



Research article

Variable-order Caputo fractional optimal control problems: optimality conditions and a forward–backward numerical method

Huaqing Ma, Wenrui Wang and Changyuan Chen*

School of Qilu Transportation, Shandong University, Jinan 250002, China

* **Correspondence:** Email: changyuan.chen@sdu.edu.cn.

Abstract: This paper investigates an optimal control problem governed by variable-order Caputo fractional dynamics. The main difficulty arises from the fact that, for a type-I variable-order Caputo derivative, the corresponding adjoint operator is not obtained by a direct replacement of the constant fractional order in the standard right-sided Riemann–Liouville derivative. Instead, it must be derived from the dual kernel induced by the variable-order Caputo operator. Based on this observation, we formulate a right-sided variable-order adjoint operator and establish the associated integration-by-parts formula. Under suitable regularity assumptions, first-order necessary optimality conditions are derived for a class of quadratic optimal control problems with nonlinear variable-order fractional dynamics. The resulting optimality system consists of a forward variable-order Caputo state equation, a backward adjoint equation involving the dual-kernel operator, a stationarity condition, and a natural transversality condition. A predictor–corrector forward–backward sweep method is then developed for the numerical approximation of the coupled system. The adjoint equation is treated by a backward Volterra-type approximation motivated by the fractional memory structure. Numerical examples, including linear-quadratic, nonlinear scalar, and coupled nonlinear systems, illustrate the feasibility and mesh-refinement behavior of the proposed approach.

Keywords: variable-order fractional calculus; Caputo derivative; fractional optimal control; adjoint equation; integration by parts; forward–backward sweep method; predictor–corrector method

1. Introduction

Fractional calculus has become a powerful mathematical framework for modeling memory-dependent and hereditary phenomena. Compared with integer-order differential equations, fractional differential equations are able to describe nonlocal temporal effects and long-range memory by means of convolution-type kernels. Fractional models have been widely used in viscoelasticity, diffusion processes, anomalous transport, biological systems, signal processing, and control theory.

Classical references on fractional calculus and fractional differential equations include [1–5].

In parallel, numerical modeling of complex dynamical systems has continued to advance in various applied fields. For example, improved computational fluid dynamics–discrete element method–volume of fluid (CFD–DEM–VOF) models have been used to simulate wet milling processes [6], while Lagrangian mesh-free multiphase models have been developed for complex bubbly-flow dynamics [7]. These studies illustrate the broad demand for accurate and efficient computational approaches, which also motivates the development of numerical methods for fractional and variable-order dynamical systems.

Among fractional derivatives, the Caputo derivative is particularly important in applications because it allows the use of classical initial conditions. For a constant order $\alpha \in (0, 1)$, the Caputo derivative is defined by

$${}^c D_{a+}^{\alpha} y(t) = \frac{1}{\Gamma(1-\alpha)} \int_a^t (t-s)^{-\alpha} y'(s) ds.$$

In many applications, however, the memory intensity is not fixed during the whole evolution process. For example, material properties may change with time due to aging, thermal effects, or damage accumulation; biological systems may display stage-dependent memory; and controlled processes may exhibit different memory mechanisms at different time intervals. These observations motivate the use of variable-order fractional derivatives, where the fractional order is a function of time or space.

Variable-order fractional calculus was introduced to describe memory processes whose order changes during the evolution. Related models have been studied in anomalous diffusion, wave propagation, optimal control, numerical analysis, and inverse problems; see, for instance, [8–12]. In the variable-order setting, one usually replaces the constant order α by an order function $\alpha(t)$ satisfying

$$0 < \alpha_- \leq \alpha(t) \leq \alpha_+ < 1.$$

This seemingly simple generalization leads to substantial analytical and numerical difficulties. For example, the semigroup property of fractional integrals generally fails, the kernel depends on the observation time, and the adjoint equation no longer has the same simple form as in the constant-order case.

A substantial body of research has been devoted to the analysis and numerical approximation of variable-order fractional differential equations. For variable-order time-fractional diffusion equations, optimal-order finite element error estimates without imposing regularity assumptions on the exact solution were established in [13]. These studies demonstrate that the changing fractional order affects not only the underlying memory mechanism but also solution regularity, stability analysis, and the construction of accurate numerical methods.

In recent years, variable-order fractional optimal control problems have attracted increasing attention. Theoretical and numerical methods have subsequently been investigated for variable-order diffusion control, delay systems, distributed-order models, and multidimensional optimal control problems [14–19]. These contributions demonstrate the importance of variable-order models in describing controlled systems with time-dependent memory.

Fractional optimal control has also been extended to collective and transport models. In particular, Zheng et al. developed a time-fractional mean field control model for subdiffusive advective transport [20]. Although that model uses a time-fractional operator rather than the type-I variable-order Caputo operator considered here, it illustrates how nonlocal memory can be

incorporated into the optimal control of anomalous transport processes. More broadly, fractional and variable-order optimal control models have potential applications in engineering, economics, data science, transport processes, and biomedical systems. Time-varying memory effects may appear in viscoelastic materials with aging properties, anomalous diffusion in heterogeneous media, macroeconomic systems with changing memory intensity, fractional neural networks, drug delivery optimization, and pharmacokinetic models.

Despite these advances, the continuous-level derivation of optimality conditions for variable-order Caputo systems remains delicate. In particular, the adjoint operator associated with the type-I variable-order Caputo derivative is not obtained by a direct formal replacement of the constant fractional order by a time-dependent one. Since the kernel depends on the observation time, the correct right-sided adjoint operator must be derived from the exact duality pairing. This distinction is essential because different definitions of variable-order fractional derivatives generally induce different dual kernels and, consequently, different adjoint equations.

Optimal control theory seeks a control input that optimizes a prescribed performance index subject to a dynamical system. Classical optimal control has been extensively studied in [21–24]. Fractional optimal control problems have received increasing attention in recent decades because many controlled systems exhibit memory effects that are more accurately described by fractional dynamics. Necessary optimality conditions and numerical methods for fractional optimal control problems can be found in [25–31].

Related optimization and variational control problems have also been investigated in the context of generalized convexity, variational inequalities, and fractional variational models. These studies show that convexity, invexity, pseudo-convexity, quasi-convexity, and related generalized convexity notions play an important role in deriving sufficient optimality conditions, duality results, and global optimality criteria. Such ideas are particularly relevant when one seeks sufficient conditions beyond first-order necessary optimality systems. In the present work, however, we focus on the derivation of necessary optimality conditions for variable-order Caputo systems. Extensions to sufficient optimality conditions under generalized convexity assumptions will be considered in future work.

The aim of this paper is to develop a rigorous formulation for variable-order Caputo fractional optimal control problems. We consider a quadratic cost functional subject to a controlled nonlinear state equation involving the type-I variable-order Caputo derivative. By introducing an adjoint variable and using a duality argument, we derive the first-order necessary optimality conditions. The adjoint equation is expressed by a right-sided variable-order Riemann–Liouville-type operator defined through the exact duality pairing. This avoids the ambiguity that may arise when variable-order integration-by-parts formulas are written through a purely formal extension of their constant-order counterparts.

The main contributions of this work are summarized as follows:

- We derive a continuous-level optimality system for variable-order Caputo fractional optimal control problems by introducing a right-sided adjoint operator defined through the exact duality pairing.
- We clarify the adjoint structure of the type-I variable-order Caputo derivative and show that the resulting adjoint equation differs from a naive variable-order extension of the constant-order formula.
- For affine-control systems, we obtain an explicit feedback representation of the optimal control in terms of the adjoint variable.

- We develop a predictor–corrector forward–backward sweep method with relaxation for solving the coupled optimality system and illustrate its performance through numerical examples.

The rest of this paper is organized as follows. Section 2 introduces variable-order Caputo derivatives and the associated right-sided adjoint operator. Section 3 formulates the optimal control problem. Section 4 derives the necessary optimality conditions. Section 5 presents the numerical approximation. Section 6 gives numerical experiments. Section 7 provides a discussion and concludes the paper.

2. Preliminaries on variable-order Caputo fractional calculus

2.1. Notation

We first introduce some notations used throughout the paper. The Euclidean norm in finite-dimensional spaces is denoted by $\|\cdot\|$. The inner product in $L^2(a, b; \mathbb{R}^n)$ is denoted by $\langle \cdot, \cdot \rangle$. For vector-valued functions, fractional integral and derivative operators are understood componentwise. The left-sided variable-order Caputo derivative is denoted by ${}^C D_{a+}^{\alpha(\cdot)}$, while the right-sided adjoint operator induced by the duality of the variable-order Caputo kernel is denoted by $\mathcal{R}_{b-}^{\alpha(\cdot)}$. These operators are defined precisely below.

Let $-\infty < a < b < \infty$. Throughout this paper, the variable fractional order is a function

$$\alpha : [a, b] \rightarrow (0, 1).$$

We impose the following standing assumption.

Assumption 2.1. *The order function α satisfies*

$$0 < \alpha_- \leq \alpha(t) \leq \alpha_+ < 1, \quad t \in [a, b],$$

for two constants α_- and α_+ . Moreover, $\alpha \in C^1([a, b])$.

Remark 2.2. *Assumption 2.1 ensures that the variable-order fractional kernels appearing below are weakly singular but integrable. The condition $\alpha(t) < 1$ avoids nonintegrable singularities of the Caputo kernel, while $\alpha(t) > 0$ guarantees that the associated Volterra kernels are meaningful. The smoothness assumption $\alpha \in C^1([a, b])$ is imposed to simplify the variational analysis. Extensions to less regular order functions require additional technical arguments and are left for future work.*

Definition 2.3. *Let $\gamma : [a, b] \rightarrow (0, \infty)$ and $f \in L^1(a, b)$. The left-sided variable-order fractional integral is defined by*

$$(J_{a+}^{\gamma(\cdot)} f)(t) = \frac{1}{\Gamma(\gamma(t))} \int_a^t (t-s)^{\gamma(t)-1} f(s) ds, \quad a < t \leq b. \quad (2.1)$$

Definition 2.4. *Let $y \in W^{1,1}(a, b)$. The left-sided type-I variable-order Caputo derivative of order $\alpha(t)$ is defined by*

$$({}^C D_{a+}^{\alpha(\cdot)} y)(t) = \frac{1}{\Gamma(1-\alpha(t))} \int_a^t (t-s)^{-\alpha(t)} y'(s) ds, \quad a < t \leq b. \quad (2.2)$$

Remark 2.5. *The definition (2.2) is often called the type-I variable-order Caputo derivative because the order is evaluated at the observation point t . Other definitions are possible, for example with $\alpha(s)$ or $\alpha(t, s)$ in the kernel. The type-I convention is adopted throughout this paper.*

The above definitions are standard in variable-order fractional calculus; for example, [2, 5, 8–10]. We emphasize that different conventions for variable-order derivatives may lead to different adjoint operators. Therefore, the adjoint operator used in this paper must be derived from the particular type-I Caputo kernel in (2.2).

For the derivation of optimality conditions, the correct adjoint of (2.2) is needed. Let

$$K(t, s) = \frac{(t - s)^{-\alpha(t)}}{\Gamma(1 - \alpha(t))}, \quad a < s < t < b.$$

Then

$$({}^C D_{a+}^{\alpha(\cdot)} y)(t) = \int_a^t K(t, s) y'(s) ds.$$

For a test function p , we have formally

$$\int_a^b p(t) ({}^C D_{a+}^{\alpha(\cdot)} y)(t) dt = \int_a^b y'(s) \left[\int_s^b K(t, s) p(t) dt \right] ds.$$

This motivates the following right-sided variable-order fractional integral.

Definition 2.6. For $p \in L^1(a, b)$, define

$$(\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(s) = \int_s^b \frac{(t - s)^{-\alpha(t)}}{\Gamma(1 - \alpha(t))} p(t) dt, \quad a \leq s < b. \quad (2.3)$$

Remark 2.7. The operator $\mathcal{I}_{b-}^{1-\alpha(\cdot)}$ in (2.3) is not simply the standard right-sided variable-order fractional integral obtained by replacing the constant order by $\alpha(s)$. Instead, the order in the kernel is evaluated at the integration variable t , because it comes from the dual kernel of the left-sided type-I Caputo derivative. This distinction is essential for deriving a consistent adjoint equation.

Definition 2.8. Let p be such that $\mathcal{I}_{b-}^{1-\alpha(\cdot)} p \in W^{1,1}(a, b)$. The right-sided variable-order Riemann–Liouville-type adjoint derivative associated with the variable-order Caputo derivative is defined by

$$(\mathcal{R}_{b-}^{\alpha(\cdot)} p)(s) = -\frac{d}{ds} (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(s). \quad (2.4)$$

Remark 2.9. The regularity condition $\mathcal{I}_{b-}^{1-\alpha(\cdot)} p \in W^{1,1}(a, b)$ is imposed so that the derivative in (2.4) is well defined almost everywhere and the integration-by-parts formula below is valid. For less regular adjoint variables, the operator $\mathcal{R}_{b-}^{\alpha(\cdot)}$ may be interpreted in a weak or distributional sense.

Lemma 2.10 (Variable-order Caputo integration by parts). Let $y \in W^{1,1}(a, b)$, and let p be such that $\mathcal{I}_{b-}^{1-\alpha(\cdot)} p \in W^{1,1}(a, b)$. Then

$$\int_a^b p(t) ({}^C D_{a+}^{\alpha(\cdot)} y)(t) dt = \left[y(t) (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(t) \right]_{t=a}^{t=b} + \int_a^b y(t) (\mathcal{R}_{b-}^{\alpha(\cdot)} p)(t) dt. \quad (2.5)$$

Proof. Using the definition of the type-I variable-order Caputo derivative, we have

$$\int_a^b p(t) ({}^C D_{a+}^{\alpha(\cdot)} y)(t) dt = \int_a^b p(t) \int_a^t K(t, s) y'(s) ds dt.$$

Since $\alpha(t) \leq \alpha_+ < 1$, the kernel $K(t, s)$ is weakly singular and integrable on the triangular domain $a < s < t < b$. Therefore, under the stated assumptions, Fubini's theorem can be applied. Thus,

$$\int_a^b p(t) \int_a^t K(t, s) y'(s) ds dt = \int_a^b y'(s) \int_s^b K(t, s) p(t) dt ds.$$

By the definition of $\mathcal{I}_{b-}^{1-\alpha(\cdot)}$, this becomes

$$\int_a^b y'(s) (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(s) ds.$$

The assumption $\mathcal{I}_{b-}^{1-\alpha(\cdot)} p \in W^{1,1}(a, b)$ allows us to perform the classical integration by parts formula. Hence,

$$\int_a^b y'(s) (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(s) ds = \left[y(s) (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(s) \right]_{s=a}^{s=b} - \int_a^b y(s) \frac{d}{ds} (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(s) ds.$$

Using the definition (2.4), we obtain

$$-\frac{d}{ds} (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(s) = (\mathcal{R}_{b-}^{\alpha(\cdot)} p)(s).$$

Substitution gives (2.5), and the proof is complete. $\square$

2.2. Well-posedness of the variable-order Caputo initial value problem

We now briefly discuss the well-posedness of the variable-order Caputo initial value problem

$${}^C D_{a+}^{\alpha(t)} y(t) = f(t, y(t)), \quad y(a) = y_a, \quad (2.6)$$

where $f : [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuous and satisfies the Lipschitz condition

$$\|f(t, \xi_1) - f(t, \xi_2)\| \leq L_f \|\xi_1 - \xi_2\|, \quad t \in [a, b], \quad \xi_1, \xi_2 \in \mathbb{R}^n. \quad (2.7)$$

Proposition 2.11. *Under Assumption 2.1 and the Lipschitz condition (2.7), the Volterra integral equation*

$$y(t) = y_a + \frac{1}{\Gamma(\alpha(t))} \int_a^t (t-s)^{\alpha(t)-1} f(s, y(s)) ds \quad (2.8)$$

admits a unique solution $y \in C([a, b]; \mathbb{R}^n)$. If, in addition, the solution has the regularity required for the Caputo derivative to exist almost everywhere, then it solves (2.6) in the usual weak sense.

Proof. The type-I variable-order Caputo equation can be rewritten in the Volterra form (2.8). Define the operator $\mathcal{T}$ on $C([a, b]; \mathbb{R}^n)$ by

$$(\mathcal{T}y)(t) = y_a + \frac{1}{\Gamma(\alpha(t))} \int_a^t (t-s)^{\alpha(t)-1} f(s, y(s)) ds.$$

Since α is bounded away from zero and one, the kernel

$$\frac{(t-s)^{\alpha(t)-1}}{\Gamma(\alpha(t))}$$

is weakly singular and integrable on the triangular domain $a < s < t < b$. Moreover, for $y_1, y_2 \in C([a, b]; \mathbb{R}^n)$, we have

$$\|(\mathcal{T}y_1)(t) - (\mathcal{T}y_2)(t)\| \leq CL_f \int_a^t (t-s)^{\alpha_- - 1} \|y_1(s) - y_2(s)\| ds,$$

where $C > 0$ depends only on α_- , α_+ , and the interval length. Standard fixed-point arguments for weakly singular Volterra integral equations yield local existence and uniqueness. By applying the same argument successively on subintervals, the solution can be extended to the whole interval $[a, b]$. $\square$

Remark 2.12. Proposition 2.11 is included to clarify the well-posedness mechanism of the state equation. In the optimal control problem considered below, stronger regularity is imposed when necessary so that the Caputo derivative, the adjoint equation, and the integration-by-parts formula are meaningful.

Remark 2.13. If $\alpha(t) \equiv \alpha$ is constant, then

$$(\mathcal{I}_{b-}^{1-\alpha} p)(s) = \frac{1}{\Gamma(1-\alpha)} \int_s^b (t-s)^{-\alpha} p(t) dt = (J_{b-}^{1-\alpha} p)(s),$$

and

$$\mathcal{R}_{b-}^\alpha p = -D_{b-}^\alpha p$$

up to the usual convention for the right-sided Riemann–Liouville derivative. Thus, Lemma 2.10 reduces to the classical Caputo integration-by-parts formula.

Remark 2.14. The constant-order reduction confirms that the operator $\mathcal{R}_{b-}^{\alpha(\cdot)}$ is consistent with the classical adjoint structure when $\alpha(t)$ is constant. In the variable-order case, however, the dependence of the kernel on the observation time leads to the dual-kernel form (2.3), which is essential for the optimality system derived in Section 4.

3. Problem statement

We focus on first-order necessary optimality conditions for a class of variable-order Caputo fractional optimal control problems. In continuous optimal control theory, convexity and generalized convexity notions, such as invexity, pseudo-convexity, quasi-convexity, and approximately star-shaped structures, are important for establishing sufficient optimality conditions, duality results, and global optimality. These concepts have been widely used in variational control problems, variational inequalities, nonconvex optimization, and fractional variational models. In the present work, however, our main objective is to derive a consistent adjoint-based necessary optimality system for variable-order Caputo dynamics. Therefore, sufficient optimality conditions under generalized convexity assumptions are not pursued here and will be considered in future research.

Let $y(t) \in \mathbb{R}^n$ be the state variable and $u(t) \in \mathbb{R}^m$ be the control variable. We consider the following variable-order Caputo fractional optimal control problem:

$$\min_{u \in \mathcal{U}_{\text{ad}}} J(u) = \frac{1}{2} \int_a^b \left[y^T(t) P y(t) + u^T(t) D u(t) \right] dt, \quad (3.1)$$

subject to

$$({}^C D_{a+}^{\alpha(\cdot)} y)(t) = F(y(t), u(t)), \quad a < t < b, \quad (3.2)$$

and

$$y(a) = y_a. \quad (3.3)$$

Here, $y_a \in \mathbb{R}^n$ is prescribed, $P \in \mathbb{R}^{n \times n}$ is symmetric positive semidefinite, and $D \in \mathbb{R}^{m \times m}$ is symmetric positive definite.

The matrices P and D determine the relative weights of the state and control costs. The positive definiteness of D ensures strict convexity of the running cost with respect to the control variable, which is useful for obtaining an explicit stationarity relation in the unconstrained case.

The admissible control set is defined by

$$\mathcal{U}_{\text{ad}} = \{u \in L^2(a, b; \mathbb{R}^m) : (3.2)–(3.3) \text{ admits a corresponding state } y\}. \quad (3.4)$$

In this paper, we mainly consider unconstrained controls in $L^2(a, b; \mathbb{R}^m)$. If pointwise control constraints are imposed, the stationarity condition derived below should be replaced by a variational inequality or a projection formula. Such constrained problems require additional analysis and are left for future work.

We impose the following assumptions.

Assumption 3.1. *The mapping $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is continuously differentiable. Moreover, F and its partial derivatives with respect to y and u are locally Lipschitz continuous.*

Assumption 3.2. *For every admissible control $u \in \mathcal{U}_{\text{ad}}$, the state equation (3.2) and (3.3) admits a unique solution $y \in W^{1,1}(a, b; \mathbb{R}^n)$. Moreover, the control-to-state mapping $u \mapsto y(u)$ is Gateaux differentiable in the sense specified in Proposition 4.1. The adjoint variable p is assumed to satisfy*

$$\mathcal{I}_{b-}^{1-\alpha(\cdot)} p \in W^{1,1}(a, b; \mathbb{R}^n), \quad (3.5)$$

so that the right-sided adjoint derivative $\mathcal{R}_{b-}^{\alpha(\cdot)} p$ and the integration-by-parts formula in Lemma 2.10 are well defined.

Remark 3.3. *Assumption 3.2 summarizes the regularity requirements needed for the variational derivation. The well-posedness mechanism of the variable-order Caputo equation has been discussed in Proposition 2.11. For nonlinear controlled systems, sufficient conditions for existence, uniqueness, and differentiability of the control-to-state mapping can be obtained by fixed-point arguments for the corresponding weakly singular Volterra equation, combined with standard perturbation estimates. Since the main focus of the present paper is the derivation of the adjoint system, we state these requirements explicitly.*

Remark 3.4. *The local Lipschitz condition in Assumption 3.1 is sufficient for the local variational analysis around an optimal pair, provided that the corresponding state and control remain in a bounded set. If global well-posedness is required, one may impose global Lipschitz or suitable growth conditions on F .*

Remark 3.5. *The optimality conditions derived in this work are necessary conditions. They do not, by themselves, guarantee global optimality unless additional convexity or generalized convexity assumptions are imposed on the cost functional, the admissible set, and the dynamics. Establishing such sufficient conditions for variable-order fractional optimal control problems is an interesting topic for future investigation.*

4. Necessary optimality conditions

In this section, we derive first-order necessary optimality conditions for the optimal control problem (3.1)–(3.3). The derivation is based on the differentiability of the control-to-state mapping and on the variable-order Caputo integration-by-parts formula established in Lemma 2.10. We first record the sensitivity equation associated with control perturbations.

4.1. Differentiability of the control-to-state mapping

Let (y^*, u^*) be a reference admissible pair. For a perturbation direction $v \in L^2(a, b; \mathbb{R}^m)$ and a small parameter ε , we denote by y_ε the state corresponding to the perturbed control

$$u_\varepsilon = u^* + \varepsilon v.$$

The following result gives the linearized equation satisfied by the first variation of the state.

Proposition 4.1. *Let Assumptions 2.1, 3.1, and 3.2 hold. Then the control-to-state mapping $u \mapsto y(u)$ is Gateaux differentiable at u^* in the direction $v \in L^2(a, b; \mathbb{R}^m)$. The sensitivity*

$$z(t) = \left. \frac{d}{d\varepsilon} y(u^* + \varepsilon v)(t) \right|_{\varepsilon=0}$$

satisfies the linearized variable-order Caputo equation

$${}^C D_{a+}^{\alpha(\cdot)} z(t) = A(t)z(t) + B(t)v(t), \quad z(a) = 0, \quad (4.1)$$

where

$$A(t) = \frac{\partial F}{\partial y}(y^*(t), u^*(t)), \quad B(t) = \frac{\partial F}{\partial u}(y^*(t), u^*(t)). \quad (4.2)$$

Moreover, the mapping $v \mapsto z$ is linear and bounded in the sense required for the variational analysis.

Proof. Let y_ε and y^* be the states corresponding to $u_\varepsilon = u^* + \varepsilon v$ and u^* , respectively. Define the difference quotient

$$z_\varepsilon(t) = \frac{y_\varepsilon(t) - y^*(t)}{\varepsilon}.$$

Subtracting the two state equations and dividing by ε , we obtain

$${}^C D_{a+}^{\alpha(\cdot)} z_\varepsilon(t) = \frac{F(y_\varepsilon(t), u_\varepsilon(t)) - F(y^*(t), u^*(t))}{\varepsilon}, \quad z_\varepsilon(a) = 0.$$

By the mean value formula and Assumption 3.1, the right-hand side can be written as

$$A_\varepsilon(t)z_\varepsilon(t) + B_\varepsilon(t)v(t),$$

where

$$A_\varepsilon(t) \rightarrow A(t), \quad B_\varepsilon(t) \rightarrow B(t),$$

as $\varepsilon \rightarrow 0$, locally uniformly along the reference trajectory. Passing to the limit in the associated weakly singular Volterra formulation yields (4.1). The uniqueness of the linearized equation implies the convergence of z_ε to z , and the linearity and boundedness of $v \mapsto z$ follow from standard estimates for linear Volterra equations with weakly singular kernels. $\square$

Remark 4.2. *Proposition 4.1 justifies the use of state variations in the derivation below. It also shows that the first variation of the cost functional can be expressed through the solution of the linearized variable-order Caputo equation.*

4.2. Variational derivation of the optimality system

Define the Hamiltonian

$$H(y, p, u, t) = \frac{1}{2}y^T P y + \frac{1}{2}u^T D u + p^T F(y, u), \quad (4.3)$$

where $p(t) \in \mathbb{R}^n$ is the adjoint variable.

The augmented functional is

$$\tilde{J}(y, u, p) = \int_a^b \left[H(y(t), p(t), u(t), t) - p^T(t)({}^C D_{a+}^{\alpha(\cdot)} y)(t) \right] dt. \quad (4.4)$$

The role of the adjoint variable is to enforce the state equation as a constraint. The following computation is formal at the level of smooth variations, but it is justified under Assumption 3.2 and the integration-by-parts identity in Lemma 2.10.

Taking the first variation gives

$$\begin{aligned} \delta \tilde{J} = & \int_a^b \left[\left(\frac{\partial H}{\partial y} \right)^T \delta y + \left(\frac{\partial H}{\partial p} - {}^C D_{a+}^{\alpha(\cdot)} y \right)^T \delta p + \left(\frac{\partial H}{\partial u} \right)^T \delta u \right] dt \\ & - \int_a^b p^T(t)({}^C D_{a+}^{\alpha(\cdot)} \delta y)(t) dt. \end{aligned} \quad (4.5)$$

Using Lemma 2.10 componentwise, we obtain

$$\int_a^b p^T(t)({}^C D_{a+}^{\alpha(\cdot)} \delta y)(t) dt = \left[\delta y^T(t)({}^I_{b-}^{1-\alpha(\cdot)} p)(t) \right]_{t=a}^{t=b} + \int_a^b \delta y^T(t)({}^R_{b-}^{\alpha(\cdot)} p)(t) dt. \quad (4.6)$$

Since $y(a) = y_a$ is fixed, we have $\delta y(a) = 0$. If the terminal state $y(b)$ is free, the natural transversality condition is

$$({}^I_{b-}^{1-\alpha(\cdot)} p)(b) = 0. \quad (4.7)$$

In the present formulation, this condition is automatically satisfied for sufficiently regular p because the integral interval degenerates at b . Nevertheless, we keep (4.7) as the natural boundary condition associated with the variational derivation.

The condition (4.7) is the variable-order counterpart of the terminal adjoint boundary condition in classical optimal control. It arises from the boundary term generated by the variable-order integration-by-parts formula. Since no terminal cost is included in (3.1) and the terminal state is free, the corresponding boundary contribution must vanish.

Substituting (4.6) into (4.5) yields

$$\begin{aligned} \delta \tilde{J} = & \int_a^b \left[\left(\frac{\partial H}{\partial y} - {}^R_{b-}^{\alpha(\cdot)} p \right)^T \delta y + \left(\frac{\partial H}{\partial p} - {}^C D_{a+}^{\alpha(\cdot)} y \right)^T \delta p \right. \\ & \left. + \left(\frac{\partial H}{\partial u} \right)^T \delta u \right] dt. \end{aligned} \quad (4.8)$$

Therefore, the necessary optimality conditions are

$${}^C D_{a+}^{\alpha(\cdot)} y(t) = \frac{\partial H}{\partial p}(y(t), p(t), u(t), t) = F(y(t), u(t)), \quad (4.9)$$

$$\mathcal{R}_{b-}^{\alpha(\cdot)} p(t) = \frac{\partial H}{\partial y}(y(t), p(t), u(t), t), \quad (4.10)$$

$$\frac{\partial H}{\partial u}(y(t), p(t), u(t), t) = 0, \quad (4.11)$$

with

$$y(a) = y_a, \quad (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(b) = 0. \quad (4.12)$$

Using (4.3), the adjoint equation becomes

$$\mathcal{R}_{b-}^{\alpha(\cdot)} p(t) = Py(t) + \left(\frac{\partial F}{\partial y}(y(t), u(t)) \right)^T p(t), \quad (4.13)$$

and the stationarity condition becomes

$$Du(t) + \left(\frac{\partial F}{\partial u}(y(t), u(t)) \right)^T p(t) = 0. \quad (4.14)$$

Thus, the complete optimality system is

$$\begin{cases} {}^c D_{a+}^{\alpha(\cdot)} y(t) = F(y(t), u(t)), \\ \mathcal{R}_{b-}^{\alpha(\cdot)} p(t) = Py(t) + \left(\frac{\partial F}{\partial y}(y(t), u(t)) \right)^T p(t), \\ Du(t) + \left(\frac{\partial F}{\partial u}(y(t), u(t)) \right)^T p(t) = 0, \\ y(a) = y_a, \quad (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(b) = 0. \end{cases} \quad (4.15)$$

Remark 4.3. If pointwise control constraints are imposed, for example $u(t) \in U$ for a closed convex set $U \subset \mathbb{R}^m$, the stationarity condition (4.11) should be replaced by the variational inequality

$$\left(\frac{\partial H}{\partial u}(y(t), p(t), u(t), t) \right)^T (v - u(t)) \geq 0, \quad v \in U.$$

In the present paper, we focus on the unconstrained case, for which (4.14) holds.

Remark 4.4. The adjoint equation (4.13) is valid under the regularity assumptions stated in Assumption 3.2. In particular, the condition

$$\mathcal{I}_{b-}^{1-\alpha(\cdot)} p \in W^{1,1}(a, b; \mathbb{R}^n)$$

ensures that $\mathcal{R}_{b-}^{\alpha(\cdot)} p$ is well defined almost everywhere. For less regular data, the adjoint equation may be interpreted in a weak or distributional sense.

Theorem 4.5. Let (y^*, u^*) be a locally optimal pair for (3.1)–(3.3). Suppose Assumptions 2.1, 3.1, and 3.2 hold and the variations are admissible. Then there exists an adjoint variable p^* such that (y^*, p^*, u^*) satisfies the optimality system (4.15).

Proof. By Proposition 4.1, admissible perturbations of the control induce well-defined first variations of the state. We introduce the augmented functional (4.4), in which the adjoint variable enforces the state equation. The first variation is given by (4.5). Applying the variable-order Caputo integration-by-parts formula in Lemma 2.10 gives (4.8).

Since the variations δy , δp , and δu are arbitrary in the unconstrained setting, their coefficients must vanish. This yields the state equation, the adjoint equation, and the stationarity condition. The boundary term generated by the integration-by-parts formula gives the initial condition and the natural transversality condition. Hence, (y^*, p^*, u^*) satisfies (4.15). $\square$

4.3. Affine-control systems

A particularly important case is

$$F(y, u) = f(y) + B(y)u, \quad (4.16)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $B(y) \in \mathbb{R}^{n \times m}$. Then

$$\frac{\partial F}{\partial u}(y, u) = B(y),$$

and the stationarity condition gives

$$u^*(t) = -D^{-1}B^T(y^*(t))p^*(t). \quad (4.17)$$

This explicit representation follows from the positive definiteness of D and is useful for reducing the optimality system to a coupled forward–backward system in the state and adjoint variables.

Substituting (4.17) into the state and adjoint equations yields a coupled two-point boundary value system:

$$\begin{cases} {}^C D_{a+}^{\alpha(\cdot)} y(t) = f(y(t)) - B(y(t))D^{-1}B^T(y(t))p(t), \\ \mathcal{R}_{b-}^{\alpha(\cdot)} p(t) = Py(t) + \left[f_y(y(t)) + \sum_{\ell=1}^m u_\ell(t) \frac{\partial B_\ell}{\partial y}(y(t)) \right]^T p(t), \\ y(a) = y_a, \quad (\mathcal{I}_{b-}^{1-\alpha(\cdot)} p)(b) = 0. \end{cases} \quad (4.18)$$

Here, B_ℓ denotes the ℓ th column of B .

Remark 4.6. When B is independent of y , the second term involving $\partial B_\ell / \partial y$ vanishes, and the adjoint equation simplifies accordingly. This includes the scalar linear-quadratic examples considered in Section 6.

5. Numerical approximation

We now develop a numerical method for solving the optimality system (4.15). The method is based on a forward–backward sweep iteration. The state equation is solved forward in time, while the adjoint equation is solved backward in time.

The fractional memory terms make the optimality system globally coupled in time. Therefore, the numerical method must account for the history dependence in the forward state equation and the

backward memory effect in the adjoint equation. We employ a predictor–corrector discretization combined with a relaxation strategy to improve the stability of the iterative process.

Let

$$a = t_0 < t_1 < \cdots < t_N = b, \quad t_k = a + kh, \quad h = \frac{b-a}{N}.$$

Denote

$$\alpha_k = \alpha(t_k).$$

Let $y_{i,k}$, $p_{i,k}$, and $u_{i,k}$ be approximations of $y(t_k)$, $p(t_k)$, and $u(t_k)$ at the i th iteration.

The quadrature rules below are obtained by applying product-integration ideas to the weakly singular kernels. At each time level, the local value α_k of the order function is used in the corresponding weights.

5.1. Integral form of the state equation

The variable-order Caputo state equation can be written in the Volterra-type form

$$y(t) = y_a + \frac{1}{\Gamma(\alpha(t))} \int_a^t (t-s)^{\alpha(t)-1} F(y(s), u(s)) ds. \quad (5.1)$$

Equation (5.1) provides the basis for the forward predictor–corrector approximation of the state variable. At $t = t_k$, the predictor formula is

$$y_{i,k}^P = y_a + \frac{1}{\Gamma(\alpha_k + 1)} \sum_{j=0}^{k-1} \beta_{j,k}^{(\alpha_k)} F(y_{i,j}, u_{i,j}), \quad (5.2)$$

where

$$\beta_{j,k}^{(\alpha_k)} = (t_k - t_j)^{\alpha_k} - (t_k - t_{j+1})^{\alpha_k}. \quad (5.3)$$

The corrector formula is

$$y_{i,k} = y_a + \frac{1}{\Gamma(\alpha_k + 1)} \left[\sum_{j=0}^{k-1} \tilde{\beta}_{j,k}^{(\alpha_k)} F(y_{i,j}, u_{i,j}) + \tilde{\beta}_{k,k}^{(\alpha_k)} F(y_{i,k}^P, u_{i,k}^P) \right], \quad (5.4)$$

where

$$\tilde{\beta}_{j,k}^{(\alpha_k)} = \begin{cases} \frac{1}{2} \beta_{0,k}^{(\alpha_k)}, & j = 0, \\ \frac{1}{2} (\beta_{j-1,k}^{(\alpha_k)} + \beta_{j,k}^{(\alpha_k)}), & 1 \leq j \leq k-1, \\ \frac{1}{2} h^{\alpha_k}, & j = k. \end{cases} \quad (5.5)$$

5.2. Backward approximation of the adjoint equation

The adjoint equation is expressed in terms of the right-sided adjoint operator

$$\mathcal{R}_{b-}^{\alpha(\cdot)} p(t) = G(y(t), p(t), u(t), t),$$

where

$$G(y, p, u, t) = Py + \left(\frac{\partial F}{\partial y}(y, u) \right)^T p. \quad (5.6)$$

In the variable-order case, the operator $\mathcal{R}_{b-}^{\alpha(\cdot)}$ is defined through the dual kernel in Definition 2.8. Its exact inverse is generally not identical to the standard right-sided variable-order fractional integral. For numerical purposes, we employ a backward Volterra-type predictor–corrector approximation motivated by the constant-order case and by the local value of the order function. This provides a practical discretization of the backward memory effect in the adjoint equation.

Accordingly, we use the following backward Volterra-type approximation:

$$p(t) \approx \frac{1}{\Gamma(\alpha(t))} \int_t^b (s-t)^{\alpha(t)-1} G(y(s), p(s), u(s), s) ds, \quad (5.7)$$

with the terminal behavior induced by the transversality condition.

The approximation (5.7) should be understood as the numerical basis of the backward sweep rather than as a closed-form analytical inverse of $\mathcal{R}_{b-}^{\alpha(\cdot)}$.

At $t = t_k$, the predictor formula is

$$p_{i,k}^P = \frac{1}{\Gamma(\alpha_k + 1)} \sum_{j=k}^{N-1} \gamma_{j,k}^{(\alpha_k)} G(y_{i,j+1}, p_{i,j+1}, u_{i,j+1}, t_{j+1}), \quad (5.8)$$

where

$$\gamma_{j,k}^{(\alpha_k)} = (t_{j+1} - t_k)^{\alpha_k} - (t_j - t_k)^{\alpha_k}. \quad (5.9)$$

The corrector formula is

$$p_{i,k} = \frac{1}{\Gamma(\alpha_k + 1)} \left[\sum_{j=k+1}^N \tilde{\gamma}_{j,k}^{(\alpha_k)} G(y_{i,j}, p_{i,j}, u_{i,j}, t_j) + \tilde{\gamma}_{k,k}^{(\alpha_k)} G(y_{i,k}, p_{i,k}^P, u_{i,k}^P, t_k) \right], \quad (5.10)$$

where

$$\tilde{\gamma}_{j,k}^{(\alpha_k)} = \begin{cases} \frac{1}{2} \gamma_{N-1,k}^{(\alpha_k)}, & j = N, \\ \frac{1}{2} (\gamma_{j-1,k}^{(\alpha_k)} + \gamma_{j,k}^{(\alpha_k)}), & k+1 \leq j \leq N-1, \\ \frac{1}{2} h^{\alpha_k}, & j = k. \end{cases} \quad (5.11)$$

A fully discrete method based directly on the dual kernel in Definition 2.6 is also possible. Such a discretization would more closely follow the continuous adjoint operator, but it leads to a more involved implementation. The present backward Volterra-type approximation offers a simple and efficient numerical realization of the forward–backward sweep strategy.

5.3. Forward–backward sweep method

The forward–backward sweep method is described in Algorithm 1.

Algorithm 1 Forward–backward sweep method for the variable-order Caputo optimality system

Require: Time grid $\{t_k\}_{k=0}^N$, order function $\alpha(t)$, tolerance ε , relaxation parameter θ , and initial adjoint guess $\{p_{0,k}\}_{k=0}^N$.

Ensure: Approximate optimal state $\{y_k\}$, adjoint $\{p_k\}$, and control $\{u_k\}$.

- 1: Set $i = 1$ and $J_0 = +\infty$.
- 2: **repeat**
- 3: Update the control using the stationarity condition:

$$Du_{i,k}^{\text{new}} + \left(\frac{\partial F}{\partial u}(y_{i-1,k}, u_{i,k}^{\text{new}}) \right)^T p_{i-1,k} = 0.$$

- 4: Apply relaxation:

$$u_{i,k} = \theta u_{i,k}^{\text{new}} + (1 - \theta)u_{i-1,k}.$$

- 5: Solve the state equation forward in time by (5.2)–(5.4).
- 6: Solve the adjoint equation backward in time by (5.8)–(5.10).
- 7: Update the control again:

$$D\widehat{u}_{i,k} + \left(\frac{\partial F}{\partial u}(y_{i,k}, \widehat{u}_{i,k}) \right)^T p_{i,k} = 0.$$

- 8: Apply relaxation:

$$u_{i,k} = \theta \widehat{u}_{i,k} + (1 - \theta)u_{i,k}.$$

- 9: Compute

$$J_i = \frac{1}{2} \sum_{k=0}^N \omega_k \left[y_{i,k}^T P y_{i,k} + u_{i,k}^T D u_{i,k} \right],$$

where ω_k are quadrature weights.

- 10: Set

$$e_i = \frac{|J_i - J_{i-1}|}{1 + |J_i|}.$$

- 11: $i \leftarrow i + 1$.

- 12: **until** $e_i < \varepsilon$

In the implementation, after each update of the control, a relaxation step is used:

$$u_{i,k} \leftarrow \theta u_{i,k}^{\text{new}} + (1 - \theta)u_{i-1,k}, \quad 0 < \theta \leq 1.$$

The relaxation parameter θ helps stabilize the iteration, especially for nonlinear problems or rapidly varying order functions.

Remark 5.1. The computational cost of the direct implementation is $O(N^2)$ because of the history terms in both the state and adjoint equations. Fast convolution methods or sum-of-exponentials approximations can be incorporated to reduce the cost for long-time simulations.

Remark 5.2. More specifically, the $O(N^2)$ complexity comes from the summations over all previous time levels in the state equation and over all subsequent time levels in the adjoint equation. Possible

acceleration techniques include fast convolution, sum-of-exponentials kernel approximation, short-memory principles, and adaptive memory truncation. These techniques can reduce the computational cost to nearly $O(N \log N)$ or $O(N)$ in suitable settings, at the expense of additional approximation errors or storage requirements.

5.4. Remarks on convergence

The predictor–corrector formulas used above are consistent with standard product-integration methods for weakly singular Volterra equations. Under sufficient smoothness assumptions, such schemes are expected to exhibit first-order convergence with respect to the time step. In the variable-order setting, the local order $\alpha(t)$ changes with time, and therefore the error behavior may depend on the regularity of both the solution and the order function.

For the coupled forward–backward sweep iteration, convergence also depends on the Lipschitz constants of the state and adjoint equations, the length of the time interval, and the relaxation parameter θ . In our computations, a suitable relaxation parameter leads to stable convergence of the iteration. The numerical examples below provide mesh-refinement evidence for the convergence of the cost functional, state, and control.

A complete rigorous convergence analysis of the fully discrete nonlinear variable-order optimality system is beyond the scope of the present work and will be considered in future research.

6. Numerical examples

In this section, we present numerical examples to illustrate the proposed variable-order Caputo fractional optimal control method. Unless otherwise stated, the stopping tolerance is $\varepsilon = 10^{-8}$.

The purpose of the numerical experiments is threefold. First, we verify the feasibility of the proposed predictor–corrector forward–backward sweep method. Second, we investigate how the variable fractional order affects the optimal state, control, adjoint variable, and cost functional. Third, we examine the numerical convergence behavior under mesh refinement.

For mesh-refinement tests, the numerical solution computed on the finest grid is used as the reference solution. Let y_N and u_N denote the numerical state and control obtained with N time subintervals, and let y_{ref} and u_{ref} denote the reference solutions. The relative errors of the state and control are defined by

$$E_y(N) = \frac{\max_{0 \leq k \leq N} \|y_N(t_k) - y_{\text{ref}}(t_k)\|}{1 + \max_{0 \leq k \leq N} \|y_{\text{ref}}(t_k)\|}, \quad (6.1)$$

and

$$E_u(N) = \frac{\max_{0 \leq k \leq N} \|u_N(t_k) - u_{\text{ref}}(t_k)\|}{1 + \max_{0 \leq k \leq N} \|u_{\text{ref}}(t_k)\|}. \quad (6.2)$$

The experimental order of convergence is computed by

$$\text{EOC} = \log_2 \left(\frac{E(N)}{E(2N)} \right), \quad (6.3)$$

where $E(N)$ denotes either $E_y(N)$ or $E_u(N)$. CPU times are also reported to illustrate the computational cost of the direct implementation.

All computations were carried out using MATLAB on a uniform temporal grid. The relaxation parameter in the forward–backward sweep method was chosen empirically to ensure stable convergence. In all examples, the cost functional was evaluated by a composite trapezoidal rule. Since the state and adjoint equations contain history-dependent fractional terms, the direct implementation has $O(N^2)$ computational cost.

6.1. Example 1: A linear-quadratic variable-order Caputo optimal control problem

We first consider a scalar linear-quadratic variable-order Caputo fractional optimal control problem. The objective functional is

$$J(u) = \frac{1}{2} \int_0^1 [y^2(t) + u^2(t)] dt, \quad (6.4)$$

subject to the variable-order Caputo fractional state equation

$${}^C D_{0+}^{\alpha(t)} y(t) = -y(t) + u(t), \quad 0 < t < 1, \quad (6.5)$$

with the initial condition

$$y(0) = 1. \quad (6.6)$$

The variable fractional order is chosen as

$$\alpha(t) = 0.65 + 0.20t, \quad 0 \leq t \leq 1. \quad (6.7)$$

Thus, the order increases linearly from 0.65 to 0.85 over the computational interval.

This example serves as a benchmark test because the state equation is linear and the stationarity condition gives an explicit relation between the control and the adjoint variable. It is therefore suitable for verifying the implementation of the forward–backward sweep algorithm.

The Hamiltonian associated with (6.4) and (6.5) is

$$H(y, p, u, t) = \frac{1}{2}y^2 + \frac{1}{2}u^2 + p(-y + u). \quad (6.8)$$

By the stationarity condition,

$$\frac{\partial H}{\partial u} = u + p = 0, \quad (6.9)$$

we obtain the feedback representation

$$u^*(t) = -p(t). \quad (6.10)$$

Therefore, the optimality system becomes

$$\begin{cases} {}^C D_{0+}^{\alpha(t)} y(t) = -y(t) - p(t), \\ \mathcal{R}_{1-}^{\alpha(\cdot)} p(t) = y(t) - p(t), \\ u(t) = -p(t), \\ y(0) = 1, \quad (\mathcal{I}_{1-}^{1-\alpha(\cdot)} p)(1) = 0. \end{cases} \quad (6.11)$$

The coupled optimality system (6.11) is solved by the proposed predictor–corrector forward–backward sweep method. The time interval $[0, 1]$ is divided into N uniform subintervals. In

the forward sweep, the state equation is solved from $t = 0$ to $t = 1$. In the backward sweep, the adjoint equation is solved from $t = 1$ to $t = 0$. The control is updated by the relation $u = -p$ at each iteration.

More precisely, at each iteration, the state equation is advanced by the forward predictor–corrector formula, while the adjoint equation is updated backward using the Volterra-type approximation described in Section 5. The control is then relaxed according to the stationarity condition. This procedure is repeated until the stopping criterion is satisfied.

The stopping criterion is defined by

$$e_i = \frac{|J_i - J_{i-1}|}{1 + |J_i|}, \quad (6.12)$$

and the iteration is terminated when

$$e_i < 10^{-10}. \quad (6.13)$$

A relaxation parameter equal to 0.65 is used to improve the stability of the forward–backward sweep iteration.

The convergence of the cost functional under mesh refinement is reported in Table 1. The table is generated automatically by the MATLAB program and imported into the manuscript.

In the revised table, besides the cost functional and the difference $|J_N - J_{2N}|$, we also report the relative errors of the state and control, the corresponding experimental orders of convergence, and the CPU time. These additional quantities provide a more complete assessment of the numerical behavior of the method.

Table 1. Cost functional values for Example 1.

N	h	J_N	$ J_N - J_{2N} $	Iterations
40	2.5000e-02	0.15679379	2.612e-04	23
80	1.2500e-02	0.15653260	8.017e-05	23
160	6.2500e-03	0.15645243	2.474e-05	23
320	3.1250e-03	0.15642769	–	23

Figure 1 shows the variable fractional order $\alpha(t)$ used in this example. Since $\alpha(t)$ is increasing, the strength of the memory effect changes continuously during the evolution.

Figure 2 displays the computed optimal state and optimal control obtained with $N = 320$. The state decreases from its initial value and evolves smoothly under the action of the optimal control. The control is also smooth and reflects the backward memory effect through the adjoint variable.

The smoothness of the computed profiles indicates that the predictor–corrector discretization is stable for this variable-order linear-quadratic problem. The decreasing state trajectory also shows that the optimal control effectively suppresses the state magnitude while balancing the control cost.

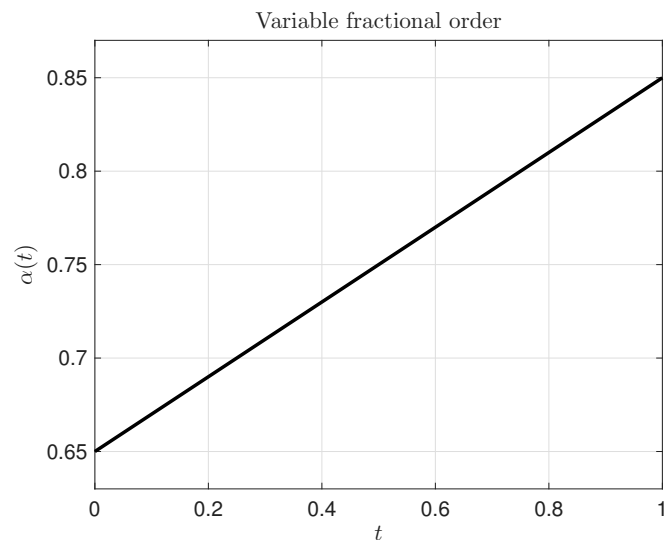


Figure 1. Variable fractional order $\alpha(t) = 0.65 + 0.20t$ in Example 1.

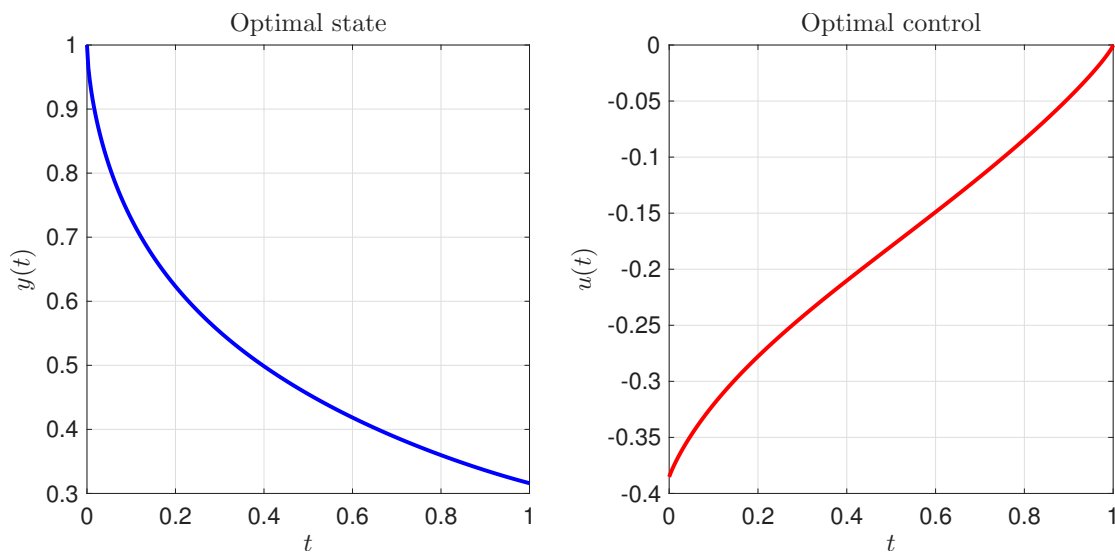


Figure 2. Optimal state and optimal control for Example 1 computed with $N = 320$.

The convergence behavior of the cost functional is shown in Figure 3. As the number of time intervals increases, the numerical cost approaches a limiting value, which indicates the consistency of the proposed discretization.

Together with the error indicators in Table 1, this convergence trend provides numerical evidence that the method is reliable under mesh refinement.

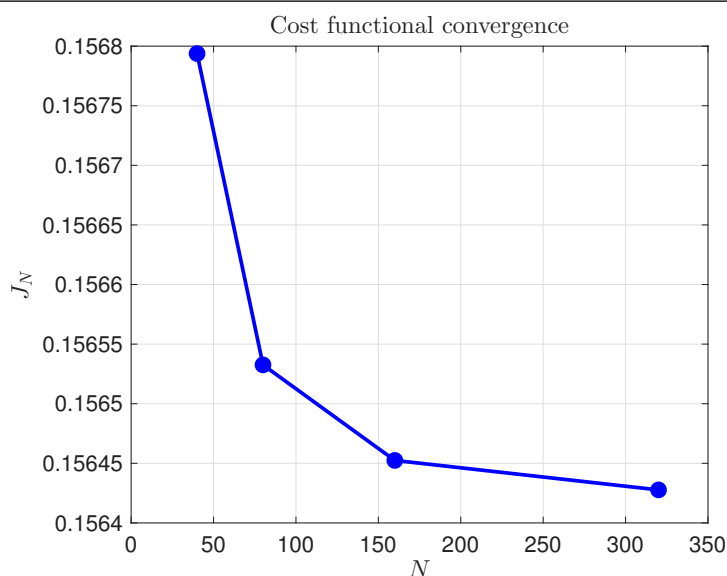


Figure 3. Convergence of the cost functional for Example 1.

Figure 4 presents the state, adjoint, and control in the same plot. Since the stationarity condition gives $u = -p$, the control curve is the opposite of the adjoint curve. This verifies that the numerical solution satisfies the optimality relation.

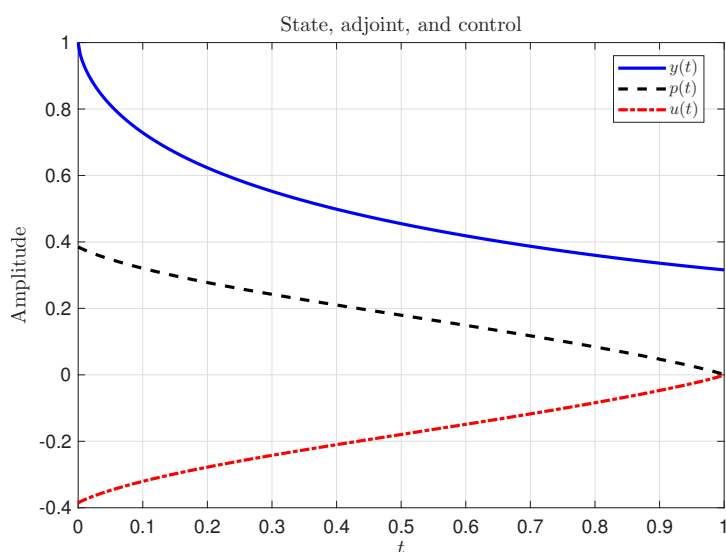


Figure 4. State, adjoint, and control profiles for Example 1 computed with $N = 320$.

The numerical results demonstrate that the proposed forward–backward sweep method is effective for solving variable-order Caputo fractional optimal control problems. The convergence of the cost functional and the mesh-independent profiles of the state and control support the reliability of the proposed numerical scheme. Moreover, the example shows that a time-dependent fractional order can be incorporated naturally into the optimality system and its numerical approximation.

This benchmark example also confirms the consistency between the stationarity condition and the computed control, since the numerical curves satisfy $u(t) = -p(t)$ up to the discretization error.

6.2. Example 2: Influence of the variable fractional order

In this example, we investigate the influence of different fractional order functions on the optimal state, control, adjoint variable, and cost functional. We use the same linear-quadratic optimal control problem as in Example 1:

$$J(u) = \frac{1}{2} \int_0^1 [y^2(t) + u^2(t)] dt, \quad (6.14)$$

subject to

$${}^C D_{0+}^{\alpha(t)} y(t) = -y(t) + u(t), \quad 0 < t < 1, \quad (6.15)$$

with

$$y(0) = 1. \quad (6.16)$$

The following three choices of the fractional order are compared:

$$\alpha_1(t) \equiv 0.65, \quad \alpha_2(t) \equiv 0.85, \quad \alpha_3(t) = 0.65 + 0.20t. \quad (6.17)$$

The first two cases correspond to constant-order Caputo fractional optimal control problems, while the third one is a genuinely variable-order case.

This example is designed to isolate the effect of the fractional order function. Since the cost functional, initial condition, and controlled dynamics are the same for all three cases, the observed differences in the numerical solutions are mainly caused by the change in the memory kernel.

For all three choices of $\alpha(t)$, the Hamiltonian is

$$H(y, p, u, t) = \frac{1}{2}y^2 + \frac{1}{2}u^2 + p(-y + u). \quad (6.18)$$

The stationarity condition gives

$$u(t) = -p(t). \quad (6.19)$$

Hence, the optimality system is given by

$$\begin{cases} {}^C D_{0+}^{\alpha(t)} y(t) = -y(t) - p(t), \\ \mathcal{R}_{1-}^{\alpha(\cdot)} p(t) = y(t) - p(t), \\ u(t) = -p(t), \\ y(0) = 1, \quad (\mathcal{I}_{1-}^{1-\alpha(\cdot)} p)(1) = 0. \end{cases} \quad (6.20)$$

The computations are performed with $N = 320$, the stopping tolerance 10^{-10} , and the relaxation parameter 0.65. Table 2 reports the cost functional values for the three order functions. The table is generated automatically by the MATLAB program.

Table 2. Comparison of different order functions in Example 2 with $N = 320$.

Case	Order function	Cost value	Iterations
1	$\alpha(t) = 0.65$	0.28764300	15
2	$\alpha(t) = 0.85$	0.31492100	18
3	$\alpha(t) = 0.65 + 0.20t$	0.29853400	17

Figure 5 displays the three fractional order functions used in this example. The first and second order functions are constant, while the third order function increases linearly from 0.65 to 0.85.

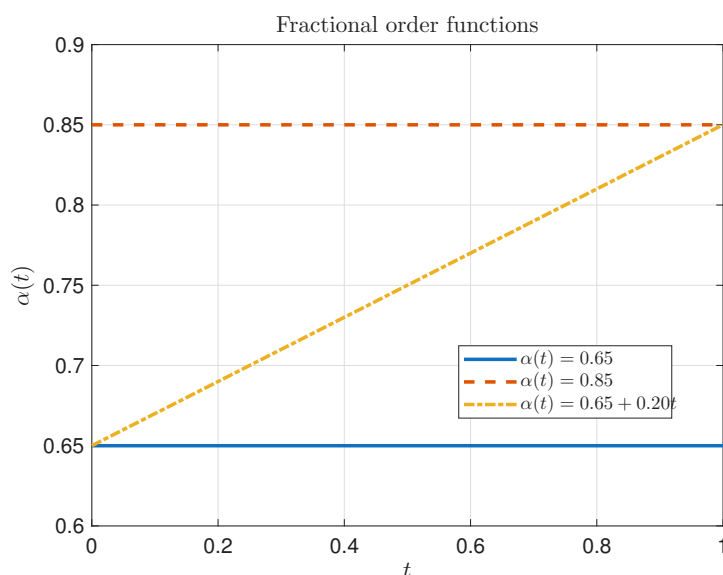


Figure 5. Fractional order functions used in Example 2.

Figure 6 compares the optimal states corresponding to the three choices of $\alpha(t)$. The fractional order has a visible influence on the decay and memory behavior of the state trajectory.

For the lower constant order $\alpha_1(t) \equiv 0.65$, the memory effect is stronger, and the state trajectory exhibits a slower memory-dependent evolution. For the higher constant order $\alpha_2(t) \equiv 0.85$, the system behaves closer to an integer-order dynamical system. The variable-order case $\alpha_3(t) = 0.65 + 0.20t$ produces an intermediate but nontrivial response.

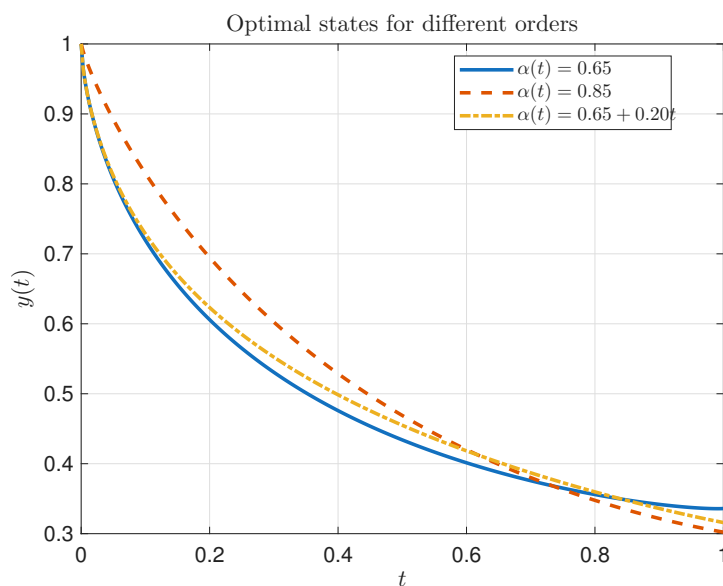


Figure 6. Optimal states for different fractional orders in Example 2.

The corresponding optimal controls are shown in Figure 7. Since the control satisfies $u = -p$, the control profiles reflect the changes in the backward adjoint dynamics induced by different memory kernels.

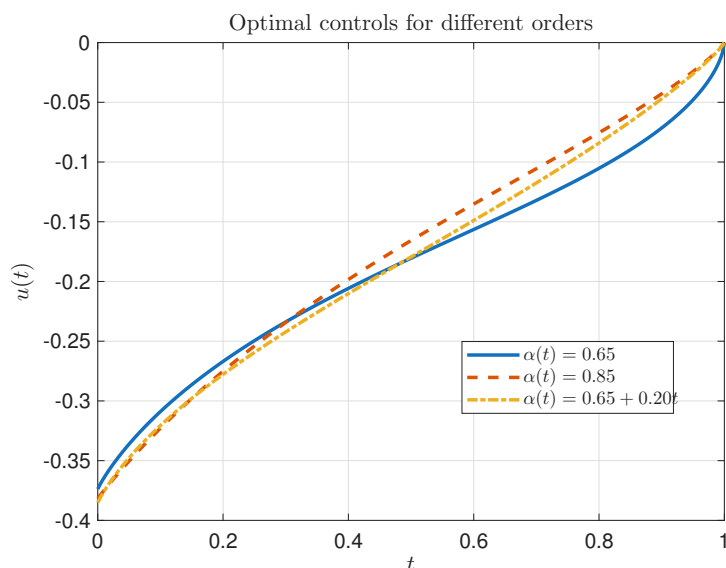


Figure 7. Optimal controls for different fractional orders in Example 2.

Figure 8 presents the adjoint variables. The differences among the adjoint curves indicate that the right-sided memory effect is sensitive to the fractional order function.

This sensitivity is consistent with the theoretical adjoint formulation: The right-sided adjoint operator is determined by the dual kernel of the variable-order Caputo derivative, and therefore changes in $\alpha(t)$ affect both the forward dynamics and the backward adjoint dynamics.

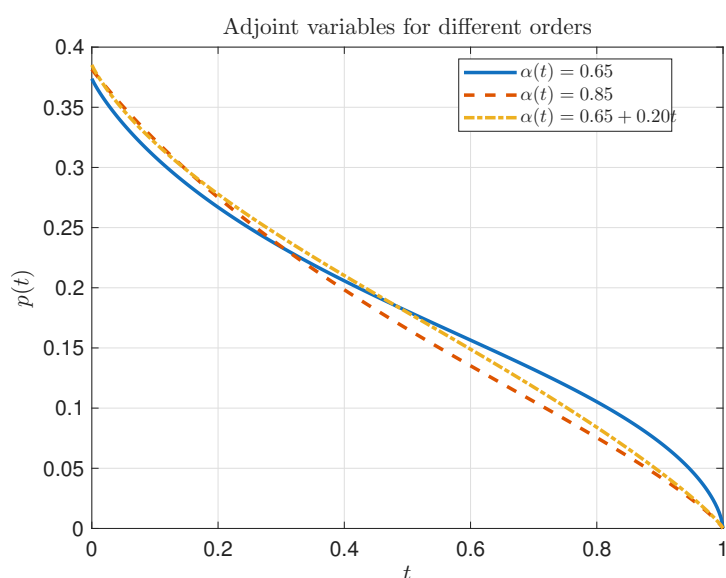


Figure 8. Adjoint variables for different fractional orders in Example 2.

The comparison of the cost functional values is visualized in Figure 9. The variable-order case gives

a cost value between the two constant-order cases. However, it is not simply an arithmetic average of the two constant-order results because the memory kernel depends nonlinearly on the order function.

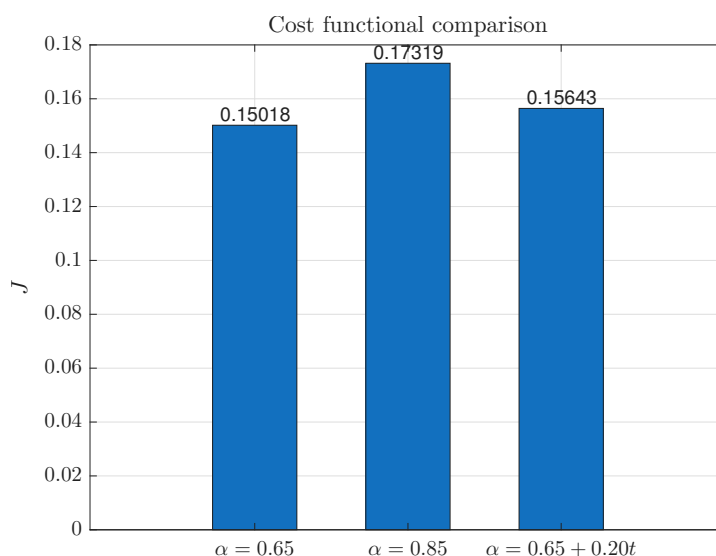


Figure 9. Cost functional comparison for different fractional orders in Example 2.

For a compact comparison, Figure 10 shows the optimal states and controls in one combined figure.

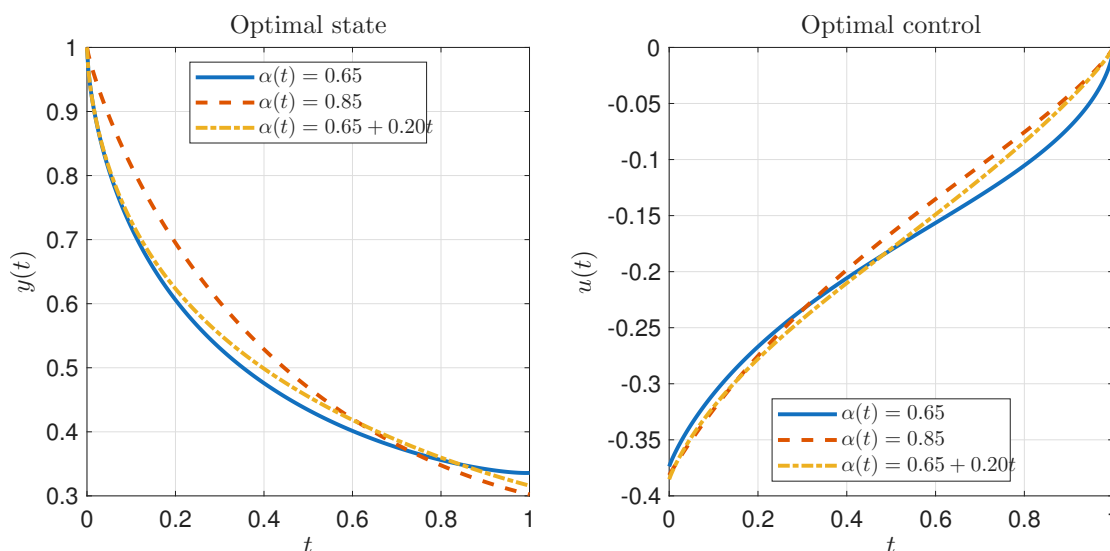


Figure 10. Comparison of optimal states and controls for different fractional orders in Example 2.

The numerical results show that the fractional order function significantly affects the optimal state, adjoint variable, and control. Compared with constant-order models, the variable-order formulation provides a more flexible mechanism for describing time-dependent memory intensity in controlled dynamical systems.

Overall, this example confirms that the fractional order should not be viewed as a merely technical parameter. It directly influences the memory kernel and therefore changes the computed optimal

control strategy.

6.3. Example 3: A nonlinear variable-order Caputo optimal control problem

We now consider a nonlinear variable-order Caputo fractional optimal control problem. The objective functional is

$$J(u) = \frac{1}{2} \int_0^1 [y^2(t) + u^2(t)] dt. \quad (6.21)$$

The state equation is given by

$${}^C D_{0+}^{\alpha(t)} y(t) = -y(t) + \frac{1}{2} y^2(t) + u(t), \quad 0 < t < 1, \quad (6.22)$$

with the initial condition

$$y(0) = 0.5. \quad (6.23)$$

The variable fractional order is chosen as

$$\alpha(t) = 0.7 + 0.1 \sin(\pi t), \quad 0 \leq t \leq 1. \quad (6.24)$$

Thus, $\alpha(t)$ varies smoothly between 0.7 and 0.8.

This example is used to test the proposed method for nonlinear dynamics. The quadratic term $\frac{1}{2}y^2(t)$ introduces a state-dependent nonlinear effect and changes the coupling between the state and adjoint equations.

The Hamiltonian is

$$H(y, p, u, t) = \frac{1}{2}y^2 + \frac{1}{2}u^2 + p \left(-y + \frac{1}{2}y^2 + u \right). \quad (6.25)$$

The stationarity condition gives

$$\frac{\partial H}{\partial u} = u + p = 0, \quad (6.26)$$

and therefore

$$u(t) = -p(t). \quad (6.27)$$

Substituting the control expression into the optimality system yields

$$\begin{cases} {}^C D_{0+}^{\alpha(t)} y(t) = -y(t) + \frac{1}{2}y^2(t) - p(t), \\ \mathcal{R}_{1-}^{\alpha(\cdot)} p(t) = y(t) + (-1 + y(t))p(t), \\ u(t) = -p(t), \\ y(0) = 0.5, \quad (\mathcal{I}_{1-}^{1-\alpha(\cdot)} p)(1) = 0. \end{cases} \quad (6.28)$$

The nonlinear coupled system (6.28) is solved by the predictor–corrector forward–backward sweep method. In the numerical implementation, a relaxation parameter equal to 0.55 is used to enhance convergence. The stopping tolerance is set to 10^{-10} .

The convergence of the cost functional under mesh refinement is reported in Table 3. The table is generated automatically by the MATLAB code.

As in Example 1, the table reports mesh-refinement information for the cost functional and the computed state and control. The results show that the numerical quantities stabilize as the grid is refined, even in the presence of the nonlinear term.

Table 3. Cost functional values for Example 3.

N	h	J_N	$ J_N - J_{2N} $	Iterations
40	2.5000e-02	0.04564681	5.452e-05	29
80	1.2500e-02	0.04559229	1.628e-05	29
160	6.2500e-03	0.04557601	4.879e-06	29
320	3.1250e-03	0.04557113	–	29

Figure 11 shows the variable fractional order used in this example. The order increases from 0.7 to 0.8 and then decreases back to 0.7, which introduces a time-varying memory effect into the controlled nonlinear system.

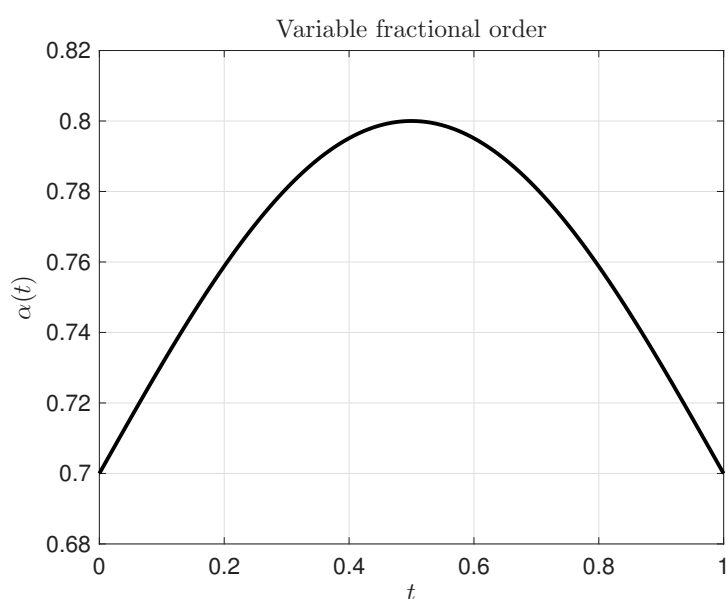


Figure 11. Variable fractional order $\alpha(t) = 0.7 + 0.1 \sin(\pi t)$ in Example 3.

The computed optimal state and optimal control for $N = 320$ are shown in Figure 12. The state remains positive and decreases smoothly, while the control acts to reduce the cost functional under the nonlinear state dynamics.

Compared with the linear case, the nonlinear term modifies the decay behavior of the state. Nevertheless, the computed profiles remain smooth, indicating that the forward–backward sweep iteration is stable for this nonlinear variable-order example under the chosen relaxation parameter.

Figure 13 presents the state, adjoint, and control in the same plot. As expected from the stationarity condition, the control is the negative of the adjoint variable.

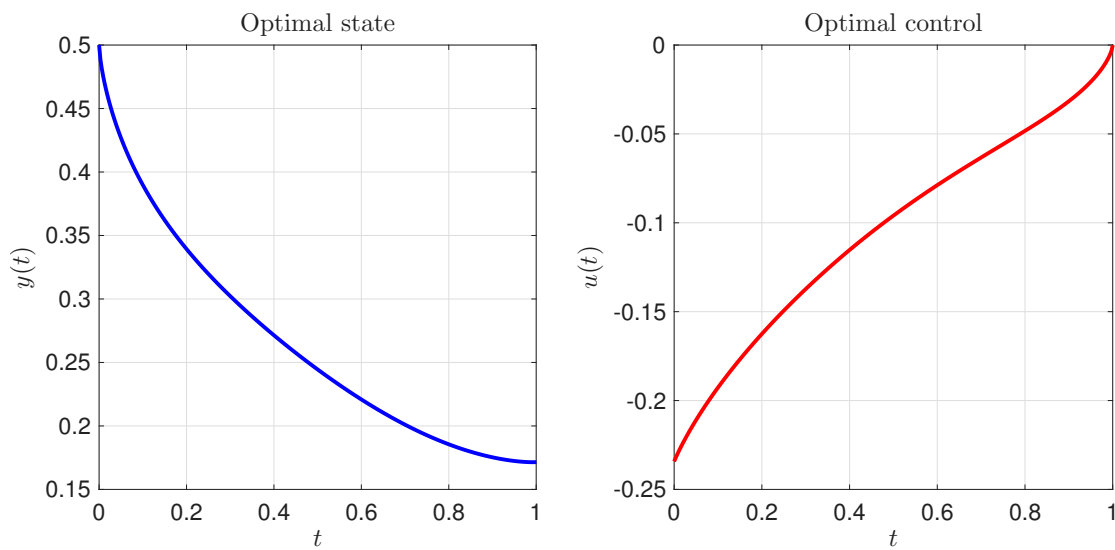


Figure 12. Optimal state and optimal control for Example 3 computed with $N = 320$.

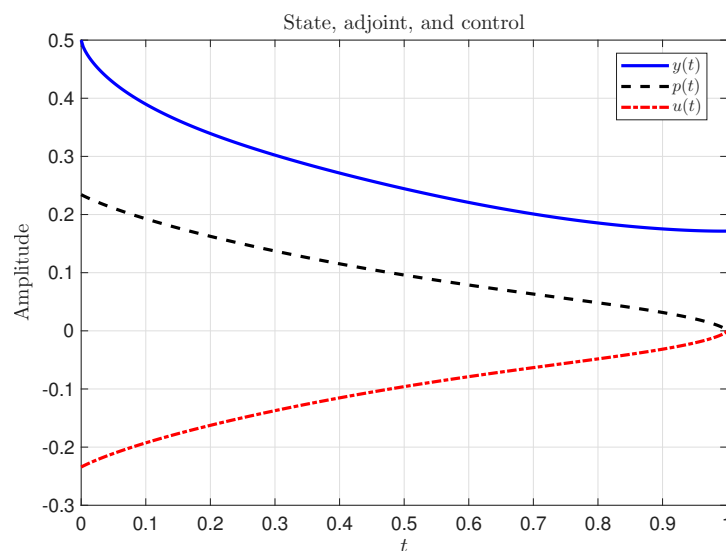


Figure 13. State, adjoint, and control profiles for Example 3 computed with $N = 320$.

The convergence of the cost functional is displayed in Figure 14. As the grid is refined, the cost functional approaches a stable limiting value.

This convergence trend provides numerical evidence that the proposed Volterra-type approximation remains effective for nonlinear variable-order Caputo dynamics.

The results of this example indicate that the proposed method can handle nonlinear state equations efficiently. The nonlinear term $\frac{1}{2}y^2(t)$ changes the state and adjoint dynamics, but the forward-backward sweep method still converges robustly when an appropriate relaxation strategy is used.

This example also confirms that the proposed method is not restricted to linear-quadratic dynamics and can be applied to nonlinear variable-order Caputo fractional optimal control problems.

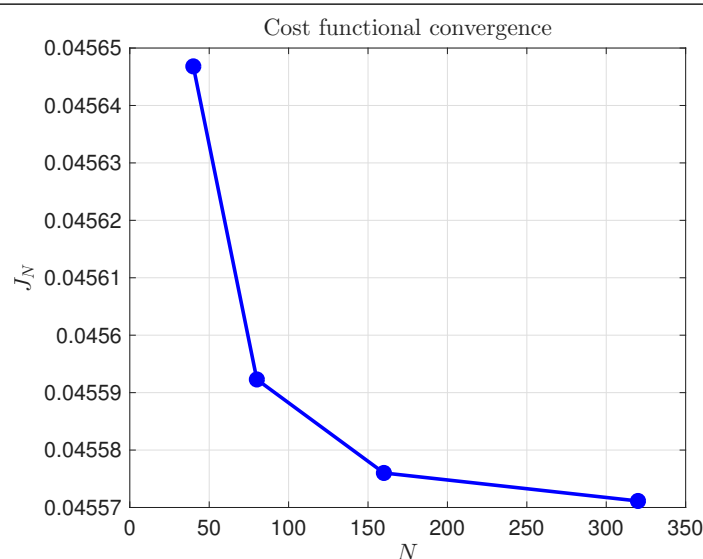


Figure 14. Convergence of the cost functional for Example 3.

7. Conclusions

In this paper, we studied a class of optimal control problems governed by variable-order Caputo fractional differential equations. The main objective was to derive a consistent adjoint-based necessary optimality system and to develop a practical numerical method for solving the resulting coupled system.

The first contribution of this work is the construction of the appropriate right-sided adjoint operator associated with the type-I variable-order Caputo derivative. Because the fractional order depends on the observation time, the dual kernel differs from the standard right-sided variable-order fractional integral. This distinction is essential for obtaining a mathematically consistent integration-by-parts formula and, consequently, a consistent adjoint equation.

The second contribution is the derivation of first-order necessary optimality conditions for the considered variable-order fractional optimal control problem. Under suitable regularity assumptions, the optimality system consists of the forward variable-order Caputo state equation, the backward right-sided adjoint equation, the stationarity condition, and the natural transversality condition. The derivation clarifies the role of the variable-order memory kernel in the adjoint dynamics.

The third contribution is the development of a predictor–corrector forward–backward sweep method for the numerical solution of the optimality system. The state equation is discretized by a forward Volterra-type product-integration scheme, while the adjoint equation is approximated by a backward Volterra-type discretization motivated by the fractional memory structure. Numerical examples, including linear-quadratic, nonlinear scalar, and coupled nonlinear systems, demonstrate the feasibility of the proposed computational approach.

The numerical experiments show that the proposed method produces stable state, adjoint, and control profiles under mesh refinement. The cost functional converges as the temporal grid is refined, and the stationarity relations are satisfied by the computed solutions. These observations provide numerical evidence for the effectiveness of the proposed method.

There are several limitations of the present work. First, the optimality conditions derived here

are necessary conditions only. Sufficient optimality conditions would require additional convexity or generalized convexity assumptions on the cost functional, the admissible set, and the dynamics. Second, the numerical approximation of the adjoint equation is based on a backward Volterra-type scheme rather than a fully discrete inversion of the exact dual-kernel adjoint operator. Third, a rigorous convergence analysis of the fully discrete nonlinear forward–backward sweep method has not been established in this paper.

Future research may proceed in several directions. One direction is to develop sufficient optimality conditions under convexity, invexity, pseudo-convexity, quasi-convexity, or other generalized convexity assumptions. Another direction is to design discretizations based directly on the dual kernel of the variable-order adjoint operator. Such methods may provide a closer correspondence between the continuous and discrete optimality systems. It is also important to investigate rigorous convergence and stability estimates for fully discrete schemes, especially for nonlinear systems and rapidly varying order functions.

Further extensions include optimal control problems with pointwise control constraints, state constraints, terminal costs, free terminal time, distributed-parameter systems, and inverse problems governed by variable-order fractional dynamics. From the computational perspective, fast convolution algorithms, sum-of-exponentials approximations, adaptive memory truncation, and parallel implementations are promising tools for reducing the cost of long-time simulations.

Overall, the results of this paper provide a theoretical and computational framework for studying variable-order Caputo fractional optimal control problems and highlight the importance of using the correct adjoint structure induced by the variable-order memory kernel.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was partially supported by the National Key R & D Program of China (No. 2024YFF0507904), the National Natural Science Foundation of China (No. 52509151), the China Postdoctoral Science Foundation (No. 2025M783290), the Taishan Scholars Program of Shandong Province (Nos. tsqn202312053 and tsqn202408001), the Shandong Provincial Natural Science Foundation (Grant No. ZR2025MS743), and the Shandong University Undergraduate Education Reform Project (Grant No. 2025Z31). The authors would like to express their gratitude to Zhiwei Yang for his invaluable guidance and unwavering support throughout this research.

Conflict of interest

The authors declare there are no conflicts of interest.

References

1. K. Diethelm, *The Analysis of Fractional Differential Equations*, Springer, Berlin, 2010. <https://doi.org/10.1007/978-3-642-14574-2>

2. A. A. Kilbas, H. M. Srivastava, J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, Elsevier, Amsterdam, 2006.
3. K. B. Oldham, J. Spanier, *The Fractional Calculus*, Academic Press, New York, 1974.
4. I. Podlubny, *Fractional Differential Equations*, Academic Press, San Diego, 1999.
5. S. G. Samko, A. A. Kilbas, O. I. Marichev, *Fractional Integrals and Derivatives: Theory and Applications*, Gordon and Breach, London, 1993.
6. H. Q. Ma, C. H. Xie, W. R. Wang, L. Y. Zhou, Z. H. Liu, T. Song, et al., Numerical modeling of wet milling process by an improved CFD–DEM–VOF model, *MetaResource*, **2** (2025), 33–55. <https://doi.org/10.23919/METAR.2025.000003>
7. Y. Tian, X. Y. Guo, Q. Zhang, Z. C. Shen, X. W. Dong, C. Z. Li, et al., A Lagrangian mesh-free multi-phase model for simulating complex dynamics of bubbly flows, *MetaResource*, **2** (2025), 11–32. <https://doi.org/10.23919/METAR.2025.000002>
8. C. F. M. Coimbra, Mechanics with variable-order differential operators, *Ann. Phys.*, **515** (2003), 692–703. <https://doi.org/10.1002/andp.200351511-1203>
9. C. F. Lorenzo, T. T. Hartley, Variable order and distributed order fractional operators, *Nonlinear Dyn.*, **29** (2002), 57–98. <https://doi.org/10.1023/A:1016586905654>
10. H. G. Sun, W. Chen, Y. Q. Chen, Variable-order fractional differential operators in anomalous diffusion modeling, *Physica A*, **388** (2009), 4586–4592. <https://doi.org/10.1016/j.physa.2009.07.024>
11. H. Hassani, Z. Avazzadeh, A. Turan-Dincel, P. Rahimkhani, A novel optimisation strategy for solving optimal control of variable-order fractional dynamic systems with nonlocal boundary conditions, *Int. J. Syst. Sci.*, **57** (2026), 1703–1717. <https://doi.org/10.1080/00207721.2025.2536210>
12. M. Alipour, S. Soradi-Zeid, Optimal control governed by nonlinear variable-order fractional integro-differential equations through generalized Dickson polynomial expansions, *J. Comput. Appl. Math.*, **473** (2026), 116917. <https://doi.org/10.1016/j.cam.2025.116917>
13. X. Zheng, H. Wang, Optimal-order error estimates of finite element approximations to variable-order time-fractional diffusion equations without regularity assumptions of the true solutions, *IMA J. Numer. Anal.*, **41** (2021), 1522–1545. <https://doi.org/10.1093/imanum/draa013>
14. M. H. Heydari, Z. Avazzadeh, A hybrid method for variable-order fractional 2D optimal control problems on an unbounded domain, *Eng. Comput.*, **38** (2022), 3237–3249. <https://doi.org/10.1007/s00366-021-01287-w>
15. M. M. A. Hasan, Optimal control problem of variable-order delay system of advertising procedure: numerical treatment, *Discrete Contin. Dyn. Syst. - Ser. S*, **15** (2022), 1247–1268. <https://doi.org/10.3934/dcdss.2021085>
16. N. Kumar, M. Mehra, Legendre wavelet method for solving variable-order nonlinear fractional optimal control problems with variable-order fractional Bolza cost, *Asian J. Control*, **25** (2023), 2122–2138. <https://doi.org/10.1002/asjc.2856>

17. F. Soufivand, F. Soltanian, K. Mamehrashi, A numerical approach for solving a class of two-dimensional variable-order fractional optimal control problems using Gegenbauer operational matrix, *IMA J. Math. Control Inf.*, **40** (2023), 1–19. <https://doi.org/10.1093/imamci/dnac031>
18. A. Singh, A. Kanaujiya, J. Mohapatra, Chelyshkov wavelet method for solving multidimensional variable order fractional optimal control problem, *J. Appl. Math. Comput.*, **70** (2024), 3135–3160. <https://doi.org/10.1007/s12190-024-02083-7>
19. N. H. Sweilam, F. Megahed, S. A. Shatta, D. Baleanu, Optimal control for a variable-order diffusion-wave equation with a reaction term: a numerical study, *Partial Differ. Equations Appl. Math.*, **10** (2024), 100658. <https://doi.org/10.1016/j.padiff.2024.100658>
20. X. Zheng, Z. Yang, W. Li, H. Wang, A time-fractional mean field control modeling subdiffusive advective transport, *SIAM J. Sci. Comput.*, **45** (2023), B884–B905. <https://doi.org/10.1137/22M1527726>
21. M. Athans, P. L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*, Dover Publications, New York, 2007.
22. A. E. Bryson, *Applied Optimal Control*, Routledge, Cambridge, 1975.
23. D. E. Kirk, *Optimal Control Theory: An Introduction*, Dover Publications, New York, 2004.
24. D. Liberzon, *Calculus of Variations and Optimal Control Theory: A Concise Introduction*, Princeton University Press, Princeton, 2012. <https://doi.org/10.1515/9781400842643>
25. O. P. Agrawal, General formulation for the numerical solution of optimal control problems, *Int. J. Control*, **50** (1989), 627–638. <https://doi.org/10.1080/00207178908953385>
26. O. P. Agrawal, O. Defterli, D. Baleanu, Fractional optimal control problems with several state and control variables, *J. Vib. Control*, **16** (2010), 1967–1976. <https://doi.org/10.1177/1077546309353361>
27. T. Akbarian, M. Keyanpour, A new approach to the numerical solution of fractional order optimal control problems, *Appl. Appl. Math.*, **8** (2013), 523–534. Available from: <https://digitalcommons.pvamu.edu/aam/vol8/iss2/12>.
28. A. Alizadeh, S. Effati, An iterative approach for solving fractional optimal control problems, *J. Vib. Control*, **24** (2018), 18–36. <https://doi.org/10.1177/1077546316633391>
29. A. Jajarmi, D. Baleanu, On the fractional optimal control problems with a general derivative operator, *Asian J. Control*, **23** (2021), 1062–1071. <https://doi.org/10.1002/asjc.2282>
30. P. Sahu, R. S. Saha, Comparison on wavelets techniques for solving fractional optimal control problems, *J. Vib. Control*, **24** (2018), 1185–1201. <https://doi.org/10.1177/1077546316659611>
31. N. Singha, C. Nahak, An efficient approximation technique for solving a class of fractional optimal control problems, *J. Optim. Theory Appl.*, **174** (2017), 785–802. <https://doi.org/10.1007/s10957-017-1143-y>

