32 |
|---|---|---|---|---|---|
| 1 | 0.59232919871 | 0.758986537 | 0.592319871 | 0.758986537 | 0.833609761 |
| 0.999 | 0.592609998 | 0.759276665 | 0.592610281 | 0.759276948 | |
| 0.99 | 0.595224540 | 0.761891207 | 0.595227150 | 0.761893817 | |
| 0.98 | 0.598136364 | 0.764803031 | 0.598141037 | 0.764807703 | |
| 0.97 | 0.601054736 | 0.767721403 | 0.601060817 | 0.767727484 | |
| 0.96 | 0.603979032 | 0.770645699 | 0.603985765 | 0.770652432 | |
| 0.95 | 0.606908605 | 0.773575272 | 0.606915143 | 0.773581809 | |
| 0.94 | 0.609842787 | 0.776509454 | 0.609848197 | 0.776514863 | |
| 0.93 | 0.612780884 | 0.779447551 | 0.612784163 | 0.779450830 | |
| 0.92 | 0.615722181 | 0.782388847 | 0.615722265 | 0.78238932 | |
| 0.91 | 0.618665935 | 0.785332602 | 0.618661716 | 0.785328382 | |
| 0.9 | 0.621611381 | 0.788278047 | 0.621601717 | 0.788268384 | |
| 0.8 | 0.650973598 | 0.817640265 | 0.650850308 | 0.817516974 | |
| 0.7 | 0.679517928 | 0.846184595 | 0.679227593 | 0.845894260 | |
| 0.6 | 0.706304327 | 0.872900994 | 0.705895247 | 0.872561914 | |

4.6. Example 6. Now we consider a complex nonlinear system as a real-world application of time-delay optimal control theory as a cascade chemical system with two reactors arises in the industry, see [14]. The plant is described by

$$
\mathcal{D}^\alpha x_1(t) = -k_1 x_1(t) - \frac{1}{\theta_1} x_1(t) - \frac{1}{\theta_1} x_1(t - h_1) + \frac{1 - R_2}{V_1} x_2(t) + \delta_1(t, x_1(t - h_1)),
$$

$$
\mathcal{D}^\alpha x_2(t) = -k_2 x_2(t) - \frac{1}{\theta_2} x_2(t) + \frac{R_1}{V_2} x_1(t - h_1) - \frac{1}{\theta_2} x_2(t) + \frac{R_2}{V_2} x_2(t - h_2)
$$

$$
+ \frac{F}{V_2} u(t) + \delta_2(t, x_2(t - h_2)),
$$

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

$$
J = \int_0^t \{ x^T(t) x(t) + 0.01 u^2(t) \} dt.
$$

We have $R_1 = 0.5, k_1 = 0.5, F = 0.5, V_1 = 0.5, \delta_1(t, x_1(t - h_1)) = \theta_3 x_1(t - h_1), h_i = 0.25, t_f = 5, \delta_2(x_2(t - h_2)) = 0.5 \theta_4 x_2^2(t - h_2) e^{0.01 x_2(t - h_2)}, \quad \Theta = [8, -8]^T, \quad \theta_1 = \theta_2 = 2 \quad \text{and} \quad \theta_3 = \theta_4 = 1;$ the system is constrained by the following constraint

$$
x_2^2(t) - x_2^2(t) + u(t) \leq 0, \quad 0.25 \leq t \leq 5.
$$

For more details, see [31].

We take $\epsilon = 0.00000000001$ and use the iterative method until we find $|J[i+1] - J[i]| \leq \epsilon.$ We solve the problem with the use of the method given for the both wavelets family and the result are given in Table 7. As can be seen, we reach the desired accuracy by a reasonable number of iterations $i_{max}$. It is usually very difficult or even impossible to obtain analytically the optimal control law of a nonlinear system.
fractional system. Here we present a suboptimal control law. Assuming the suboptimal control law as a function of the state and the delayed state vectors, the values of (negative) suboptimal feedback gains are given in Table 7, where $K_x$ is designed for the state vector in $[0, 5]$ and $K_{x_h}$ for the delayed state vector in $[0, 25, 5]$. The obtained optimal controls from the methods of Chebyshev and Legendre wavelets are given in Figure 4.

**Table 7.** Comparison of results for Example 6.

| Method     | $t_{\text{max}}$ | $J^*$     | $K_x$              | $K_{x_h}$            |
|------------|-------------------|-----------|--------------------|----------------------|
| [31]       | 30                | 46.190119 | 1.114775 11.230130 | [8.331317 1.695857]  |
| This work, CW | 29               | 46.190067 | 1.144505 11.220603 | [8.297838 1.589278]  |
| This work, LW | 29              | 46.190014 | 1.144826 11.220515 | [8.297552 1.589329]  |

**Figure 4.** Optimal control for Example 6.

**Conclusion.** The Caputo derivative operational matrix of Legendre and Chebyshev wavelets have been presented for the first time. By using the direct method, these operational matrices have been applied for the optimal control systems having multifractional orders or integer orders state equations. Different types of control systems have been considered. The direct method given in this study can be applied to the mentioned problems by using other wavelets.

**Pros and Cons of the proposed method:**

- This method is more applicable to multifractional systems compared to the method of using the integral operator method.
- The method has the great advantage over the previous method in Type II systems. Since the both methods are based on QP, we can apply them on the constrained (fractional or integer order) quadratic optimal control systems. Similarly, the method can be applied on fractional optimal control of delay-free linear and nonlinear systems with quadratic performance indices.
- We observe a drawback of the method for systems of Type III.
- The method by using the integral operator is more robust in handling some problems.

**Acknowledgement.** I would also like to thank the Editor and Reviewers for their suggestions and comments which helped to improve the quality of the article.
REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications: New York, NY, USA, 1972.

[2] H. T. Banks and J. A. Burns, *Hereditary control problem: numerical methods based on averaging approximations*, SIAM Journal on Control and Optimization, 16 (1978), 169–208.

[3] M. E. Benattia and B. Kacem, Numerical solution for solving fractional differential equations using shifted Chebyshev wavelet, *General Letters in Mathematics*, 3 (2017), 101–110.

[4] M. Caputo and F. Mainardi, *A new dissipation model based on memory mechanism*, Pure and Applied Geophysics, 91 (1971), 134–147.

[5] P. Chang and A. Isah, Legendre wavelet operational matrix of fractional derivative through wavelet-polynomial transformation and its applications in solving fractional order Brusselator system, *Journal of Physics: Conference Series*, 693 (2016), IOP Publishing.

[6] S.-B. Chen, S. Soradi-Zeidi, H. Jahanhashi, R. Alcaraz, J. F. Gómez-Aguilar, S. Bekiros and Y.-M. Chu, Optimal Control of time-delay fractional equations via a joint application of radial basis functions and collocation method, *Entropy*, 22 (2020), 1213.

[7] K. B. Datta and B. M. Mohan, *Orthogonal functions in systems and control*, Advanced Series in Electrical and Computer Engineering, World Scientific Publishing Co., (1995).

[8] K. Diethelm and A. D. Freed, On the solution of nonlinear fractional differential equations used in the modeling of viscoelasticity, in *Scientific Computing in Chemical Engineering II* (eds. F. Keil, W. Mackens, H. Voß, J. Werther), Springer, Heidelberg, (1999), 217–224.

[9] L. Gaul, P. Klein and S. Kempfle, Damping description involving fractional operators, *Mechanical Systems and Signal Processing*, 5 (1991), 81–88.

[10] W. G. Glöckle and T. F. Nonnenmacher, A fractional calculus approach to self-similar protein, *Biophysical Journal*, 68 (1995), 46–53.

[11] A. Graps, An introduction to wavelets, *IEEE Computational Science and Engineering*, 2 (1995), 50–61.

[12] R. W. Hamming, *Numerical Methods for Scientists and Engineers*, McGraw Hill Book Company, Inc., ISBN 0-486-65241-6, USA, 1962.

[13] S. Hosseinpour, A. Nazemi and E. Tohidi, Müntz-Legendre spectral collocation method for solving delay fractional optimal control problems, *Journal of Computational and Applied Mathematics*, 351 (2019), 344–363.

[14] C. Hua, P. X. Liu and X. Guan, Backstepping control for nonlinear systems with time delays and applications to chemical reactor systems, *IEEE Transactions on Industrial Electronics*, 56 (2009), 3723–3732.

[15] A. Isah and P. Chang, *Chebyshev Wavelet Operational Matrix of Fractional Derivative Through Wavelet-Polynomial Transformation and Its Applications on Fractional Order Differential Equations*, Proceedings of the International Conference on Computing, Mathematics and Statistics, Springer, Singapore, 2017.

[16] M. Ishteva, Properties and Applications of the Caputo Fractional Operator, Department of Mathematics, University of Karlsruhe, Karlsruhe 5, 2005.

[17] M. A. Johnson and F.C. Moon, Experimental characterization of quasi periopdicity and chaos in a mechanical system with delay, *International Journal of Bifurcation and Chaos*, 9 (1999), 49–65.

[18] D. E. Kirk, *Optimal Control Theory: An Introduction*, Courier Corporation, 2004.

[19] K. C. Kiwiel, A dual method for certain positive semidefinite quadratic programming problems, *SIAM Journal on Scientific and Statistical Computing*, 10 (1989), 175–186.

[20] D. Kraft, A Software Package for Sequential Quadratic Programming, Wiss. Berichtswesen d. DFVLR, 1988.

[21] L. Li and J.-G. Liu, A generalized definition of Caputo derivatives and its application to fractional ODEs, *SIAM Journal on Mathematical Analysis*, 50 (2018), 2867–2900.

[22] W. Li, S. Wang and V. Rehbock, A 2nd-order one-step numerical integration scheme for a fractional differential equation, *Numerical Algebra, Control and Optimization*, 7 (2017), 273–287.

[23] W. Li, S. Wang and V. Rehbock, Numerical solution of fractional optimal control, *Journal of Optimization Theory and Applications*, 180 (2019), 556–573.

[24] M. Malek-Zavarei and M. Jamshidi, *Time-Delay Systems: Analysis, Optimization and Applications*, Elsevier Science Inc., North-Holland, 1978.
I. Malmir, Optimal control of linear time-varying systems with state and input delays by Chebyshev wavelets, *Statistics, Optimization & Information Computing, 5* (2017), 302–324.

I. Malmir, A novel wavelet-based optimal linear quadratic tracker for time-varying systems with multiple delays, preprint, arXiv:1802.05618.

I. Malmir, Novel Chebyshev wavelets algorithms for optimal control and analysis of general linear delay models, *Applied Mathematical Modelling, 69* (2019), 621–647.

I. Malmir, Legendre wavelets with scaling in time-delay systems, *Statistics, Optimization & Information Computing, 7* (2019), 235–253.

I. Malmir, A new fractional integration operational matrix of chebyshev wavelets in fractional delay systems, *Fractal and Fractional, 3* (2019), 46.

I. Malmir and S. H. Sadati, Transforming linear time-varying optimal control problems with quadratic criteria into quadratic programming ones via wavelets, *Journal of Applied Analysis, 26* (2020), 131–151.

I. Malmir, A general framework for optimal control of fractional nonlinear delay systems by wavelets, *Statistics, Optimization & Information Computing, 8* (2020), 858–875.

H. R. Marzban and H. Pirmoradian, A novel approach for the numerical investigation of optimal control problems containing multiple delays, *Optimal Control Applications and Method, 2017*, 1–24.

H. R. Marzban, Solution of a specific class of nonlinear fractional optimal control problems including multiple delays, *Optimal Control Applications and Method, 2020*, 1–28.

J. C. Mason and D. C. Handscomb, *Chebyshev Polynomials*, Chapman & Hall/CRC, Boca Raton, 2003.

K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, John Wiley & Sons, Hoboken, NJ, USA, 1993.

D. S. Naidu, *Optimal Control Systems*, Idaho State University, Pocatello, Idaho, USA, CRC PRESS, 2003.

J. Nocedal and S. J. Wright, *Sequential Quadratic Programming*, Numerical optimization, (2006), 529–562.

I. Podlubny, *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*, Elsevier, 1998.

E. Safaie, M. Farahi and M. Farmani-Ardehaie, An approximate method for numerically solving multi-dimensional delay fractional optimal control problems by Bernstein polynomials, *Computational and Applied Mathematics, 34* (2015), 831–846.

A. Secer and S. Altun, A new operational matrix of fractional derivatives to solve systems of fractional differential equations via Legendre wavelets, *Mathematics, 6* (2018), 238.

R. Serban and L. R. Petzold, COOPT-a software package for optimal control of large-scale differential-algebraic equation systems, *Mathematics and Computers in Simulation, 56* (2001), 187–203.

Y. Wang and Y. Chen, Shifted Legendre polynomials algorithm used for the dynamic analysis of viscoelastic pipes conveying fluid with variable fractional order model, *Applied Mathematical Modelling, 81* (2020), 159–176.

Y. Zhou and Z. Wang, Optimal feedback control for linear systems with input delays revisited, *Journal of Optimization Theory and Applications, 2014*, 989–1017.

Received August 2020; 1st revision January 2021; Final revision March 2021.

E-mail address: iman.malmir@outlook.com