



ARTICLE

Framework for the Structural Analysis of Fractional Differential Equations via Optimized Model Reduction

Inga Telksniene¹, Tadas Telksnys², Romas Marcinkevičius³, Zenonas Navickas², Raimondas Čiegis¹ and Minvydas Ragulskis^{2,*}

¹Mathematical Modelling Department, Faculty of Fundamental Sciences, Vilnius Gediminas Technical University, Saulėtekio al. 11, Vilnius, LT-10223, Lithuania

²Department of Mathematical Modelling, Kaunas University of Technology, Studentu 50-147, Kaunas, LT-51368, Lithuania

³Department of Software Engineering, Kaunas University of Technology, Studentu 50-415, Kaunas, LT-51368, Lithuania

*Corresponding Author: Minvydas Ragulskis. Email: minvydas.ragulskis@ktu.lt

Received: 07 September 2025; Accepted: 20 October 2025; Published: 26 November 2025

ABSTRACT: Fractional differential equations (FDEs) provide a powerful tool for modeling systems with memory and non-local effects, but understanding their underlying structure remains a significant challenge. While numerous numerical and semi-analytical methods exist to find solutions, new approaches are needed to analyze the intrinsic properties of the FDEs themselves. This paper introduces a novel computational framework for the structural analysis of FDEs involving iterated Caputo derivatives. The methodology is based on a transformation that recasts the original FDE into an equivalent higher-order form, represented as the sum of a closed-form, integer-order component $G(y)$ and a residual fractional power series $\Psi(x)$. This transformed FDE is subsequently reduced to a first-order ordinary differential equation (ODE). The primary novelty of the proposed methodology lies in treating the structure of the integer-order component $G(y)$ not as fixed, but as a parameterizable polynomial whose coefficients can be determined via global optimization. Using particle swarm optimization, the framework identifies an optimal ODE architecture by minimizing a dual objective that balances solution accuracy against a high-fidelity reference and the magnitude of the truncated residual series. The effectiveness of the approach is demonstrated on both a linear FDE and a nonlinear fractional Riccati equation. Results demonstrate that the framework successfully identifies an optimal, low-degree polynomial ODE architecture that is not necessarily identical to the forcing function of the original FDE. This work provides a new tool for analyzing the underlying structure of FDEs and gaining deeper insights into the interplay between local and non-local dynamics in fractional systems.

KEYWORDS: Fractional differential equations; Caputo derivative; fractional power series; ordinary differential equation; model reduction; structural optimization; particle swarm optimization

1 Introduction

Fractional differential equations (FDEs) have emerged as indispensable tools in mathematical modeling, extending the classical concepts of differentiation and integration to non-integer orders [1]. This generalization allows FDEs to capture complex phenomena exhibiting memory effects and non-local dynamics, which are often inadequately described by traditional integer-order models. The inherent ability of fractional operators, such as the widely used Caputo derivative [1], to incorporate system history makes FDEs particularly suitable for applications across diverse fields, including viscoelasticity [2,3], finance [4,5], biology [6,7], health sciences [8,9], control theory [10,11], signal processing [12,13] and others.



Despite their descriptive power, finding and analyzing solutions to FDEs poses significant challenges. While analytical solutions exist only for specific classes of FDEs [14,15], a broad landscape of methodologies has been developed to address various scientific objectives within the field of fractional calculus. A large portion of research includes numerical solvers, which aim to compute high-fidelity approximations of the solution with a focus on accuracy and stability [16,17]. Such methods are essential for simulating complex physical systems, for example in studies on the phase dynamics of Josephson junctions [18] or the analysis of heat transport in porous media [19], where a precise numerical result is the primary goal. Alongside these are semi-analytical methods, which often aim to bridge the gap between exact solutions and purely numerical approximations, offering computational efficiency while retaining a degree of analytical insight [20,21].

Other research directions have also been pursued. Model order reduction techniques, such as proper orthogonal decomposition [22], are data-driven methods designed to replace high-dimensional, computationally expensive mathematical models with more efficient low-dimensional surrogate models. The fundamental idea of such techniques is that the dynamics of many complex systems with large degrees of freedom often evolve on a much lower-dimensional manifold, which these methods aim to identify. The primary motivation is efficiency, particularly in multi-query or real-time applications, where repeated evaluation of the full-order model is prohibitive. Another class of methods that address a similar problem involves computational transformations, such as the use of rational approximations [23] to replace the fractional operator with a system of ODEs for the purpose of accelerating the numerical solution process. In contrast, system identification methods address an inverse problem, where the goal is to discover the governing equations of a system from measurement data. A prominent example is the sparse identification of nonlinear dynamics (SINDy) framework [24], which offers a systematic way to identify parsimonious and interpretable models from time-series data. The objective of the present work is distinct from these approaches. It is not intended to compete with numerical solvers on precision, nor with data-driven methods on model discovery. Instead, it introduces a framework for the structural analysis of a known FDE, with the goal of gaining a deeper understanding of its mathematical properties.

Among existing semi-analytical approaches, methods based on fractional power series (FPS) expansions have proven effective [25,26]. Previous research demonstrated that certain types of FDEs can be effectively analyzed using FPS expansions combined with operator calculus. Previous work [26] has shown that an FDE involving a single Caputo derivative of order $1/n$, denoted $C_{\mathbf{D}^{(1/n)}}$,

$$C_{\mathbf{D}^{(1/n)}} y(x) = F(y(x)), \quad (1)$$

can be transformed into an equivalent FDE involving the n -th iteration of this operator, $(C_{\mathbf{D}^{(1/n)}})^n$:

$$(C_{\mathbf{D}^{(1/n)}})^n y(x) = G(y(x)) + \Psi(x), \quad (2)$$

where functions $G(y(x))$ and $\Psi(x)$ depend on the structure of $F(y(x))$ and the function $\Psi(x)$ is given by the FPS with coefficients defined via a recurrence relation. This transformation is also shown to be reducible, via a nonlinear variable substitution $t = x^{1/n}$, to a first-order ordinary differential equation (ODE) that encapsulates the dynamics of the original FDE:

$$\frac{d\hat{y}}{dt} = H(t, \hat{y}(t)) + \Psi(t^n), \quad y(x) = \hat{y}(\sqrt[n]{x}). \quad (3)$$

These derivations recast the fractional problem into the domain of ODEs, allowing the application of well-established analytical or numerical techniques for ODEs. It should be noted that these transformations, while powerful, involve the emergence of auxiliary functions in the form of infinite series that require truncation for practical computation, influencing the accuracy of the resulting ODE approximation [26].

The work presented in [26] established this transformation as a deterministic, analytical procedure for deriving a single, fixed higher-order representation to facilitate a solution. The primary contribution of the present study is a conceptual extension of that work. It is important to note that the separation of the system's dynamics into an integer-order component $G(y(x))$ and a residual series $\Psi(x)$ is not necessarily unique. This ambiguity can be utilized by reframing the problem: instead of a fixed derivation, this study proposes an optimization framework that searches a space of possible architectures. This transforms the methodology from a tool for constructing a solution into a tool for structural analysis. The primary motivation for the present work is therefore to apply this transformation framework as a tool for gaining deeper structural insights into the FDE itself. In this paper, the following $(C_{\mathbf{D}(1/n)})^k$ -type FDEs are investigated:

$$(C_{\mathbf{D}(1/n)})^k y(x) = F(y(x)). \quad (4)$$

The original FDE (4) can be transformed, similarly to the derivations presented in [26], into a higher-order FDE (2), represented as the sum of a closed-form, integer-order dynamic component, denoted $G(y(x))$, and an infinite FPS, $\Psi(x)$, which captures the residual non-local effects. The resulting $(C_{\mathbf{D}(1/n)})^n$ -type FDE can subsequently be reduced to a first-order ODE (3).

The central hypothesis of this paper is that by treating the structure of $G(y(x))$ not as fixed, but as parameterizable and optimizable, it is possible to uncover the most fitting integer-order dynamics for the original fractional system. The optimization seeks an optimal ODE architecture that minimizes a dual objective: the error between the approximate and reference solutions, and the magnitude of the residual fractional series $\Psi(x)$. Analyzing the structure of the optimized ODE is expected to yield valuable insights into the interplay between local dynamics and non-local memory effects within the fractional system, revealing a more fundamental understanding of its behavior.

This paper is organized as follows. Section 2 reviews the preliminaries, including definitions of fractional power series and the transformation framework that reduces an FDE to an ODE. Section 3 details the proposed methodology, outlining the parameterization of the integer-order component and the formulation of the optimization problem. Section 4 presents the results of numerical experiments on both linear and nonlinear FDEs, including detailed parameter and sensitivity analyses. Finally, Sections 5 and 6 provide concluding remarks, summarize the findings, and discuss the implications for the structural analysis of fractional systems.

2 Preliminaries

This section establishes the mathematical foundation necessary for the subsequent analysis. Key definitions related to fractional calculus, FPS, and the specific type of FDEs under investigation are presented.

2.1 Fractional Power Series Algebra

FPS provide a convenient framework to analyze the solutions of the FDEs considered in this paper [25]. Solutions $y(x)$ are represented as series expansions involving fractional powers of the independent variable x .

2.1.1 Fractional Power Series

A FPS centered at $x = 0$ with a fractional exponent step $1/n$ ($n \in \mathbb{N}$) is defined as:

$$y(x) = \sum_{j=0}^{\infty} c_j x^{j/n}, \quad c_j \in \mathbb{R}. \quad (5)$$

This series is assumed to converge for $0 \leq x < R$ for some radius of convergence $R > 0$. For operational convenience, particularly when applying the Caputo derivative, it is often useful to express this series using a specific basis [25]:

$$y(x) = \sum_{j=0}^{\infty} v_j w_j^{(n)}, \quad \text{where} \quad w_j^{(n)} = \frac{x^{j/n}}{\Gamma(1 + j/n)}. \quad (6)$$

The coefficients are related by $v_j = c_j \Gamma(1 + j/n)$. The set of such convergent series is denoted by ${}^C\mathbb{F}^{(1/n)}$.

2.1.2 Fractional Power Series Operations

Addition of two FPS as well as multiplication of FPS by a scalar is defined in a standard way, while multiplication of two FPS $f(x) = \sum b_j w_j^{(n)} \in {}^C\mathbb{F}^{(1/n)}$ and $g(x) = \sum c_j w_j^{(n)} \in {}^C\mathbb{F}^{(1/n)}$ is defined in the Cauchy sense:

$$f(x) \cdot g(x) = \sum_{j=0}^{\infty} \left(\sum_{k=0}^j \binom{j/n}{k/n} b_k c_{j-k} \right) w_j^{(n)}, \quad (7)$$

using the property

$$w_j^{(n)} w_k^{(n)} = \binom{(j+k)/n}{k/n} w_{j+k}^{(n)}. \quad (8)$$

The binomial coefficient of non-integer arguments α, β is defined as:

$$\binom{\alpha}{\beta} = \frac{\Gamma(\alpha + 1)}{\Gamma(\beta + 1) \Gamma(\alpha - \beta + 1)}. \quad (9)$$

A Caputo differentiation operator of order $\frac{1}{n}$ acts on basis elements $w_j^{(n)}$ as [25]:

$${}_{\mathcal{D}}^{(1/n)} w_j^{(n)} = \begin{cases} 0, & \text{if } j = 0 \\ w_{j-1}^{(n)}, & \text{if } j \geq 1. \end{cases} \quad (10)$$

Thus, the iterated application of the Caputo derivative of order $\frac{1}{n}$ follows naturally:

$$({}_{\mathcal{D}}^{(1/n)})^k w_j^{(n)} = \begin{cases} 0, & \text{if } j < k \\ w_{j-k}^{(n)}, & \text{if } j \geq k. \end{cases} \quad (11)$$

This property simplifies the application of fractional operators to series solutions.

2.2 Transformation of the FDE and Reduction to the ODE

2.2.1 Transformation Methodology

Previous work demonstrated a methodology for transforming FDEs involving the operator ${}_{\mathcal{D}}^{(1/n)}$

$${}_{\mathcal{D}}^{(1/n)} y(x) = F(y(x)), \quad (12)$$

subject to $y(0) = v_0$, into an equivalent higher-order FDE involving the n -th iteration of the operator, $({}_{\mathcal{D}}^{(1/n)})^n$ [26].

This transformation utilizes the properties of FPS and introduces auxiliary relations to manage the fractional differentiation of nonlinear functions of $y(x)$. For example, when dealing with a quadratic term like y^2 and $n = 2$, the application of the operator requires a specific relation:

$$C_{\mathbf{D}^{(1/2)}}(z^2) = 2z(C_{\mathbf{D}^{(1/2)}}z) + \Theta_z(x), \quad (13)$$

where $\Theta_z(x)$ is an auxiliary FPS whose coefficients depend on the coefficients of $z(x)$. This relation accounts for the deviation from the integer-order chain rule. Using such auxiliary relations derived within the FPS framework, the original FDE (12) can be systematically transformed into the $(C_{\mathbf{D}^{(1/n)}})^n$ -type FDE:

$$(C_{\mathbf{D}^{(1/n)}})^n y(x) = G(y(x)) + \Psi(x). \quad (14)$$

Here, $G(y(x))$ is a closed-form function, while $\Psi(x)$ is the resulting infinite FPS that encapsulates the residual effects [26].

For a general analytic nonlinearity $F(y)$, where $y(x) = \sum v_j w_j^{(n)}$, the function $F(y(x))$ also has an FPS representation. The auxiliary relations are derived by applying the operator $C_{\mathbf{D}^{(1/n)}}$ to this series and relating it to the derivative of the argument, $C_{\mathbf{D}^{(1/n)}} y(x)$. This process, which utilizes the Cauchy product rule for FPS multiplication, systematically separates the resulting expression into terms that depend on $y(x)$ and its fractional derivatives, and a residual FPS, such as $\Theta_z(x)$ in (13). While the specific form of the auxiliary relation depends on the function F , the procedure is generalizable to any analytic function, thereby providing a deterministic method to perform the transformation for a given FDE.

2.2.2 Reduction to a First-Order ODE

Subsequently, the resulting $(C_{\mathbf{D}^{(1/n)}})^n$ -type FDE is reducible to a first-order ODE via the substitution:

$$t = x^{1/n}. \quad (15)$$

Letting $\widehat{y}(t) = y(t^n)$, the equivalent ODE takes the form:

$$\frac{d\widehat{y}}{dt} = nt^{n-1} (G(\widehat{y}(t)) + \Psi(t^n)) + \sum_{j=1}^{n-1} \frac{jv_j}{\Gamma(1 + \frac{j}{n})} t^{j-1}, \quad (16)$$

with $\widehat{y}(0) = v_0$. The summation term arises from the initial conditions associated with the lower-order fractional derivatives $v_j = (C_{\mathbf{D}^{(1/n)}})^j y(x)|_{x=0}$ for $j = 1, \dots, n-1$ [26].

For computation, the infinite series $\Psi(t^n)$ is truncated to

$$\Psi_M(t^n) = \sum_{j=0}^M \psi_j t^j / \Gamma(1 + j/n), \quad (17)$$

which introduces an approximation error, the magnitude of which depends on the truncation order M and the convergence properties of $\Psi(x)$.

3 Methodology

Building upon the established concepts, this section outlines the methodology for analyzing the more general $(C_{\mathbf{D}^{(1/n)}})^k$ -type FDE, focusing on the exploration of optimal ODE architectures.

3.1 FDE Transformation

Consider the FDE involving the k -th iteration of the Caputo derivative operator of order $\frac{1}{n}$ ($1 \leq k < n$):

$$(C_{\mathbf{D}^{(1/n)}})^k y(x) = F(y(x)), \quad (18)$$

subject to initial conditions

$$y(0) = v_0, (C_{\mathbf{D}^{(1/n)}})^j y(x)|_{x=0} = v_j, \quad (19)$$

for $j = 1, \dots, k-1$. Here, F is assumed to be an analytic function of y .

This equation can be transformed into the $(C_{\mathbf{D}^{(1/n)}})^n$ -type form by applying the $C_{\mathbf{D}^{(1/n)}}$ operator $n-k$ times, similarly to the $k=1$ case described earlier. This process again involves managing the fractional differentiation of $F(y(x))$, potentially requiring the derivation and use of auxiliary relations analogous to (13) for the specific nonlinearities present in $F(y)$.

Performing this transformation leads to an equation structure:

$$(C_{\mathbf{D}^{(1/n)}})^n y(x) = G(y(x)) + \Psi(x). \quad (20)$$

This transformation process, however, reveals an ambiguity. The separation of the system's dynamics into a closed-form, integer-order component $G(y)$ and a residual fractional power series $\Psi(x)$ is not necessarily unique. For instance, consider a hypothetical transformed FDE where $(C_{\mathbf{D}^{(1/n)}})^n y(x) = y(x)^2 + \sin(x)$. One could make the straightforward choice of $G(y) = y^2$ and $\Psi(x) = \sin(x)$. However, an alternative choice, such as $G(y) = y^2 + C$ and $\Psi(x) = \sin(x) - C$ for some constant C , is equally valid. This inherent flexibility in the separation of dynamics into the closed-form term $G(y)$ and the infinite power series term $\Psi(x)$ suggests that multiple valid architectures (G, Ψ) might exist that represent the original FDE (18).

Previous work in this area has focused on deriving a single, valid transformation to facilitate a solution, without exploring the implications of this non-uniqueness. For instance, the methodology presented in [26] was fundamentally analytical and relied on deriving specific auxiliary relations to establish a deterministic, step-by-step procedure for converting an FDE into a higher-order form. The primary objective of that approach was to prove the transformation's viability for the construction of the solution. Consequently, the resulting integer-order component $G(y)$ and the residual series $\Psi(x)$ were treated as fixed outcomes of this analytical derivation. The focus was on the integrity of the transformation itself, viewing the higher-order equation as a necessary intermediate step towards finding a solution. The structural properties of that intermediate equation were not the object of study.

This paper shifts that perspective by exploiting the aforementioned ambiguity. Instead of manually deriving one possible architecture based on preselected analytical relations, this work treats the structure of $G(y)$ not as a fixed result, but as a flexible and optimizable term. This approach allows for the use of global optimization to search the space of possible architectures and identify one (G^*, Ψ^*) that, when used in the corresponding truncated ODE approximation, most accurately reflects the dynamics of the original FDE (18).

3.2 Reduction to ODE and Truncation

Regardless of the specific choice of $G(y)$ and its corresponding $\Psi(x)$ in (20), the reduction to a first-order ODE proceeds via the substitution $t = x^{1/n}$. Letting $\widehat{y}(t) = y(t^n)$, the resulting ODE is:

$$\frac{d\widehat{y}}{dt} = nt^{n-1} (G(\widehat{y}(t)) + \Psi(t^n)) + \sum_{j=1}^{n-1} \frac{jv_j}{\Gamma(1 + \frac{j}{n})} t^{j-1}, \quad \widehat{y}(0) = v_0. \quad (21)$$

The initial conditions $v_1, \dots, v_{k-1}$ are given by (19), while $v_k, \dots, v_{n-1}$ are determined during the transformation process (e.g., $v_k = F(v_0)$, etc.).

For computation, the infinite series $\Psi(t^n)$ is truncated at an order M :

$$\Psi_M(t^n) = \sum_{j=0}^M \psi_j \frac{t^j}{\Gamma(1 + j/n)}, \quad (22)$$

leading to the approximate ODE:

$$\frac{d\widetilde{y}}{dt} \approx nt^{n-1} (G(\widetilde{y}(t)) + \Psi_M(t^n)) + \sum_{j=1}^{n-1} \frac{jv_j}{\Gamma(1 + \frac{j}{n})} t^{j-1}, \quad \widetilde{y}(0) = v_0, \quad (23)$$

where $\widetilde{y}(t) \approx y(t^n)$.

This truncation is a primary source of approximation error. The core idea is that by optimizing the structure of $G(y)$, one can identify a representation that both possesses a smaller residual fractional series $\Psi(x)$, and is more accurate, yielding a smaller error in the final solution upon truncation.

The choice of the truncation order M used in this study is based on both the theoretical properties of the residual series and practical considerations of convergence. Theoretically, the coefficients ψ_j of the residual series $\Psi(x)$ are constructed from the coefficients v_j of the solution's fractional power series. Given that the solution $y(x)$ to the FDE is assumed to be analytic on the domain of interest, its fractional power series representation is convergent. This implies that the coefficients v_j , and consequently the derived coefficients ψ_j , must decay sufficiently rapidly to ensure that the truncation error diminishes as M increases. From a practical standpoint, the specific value of $M = 15$ was used throughout this study which was determined empirically during preliminary investigations. It was consistently observed across the tested FDEs that increasing the truncation order beyond this value resulted in negligible improvements to the quality of the optimization. This indicates that the series had effectively converged for the purposes of the performed experiments. Therefore, $M = 15$ was selected as a computationally efficient choice that robustly balances the accuracy of the approximation with the complexity of the resulting ordinary differential equation.

The step-by-step process described in this section is illustrated in Fig. 1.

3.3 Optimization Framework

To explore the range of possible architectures and find an optimal one, the function $G(y)$ is parameterized as a polynomial of degree N_G :

$$G(y; \mathbf{b}) = \sum_{p=0}^{N_G} b_p y^p, \quad (24)$$

where $\mathbf{b} = (b_0, b_1, \dots, b_{N_G})$ is the vector of coefficients to be optimized. The choice of a polynomial basis for parameterizing $G(y)$ is motivated by several factors. From a theoretical standpoint, the Weierstrass

approximation theorem suggests that polynomials can uniformly approximate any continuous function on a closed interval, making them a versatile choice for representing the unknown integer-order dynamics. From a practical standpoint, polynomials offer a straightforward parameterization via their coefficients and are computationally tractable, simplifying both the construction of the approximate ODE and the evaluation of the objective function. While other basis functions, such as rational functions or trigonometric series, could be employed in future extensions of this work, polynomials provide a simple and interpretable starting point for this foundational study, where the resulting low-degree models offer clear insights into the system's dominant dynamics.

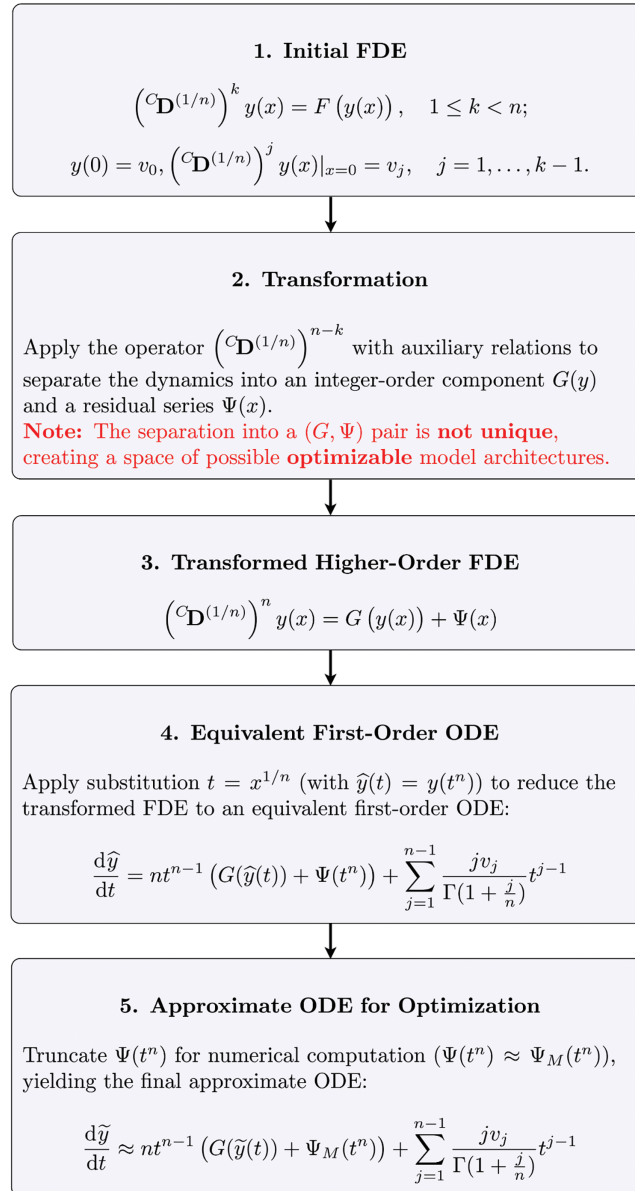


Figure 1: The step-by-step process described in [Section 3.2](#) which is used in obtaining an approximate ODE for a given FDE

The overall optimization procedure is structured as follows:

1. **Define the search space:** For a given polynomial degree N_G , the coefficients $\mathbf{b}$ define a potential ODE architecture. For any chosen $\mathbf{b}$, the corresponding infinite power series term $\Psi(x; \mathbf{b})$ is implicitly defined by:

$$\Psi(x; \mathbf{b}) = (C_{D(1/n)})^n y(x) - G(y(x); \mathbf{b}) = \sum_{j=0}^{\infty} \psi_j(\mathbf{b}) w_j^{(n)}. \quad (25)$$

The coefficients $\psi_j(\mathbf{b})$ of this series depend on the coefficients v_j of the FPS solution $y(x)$, which are determined by the original FDE (18), and the chosen parameters $\mathbf{b}$. Specifically, if

$$(C_{D(1/n)})^n y(x) = \sum_{j=0}^{\infty} v_{j+n} w_j^{(n)}, \quad (26)$$

and

$$G(y; \mathbf{b}) = \sum_{j=0}^{\infty} g_j(\mathbf{b}) w_j^{(n)}, \quad (27)$$

then

$$\psi_j(\mathbf{b}) = v_{j+n} - g_j(\mathbf{b}). \quad (28)$$

2. **Compute quality metrics:** The goal is to find a parameter vector $\mathbf{b}^*$ that simultaneously minimizes two often competing objectives: solution accuracy and structural simplicity. This constitutes a multi-objective optimization problem, also known as a Pareto optimization problem. A common and effective method for solving such problems, adopted in this work, is the weighted sum method, which incorporates the multiple objectives into a single target function. The two objectives are defined as follows:

- The solution accuracy, quantified by the Root Mean Square Error (RMSE) between the approximate solution $y(x; \mathbf{b}, M)$ and a high-fidelity reference solution $y_{\text{ref}}(x)$:

$$\text{RMSE}(\mathbf{b}) = \sqrt{\frac{1}{P} \sum_{j=1}^P (y(x_j; \mathbf{b}, M) - y_{\text{ref}}(x_j))^2}, \quad (29)$$

where the error is evaluated over a discrete set of P points, $\{x_j\}_{j=1}^P$, within the solution interval. The reference $y_{\text{ref}}(x)$ can be an exact solution if known, or a numerical solution computed using established numerical methods.

- The model's structural simplicity, quantified by the magnitude of the truncated residual series, measured by its L^2 -norm, $\|\Psi_M(\mathbf{b})\|_2$:

$$\|\Psi_M(\mathbf{b})\|_2 = \sqrt{\int_0^{t_{\max}} (\Psi_M(t; \mathbf{b}))^2 dt}, \quad (30)$$

where $t_{\max} = X^{1/n}$ is the upper limit of the integration interval corresponding to the FDE's domain $[0, X]$.

3. **Establish a baseline for normalization:** To create a balanced optimization problem from these two potentially different-scaled objectives, a dynamic normalization strategy is employed. First, a baseline

parameter vector, $\mathbf{b}_{\text{baseline}}$, is defined to be identical to the analytical structure of the original FDE's forcing function, $F(y)$. The objectives are evaluated once at this baseline to obtain reference values:

$$\text{RMSE}_{\text{ref}} = \text{RMSE}(\mathbf{b}_{\text{baseline}}); \quad \|\Psi_M\|_{\text{ref}} = \|\Psi_M(\mathbf{b}_{\text{baseline}})\|_2. \quad (31)$$

For example, given a linear forcing function of the form $F(y) = a_0 + a_1 y$, the baseline is chosen as $\mathbf{b}_{\text{baseline}} = [a_0, a_1, 0, \dots, 0]$, fixing the initial guess for the polynomial $G(y)$ to be identical to $F(y)$. The metrics defined in the previous point are evaluated to obtain reference values, RMSE_{ref} and $\|\Psi_M\|_{\text{ref}}$.

4. **Define the final objective function:** The final, normalized scalar objective function $T_{\text{norm}}(\mathbf{b})$ to be minimized is a weighted sum of the normalized components:

$$\text{Minimize}_{\mathbf{b}} \quad T_{\text{norm}}(\mathbf{b}) = \kappa_1 \frac{\text{RMSE}(\mathbf{b})}{\text{RMSE}_{\text{ref}}} + \kappa_2 \frac{\|\Psi_M(\mathbf{b})\|_2}{\|\Psi_{M,\text{ref}}\|_2}. \quad (32)$$

In this study, balanced weights $\kappa_1 = \kappa_2 = 0.5$ are used. Since each term in the objective function is normalized by its respective baseline value, the two components are rendered dimensionless and of a comparable magnitude. This makes equal weighting a natural and unbiased choice, ensuring that the optimization seeks a balanced improvement relative to a consistent, problem-derived baseline without a priori favoring either solution accuracy or structural simplicity.

5. **Perform global optimization:** A global optimization algorithm, Particle Swarm Optimization (PSO) [27], is used to find the optimal parameter vector $\mathbf{b}^*$ that minimizes $T_{\text{norm}}(\mathbf{b})$. PSO algorithm is widely recognized for its strong practical performance in finding high-quality solutions and its effectiveness in navigating complex, non-convex search spaces often found in engineering problems [28]. Its implementation is relatively straightforward and it requires tuning fewer parameters compared to other metaheuristics, making it a robust and practical choice for this foundational work. Note that as a stochastic heuristic, PSO does not mathematically guarantee convergence to the global optimum, and a formal proof of convergence for such methods is a topic of ongoing research beyond the scope of this foundational study. Moreover, given the likely non-convex nature of the optimization landscape defined by the objective function $T_{\text{norm}}(\mathbf{b})$, the existence of multiple local optima is possible. However, the framework does not aim to prove the existence of a single, unique optimal integer-order structure. Instead, its objective is to identify a parsimonious and high-performing ODE architecture that provides valuable structural insight. The consistency of the results across different parameter spaces, as demonstrated in following Sections, suggests that the method robustly identifies meaningful structures rather than spurious local minima.

While other algorithms like genetic algorithms or simulated annealing could also be applied, a comparative study of different optimizers was considered beyond the scope of this work, which aims to demonstrate the feasibility of the overall framework. It is important to note that the normalized objective function $T_{\text{norm}}(\mathbf{b})$ serves as a balanced internal metric to guide the optimizer within a specific problem configuration. However, for the purpose of comparing the method's performance across different FDE configurations (e.g., with different parameters or initial conditions), the final, un-normalized physical metrics—the RMSE and the L^2 -norm of the truncated residual fractional series—are used as the primary basis for assessment and comparison.

To summarize the procedure presented in section succinctly, see the steps outlined in Fig. 2.

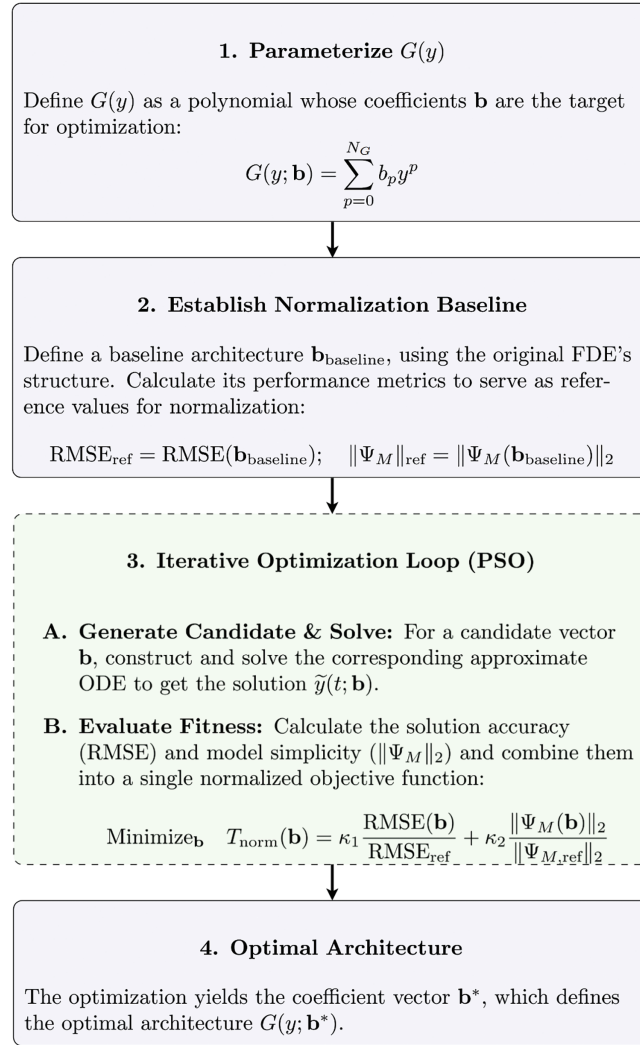


Figure 2: The step-by-step process described in Section 3.3, outlining the procedure for determining optimal $\mathbf{b}^*$ values which define part of the architecture of the FDE

3.4 Computational Complexity

The computational complexity of the proposed framework is primarily determined by the global optimization process, as the initial FDE-to-ODE transformation is performed only once. The dominant source of computational cost is the iterative evaluation of the objective function $T_{\text{norm}}(\mathbf{b})$ within the PSO algorithm. The total cost can be approximated as the product of the number of PSO particles, the number of iterations, and the cost of a single objective function evaluation.

The complexity of a single evaluation for a candidate parameter vector $\mathbf{b}$ is composed of three main parts. First, the approximate ODE is constructed by calculating the coefficients of the truncated residual series $\Psi_M(t''; \mathbf{b})$. The cost of this step depends on the polynomial degree of the integer-order component N_G and the truncation order of the series M . Second, the approximate ODE (23) is solved using a standard numerical integrator, where the cost is dependent on the solver's efficiency, the required tolerance, and the stiffness of the resulting ODE over the integration interval. Third, the quality metrics are calculated, where

the RMSE is computed over P discrete points, resulting in a cost proportional to P , and the L^2 -norm of the residual series requires a numerical integration.

It is important to contextualize this computational cost. The optimization procedure is a computationally intensive, offline analysis performed once for a given FDE structure. The objective of this one-time investment is not to produce a single numerical solution more efficiently, but rather to obtain a simplified, integer-order ODE architecture that reveals fundamental properties of the FDE's dynamics. The analytical advantage gained from this process—a deeper insight into the interplay between local and non-local behaviors—is the primary justification for the proposed computational framework.

4 Results

This section presents numerical experiments designed to illustrate the proposed methodology for finding optimal ODE architectures. The approach is applied to two distinct cases: a linear FDE with a known exact solution and a nonlinear (quadratic) FDE where a high-fidelity numerical solution serves as the reference. The primary goal is to demonstrate the optimization process and analyze the resulting ODE structures, particularly the optimized polynomial $G(y; \mathbf{b}^*)$ and its relationship to the original FDE dynamics.

4.1 Case 1: Linear FDE

First, a linear FDE, involving the iterated Caputo derivative $(C_{\mathbf{D}(1/3)})^2$, corresponding to $n = 3, k = 2$ in the general formulation (18), is considered:

$$(C_{\mathbf{D}(1/3)})^2 y(x) = a_0 + a_1 y(x), \quad (33)$$

subject to initial conditions $y(0) = v_0$ and $(C_{\mathbf{D}(1/3)} y)(0) = v_1$.

This FDE possesses an exact analytical solution expressed using the Mittag-Leffler functions [29]:

$$y_{\text{ref}}(x) = v_0 E_{\frac{2}{3}, 1}(a_1 x^{\frac{2}{3}}) + v_1 x^{\frac{1}{3}} E_{\frac{2}{3}, \frac{4}{3}}(a_1 x^{\frac{2}{3}}) + a_0 x^{\frac{2}{3}} E_{\frac{2}{3}, \frac{5}{3}}(a_1 x^{\frac{2}{3}}). \quad (34)$$

This exact solution serves as the high-fidelity reference $y_{\text{ref}}(x)$ for calculating the RMSE in the optimization process.

The methodology transforms the FDE (33) into the $(C_{\mathbf{D}(1/3)})^3$ -type FDE

$$(C_{\mathbf{D}(1/3)})^3 y(x) = G(y(x); \mathbf{b}) + \Psi(x; \mathbf{b}), \quad (35)$$

which is then reduced to the approximate ODE (23) using the parameterized polynomial $G(y; \mathbf{b}) = \sum_{p=0}^{N_G} b_p y^p$ and the truncated term $\Psi_M(t^n; \mathbf{b})$. The optimization seeks the coefficients $\mathbf{b}^*$ for the polynomial $G(y; \mathbf{b})$ that minimize the normalized target function (32). Key parameters used in the experiments are summarized in Table 1.

Table 1: Parameters for computational experiments

Parameter	Value
<i>Problem and method parameters</i>	
Domain	$x \in [0, 1]$
Polynomial degrees tested (N_G)	$\{1, 2, 3, 4\}$
Residual fractional series truncation order (M)	15

(Continued)

Table 1 (continued)

Parameter	Value
RMSE evaluation points (P)	1000
Optimization weights (κ_1, κ_2)	0.5, 0.5
<i>Numerical solver parameters</i>	
ODE solver tolerances (Rel/Abs)	1.0×10^{-13}
PSO swarm size	100
PSO max iterations	300
PSO function tolerance	1.0×10^{-13}
PSO search bounds for $\mathbf{b}$	$[-10, 10]$

First, the case

$$a_0 = -1.5, a_1 = 0.5, v_0 = 1.5, v_1 = 2 \quad (36)$$

is analyzed. The results, summarized in Table 2, show that the case of $N_G = 2$ represents the optimal trade-off in model complexity for this problem configuration. It provides a substantial improvement over the $N_G = 1$ model, reducing the normalized target value by over 90% and the un-normalized RMSE by a factor of approximately 50. This demonstrates that a quadratic model for $G(y)$ is more effective at capturing the intrinsic integer-order dynamics of the fractional system.

Table 2: Optimization results for different polynomial degrees N_G

N_G	$T_{\text{norm}}(\mathbf{b}^*)$	Optimal $\mathbf{b}^*$	RMSE($\mathbf{b}^*$)	$\ \Psi_M(\mathbf{b}^*)\ _2$
1	3.82×10^{-1}	$[1.243, -0.107]$	6.70×10^{-7}	1.19×10^{-1}
2	2.41×10^{-2}	$[0.824, 0.050, 0.002]$	1.28×10^{-8}	4.31×10^{-2}
3	2.63×10^0	$[-8.245, 7.660, -1.346, -0.049]$	4.56×10^{-6}	8.81×10^{-1}
4	1.76×10^1	$[-9.022, 5.712, 4.119, -1.458, -0.073]$	2.84×10^{-5}	8.47×10^0

Conversely, increasing the complexity further to $N_G = 3$ and $N_G = 4$ results in a significant degradation in performance. The final target values for these models are two and three orders of magnitude worse than that of the $N_G = 2$ model, respectively, with corresponding increases in both the final RMSE and residual fractional series norm. This indicates that higher-degree polynomials do not yield a better representation for this system. Such a decline in performance suggests that the optimization process either encounters a more challenging, high-dimensional parameter space or the models begin to overfit the reference solution.

Moreover, Fig. 3 provides a visual representation of the optimized residual fractional series, which corroborates the numerical results in Table 2. The plot clearly shows that the residual fractional series for the $N_G = 1$ and $N_G = 2$ models are well-behaved, remaining close to zero across the entire interval. In stark contrast, the residual series for the higher-complexity models ($N_G = 3$ and $N_G = 4$) are substantially larger and display significant growth, which is consistent with their much larger L^2 -norms. This visualization reinforces the conclusion that the quadratic architecture is superior, as it not only minimizes the solution error but also produces the most compact residual fractional series.

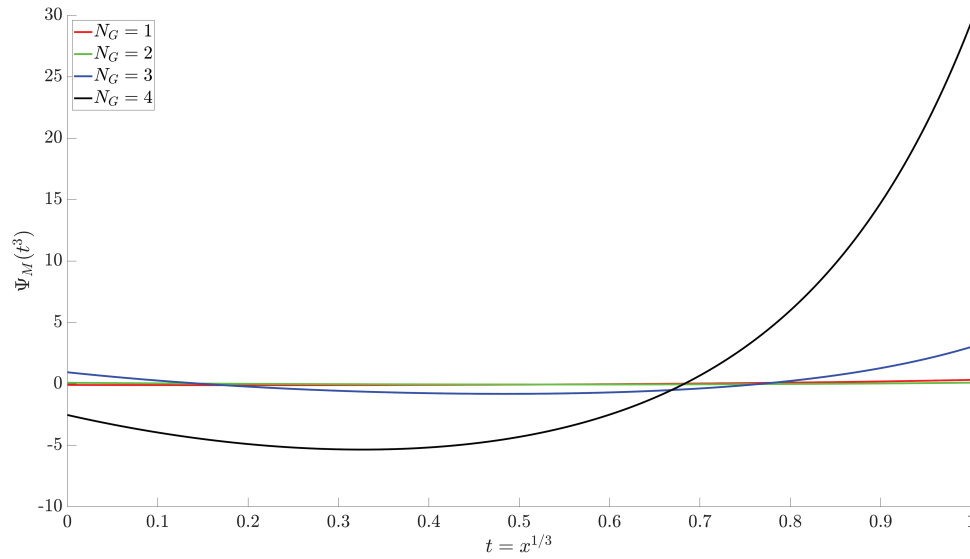


Figure 3: Optimized residual fractional series $\Psi_M(t^3)$ for the linear FDE for polynomial degrees $N_G = 1, 2, 3$, and 4 . The plot shows that the quadratic model ($N_G = 2$) yields a small and well-behaved residual series. The significant growth of the residuals for higher-degree models ($N_G \geq 3$) could be an indicator of overfitting, which corresponds to the poor performance metrics reported in [Table 2](#)

These findings are significant, as they demonstrate the framework's capacity to identify an optimal level of complexity, beyond which adding more parameters is detrimental. The analysis reveals that the underlying integer-order dynamics of this specific fractional system are most effectively represented not by a simple linear model, but by a quadratic one.

4.1.1 Parameter Space Analysis of Optimal Model Complexity

To further investigate the robustness of the optimal model architecture, a parameter space analysis is conducted. The optimization procedure is systematically applied across a grid of FDE coefficients, with a_0 and a_1 each varying over the range $[-1, 1]$. For each pair of (a_0, a_1) coefficients, the optimal polynomial degree $N_G \in \{1, 2, 3, 4\}$ is determined by identifying the model that yields the minimum normalized target value, T_{norm} . The initial conditions are held constant at $v_0 = 1.5$ and $v_1 = 2$, and all other computational parameters are maintained as specified in [Table 1](#). Note that the objective of this analysis is to demonstrate the stability of the optimal architecture across a representative, bounded domain of system parameters, rather than to conduct a computationally exhaustive global exploration. Naturally, a broader analysis of the parameter space would be a valuable direction for future work.

The results of this meta-analysis are summarized in [Table 3](#), which presents a map of the optimal polynomial degree, N_G^* , for each (a_0, a_1) pair. The analysis reveals a consistent pattern: the quadratic model ($N_G = 2$) is the optimal choice in 28 out of the 36 cases tested, representing over 77% of the parameter space.

A closer examination of the cases where $N_G = 2$ is not the optimal choice provides further insight. In the instances where a linear model ($N_G = 1$) is selected, a trade-off between solution accuracy and the magnitude of the residual series is typically observed. For example, for the configuration $(a_0, a_1) = (-0.2, -1.0)$, the optimal $N_G = 1$ model produced a smaller residual series norm ($\|\Psi_M\|_2 = 6.55 \times 10^{-1}$) compared to the $N_G = 2$ model ($\|\Psi_M\|_2 = 6.92 \times 10^{-1}$). Conversely, the $N_G = 2$ model achieved a slightly lower RMSE (4.15×10^{-6}) than the $N_G = 1$ model (4.86×10^{-6}). This indicates that even in regions where the simpler

linear model is marginally preferred by the normalized target function, the quadratic model remains a highly competitive alternative.

Table 3: Optimal model complexity N_G^* as a function of the FDE coefficients a_0 and a_1

a_0	a_1					
	-1.0	-0.6	-0.2	0.2	0.6	1.0
-1.0	2	1	3	2	2	2
-0.6	2	1	2	2	2	2
-0.2	1	1	2	2	2	2
0.2	2	2	2	2	2	2
0.6	2	2	2	2	2	2
1.0	2	2	2	2	2	2

The single instance where $N_G = 3$ is optimal, at $(a_0, a_1) = (-1.0, -0.2)$, also warrants comment. For this configuration, the final metrics for the $N_G = 3$ model ($\text{RMSE} = 6.353 \times 10^{-9}$, $\|\Psi_M\|_2 = 5.229 \times 10^{-2}$) are nearly identical to those of the $N_G = 2$ model ($\text{RMSE} = 6.356 \times 10^{-9}$, $\|\Psi_M\|_2 = 5.229 \times 10^{-2}$), confirming that the performance difference is negligible.

Collectively, these findings strongly support the conclusion that a quadratic ODE architecture provides a robust and broadly applicable representation for the linear fractional differential equation under study. The stability of this result across a wide range of system parameters suggests that the method is not merely finding isolated numerical artifacts but is successfully identifying a fundamental, underlying integer-order structure inherent to the fractional system.

To better understand the structure of the quadratic architecture ($N_G = 2$), we now examine how its optimal coefficients $\mathbf{b}^* = (b_0, b_1, b_2)$ depend on the original FDE parameters. Fig. 4 visualizes this relationship across the same intervals for a_0 and a_1 as given in Table 3. Even though the discretization is coarse, it can be seen that the values of b_0, b_1, b_2 do not oscillate wildly and depend smoothly on a_0 and a_1 .

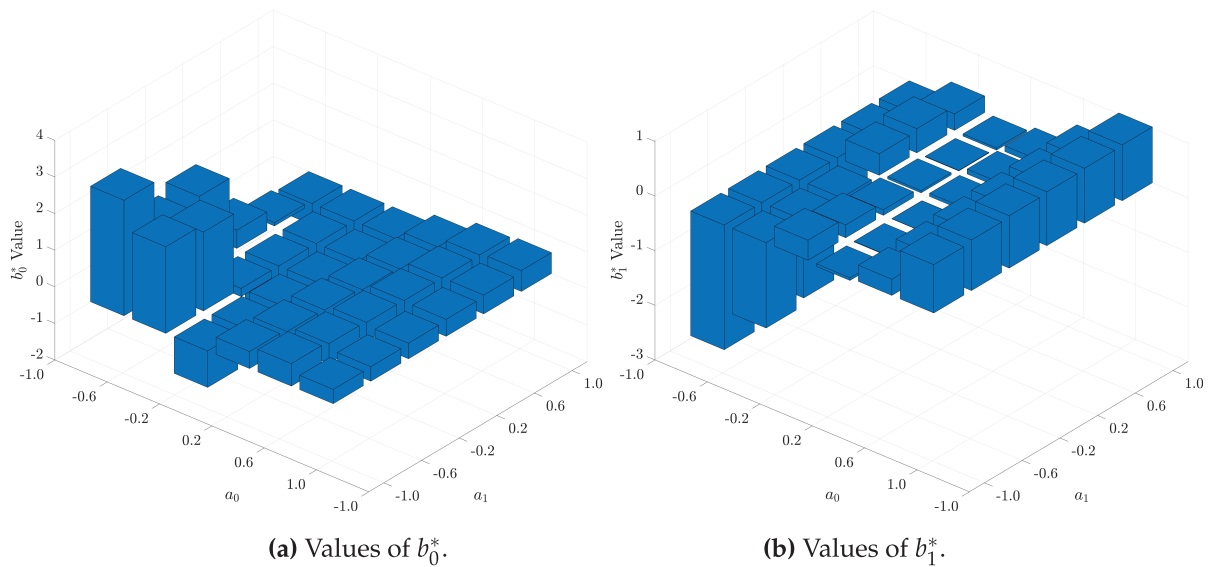


Figure 4: (Continued)

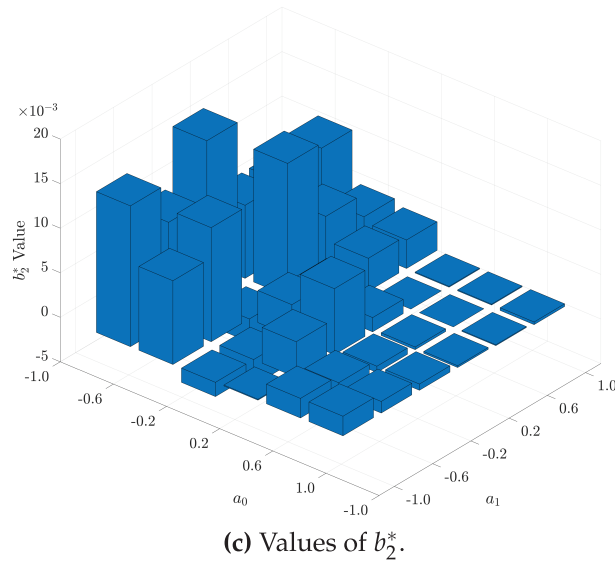


Figure 4: The optimal coefficients $\mathbf{b}^* = (b_0^*, b_1^*, b_2^*)$ for the quadratic ($N_G = 2$) architecture as the FDE coefficients a_0, a_1 vary in the interval $[-1, 1]$. It can be seen that the optimal coefficients act as smooth, continuous functions of the FDE parameters, confirming the stability of the identified integer-order structure

4.1.2 Effects of Initial Condition on the Complexity of the Model

To assess whether the optimal ODE architecture is an intrinsic property of the FDE or is dependent on the specific solution trajectory, a second meta-analysis has been performed. In this set of experiments, the FDE coefficients are held constant at $a_0 = -1.5$ and $a_1 = 0.5$, while the initial conditions $v_0 = y(0)$ and $v_1 = (C_D^{(1/3)} y)(0)$ are varied over the range $[-2, 2]$. All other computational parameters are maintained as specified in Table 1. The same optimization and dynamic normalization procedure is applied to find the optimal polynomial degree N_G^* for each pair of initial conditions.

The results, presented in Table 4, demonstrate a high degree of stability in the optimal model architecture. Across the 36 different initial conditions tested, the quadratic model ($N_G = 2$) is identified as the optimal choice in 33 cases, representing approximately 92% of the tested space. This consistency suggests that the optimal ODE representation is largely independent of the specific solution trajectory and is instead determined by the inherent structure of the FDE itself.

Table 4: Optimal model complexity N_G^* as a function of the initial conditions v_0 and v_1

v_0	v_1					
	-2.0	-1.2	-0.4	0.4	1.2	2.0
-2.0	2	2	2	2	2	1
-1.2	2	2	2	2	2	1
-0.4	2	2	2	2	2	2
0.4	2	2	2	2	1	2
1.2	2	2	2	2	2	2
2.0	2	2	2	1	2	2

Notably, the few instances where $N_G = 1$ has been selected as optimal occurred at or close to the boundaries of the tested parameter space. In these cases, the linear model is found to be genuinely superior across both performance metrics. For example, for the initial conditions $(v_0, v_1) = (-2.0, 2.0)$, the winning $N_G = 1$ model produces both a lower RMSE and a smaller residual series norm than the $N_G = 2$ model. However, the overwhelming dominance of the quadratic model across the vast majority of the parameter space remains the principal finding. Furthermore, the higher-complexity models ($N_G = 3$ and $N_G = 4$) consistently underperform, often yielding results that are orders of magnitude worse than the simpler models.

The combined results from the parameter space analyses of both the FDE coefficients and the initial conditions provide strong evidence for the central hypothesis. The methodology consistently identifies a quadratic ODE as the most suitable integer-order representation for the given linear FDE. This stability across different system parameters and solution trajectories indicates that the optimization framework is successfully uncovering a fundamental, intrinsic property of the fractional system's dynamics, rather than artifacts of a specific configuration.

4.1.3 Sensitivity Analysis on the Truncation Order of the Residual of the Fractional Series

A sensitivity analysis has been conducted to evaluate the influence of the residual fractional series truncation order, M , on the performance of the optimization framework. For this study, the FDE coefficients and initial conditions are held constant at the values used in the first example ($a_0 = -1.5$, $a_1 = 0.5$, $v_0 = 1.5$, $v_1 = 2$), while the parameter M is varied from 5 to 20. All other computational parameters are maintained as specified in Table 1. The primary objective is to determine if the choice of the optimal model complexity, N_G^* , is stable with respect to the accuracy of the residual fractional series approximation.

The results, summarized in Table 5, reveal a distinct trend. For lower values of M (specifically, $M \leq 7$), the optimization procedure has selected a cubic polynomial ($N_G = 3$) as the optimal architecture. However, as M increases to 8 and beyond, the optimal model complexity consistently stabilizes at the quadratic model ($N_G = 2$).

Table 5: Optimal model complexity N_G^* and corresponding final un-normalized metrics as a function of the residual fractional series truncation order M

M	Optimal N_G^*	Final RMSE	Final $\ \Psi_M\ _2$
5	3	5.37×10^{-6}	3.26×10^{-2}
6	3	1.71×10^{-6}	3.68×10^{-2}
7	3	5.87×10^{-7}	3.72×10^{-2}
8	2	2.19×10^{-6}	4.15×10^{-2}
9	2	6.39×10^{-7}	4.04×10^{-2}
10	2	3.09×10^{-7}	4.24×10^{-2}
11	2	9.20×10^{-8}	4.18×10^{-2}
12	2	4.32×10^{-8}	4.30×10^{-2}
13	2	1.76×10^{-8}	4.26×10^{-2}
14	2	1.38×10^{-8}	4.33×10^{-2}
15	2	1.28×10^{-8}	4.31×10^{-2}
16	2	1.27×10^{-8}	4.35×10^{-2}
17	2	1.27×10^{-8}	4.34×10^{-2}
18	2	1.27×10^{-8}	4.37×10^{-2}

(Continued)

Table 5 (continued)

M	Optimal N_G^*	Final RMSE	Final $\ \Psi_M\ _2$
19	2	1.27×10^{-8}	4.36×10^{-2}
20	2	1.27×10^{-8}	4.38×10^{-2}

This transition in the optimal model complexity is a significant finding. It suggests that when the residual fractional series $\Psi_M(t^n)$ is poorly approximated (i.e., when M is small), the optimization framework compensates by selecting a more complex polynomial $G(y)$ to capture the dynamics that are missing from the truncated series. As the series approximation becomes more accurate with increasing M , this compensation is no longer necessary, and the framework is able to identify the underlying quadratic structure of the system.

Furthermore, an examination of the final un-normalized metrics reveals the practical limits of increasing the residual series order. While the final RMSE of the optimal solution improves substantially as M increases from 8 to 15, it converges to approximately 1.27×10^{-8} for $M \geq 15$. Similarly, the final residual series norm, $\|\Psi_M\|_2$, stabilizes around a value of 4.3×10^{-2} . This indicates that beyond a certain point ($M \approx 15$ for this problem), further increasing the number of terms in the residual fractional series approximation yields no discernible improvement in the quality of the final solution.

In conclusion, this sensitivity analysis demonstrates the robustness of the proposed methodology. It shows that provided the residual fractional series is approximated with a sufficient number of terms, the optimization consistently converges to the same optimal ODE architecture. This reinforces the conclusion that the framework is successfully identifying an intrinsic and stable integer-order representation of the fractional system.

4.2 Case 2: Nonlinear Riccati FDE

To demonstrate the methodology's applicability to nonlinear problems without known analytical solutions, the framework is applied to a fractional Riccati FDE. The equation, corresponding to $n = 3, k = 2$ in the general formulation, is given by:

$$(C_{\mathbf{D}^{(1/3)}})^2 y(x) = a_0 + a_1 y(x) + a_2 y(x)^2, \quad (37)$$

subject to initial conditions $y(0) = v_0$ and $(C_{\mathbf{D}^{(1/3)}} y)(0) = v_1$. An exact analytical solution for (37) is not known. Therefore, a high-fidelity numerical solution obtained using Garrappa's method [30] serves as the reference solution $y_{\text{ref}}(x)$ for the RMSE calculation during the optimization.

The same optimization framework as in the linear case is used. The specific FDE parameters and computational settings for this experiment are detailed in Table 6.

The optimization results for polynomial degrees N_G from 1 to 5 are presented in Table 7 and Fig. 5. The analysis reveals a different pattern of results compared to the linear FDE. In this case, the optimization identifies the cubic model ($N_G = 3$) as the optimal architecture, achieving the lowest target value of 3.43×10^{-2} . This optimal result is driven by the model's exceptional solution accuracy; its final RMSE of 1.27×10^{-8} is more than two orders of magnitude lower than that of the next best competitor, the $N_G = 2$ model.

Table 6: Parameters for the Nonlinear Riccati FDE Experiment

Parameter	Value
<i>Problem parameters</i>	
FDE coefficients (a_0, a_1, a_2)	$2/5, -3/5, 1/5$
Initial conditions (v_0, v_1)	$1/10, 1/2$
Domain	$x \in [0, 1]$
<i>Method and solver parameters</i>	
Reference solver step size (h)	1.0×10^{-6}
ODE solver tolerances (Rel/Abs)	1.0×10^{-13}
RMSE evaluation points (P)	1000
Polynomial degrees tested (N_G)	$\{1, 2, 3, 4, 5\}$
Residual fractional series truncation order (M)	15
Optimization weights (κ_1, κ_2)	0.5, 0.5
PSO swarm size/Max iterations	100/300
PSO function tolerance	1.0×10^{-13}
PSO search bounds for $\mathbf{b}$	$[-10, 10]$

Table 7: Optimization results for the nonlinear Riccati FDE

N_G	$T_{\text{norm}}(\mathbf{b}^*)$	Optimal $\mathbf{b}^*$	RMSE($\mathbf{b}^*$)	$\ \Psi_M(\mathbf{b}^*)\ _2$
1	4.41×10^{-1}	$[-0.326, 0.096]$	3.42×10^{-4}	1.11×10^{-2}
2	4.47×10^{-2}	$[-0.411, 0.622, -0.625]$	7.44×10^{-6}	3.14×10^{-2}
3	3.43×10^{-2}	$[-0.397, 0.511, -0.373, -0.172]$	1.27×10^{-8}	3.04×10^{-2}
4	1.31×10^{-1}	$[-0.739, 2.231, -1.782, -0.735, -0.217]$	2.13×10^{-6}	1.14×10^{-1}
5	2.80×10^{-1}	$[0.376, -5.323, 9.998, -1.223, -3.719, -2.874]$	1.52×10^{-4}	7.93×10^{-2}

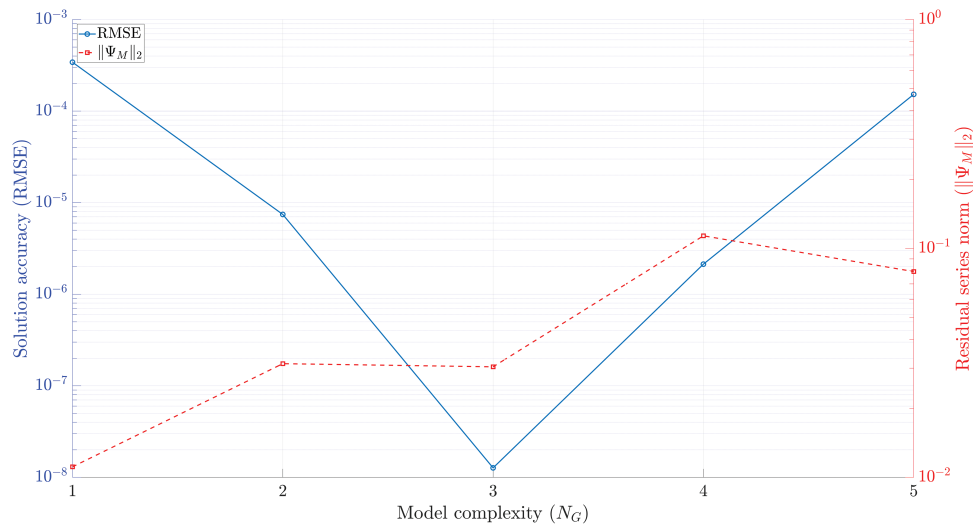


Figure 5: The complexity-accuracy trade-off curve for the nonlinear Riccati FDE. The plot shows the final solution accuracy (RMSE, left axis) and the norm of the residual series ($\|\Psi_M\|_2$, right axis) as a function of the model complexity (N_G). The optimal performance is achieved at $N_G = 3$

The emergence of an optimal architecture that is of a higher polynomial degree than the original FDE's nonlinearity is a significant finding. This suggests that the transformation from the FDE to the ODE, which involves the action of a fractional operator on the nonlinear term, generates a richer set of dynamics. The optimization framework discovers that it is more effective to capture the dominant part of these emergent dynamics within the integer-order component $G(y)$ using a cubic term, rather than relegating them to the residual series $\Psi(x)$. This choice results in a model that achieves a substantially lower solution error, providing insight into the complex interplay between the fractional operator and the system's inherent nonlinearity.

A trade-off between solution accuracy and the magnitude of the residual fractional series is also observed. While the linear model ($N_G = 1$) produced the smallest residual fractional series, its significantly higher RMSE resulted in a much poorer overall target value. This illustrates a case where the optimizer correctly favored the model with superior accuracy, even at the cost of a marginally larger residual series. The optimized residual fractional series are plotted in Fig. 6. The visualization confirms that the residual series for the $N_G = 1$ model has the smallest overall magnitude, consistent with its minimal L^2 -norm, while the residual series for the more accurate $N_G = 2$ and $N_G = 3$ models are similarly well-behaved and small in magnitude. In contrast, the residual series for $N_G = 4$ and $N_G = 5$ are significantly larger and more oscillatory, providing a visual explanation for their poor performance.

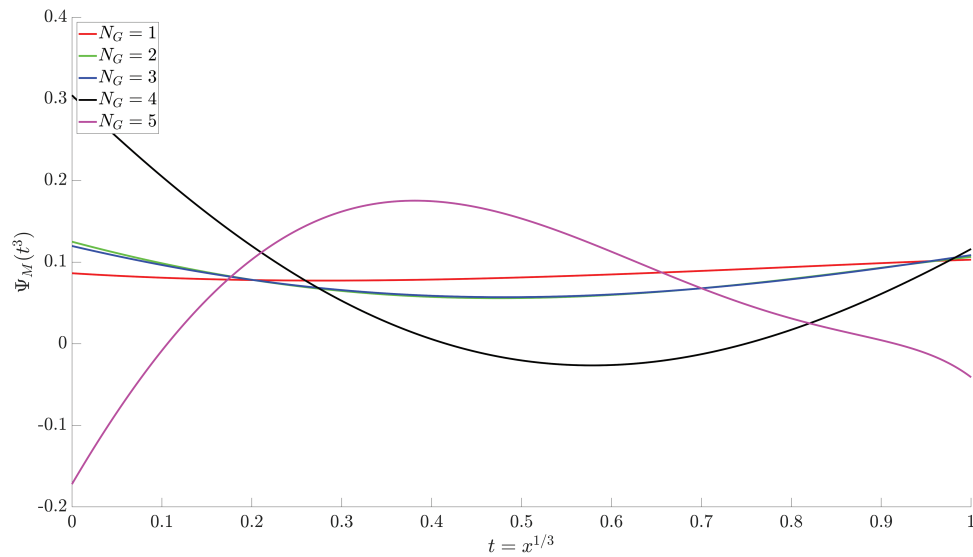


Figure 6: Optimized residual fractional series $\Psi_M(t^3)$ for the Riccati FDE for polynomial degrees $N_G = 1, \dots, 5$. The visualization confirms that while the $N_G = 1$ model yields the smallest residual, the optimal $N_G = 3$ model remains well-behaved. The large and oscillatory residuals for the higher-degree models ($N_G \geq 4$) provide a visual explanation for their suboptimal performance

Consistent with the findings from the linear analysis, increasing the model complexity beyond an optimal point ($N_G > 3$) is detrimental. This result reinforces the conclusion that the methodology is capable of identifying an optimal level of complexity for the approximating ODE, which is sensitive to the underlying structure of the FDE. The primary significance of this experiment is the demonstration that the framework is not limited to linear FDEs, but serves as a versatile tool for the structural analysis of a broader class of fractional differential equations.

4.2.1 Effect of the Integration Horizon

To verify the effectiveness and generality of the proposed framework over different integration intervals, the experiment on the nonlinear Riccati FDE was repeated for two longer domains: $x \in [0, 2]$ and $x \in [0, 3]$. The objective of this analysis is to determine if the optimal model architecture identified in the previous section is a robust structural property of the FDE or an artifact of the specific integration horizon. All FDE and computational parameters are kept identical to those in Table 6, with only the domain of the problem being extended. The reference solution for the RMSE calculation on the extended domains was again obtained using Garrappa's method.

The optimization results for the extended domains are presented in Tables 8 and 9. The analysis reveals a consistency in the optimal architecture. For both the $x \in [0, 2]$ and $x \in [0, 3]$ intervals, the cubic model ($N_G = 3$) is unambiguously identified as the superior choice, achieving the lowest normalized target value in each case. While the absolute RMSE values naturally increase as the integration interval becomes longer, the optimizer's structural preference remains stable. This finding strongly supports the conclusion that the identified cubic architecture is not a feature of a specific problem configuration but rather an intrinsic property of the underlying dynamics of the fractional Riccati equation.

Table 8: Optimization results for the nonlinear Riccati FDE on the domain $x \in [0, 2]$

N_G	$T_{\text{norm}}(\mathbf{b}^*)$	Optimal $\mathbf{b}^*$	RMSE($\mathbf{b}^*$)	$\ \Psi_M(\mathbf{b}^*)\ _2$
1	5.26×10^{-1}	$[-0.303, 0.026]$	2.67×10^{-2}	4.30×10^{-2}
2	7.14×10^{-2}	$[-0.480, 0.872, -0.745]$	6.57×10^{-4}	5.47×10^{-2}
3	6.07×10^{-2}	$[-0.443, 0.705, -0.518, -0.097]$	3.05×10^{-4}	5.07×10^{-2}
4	1.48×10^{-1}	$[-0.142, -1.288, 2.887, -1.724, -0.040]$	5.85×10^{-3}	3.98×10^{-2}
5	6.22×10^{-1}	$[0.813, -7.537, 9.523, 6.045, -5.619, -5.423]$	2.12×10^{-2}	2.21×10^{-1}

Table 9: Optimization results for the nonlinear Riccati FDE on the domain $x \in [0, 3]$

N_G	$T_{\text{norm}}(\mathbf{b}^*)$	Optimal $\mathbf{b}^*$	RMSE($\mathbf{b}^*$)	$\ \Psi_M(\mathbf{b}^*)\ _2$
1	6.13×10^{-1}	$[-1.976, 2.964]$	9.44×10^{-3}	9.37×10^{-1}
2	6.15×10^{-2}	$[-0.532, 1.056, -0.848]$	8.64×10^{-3}	7.52×10^{-2}
3	3.91×10^{-2}	$[-0.459, 0.642, -0.296, -0.178]$	6.12×10^{-4}	5.98×10^{-2}
4	1.79×10^{-1}	$[-0.212, -0.787, 1.662, -1.069, -0.066]$	6.82×10^{-2}	1.14×10^{-1}
5	3.79×10^{-1}	$[1.103, -7.764, 9.728, 1.800, -2.486, -2.772]$	7.90×10^{-2}	4.01×10^{-1}

4.2.2 Comparison with an Alternative Optimization Algorithm

To ensure that the findings of this study are not specific to the choice of the optimization algorithm, the experiment on the nonlinear Riccati FDE presented at the beginning of this section (with the integration interval $x \in [0, 1]$) was repeated using an alternative global optimizer. The Genetic Algorithm (GA), which is another well-established, population-based metaheuristic, was chosen for this comparative analysis. The GA was configured with parameters comparable to those used for the PSO; specifically, the population size was set to 100 and the maximum number of generations to 300, mirroring the swarm size and maximum iterations of the PSO, respectively. All other FDE and method parameters were kept identical to those listed in Table 6.

The optimization results obtained using the GA are presented in Table 10 and are consistent with the PSO results from Table 7. GA also unambiguously identifies the cubic model ($N_G = 3$) as the optimal architecture, achieving the lowest normalized target value. While the final numerical metrics differ slightly between the two algorithms, which is a common occurrence given their stochastic nature, the central conclusion remains the same. This successful replication of the result with a different optimizer provides strong evidence that the identified cubic architecture is a genuine feature of the FDE's underlying dynamics and not an artifact of the optimization method used to find it.

Table 10: Optimization results for the nonlinear Riccati FDE using the GA

N_G	$T_{\text{norm}}(\mathbf{b}^*)$	Optimal $\mathbf{b}^*$	RMSE($\mathbf{b}^*$)	$\ \Psi_M(\mathbf{b}^*)\ _2$
1	4.41×10^{-1}	$[-0.326, 0.096]$	3.42×10^{-4}	1.11×10^{-2}
2	4.47×10^{-2}	$[-0.411, 0.622, -0.625]$	7.44×10^{-6}	3.14×10^{-2}
3	3.64×10^{-2}	$[-0.400, 0.533, -0.424, -0.137]$	1.51×10^{-6}	3.06×10^{-2}
4	1.50×10^{-1}	$[-0.633, 1.551, -0.463, -1.469, -0.197]$	2.91×10^{-5}	1.01×10^{-1}
5	6.84×10^{-1}	$[-0.617, 0.571, 1.707, 3.033, -5.874, -4.142]$	3.41×10^{-4}	2.27×10^{-1}

4.2.3 Sensitivity Analysis of the Optimal Solution

To assess the robustness of the optimization outcomes, a sensitivity analysis was performed on the optimal solution found by the PSO algorithm for the nonlinear Riccati FDE (see Table 7). The objective of this analysis is to determine if the identified optimal architecture is stable with respect to small variations in the FDE's parameters. For this study, the optimal cubic architecture ($N_G = 3$) and its corresponding coefficient vector $\mathbf{b}^*$ were held fixed. The performance of this fixed model was then evaluated on a set of perturbed FDEs, where the coefficient a_1 was varied in increments of 1% over the range of $\pm 5\%$ from its original value. For each perturbed FDE, a new high-fidelity reference solution was computed to ensure an accurate RMSE calculation.

The results of this analysis are presented in Table 11. The results show that the optimal solution exhibits structural stability, with performance metrics degrading continuously with respect to parameter perturbations. Both the final RMSE and the norm of the residual series are minimized for the unperturbed, original problem and increase monotonously and symmetrically as the perturbation magnitude grows. This demonstrates that the optimal solution represents a stable optimum within a local neighborhood of the parameter space, confirming the robustness of the optimization outcome.

Table 11: Performance of the fixed optimal cubic model on perturbed Riccati FDEs

Perturbation	Value of a_1	RMSE($\mathbf{b}^*$)	$\ \Psi_M(\mathbf{b}^*)\ _2$
-5%	-0.5700	1.25×10^{-6}	3.49×10^{-2}
-4%	-0.5760	1.03×10^{-6}	3.34×10^{-2}
-3%	-0.5820	7.94×10^{-7}	3.22×10^{-2}
-2%	-0.5880	5.44×10^{-7}	3.12×10^{-2}
-1%	-0.5940	2.79×10^{-7}	3.07×10^{-2}
Original	-0.6000	1.27×10^{-8}	3.04×10^{-2}
+1%	-0.6060	3.00×10^{-7}	3.05×10^{-2}
+2%	-0.6120	6.14×10^{-7}	3.09×10^{-2}
+3%	-0.6180	9.46×10^{-7}	3.16×10^{-2}

(Continued)

Table 11 (continued)

Perturbation	Value of a_1	RMSE($\mathbf{b}^*$)	$\ \Psi_M(\mathbf{b}^*)\ _2$
+4%	−0.6240	1.30×10^{-6}	3.27×10^{-2}
+5%	−0.6300	1.66×10^{-6}	3.39×10^{-2}

5 Discussion

The effectiveness of the presented approach was demonstrated on both a linear FDE and a nonlinear fractional Riccati equation. These cases were selected as representative examples to establish the feasibility of the proposed structural analysis on both a linear problem with a known analytical solution and a nonlinear problem requiring a numerical reference solution. The fractional Riccati equation, in particular, serves as a relevant nonlinear model due to its importance in applied mathematics for modeling real-world phenomena, including population dynamics and tumor growth [31], where the inclusion of memory effects via fractional operators is an active area of research to improve model realism [32]. The numerical results consistently show that the method can identify an optimal, low-degree polynomial ODE that accurately captures the dynamics of the fractional system. For the linear FDE, a quadratic architecture has been found to be robustly optimal across a wide range of system parameters and initial conditions, revealing an underlying structure more complex than the original linear forcing term. For the nonlinear Riccati FDE, a cubic architecture has been identified as optimal.

Notably, the results often indicate a significant decline in performance when the model complexity is increased beyond the optimal degree which could indicate the model's sensitivity to over-parameterization. The primary goal of this work is to introduce the novel concept of structural optimization for FDEs and demonstrate its feasibility on representative examples. Thus, in this foundational study, objective function is designed to measure only the core metrics of interest: solution accuracy and the magnitude of the residual series. It does not contain an explicit regularization term that would penalize model complexity. However, while the framework minimizes the error against the reference, the dual nature of the objective function, which also penalizes $\|\Psi_M(\mathbf{b})\|_2$, acts as an implicit regularizer promoting structural simplicity. Thus, while the absence of explicit regularization can make the optimization landscape for over-parameterized models more challenging, this sensitivity is advantageous for model identification. The resulting performance decline acts as a clear and unambiguous signal against overfitting, effectively confirming that the identified low-degree architecture is not merely a good approximation, but the most robust representation of the system's underlying integer-order dynamics. This paves the way for future research. A natural extension of this work would be to incorporate regularization techniques into the framework. Doing so would likely stabilize the optimization for higher-degree polynomials by promoting sparse solutions, potentially transforming the proposed tool into an even more powerful predictive one.

A crucial distinction must be drawn between the proposed framework and other established methodologies. While high-fidelity numerical methods provide accurate solutions, they yield a numerical result without offering deeper insight into the underlying structure of the FDE. It is not the goal of this study to compete with these methods in terms of raw numerical precision, but to introduce a novel tool for structural analysis. By identifying an optimal, low-degree integer-order ODE, our framework provides a simplified yet powerful lens through which to understand the intrinsic dynamics of the fractional system. It emphasizes the interplay between local, integer-order behavior (captured by $G(y)$) and non-local, memory-dependent effects (captured by $\Psi(x)$). Moreover, it addresses a forward problem of equation analysis, which is fundamentally different from the inverse problem solved by data-driven discovery methods like SINDy,

where the governing equation itself is the unknown to be found from data. This structural understanding provided by the proposed method is a valuable analytical complement to purely numerical solutions, offering a new perspective on the fundamental properties of fractional systems.

6 Conclusion

This paper introduces a novel computational framework for identifying optimal integer-order ODE representations for FDEs containing iterated Caputo derivatives. The methodology utilizes a transformation that recasts the FDE into a higher-order form composed of an integer-order dynamic component $G(y)$ and a residual fractional power series $\Psi(x)$. By parameterizing $G(y)$ as a polynomial and employing a global optimization algorithm, the framework seeks an architecture that simultaneously minimizes the solution error and the magnitude of the residual series. Presented findings underscore the framework's ability to discern an appropriate level of complexity and avoid overfitting, thereby providing genuine insight into the system's intrinsic integer-order dynamics.

In summary, this work provides a new tool for analyzing the underlying structure of FDEs and gaining deeper insights into the interplay between local and non-local dynamics. The primary limitations of this foundational study are its focus on a specific class of FDEs with iterated Caputo derivatives and the use of a polynomial basis for parameterization, which may not be optimal for all systems.

Future research could extend this framework to broader classes of FDEs, including its scalability to high-dimensional and coupled systems of equations or those involving other fractional operators (such as the Atangana-Baleanu or Riemann-Liouville operators, etc.). Furthermore, exploring alternative parameterizations for the integer-order component, such as rational functions, may reveal even richer structural properties of fractional systems.

Acknowledgement: This research was funded by the Research Council of Lithuania.

Funding Statement: This project has received funding from the Research Council of Lithuania (LMTLT), agreement No. S-PD-24-120.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Tadas Telksnys, Inga Telksniene, Minvydas Ragulskis and Raimondas Čiegis; methodology, Tadas Telksnys, Inga Telksniene, Minvydas Ragulskis and Romas Marcinkevičius; software, Inga Telksniene and Romas Marcinkevičius; validation, Inga Telksniene, Tadas Telksnys and Romas Marcinkevičius; formal analysis, Minvydas Ragulskis, Zenonas Navickas, Tadas Telksnys, Inga Telksniene and Raimondas Čiegis; investigation, Inga Telksniene, Tadas Telksnys, Romas Marcinkevičius, Zenonas Navickas, Minvydas Ragulskis and Raimondas Čiegis; writing—original draft preparation, Inga Telksniene, Tadas Telksnys and Minvydas Ragulskis; writing—review and editing, Inga Telksniene, Tadas Telksnys, Romas Marcinkevičius, Zenonas Navickas, Minvydas Ragulskis and Raimondas Čiegis; supervision, Minvydas Ragulskis and Raimondas Čiegis; funding acquisition, Inga Telksniene and Raimondas Čiegis. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Not applicable.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

Abbreviations

FDE	Fractional Differential Equation
ODE	Ordinary Differential Equation

FPS	Fractional Power Series
RMSE	Root Mean Square Error
PSO	Particle Swarm Optimization
GA	Genetic Algorithm

References

- Podlubny I. Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications. San Diego, CA, USA: Academic Press; 1999.
- Suzuki JL, Kharazmi E, Varghaei P, Naghibolhosseini M, Zayernouri M. Anomalous nonlinear dynamics behavior of fractional viscoelastic beams. *J Comput Nonlinear Dyn*. 2021;16(11):111005. doi:10.1115/1.4052286.
- Loghman E, Bakhtiari-Nejad F, Kamali A, Abbaszadeh M, Amabili M. Nonlinear vibration of fractional viscoelastic micro-beams. *Int J Non Linear Mech*. 2021;137(1–4):103811. doi:10.1016/j.ijnonlinmec.2021.103811.
- Ali I, Khan SU. A dynamic competition analysis of stochastic fractional differential equation arising in finance via pseudospectral method. *Mathematics*. 2023;11(6):1328. doi:10.3390/math11061328.
- Li B, Zhang T, Zhang C. Investigation of financial bubble mathematical model under fractal-fractional Caputo derivative. *Fractals*. 2023;31(05):2350050. doi:10.1142/s0218348x23500500.
- Abdeljawad T, Sher M, Shah K, Sarwar M, Amacha I, Alqudah M, et al. Analysis of a class of fractal hybrid fractional differential equation with application to a biological model. *Sci Rep*. 2024;14(1):18937. doi:10.1038/s41598-024-67158-8.
- Singh J, Ganbari B, Kumar D, Baleanu D. Analysis of fractional model of guava for biological pest control with memory effect. *J Adv Res*. 2021;32(1):99–108. doi:10.1016/j.jare.2020.12.004.
- Chatterjee AN, Ahmad B. A fractional-order differential equation model of COVID-19 infection of epithelial cells. *Chaos Solitons Fract*. 2021;147:110952.
- Srivastava H, Jan R, Jan A, Deebani W, Shutaywi M. Fractional-calculus analysis of the transmission dynamics of the dengue infection. *Chaos*. 2021;31(5):053130. doi:10.1063/5.0050452.
- Baleanu D, Sajjadi SS, Jajarmi A, Deftferli Ö. On a nonlinear dynamical system with both chaotic and non-chaotic behaviors: a new fractional analysis and control. *Adv Differ Equ*. 2021;2021(1):234. doi:10.1186/s13662-021-03393-x.
- Jajarmi A, Baleanu D. On the fractional optimal control problems with a general derivative operator. *Asian J Control*. 2021;23(2):1062–71. doi:10.1002/asjc.2282.
- Zhang X, Ding F. Optimal adaptive filtering algorithm by using the fractional-order derivative. *IEEE Signal Process Lett*. 2021;29(1):399–403. doi:10.1109/lsp.2021.3136504.
- Jackson J, Perumal R. A robust image encryption technique based on an improved fractional order chaotic map. *Nonlinear Dyn*. 2025;113(7):7277–96. doi:10.1007/s11071-024-10480-7.
- Ebaid A, Cattani C, Al Juhani AS, El-Zahar ER. A novel exact solution for the fractional Ambartsumian equation. *Adv Differ Equ*. 2021;2021(1):88. doi:10.1186/s13662-021-03235-w.
- Sene N. Analytical solutions of a class of fluids models with the Caputo fractional derivative. *Fractal Fract*. 2022;6(1):35. doi:10.3390/fractalfract6010035.
- Zabidi NA, Majid ZA, Kilicman A, Ibrahim ZB. Numerical solution of fractional differential equations with Caputo derivative by using numerical fractional predict-correct technique. *Adv Cont Disc Mod*. 2022;2022(1):26. doi:10.1186/s13662-022-03697-6.
- Dimitrov Y, Georgiev S, Todorov V. Approximation of Caputo fractional derivative and numerical solutions of fractional differential equations. *Fractal Fract*. 2023;7(10):750. doi:10.3390/fractalfract7100750.
- Ali I, Rasheed A, Anwar MS, Irfan M, Hussain Z. Fractional calculus approach for the phase dynamics of Josephson junction. *Chaos Solitons Fract*. 2021;143:110572.
- Khan M, Rasheed A, Anwar MS. Numerical analysis of nonlinear time-fractional fluid models for simulating heat transport processes in porous medium. *ZAMM-J Appl Math Mech/Zeitschrift Für Angewandte Mathematik Und Mechanik*. 2023;103(9):e202200544. doi:10.1002/zamm.202200544.

20. Alsulami M, Alyami M, Al-Mazmumy M, Alsulami A. Semi-analytical investigation for the ψ -Caputo fractional relaxation-oscillation equation using the decomposition method. *Eur J Pure Appl Math.* 2024;17(3):2311–98.
21. Iqbal N, Chughtai MT, Ullah R. Fractional study of the non-linear Burgers' equations via a semi-analytical technique. *Fractal Fract.* 2023;7(2):103. doi:10.3390/fractalfract7020103.
22. Jin B, Zhou Z. An analysis of Galerkin proper orthogonal decomposition for subdiffusion. *ESAIM Math Model Numer Anal.* 2017;51(1):89–113. doi:10.1051/m2an/2016017.
23. Khristenko U, Wohlmuth B. Solving time-fractional differential equations via rational approximation. *IMA J Numer Anal.* 2023;43(3):1263–90. doi:10.1093/imanum/drac022.
24. Brunton SL, Proctor JL, Kutz JN. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc Nat Acad Sci.* 2016;113(15):3932–7. doi:10.1073/pnas.1517384113.
25. Navickas Z, Telksnys T, Telksniene I, Marcinkevicius R, Ragulskis M. The fractal structure of analytical solutions to fractional Riccati equation. *Fractals.* 2024;32(04):2340130. doi:10.1142/s0218348x23401308.
26. Marcinkevicius R, Telksniene I, Telksnys T, Navickas Z, Ragulskis M. The construction of solutions to ${}^C D^{(1/n)}$ type FDEs via reduction to $\left({}^C D^{(1/n)}\right)^n$ type FDEs. *AIMS Math.* 2022;7(9):16536–54. doi:10.3934/math.2022905.
27. Kennedy J, Eberhart R. Particle swarm optimization. In: *Proceedings of ICNN'95-International Conference on Neural Networks*; 1995 Nov 27–Dec 1; Perth, WA, Australia. p. 1942–8.
28. Wang D, Tan D, Liu L. Particle swarm optimization algorithm: an overview. *Soft Comput.* 2018;22(2):387–408. doi:10.1007/s00500-016-2474-6.
29. Duan J. A generalization of the Mittag-Leffler function and solution of system of fractional differential equations. *Adv Differ Equ.* 2018;2018(1):239. doi:10.1186/s13662-018-1693-9.
30. Garrappa R. Numerical solution of fractional differential equations: a survey and a software tutorial. *Mathematics.* 2018;6(2):16. doi:10.3390/math6020016.
31. Timofejeva I, Telksnys T, Navickas Z, Marcinkevicius R, Mickevicius R, Ragulskis M. Solitary solutions to a metastasis model represented by two systems of coupled Riccati equations. *J King Saud Univ-Sci.* 2023;35(5):102682. doi:10.1016/j.jksus.2023.102682.
32. Arshad S, Baleanu D, Tang Y. Fractional differential equations with bio-medical applications. In: *Applications in engineering, life and social sciences, part A*. Berlin, Boston: De Gruyter; 2019. p. 1–20. doi:10.1515/9783110571905-001.