Alexandria Engineering Journal 130 (2025) 1075–1100

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

Original article

Polynomial neural network for solving Caputo-conformable fractional Volterra–Fredholm integro-differential equation with three-point non-local boundary conditions

Kawthar Alsa'di [a,1], Nik Mohd Asri B. Nik Long [a,*], Norazak Senu [a,b,2], Z.K. Eshkuvatov [c,2]

[a] Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, Serdang, 43400, Selangor, Malaysia
[b] Laboratory of Computational Sciences and Mathematical Physics, Institute for Mathematical Research, Universiti Putra Malaysia, Serdang, 43400, Selangor, Malaysia
[c] Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, Kuala Terengganu, 21030, Terengganu, Malaysia

ARTICLE INFO

Keywords:
Fractional Volterra–Fredholm fractional integro-differential equation
Caputo-conformable operator
Generalized Gronwall inequality
Polynomials neural network
Fixed-point theorems

ABSTRACT

This paper investigates a new class of fractional Volterra–Fredholm integro-differential equations (FVFIDEs) involving a new Caputo-conformable operator with three-point non-local Riemann–Liouville conformable integral boundary conditions. The Banach contraction principle and Schaefer's fixed point theorem are employed to investigate the existence and uniqueness of the solution in Banach space, and also the uniform stability is provided to prove the stability of the solution. A new generalized Gronwall inequality in the sense of the Riemann–Liouville conformable integral is established and utilized to prove the priori bounded of the solution. A hybrid technique, combining a polynomial neural network (PNN) with an extreme learning machine algorithm without using any activation functions, is developed. The Chebyshev neural network (ChebyshevNN) and Bernstein neural network (BernsteinNN) are used. To enhance the stability and control the strength of regularization numerically, L_2 regularization with hyperparameter λ is incorporated into the error function. Feed-forward and back-propagation learning algorithms are used to initialize and update the weights, while the Adam optimization method is employed to minimize the error. The error bound, convergence, computational complexity, stability, and sensitivity for parameters for the PNN method are provided. The numerical results provide the efficiency and accuracy for the proposed method.

1. Introduction

Fractional integro-differential equations (FIDEs) have been extensively applied in modeling a wide range of problems in engineering and science [1,2]. Many researchers shown their attentions for studying FIDEs extensively theoretically and numerically. For example, Santra et al. [3] proved the existence and uniqueness of multi-term time fractional partial integro-differential equations involving Volterra integral operators in the sense of Caputo-type time fractional differential operator. Das et al. [4] investigated the existence and uniqueness of the solution of fractional Volterra–Fredholm integro-differential equations (FVFIDEs) with in Caputo sense, and solved it by perturbation homotopy analysis. The same authors [5] studied the first order Volterra integro-differential equations of first kind for both initial and boundary value problems by reducing them into the fractional order Volterra- integro-differential equation of second kind using the Leibniz's rule, and also the Caputo fractional Volterra–Fredholm integro-differential equations was solved using homotopy perturbation method [6]. Das and Rana [7] studied the existence and uniqueness of the solution of the weakly singular fractional Volterra integro differential equation with Caputo-type sense based on the maximum norm and introduced an operator-based parameterized method to generate an approximate solution of the equation. Saw et al. [8] used third-kind shifted Chebyshev polynomials to approximate the solution of Pennes bioheat flow type parabolic partial differential equations in terms of Caputo-type operator, and studied the uniform convergence of the approximated solution.

Recently, many generalized fractional operators have been developed, including the Caputo–Hadamard operator [9], the Hadamard operator [10], the Hilfer operator [11], and the Katugampola operator [12], proportional fractional derivatives [13,14], among others, to address

* Corresponding author.
E-mail address: nmasri@upm.edu.my (N.M.A.B. Nik Long).

[1] PHD student
[2] Associate professor

https://doi.org/10.1016/j.aej.2025.09.022
Received 17 September 2024; Received in revised form 26 August 2025; Accepted 10 September 2025
Available online 8 October 2025

the modern real-life challenges. Khalil et al. [15] introduced the conformable operator, which was further refined by Abdeljawad [16] and later extended to the Caputo-conformable operator. Some interesting works on this operator can be found in [17], which employed H_∞ performance theory on Lipschitz nonlinear conformable fractional-order systems. A few studies have employed this operator, such as Piotrowska and Sajewski [18] used it to model the electrical circuits containing resistors. Ntouyas and Tariboon [19] explored the existence and uniqueness of solutions for Caputo-conformable fractional differential equations with four-point boundary conditions using Krasnoselskii's fixed point theorem, the Leray–Schauder nonlinear alternative, and Banach's contraction principle. Additionally, Baleanu et al. [20] investigated Caputo-conformable fractional inclusion differential equations with four-point boundary conditions based on the $(\alpha-\psi)$-contractive mappings under some considerations of the multi-dimension function α and the non-decreasing function ψ mentioned in [21]. Thabet et al. [22] examined the existence of solutions for Caputo-conformable fractional inclusion pantograph differential equations with three-point boundary conditions.

In spite of its vast available literature, only a few studies have addressed FIDEs using the Caputo-conformable operator. For instance, Rabhi et al. [23] utilized Mönch's fixed point theorem, a generalized Gronwall inequality, and the measure of non-compactness to investigate the existence of mild solutions for the non-local Caputo-conformable fractional delay integro-differential evolution equations. Amara et al. [24] investigated the existence of solutions for Caputo-conformable hybrid multi-term integro-differential equations by applying the lower solution property. Additionally, Peng et al. [25] discussed the uniqueness of solutions for nonlinear and linear Caputo-conformable Cauchy problems using the Banach fixed point theorem. They further demonstrated the extremal solutions for a nonlinear fractional Caputo-conformable p-Laplacian differential system using the monotone iterative technique.

Across all of the mentioned works, research works on the Caputo-conformable and Riemann–Liouville-conformable operator combined with FVFIDEs is limited. Combining these operators is significant because it provides powerful applications in modeling various real-world problems that exhibit memory and non-local effects, such as anomalous diffusion, viscoelastic behavior, control systems, financial modeling, biological systems, and other complex phenomena where traditional models often prove inadequate. Moreover, the properties of the Caputo-conformable operator offer a more physically meaningful approach: its consistency with Caputo derivative operators for initial conditions, its ability to capture non-local behavior and memory effects, and its modified chain and product rules make it especially suitable for applications in complex real-world dynamics compared to other fractional derivatives. Another advantage of this combined operator (Riemann–Liouville-conformable or Caputo-conformable operator), it is considered as a generalized operator. In fact, under this construction, various special cases of operators such as Hadamard fractional integral, the Riemann–Liouville operator, and the generalized fractional integral are included. To the best of our knowledge, this operator has not been extensively studied in prior works, making its investigation a significant contribution to the field.

Various numerical methods have been developed to solve FIDEs and FVFIDEs, such as the Legendre wavelet technique [26], multi-Galerkin methods [27], and the generalized Chebyshev operational method [28]. However, solving FVFIDEs with non-local boundary conditions remains challenging, as traditional numerical methods often require extensive computational time or involve complex calculations. Recent advances in artificial intelligence have contributed to modern numerical solutions that address these challenges. These methods not only minimize the errors but also reduce the computational time required for complex problems. Additionally, they have the ability for handling various initial and boundary conditions as well as different generalized fractional operators. For instance, Li et al. [29] proposed a machine learning framework combined with the degenerate kernel method to solve nonlinear integro-differential equations. Yuan et al. [30] employed the physics-informed neural networks (PINNs) to solve nonlinear partial differential and integro-differential equations. Furthermore, Allahviranloo et al. [31] applied artificial neural networks to solve Caputo FVIDEs, utilizing a multilayered feed-forward network with an iterative first-order algorithm to approximate the solution. Saneifard et al. [32] expanded the application of artificial neural networks to solve two-dimensional FVIDEs using a four-layer feed-forward neural architecture, employing the steepest descent and quasi-Newton methods to minimize the error function.

Polynomial neural network (PNN) is a type of simple neural network that utilizes the polynomial bases $P_n(u)$ instead of activation functions in the traditional hidden layers, such as Chebyshev polynomials (ChebyshevNN) and Bernstein polynomials (BernsteinNN) or other polynomials. PNN is a well-known concept in approximation theory, and they provide better convergence properties for solving differential and integro-differential equations. For example, Hajimohammadi et al. [33] employed a fractional Chebyshev deep neural network combined with the Gaussian method to solve fractional differential models. Shao et al. [34] addressed linear Fredholm integro-differential equations using a feed-forward neural network based on Legendre polynomials, where the weights were learned through the gradient descent method. Liu et al. [35] applied a Jacobi neural network to linear differential–algebraic equations using a single hidden layer. Additionally, Sun et al. [36] utilized a Bernstein neural network coupled with an extreme learning machine algorithm to solve differential equations with a single polynomial block, and Chakraverty and Mall [37] employed a single-layer Chebyshev neural network to solve ordinary differential equations.

In this paper, our contributions and the novelty of our work are summarized as follows: we develop a new FVFIDEs define in the sense of Caputo-conformable derivative, with three-points non-local Riemann–Liouville conformable integral boundary conditions:

$$\begin{cases} {}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta}\chi(u) = \mathfrak{O}\left(u, \chi(u), \mathcal{H}_1(\mathfrak{H}_1(\chi(u))), \mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right), \\ \chi(c_0) = \varrho_1,\ \varsigma_1\chi(C) + \varsigma_2 {}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\delta}\chi(\xi) = \varrho_2, \quad u \in J = [c_0, C], c_0 \geq 0, \end{cases} \tag{1}$$

where ${}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta}$ is the fractional Caputo conformable derivative of order $\theta \in (1,2]$ with $\mu \in (0,1]$ and ${}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\delta}$ is the fractional Riemann–Liouville conformable integral of order $\delta > 0$. Furthermore, $\xi \in (c_0, C]$, $\varsigma_1, \varsigma_2, \varrho_1, \varrho_2, \in \mathbb{R}$ are constants, $\mathfrak{O} : J \times \mathbb{R}^3 \to \mathbb{R}$ is a continuous function, $\mathcal{H}_1(\mathfrak{H}_1(\chi(u))) = \int_{c_0}^{C} \mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))\mathrm{d}v$, $\mathcal{H}_2(\mathfrak{H}_2(\chi(u))) = \int_{c_0}^{u} \mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))\mathrm{d}v$, such that $\mathfrak{H}_1, \mathfrak{H}_2$ represent linear or nonlinear continuous functions, and $\mathfrak{K}_1, \mathfrak{K}_2 \in (J \times J, \mathbb{R})$ with $\alpha_1 = \max_{u,v\in J}\int_{c_0}^{C} \left|\mathfrak{K}_1(u,v)\right| \mathrm{d}v$ and $\alpha_2 = \max_{u,v\in J}\int_{c_0}^{u} \left|\mathfrak{K}_2(u,v)\right| \mathrm{d}v$ are finite. We choose $\mathfrak{O}, \mathfrak{H}_1, \mathfrak{H}_2, \mathfrak{K}_1, \mathfrak{K}_2$ and χ are sufficiently smooth in $[c_0, C]$.

In addition, this study develops a new generalized Gronwall inequality within the framework of the Riemann–Liouville conformable integral, which is then utilized to investigate the existence and uniqueness of the solution for the proposed equation. Furthermore, we construct a novel approach to PNN by utilizing the Chebyshev neural network (ChebyshevNN) and Bernstein neural network (BernsteinNN), combined with the extreme learning machine algorithm. This approach differs from the traditional neural networks by using the basis of these polynomials instead of hidden layers. Choosing these two polynomials in our work refer to their abilities for solving different kinds of fractional ordinary and partial differential equations, which prove that using the non-orthogonal Bernstein polynomials provides a good approximate results and accuracy [38]. Additionally, Chebyshev polynomials are considered as an effective tool with spectral collocation method specifically in modeling partial differential equations [39]. PNN approach reduces the number of required hidden layers, trainable parameters and improving computational efficiency.

Moreover, we incorporate two additional terms: the L_2 regularization term to enhance the stability of solution and the hyper-parameter λ to control the regularization strength. These innovations lead to a novel error function that addresses complex boundary conditions, distinguishing our study from others in the literature. Consequently, we employ feed-forward and back-propagation learning algorithms to initialize and update weights, and the Adam optimization method to minimize the error.

This paper is organized as follows: In Section 2, some basic definitions and fundamental lemmas and theorems are stated. In Section 3, a new generalized Gronwall inequality in the sense of the Riemann–Liouville conformable integral is developed, and employ it to prove the priori boundedness of solution. In addition, the existence and uniqueness are investigated by using the fixed point techniques. Also, the stability is investigated using the concept of uniformly stability. In Section 4, we approximate the solutions using modern numerical techniques depends on PNN using ChebyshevNN and BernsteinNN with extreme learning machine algorithm. In Section 5, the numerical analysis for the convergence, error bound, computational complexity, computational stability and sensitivity of the parameters are discussed in more details. The validity of the results are described by some examples in Section 6. Conclusion and the summarize of the work are given in Section 7.

2. Preliminaries

In this section, we recall some fundamental and basic concepts which are needed in our study. As mention in the literature, the notion of the Riemann–Liouville fractional integral of order $\theta > 0$ of a function $\chi : [0, +\infty) \to \mathbb{R}$ is given by ${}^{\mathrm{R}}\mathcal{I}_0^\theta \chi(u) = \int_0^u \frac{(u-v)^{\theta-1}}{\Gamma(\theta)} \chi(v) \mathrm{d}v$, where $\theta \in (n-1, n)$ so that $n = [\theta]+1$, such that the value of the integral is finite [40,41]. Also, the fractional derivative of Caputo type for the function $\chi \in \mathcal{AC}_{\mathbb{R}}^{(n)}([0, +\infty))$ is given by

$$ {}^{C}\mathcal{D}_0^\theta \chi(u) = \int_0^u \frac{(u-v)^{n-\theta-1}}{\Gamma(n-\theta)} \chi^{(n)}(v) \mathrm{d}v, $$

provided that the integral is finite [40,41]. The left conformable derivative for a function $\chi : [c_0, \infty) \to \mathbb{R}$ with $\mu \in (0, 1]$ is defined as follows:

$$ \mathcal{D}_{c_0}^\mu \chi(u) = \lim_{\lambda \to 0} \frac{\chi\left(u + \lambda(u-c_0)^{1-\mu}\right) - \chi(u)}{\lambda}, $$

such that the limit exists [15]. Moreover, the definition of the left conformable integral of χ with $\mu \in (0, 1]$ is of the form

$$ \mathcal{I}_{c_0}^\mu \chi(u) = \int_{c_0}^u \chi(v) \frac{\mathrm{d}v}{(v-c_0)^{1-\mu}}, $$

where is the integral has finite values [15]. Jarad et al. [42] extended this work to the following new definitions:

Definition 2.1 ([42]). The fractional conformable integral of order $\theta \in \mathbb{C}, \mathrm{Re}(\theta) \geq 0$ and $\mu \in (0, 1]$ in Riemann–Liouville formula is defined by

$$ {}^{\mathcal{RR}}\mathfrak{I}_{c_0}^{\mu,\theta} \chi(u) = \frac{1}{\Gamma(\theta)} \int_{c_0}^u \left(\frac{(u-c_0)^\mu - (v-c_0)^\mu}{\mu} \right)^{\theta-1} \chi(v) \frac{dv}{(v-c_0)^{1-\mu}}, \tag{2} $$

whenever the value of integral is finite.

This definition coincides with the Riemann–Liouville fractional integral when $c_0 = 0$ and $\theta = 1$. It also coincides with the Hadamard fractional integral once $c_0 = 0$ and $\theta \to 0$ and with the generalized fractional integral when $c_0 = 0$.

Definition 2.2 ([42]). Let $\mathrm{Re}(\theta) \geq 0, n = [\mathrm{Re}(\theta)] + 1, \chi \in C_{\mu,c_0}^n[c_0, C]$. Then the fractional conformable derivatives of order θ and $\mu \in (0, 1]$ in the Caputo formula given as

$$ \begin{aligned} {}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta} \chi(u) &= {}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,n-\theta} \left(\mathfrak{D}_{c_0}^{\mu,n} \chi \right)(u) \\ &= \frac{1}{\Gamma(n-\theta)} \int_{c_0}^u \left(\frac{(u-c_0)^\mu - (v-c_0)^\mu}{\mu} \right)^{n-\theta-1} \frac{\mathfrak{D}_{c_0}^{\mu,n} \chi(v)}{(v-c_0)^{1-\mu}} dv. \end{aligned} \tag{3} $$

This dentition coincides with the Riemann–Liouville fractional integral when $c_0 = 0$ and $\theta = 1$. It also coincides with the Hadamard fractional integral once $c_0 = 0$ and $\theta \to 0$ and with the generalized fractional integral when $c_0 = 0$.

Lemma 2.3 ([42]). *Let $Re(\theta) > 0$, $Re(\delta) > 0$ and $Re(\eta) > 0$. Then for $\mu \in (0, 1]$ and for any $t > c_0$, we have the following identities:*

1. ${}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\theta}\left({}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\delta} \chi\right)(u) = \left({}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\theta+\delta} \chi\right)(u)$,
2. ${}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\theta}(u-c_0)^{\mu(\eta-1)}(v) = \frac{1}{\mu^\theta} \frac{\Gamma(\eta)}{\Gamma(\eta+\theta)} (v-c_0)^{\mu(\eta+\theta-1)}$,
3. ${}^{\mathcal{R}C}\mathfrak{D}_{c_0}^{\mu,\theta}(u-c_0)^{\mu(\eta-1)}(v) = \mu^\theta \frac{\Gamma(\eta)}{\Gamma(\eta-\theta)} (v-c_0)^{\mu(\eta-\theta-1)}$,
4. ${}^{\mathcal{R}C}\mathfrak{D}_{c_0}^{\mu,\theta}\left({}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\delta} \chi\right)(u) = \left({}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\delta-\theta} \chi\right)(u)$, $\left(Re(\theta) < Re(\delta)\right)$.

Lemma 2.4 ([42]). *Let $n-1 < Re(\theta) \leq n$ and $\chi \in C_{c_0,\mu}^n[c_0, C]$. Then for $\mu \in (0, 1]$, the following equation holds*

$$ {}^{\mathcal{R}C}\mathfrak{I}_{c_0}^{\mu,\theta}\left({}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta} \chi\right)(u) = \chi(u) - \sum_{r=0}^{n-1} \frac{\mathfrak{D}_{c_0}^{\mu,r} \chi(c_0)}{\mu^r r!} (u-c_0)^{r\mu}. \tag{4} $$

Notice that the series solution of the homogeneous equation $({}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta} \chi)(u) = 0$ can be expressed as

$$ \chi(u) = \zeta_0 + \zeta_1(u-c_0)^\mu + \zeta_2(u-c_0)^{2\mu} + \cdots + \zeta_{n-1}(u-c_0)^{(n-1)\mu}, $$

such that $n-1 < Re(\theta) \leq n$ and $\zeta_i, i = 0, 1, \ldots n-1 \in \mathbb{R}$.

The following theorems are useful in proving the main results.

Theorem 2.5 (*Arzelà–Ascoli Theorem*). *If a set $A \subseteq C[0,1]$ is closed, bounded, and equicontinuous then it is compact.*

Theorem 2.6 (*Banach Fixed Point Theorem [43]*). *Let $(X, \|.\|)$ be a complete normed space and let $\chi : X \to X$ be a contraction mapping. Then, χ has exactly one fixed point.*

Theorem 2.7 (*Schaefer's fixed Point Theorem [43]*). *Let $\chi : X \to X$ be a completely continuous operator. If the set $\mathfrak{E}(\chi) = \{u \in X : u = \tau\chi u$ for some $\tau \in [0,1]\}$ is bounded, then, χ has a fixed point.*

Also, we recall some basic definitions for Bernstein and Chebyshev polynomials. The Bernstein polynomial approximation of a function χ is defined on $[0,1]$ by

$$B_n(\chi)(u) = \sum_{j=0}^{n} \chi\left(\frac{j}{n}\right) B_{n,j}(u), \tag{5}$$

where the basis functions

$$B_{n,j}(u) = \binom{n}{j} u^j (1-u)^{n-j}, \quad j = 0, 1, \dots, n, \tag{6}$$

with following properties:

$$0 \leq B_{n,j}(u) \leq 1 \quad \text{and} \quad \sum_{j=0}^{n} B_{n,j}(u) = 1, \quad u \in [0,1]. \tag{7}$$

The Chebyshev polynomial of the first kind of degree n is defined by:

$$T_n(u) = \cos\left(n \cos^{-1}(u)\right), \text{ for } u \in [-1,1], \tag{8}$$

with analytic compact form of the $T_n(u)$ can be defined as follow

$$T_n(u) = n \sum_{j=0}^{n} \frac{(-2)^j (n+j-1)!}{(n-j)!(2j)!} (1-u)^j. \tag{9}$$

The expansion of the Chebyshev polynomials of the first kind into the Bernstein basis [44] is

$$T_n(u) = (2n-1) \sum_{k=0}^{n} \frac{(-1)^{n-k}}{(2k-1)!(2n-2k-1)!} B_{k,n}(u). \tag{10}$$

3. Theoretical results

In this manuscript, we let $C[c_0, C]$ be the set of all continuous functions on $[c_0, C]$, and $C^n_{c_0,\mu}[c_0, C]$ be the set of all nth continuously differentiable functions on $[c_0, C]$. Let $\mathfrak{N} = \{\chi : \chi \in C[c_0, C]\}$ be a Banach space with the norm $\|\chi\| = \|\chi\|_\infty = \sup_{u \in [c_0, C]} |\chi(u)|$. The $L2 = \|w\|_2^2 = \sum_{i=1}^{n} \sum_{j=1}^{m_i} w_{ij}^2$ will be used only on the weight represents on the training process.

We construct an equivalent integral equation corresponding to the Caputo-conformable FVFIDEs given in Eq. (1). To simplify, we define the following real constants:

$$\begin{cases} \tilde{\Psi}_1 = \varsigma_1 (C - c_0)^\mu + \varsigma_2 \dfrac{(\xi - c_0)^{\mu(1+\delta)}}{\mu^\delta \Gamma(2+\delta)} \neq 0, \\ \tilde{\Psi}_2 = \varrho_2 - \varsigma_1 \varrho_1 - \varrho_1 \frac{\varsigma_2 (\xi - c_0)^{\mu\delta}}{\mu^\delta \Gamma(1+\delta)}. \end{cases} \tag{11}$$

Lemma 3.1. *Let $\mathfrak{G} \in \mathfrak{N}$ be an arbitrary function. Then the function χ is a solution of the following Caputo-conformable FVFIDEs:*

$$\begin{cases} {}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0} \chi(u) = \mathfrak{G}(u), \\ \chi(c_0) = \varrho_1, \ \varsigma_1 \chi(C) + \varsigma_2 \, {}^{\mathcal{R}C}\mathfrak{I}^{\mu,\delta}_{c_0} \chi(\xi) = \varrho_2, \ (u \in [c_0, C], \ c_0 \geq 0), \end{cases} \tag{12}$$

if and only if χ is a solution of the following Riemann–Liouville conformable integral equation:

$$\begin{aligned} \chi(u) = \varrho_1 &+ \frac{\tilde{\Psi}_2 (u - c_0)^\mu}{\tilde{\Psi}_1} + {}^{\mathcal{R}C}\mathfrak{I}^{\mu,\theta}_{c_0} \mathfrak{G}(u) \\ &- \frac{(u - c_0)^\mu}{\tilde{\Psi}_1} \left[\varsigma_1 \, {}^{\mathcal{R}C}\mathfrak{I}^{\mu,\theta}_{c_0} \mathfrak{G}(u)\Big|_{u=C} + \varsigma_2 \, {}^{\mathcal{R}C}\mathfrak{I}^{\mu,\theta+\delta}_{c_0} \mathfrak{G}(u)\Big|_{u=\xi} \right], \end{aligned} \tag{13}$$

where $\tilde{\Psi}_1, \tilde{\Psi}_2$ are nonzero constants given in Eq. (11).

Proof. Let the function χ be a solution of Eq. (12). Apply the Riemann–Liouville conformable integral of order θ on both sides of ${}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0} \chi(u) = \mathfrak{G}(u)$, and using Lemma 2.4, gives

$$\chi(u) = \frac{1}{\Gamma(\theta)} \int_{c_0}^{u} \left(\frac{(u - c_0)^\mu - (v - c_0)^\mu}{\mu} \right)^{\theta - 1} \mathfrak{G}(v) \frac{dv}{(v - c_0)^{1-\mu}} + a_0 + a_1 (u - c_0)^\mu, \tag{14}$$

where $a_0, a_1 \in \mathbb{R}$.

By integrating both sides of Eq. (14) in the sense of Riemann–Liouville conformable operator of order δ, we obtain

$${}^{\mathcal{R}C}\mathfrak{I}^{\mu,\delta}_{c_0} \chi(u) = \frac{1}{\Gamma(\theta+\delta)} \int_{c_0}^{u} \left(\frac{(u - c_0)^\mu - (v - c_0)^\mu}{\mu} \right)^{\theta+\delta-1} \mathfrak{G}(v) \frac{dv}{(v - c_0)^{1-\mu}}$$

$$+\ a_0\frac{(u-c_0)^{\mu\delta}}{\mu^{\delta}\Gamma(1+\delta)}+a_1\frac{(u-c_0)^{\mu(1+\delta)}}{\mu^{\delta}\Gamma(2+\delta)}.$$

Using the first boundary condition, we have $a_0=\varrho_1$. Second integral boundary condition, gives

$$\begin{aligned}a_1=\frac{1}{\tilde{\Psi}_1}\Bigg[\tilde{\Psi}_2-\frac{\varsigma_1}{\Gamma(\theta)}\int_{c_0}^{C}\left(\frac{(C-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{\theta-1}\times\mathfrak{G}(v)\frac{\mathrm{d}v}{(v-c_0)^{1-\mu}}\\-\frac{\varsigma_2}{\Gamma(\theta+\delta)}\int_{c_0}^{\xi}\left(\frac{(\xi-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{\theta+\delta-1}\times\mathfrak{G}(v)\frac{\mathrm{d}v}{(v-c_0)^{1-\mu}}\Bigg].\end{aligned}$$

Substituting the values of a_0 and a_1 into Eq. (14), yields

$$\begin{aligned}\chi(u)=&\varrho_1+\frac{\tilde{\Psi}_2(u-c_0)^{\mu}}{\tilde{\Psi}_1}+{}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\mathfrak{G}(u)\\&-\frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\times\left[\varsigma_1\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\mathfrak{G}(u)|_{u=C}+\varsigma_2\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta+\delta}\mathfrak{G}(u)|_{u=\xi}\right],\end{aligned}$$

which verified that χ is a solution of the Riemann–Liouville conformable integral equation Eq. (13). Furthermore, we can easily deduce that χ is a solution of the Caputo conformable fractional Volterra–Fredholm integro-differential equation, Eq. (12), if χ is a solution of the Riemann–Liouville conformable integral equation, Eq. (13), which complete the proof. □

As a sequence of Lemma 3.1, we have the following lemma:

Lemma 3.2. *Let $\mathfrak{O}\in\mathfrak{N}$ be an arbitrary function. Then χ is a solution of the Caputo-conformable FVFIDEs, Eq. (1), if and only if χ is a solution of the following Riemann–Liouville conformable integral equation*

$$\begin{aligned}\chi(u)=&\varrho_1+\frac{\tilde{\Psi}_2(u-c_0)^{\mu}}{\tilde{\Psi}_1}\\&+{}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\\&-\frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\times\Bigg[\varsigma_1\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)|_{u=C}\\&+\varsigma_2\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta+\delta}\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)|_{u=\xi}\Bigg],\end{aligned}\tag{15}$$

where $\tilde{\Psi}_1,\tilde{\Psi}_2$ are nonzero constants given in Eq. (11).

3.1. New generalized Grönwall inequality via Riemann–Liouville conformable integral

We construct a new generalized Gronwall inequality in the sense of Riemann–Liouville conformable integral. For this objective, we need the following lemma:

Lemma 3.3 ([45,46]). *Let $\chi\in\mathfrak{N}$ be a non-negative and continuous function that satisfies the following generalized Gronwall inequality:*

$$\|\chi(u)\|\le a_1^*+a_2^*\int_0^u\|\chi(v)\|^{\beta_1}\mathrm{d}v+a_3^*\int_0^C\|\chi(v)\|^{\beta_2}\mathrm{d}v,\ u\in[0,C],\tag{16}$$

where $\beta_1,\beta_2\in[0,1)$, a_1^*,a_2^* and $a_3^*\ge 0$ are constants. Then there is a constant $\mathcal{L}^*>0$ such that

$$\|\chi(u)\|\le\mathcal{L}^*.$$

Remark 3.4. Noted that the value of $\mathcal{L}^*$ was mentioned in [45] as $a_1^*e^{(a_2^*+a_3^*)t}$ and in [46] as $\mathcal{L}^*=M\left(a_1^*+exp(\int_0^C\|\chi(v)\|^{\beta_2}\mathrm{d}v)\right)$, where $M>0$. Under the restrictions on $\beta_1,\beta_2\in[0,1)$, we ensure that the integrals remain well-defined and prevent unbounded growth or instability.

In the next result, we develop a new generalized Gronwall inequality in the sense of Riemann–Liouville conformable integral, i.e. the developed integral inequality will be modified using fractional power-law kernels (Riemann–Liouville conformable integral) instead of standard integral forms.

Lemma 3.5. *Let the following inequality holds:*

$$\begin{aligned}\|\chi(u)\|\le a_1&+a_2\int_{c_0}^{u}\left(\frac{(u-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{\theta-1}\frac{\|\chi(v)\|^{\beta}}{(v-c_0)^{1-\mu}}\mathrm{d}v\\&+a_3\int_{c_0}^{C}\left(\frac{(C-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{\theta-1}\frac{\|\chi(v)\|^{\beta}}{(v-c_0)^{1-\mu}}\mathrm{d}v\\&+a_4\int_{c_0}^{\xi}\left(\frac{(\xi-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{\theta+\delta-1}\frac{\|\chi(v)\|^{\beta}}{(v-c_0)^{1-\mu}}\mathrm{d}v,\end{aligned}\tag{17}$$

where $\chi\in\mathfrak{N}$, $\xi\in(c_0,C]$, $u\in[c_0,C]$, $c_0\ge 0$, $\theta\in(1,2]$, $\delta>0$, $\beta\in[0,1-\frac{1}{q})$ for some $q>1$ such that $q(\theta-1)+1>0$, $\frac{1}{\mu}(1-q)+q>0$ and $a_1,a_2,a_3,a_4\ge 0$ are constants. Then, there is a constant $\mathcal{L}^*>0$ such that

$$\|\chi(u)\|\le\mathcal{L}^*.$$

Proof. Set

$$\hat{\chi}(u) = \begin{cases} 1, & \|\chi(u)\| \le 1, \\ \chi(u), & \|\chi(u)\| > 1. \end{cases}$$

We have

$$\begin{aligned}\|\chi(u)\| \le \|\hat{\chi}(u)\| &\le a_1 + a_2 \int_{c_0}^{u} \left(\frac{(u-c_0)^\mu - (v-c_0)^\mu}{\mu} \right)^{\theta-1} \frac{\|\chi(v)\|^\beta}{(v-c_0)^{1-\mu}} \mathrm{d}v \\ &+ a_3 \int_{c_0}^{C} \left(\frac{(C-c_0)^\mu - (v-c_0)^\mu}{\mu} \right)^{\theta-1} \frac{\|\chi(v)\|^\beta}{(v-c_0)^{1-\mu}} \mathrm{d}v \\ &+ a_4 \int_{c_0}^{\xi} \left(\frac{(\xi-c_0)^\mu - (v-c_0)^\mu}{\mu} \right)^{\theta+\delta-1} \frac{\|\chi(v)\|^\beta}{(v-c_0)^{1-\mu}} \mathrm{d}v.\end{aligned}$$

Apply Hölder inequality inside the integral terms such that $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\begin{aligned}\|\chi(u)\| \le a_1 &+ a_2 \left(\int_{c_0}^{u} \left(\frac{(u-c_0)^\mu - (v-c_0)^\mu}{\mu} \right)^{q(\theta-1)} \frac{1}{(v-c_0)^{q(1-\mu)}} \mathrm{d}v \right)^{\frac{1}{q}} \\ &\times \left(\int_{c_0}^{u} \|\chi(v)\|^{\frac{q\beta}{q-1}} \mathrm{d}v \right)^{\frac{q-1}{q}} \\ &+ a_3 \left(\int_{c_0}^{C} \left(\frac{(C-c_0)^\mu - (v-c_0)^\mu}{\mu} \right)^{q(\theta-1)} \frac{1}{(v-c_0)^{q(1-\mu)}} \mathrm{d}v \right)^{\frac{1}{q}} \\ &\times \left(\int_{c_0}^{C} \|\chi(v)\|^{\frac{q\beta}{q-1}} \mathrm{d}v \right)^{\frac{q-1}{q}} \\ &+ a_4 \left(\int_{c_0}^{\xi} \left(\frac{(\xi-c_0)^\mu - (v-c_0)^\mu}{\mu} \right)^{q(\theta+\delta-1)} \frac{1}{(v-c_0)^{q(1-\mu)}} \mathrm{d}v \right)^{\frac{1}{q}} \\ &\times \left(\int_{c_0}^{\xi} \|\chi(v)\|^{\frac{q\beta}{q-1}} \mathrm{d}v \right)^{\frac{q-1}{q}}.\end{aligned}$$

Evaluate the integral part of the constant coefficients a_2, a_3, a_4 of the last inequality in terms of Beta function (see Appendix for details), we get

$$\begin{aligned}\|\chi(u)\| &\le a_1 + a_2 \left(\frac{(u-c_0)^{\mu\theta q-q+1}}{\mu^{q(\theta-1)+1}} \mathbb{B}\Big(q\theta - q + 1, \frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}} \times \int_{c_0}^{u} \|\chi(v)\|^{\frac{q\beta}{q-1}} \mathrm{d}v \\ &+ a_3 \left(\frac{(C-c_0)^{\mu\theta q-q+1}}{\mu^{q(\theta-1)+1}} \mathbb{B}\Big(q\theta - q + 1, x\frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}} \times \int_{c_0}^{C} \|\chi(v)\|^{\frac{q\beta}{q-1}} \mathrm{d}v \\ &+ a_4 \left(\frac{(\xi-c_0)^{q\mu(\theta+\delta)-q+1}}{\mu^{q(\theta+\delta-1)+1}} \mathbb{B}\Big(q(\theta+\delta) - q + 1, \frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}} \times \int_{c_0}^{\xi} \|\chi(v)\|^{\frac{q\beta}{q-1}} \mathrm{d}v \\ &\le a_1 + a_2 \left(\frac{(C-c_0)^{\mu\theta q-q+1}}{\mu^{q(\theta-1)+1}} \mathbb{B}\Big(q\theta - q + 1, \frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}} \times \int_{c_0}^{u} \|\chi(v)\|^{\frac{q\beta}{q-1}} \mathrm{d}v \\ &+ \Bigg[a_3 \left(\frac{(C-c_0)^{\mu\theta q-q+1}}{\mu^{q(\theta-1)+1}} \mathbb{B}\Big(q\theta - q + 1, \frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}} \\ &\quad + a_4 \left(\frac{(\xi-c_0)^{q\mu(\theta+\delta)-q+1}}{\mu^{q(\theta+\delta-1)+1}} \mathbb{B}\Big(q(\theta+\delta) - q + 1, \frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}} \Bigg] \times \int_{c_0}^{C} \|\chi(v)\|^{\frac{q\beta}{q-1}} \mathrm{d}v,\end{aligned}$$

where $\beta^* = \frac{q\beta}{q-1} \in (0,1)$ and $\mathbb{B}$ is well known Beta function. Set

$$\begin{aligned} a_1^* &:= a_1, \\ a_2^* &:= a_2 \left(\frac{(C-c_0)^{\mu\theta q-q+1}}{\mu^{q(\theta-1)+1}} \mathbb{B}\Big(q\theta - q + 1, \frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}}, \\ a_3^* &:= a_3 \left(\frac{(C-c_0)^{\mu\theta q-q+1}}{\mu^{q(\theta-1)+1}} \mathbb{B}\Big(q\theta - q + 1, \frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}}, \\ a_4^* &: a_4 \left(\frac{(\xi-c_0)^{q\mu(\theta+\delta)-q+1}}{\mu^{q(\theta+\delta-1)+1}} \mathbb{B}\Big(q(\theta+\delta) - q + 1, \frac{1}{\mu}(1-q)+q\Big) \right)^{\frac{1}{q}}, \end{aligned}$$

We have

$$\|\chi(u)\| \leq a_1^* + a_2^* \int_{c_0}^{u} \|\chi(v)\|^{\beta^*} \mathrm{d}v + a_3^* \int_{c_0}^{C} \|\chi(v)\|^{\beta^*} \mathrm{d}v + a_4^* \int_{c_0}^{\xi} \|\chi(v)\|^{\beta^*} \mathrm{d}v. \tag{18}$$

Then, using Lemma 3.3 ensures that there is a constant $\mathcal{L}^* > 0$ such that

$$\|\chi(u)\| \leq \mathcal{L}^*.$$

Hence, we proof the main goals that the inequality (17) with Caputo-conformable restriction above is bounded. Moreover, to find the upper bound for $\|\chi(u)\|$, we need to find the value of $\mathcal{L}^*$. Let $M = \sup_{c_0 \leq v \leq C} \|\chi(v)\|$. Since $\beta \in [0,1)$, we have $\|\chi(v)\|^{\beta} \leq M^{\beta}$. Thus, the integral in Eq. (18) becomes:

$$\|\chi(u)\| \leq a_1^* + a_2^* \int_{c_0}^{u} M^{\beta} \mathrm{d}v + a_3^* \int_{c_0}^{C} M^{\beta} \mathrm{d}v + a_4^* \int_{c_0}^{\xi} M^{\beta} \mathrm{d}v.$$

After simplification, we obtain

$$\|\chi(u)\| \leq a_1^* + M^{\beta}\left(a_2^*(u-c_0) + a_3^*(C-c_0) + a_4^*(\xi - c_0)\right).$$

Take supremum over $[c_0, C]$, then

$$\|\chi(u)\| \leq a_1^* + M^{\beta}\left(a_2^* + a_3^* + a_4^*\right)(C-c_0). \tag{19}$$

The generalized Gronwall's Lemma 3.3 ensures that for small β such inequality lead to exponential growth bounds of the form:

$$\|\chi(u)\| \leq \left(a_1^* + 1\right) e^{\left(a_2^*+a_3^*+a_4^*\right)(C-c_0)}.$$

Hence, $\mathcal{L}^* = \left(a_1^* + 1\right) e^{\left(a_2^*+a_3^*+a_4^*\right)(C-c_0)}$. The proof is completed. □

Remark 3.6. Note that the results of Lemma 3.5 cannot be obtained directly without following the entire process and structure outlined in its proof. Additionally, note that $\mathcal{L}^*$ is determined by the constants a_1, a_2, a_3, a_4 (which produce a_1^*, a_2^*, a_3^*, and a_4^*), the parameters $\theta, \delta, \mu, \beta$, and the integration limits c_0, C, ξ. If the exponents θ and δ are large, the integral may grow significantly, affecting the magnitude of $\mathcal{L}^*$. Therefore, we choose the value of θ belong to $(1,2]$ and $\delta > 0$. Hence, the integral terms suggest a potential recursive growth behavior, and if $\beta^* = \frac{q\beta}{q-1}$ is close to 1, then the solution may exhibit nearly exponential behavior as

$$\|\chi(u)\| \leq \mathcal{L}^* = \left(a_1^* + 1\right) e^{\left(a_2^*+a_3^*+a_4^*\right)(C-c_0)}.$$

This well-known Volterra-type inequalities lead to exponential bounds when we analyze it using the generalized Gronwall inequality stated in Lemma 3.3.

Remark 3.7. To ensure that the derivatives of the solution remain bounded, we consider the following points:

Firstly, the bounds from the integral formulation. Our solution $\chi(u)$ satisfies the integral inequality as in Lemma 3.5, which ensures that $\chi(u)$ does not grow unboundedly due to the imposed restrictions on the exponents $\theta, \beta, \mu, \delta$, and $\chi(u)$ is bounded by a constant $\mathcal{L}^*$, which directly implies boundedness of the solution. Secondly, boundedness of the derivative $\chi'(u)$ using differentiation under the integral sign. If we differentiate both sides of the integral inequality in Lemma 3.5 and apply Leibniz's Rule, we obtain:

$$\begin{aligned}
|\chi'(u)| \leq{}& a_2 \int_{c_0}^{u} \mu(\theta-1)\left(u-c_0\right)^{\mu-1}\left(\frac{\left(u-c_0\right)^{\mu}-\left(v-c_0\right)^{\mu}}{\mu}\right)^{\theta-2} \cdot \frac{|\chi(v)|^{\beta}}{\left(v-c_0\right)^{1-\mu}} dv \\
&+ a_3 \int_{c_0}^{C} \mu(\theta-1)\left(C-c_0\right)^{\mu-1}\left(\frac{\left(C-c_0\right)^{\mu}-\left(v-c_0\right)^{\mu}}{\mu}\right)^{\theta-2} \cdot \frac{|\chi(v)|^{\beta}}{\left(v-c_0\right)^{1-\mu}} dv \\
&+ a_4 \int_{c_0}^{\xi} \mu(\theta+\delta-1)\left(\xi-c_0\right)^{\mu-1}\left(\frac{\left(\xi-c_0\right)^{\mu}-\left(v-c_0\right)^{\mu}}{\mu}\right)^{\theta+\delta-2} \cdot \frac{|\chi(v)|^{\beta}}{\left(v-c_0\right)^{1-\mu}} dv.
\end{aligned}$$

Using the same technique as in the previous Lemma for the integral parts and take the supremum, this inequality shows that $\chi'(u)$ remains bounded under the imposed constraints on $\theta, \mu, \beta, \delta$, which confirms the boundedness of the derivative.

Lemma 3.8 (*Boundedness of the Caputo-Conformable Derivative*). *Let* $\mathrm{Re}(\theta) \geq 0, n = [\mathrm{Re}(\theta)]+1$ *such that* $n-\theta > 0$, *and* $\mu \in (0,1]$. *Assume* $\chi \in C^n_{\mu,c_0}\left[c_0, C\right]$ *and that its* n*th conformable derivative* $\mathfrak{D}^{\mu,n}_{c_0}\chi$ *is bounded on* $\left[c_0, C\right]$ *:*

$$\left\|\mathfrak{D}^{\mu,n}_{c_0}\chi\right\| \leq \tilde{M} < \infty.$$

Then the Caputo-conformable derivative

$${}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0}\chi(u) = \frac{1}{\Gamma(n-\theta)} \int_{c_0}^{u} \left(\frac{\left(u-c_0\right)^{\mu}-\left(v-c_0\right)^{\mu}}{\mu}\right)^{n-\theta-1} \frac{\mathfrak{D}^{\mu,n}_{c_0}\chi(v)}{\left(v-c_0\right)^{1-\mu}} dv,$$

is bounded on $u \in \left[c_0, C\right]$, *with the estimate*

$$\left\|{}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0}\chi\right\|_{L_\infty([c_0,C])} \leq \frac{\left(C-c_0\right)^{\mu(n-\theta)}}{\Gamma(n-\theta+1)\mu^{n-\theta}} \tilde{M}.$$

In particular, ${}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta}\chi$ *is continuous on* $[c_0, C]$ *and vanishes at* $u = c_0$.

Proof. For $u \in [c_0, C]$, we have

$$\left|{}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta}\chi(u)\right| = \frac{1}{\Gamma(n-\theta)}\left|\int_{c_0}^{u}\left(\frac{(u-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{n-\theta-1}\frac{\mathfrak{D}_{c_0}^{\mu,n}\chi(v)}{(v-c_0)^{1-\mu}}dv\right|. \tag{20}$$

To simplify, let

$$I^{*} = \int_{c_0}^{u}\left(\frac{(u-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{n-\theta-1}\frac{\mathfrak{D}_{c_0}^{\mu,n}\chi(v)}{(v-c_0)^{1-\mu}}dv.$$

Change the variable such that

$$t = \frac{v-c_0}{u-c_0}, \quad \text{which gives} \quad v - c_0 = t\,(u-c_0), \quad \text{and} \quad dv = (u-c_0)\,dt.$$

By changing $v := v(t) = t(u-c_0)+c_0$ inside the derivative sign, the integral I^* transforms into:

$$I^{*} = \int_{0}^{1}\left(\frac{(u-c_0)^{\mu}(1-t^{\mu})}{\mu}\right)^{n-\theta-1}\frac{(u-c_0)\,\mathfrak{D}_{c_0}^{\mu,n}\chi(v(t))dt}{t^{1-\mu}\,(u-c_0)^{1-\mu}}.$$

Next, expanding the power terms of $u-c_0$ as:

$$\left(\frac{(u-c_0)^{\mu}(1-t^{\mu})}{\mu}\right)^{n-\theta-1} = (u-c_0)^{\mu(n-\theta-1)}\,\frac{(1-t^{\mu})^{n-\theta-1}}{\mu^{n-\theta-1}}.$$

Then, the integral can be expressed as:

$$I^{*} = (u-c_0)^{\mu(n-\theta)}\,\frac{1}{\mu^{n-\theta-1}}\int_{0}^{1}(1-t^{\mu})^{n-\theta-1}\,t^{\mu-1}\mathfrak{D}_{c_0}^{\mu,n}\chi(v(t))dt.$$

By changing the variable $x = t^{\mu}$ in the integral term, such that $dx = \mu t^{\mu-1}dt$, $t = x^{1/\mu}$, and $dt = \frac{dx}{\mu}x^{(1/\mu)-1}$, we obtain

$$I^{*} = (u-c_0)^{\mu(n-\theta)}\,\frac{1}{\mu^{n-\theta}}\int_{0}^{1}(1-x)^{n-\theta-1}\mathfrak{D}_{c_0}^{\mu,n}\chi(v(x^{1/\mu}))dx, \tag{21}$$

where $v(x^{1/\mu}) = c_0 + (u-c_0)x^{1/\mu}$.

Now, substitute Eq. (21) into Eq. (20), we obtain

$$\begin{aligned}
\left|{}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta}\chi(u)\right| &= \frac{1}{\Gamma(n-\theta)}\left|\int_{c_0}^{u}\left(\frac{(u-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{n-\theta-1}\frac{\mathfrak{D}_{c_0}^{\mu,n}\chi(v)}{(v-c_0)^{1-\mu}}dv\right| \\
&\le \frac{1}{\Gamma(n-\theta)}\,(u-c_0)^{\mu(n-\theta)}\,\frac{1}{\mu^{n-\theta}}\int_{0}^{1}(1-x)^{n-\theta-1}\left|\mathfrak{D}_{c_0}^{\mu,n}\chi(v(x^{1/\mu}))\right|dx \\
&\le \frac{\tilde{M}}{\Gamma(n-\theta)\mu^{n-\theta}}\,(u-c_0)^{\mu(n-\theta)}\int_{0}^{1}(1-x)^{n-\theta-1}dx \quad \left(\text{since } \left|\mathfrak{D}_{c_0}^{\mu,n}\chi\right| \le \tilde{M}\right) \\
&\le \frac{\tilde{M}}{(n-\theta)\,\Gamma(n-\theta)\,\mu^{n-\theta}}\,(u-c_0)^{\mu(n-\theta)} \\
&\le \frac{\tilde{M}}{\Gamma(n-\theta+1)\,\mu^{n-\theta}}\,(u-c_0)^{\mu(n-\theta)}.
\end{aligned}$$

Taking the supremum over $u \in [c_0, C]$ yields

$$\left\|{}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta}\chi\right\| \le \frac{(C-c_0)^{\mu(n-\theta)}}{\Gamma(n-\theta+1)\mu^{n-\theta}}\tilde{M},$$

and also shows the continuity and vanishing at $u = c_0$, which follow from the factor $(u-c_0)^{\mu(n-\theta)}$. □

Remark 3.9. Substitute $t = (v-c_0)/(u-c_0)$, it turns into $t^{\mu-1}$, which is integrable near 0 for $\mu \in (0,1]$, so there is no singular growth provided that $\mathfrak{D}_{c_0}^{\mu,n}\chi$ is bounded. If the nth conformable derivative is not bounded (e.g., singularly perturbed/ layered solutions), the estimated expression may fail, and this is outside of the smooth class assumed in our analysis.

3.2. Existence and uniqueness

To prove the existence and uniqueness of solution of the Caputo-conformable FVFIDEs, we convert Eq. (1) into a fixed point problem by defining the operator $\Omega : \mathfrak{N} \to \mathfrak{N}$ as follows:

$$\begin{aligned}
\Omega(\chi(u)) &= \varrho_1 + \frac{\tilde{\Psi}_2(u-c_0)^{\mu}}{\tilde{\Psi}_1} + {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\mathfrak{O}\left(u, \chi(u), \mathcal{H}_1(\mathfrak{H}_1(\chi(u))), \mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right) \\
&- \frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\Big[\varsigma_1\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\mathfrak{O}\left(u, \chi(u), \mathcal{H}_1(\mathfrak{H}_1(\chi(u))), \mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)|_{u=C}
\end{aligned}$$

$$+ \varsigma_2 \ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta+\delta}\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)|_{u=\xi}\Big], \tag{22}$$

where $\tilde{\Psi}_1, \tilde{\Psi}_2$ are nonzero constants defined in Eq. (11).

Now, we state and prove the existence solution of the problem (1) by using Schaefer's fixed point theorem and a new generalized Gronwall inequality Lemma 3.5.

Theorem 3.10. *Let $\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)$ be a continuous function. If there exist constants $m_1, m_2, m_3 > 0$ and $\beta \in [0,1)$, such that:*

(i) $|\mathfrak{O}\left(u,\chi_1,\chi_2,\chi_3\right)| \le m_1\left(1+|\chi_1|^\beta+|\chi_2|+|\chi_3|\right)$,
(ii) $|\mathfrak{H}_1(\chi)| \le m_2\left(1+|\chi|^\beta\right)$
(iii) $|\mathfrak{H}_2(\chi)| \le m_3\left(1+|\chi|^\beta\right)$,

are hold for all $u \in [c_0, C]$ and each $\chi, \chi_1, \chi_2, \chi_3 \in \mathfrak{N}$. Then the Caputo-conformable FVFIDEs (1) has at least one solution.

Proof. Let the operator Ω be defined as in Eq. (22). To prove the existence of the solution, we divide the process into 4 parts as follow:
Step1. The operator Ω is continuous.

Let $\{\chi_n\}$ be a sequence such that $\chi_n \to \chi$ in $\mathfrak{N}$ as $n \to \infty$. For each $u \in J$, we have

$$\begin{aligned}
\left|\Omega(\chi_n(u)) - \Omega(\chi(u))\right| \le\ & {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\Big|\mathfrak{O}\left(u,\chi_n(u),\mathcal{H}_1(\mathfrak{H}_1(\chi_n(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi_n(u)))\right)\\
& -\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\Big|\\
& + \frac{|(u-c_0)^\mu|}{|\tilde{\Psi}_1|}\\
& \times\Bigg[|\varsigma_1|\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\Big|\mathfrak{O}\left(u,\chi_n(u),\mathcal{H}_1(\mathfrak{H}_1(\chi_n(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi_n(u)))\right)\\
& -\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\Big|_{u=C}\\
& +|\varsigma_2|\,{}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta+\delta}\Big|\mathfrak{O}\left(u,\chi_n(u),\mathcal{H}_1(\mathfrak{H}_1(\chi_n(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi_n(u)))\right)\\
& -\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\Big|_{u=\xi}\Bigg]\\
\le\ & \|\ \mathfrak{O}\left(.,\chi_n(.),\mathcal{H}_1(\mathfrak{H}_1(\chi_n(.))),\mathcal{H}_2(\mathfrak{H}_2(\chi_n(.)))\right)\\
& -\ \mathfrak{O}\left(.,\chi(.),\mathcal{H}_1(\mathfrak{H}_1(\chi(.))),\mathcal{H}_2(\mathfrak{H}_2(\chi(.)))\right)\ \||\\
& \times\left\{\frac{(C-c_0)^{\mu\theta}}{\mu^\theta\Gamma(\theta+1)} + \frac{(C-c_0)^\mu}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|\,(C-c_0)^{\mu\theta}}{\mu^\theta\Gamma(\theta+1)} + \frac{|\varsigma_2|\,(\xi-c_0)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma(\theta+\delta+1)}\right]\right\}
\end{aligned}$$

Due to the continuity of $\mathfrak{O}$, we get

$$\|\Omega\chi_n - \Omega\chi\| \to 0 \text{ as } n \to \infty.$$

Hence, Ω is a continuous operator.
Step 2. The operator Ω maps a bounded sets into a bounded sets in $\mathfrak{N}$.

Let $B_\varepsilon = \{\chi \in \mathfrak{N} : \|\chi\| \le \varepsilon\}$ be a bounded and closed set. In addition, the bounded hypothesis can be simplified as

$$\begin{aligned}
& |\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)|\\
& \le m_1\left(1+|\chi(u)|^\beta+|\mathcal{H}_1(\mathfrak{H}_1(\chi(u)))|+|\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))|\right)\\
& \le m_1\left(1+|\chi(u)|^\beta+m_2(1+\alpha_1|\chi(u)|^\beta)+m_3(1+\alpha_2|\chi(u)|^\beta)\right)\\
& \le (m_1+m_1m_2+m_1m_3)+(m_1+m_1m_2\alpha_1+m_1m_3\alpha_2)|\chi(u)|^\beta\\
& \le M_1+M_2|\chi(u)|^\beta \le M_1+M_2\|\chi\|^\beta,
\end{aligned}$$

where $M_1 = m_1 + m_1m_2 + m_1m_3$ *and* $M_2 = m_1 + m_1m_2\alpha_1 + m_1m_3\alpha_2$. For each $u \in J$, we have

$$\begin{aligned}
& |\Omega(\chi(u))| \le |\varrho_1| + \left|\frac{\tilde{\Psi}_2(u-c_0)^\mu}{\tilde{\Psi}_1}\right| + {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\left|\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\right|\\
& + \left|\frac{(u-c_0)^\mu}{\tilde{\Psi}_1}\right| \times \Bigg[|\varsigma_1|\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\left|\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\right|_{u=C}\\
& + |\varsigma_2|\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta+\delta}\left|\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\right|_{u=\xi}\Bigg]\\
& \le |\varrho_1| + \frac{|\tilde{\Psi}_2(C-c_0)^\mu|}{|\tilde{\Psi}_1|} + \frac{(C-c_0)^{\mu\theta}}{\mu^\theta\Gamma(\theta+1)}\left(M_1+M_2\|\chi\|^\beta\right)\\
& + \frac{(C-c_0)^\mu}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|\,(C-c_0)^{\mu\theta}}{\mu^\theta\Gamma(\theta+1)} + \frac{|\varsigma_2|\,(\xi-c_0)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma(\theta+\delta+1)}\right] \times \left(M_1+M_2\|\chi\|^\beta\right)\\
& \le |\varrho_1| + \frac{|\tilde{\Psi}_2(C-c_0)^\mu|}{|\tilde{\Psi}_1|} + \frac{(C-c_0)^{\mu\theta}}{\mu^\theta\Gamma(\theta+1)}\left(M_1+M_2\varepsilon^\beta\right)
\end{aligned}$$

$$+\ \frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}+\frac{|\varsigma_2|\left(\xi-c_0\right)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma(\theta+\delta+1)}\right]\times\left(M_1+M_2\varepsilon^{\beta}\right),$$

where

$$\begin{aligned}\ell :=& |\rho_1|+\frac{|\tilde{\Psi}_2(C-c_0)^{\mu}|}{|\tilde{\Psi}_1|}+\frac{\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}\left(M_1+M_2\varepsilon^{\beta}\right)\\ &+\frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}+\frac{|\varsigma_2|\left(\xi-c_0\right)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma(\theta+\delta+1)}\right]\\ &\times\left(M_1+M_2\varepsilon^{\beta}\right).\end{aligned}$$

Thus, $\|\Omega\chi\|\le\ell$.

Step 3. The operator Ω is equicontinuous.

Let $u_1,u_2\in[c_0,C], u_1\le u_2$ *and* $\chi\in B_{\varepsilon}$. By using (H_2), we get

$$\begin{aligned}\left|\Omega(\chi(u_2))-\Omega(\chi(u_1))\right| \le& \frac{|\tilde{\Psi}_2|\left|(u_2-c_0)^{\mu}-(u_1-c_0)^{\mu}\right|}{|\tilde{\Psi}_1|}\\ &+\frac{1}{\Gamma(\theta)}\int_{c_0}^{u_2}\left(\frac{(u_2-c_0)^{\mu}-(\upsilon-c_0)^{\mu}}{\mu}\right)^{\theta-1}\left|\mathfrak{O}\left(\upsilon,\chi(\upsilon),\mathcal{H}_1(\mathfrak{H}_1(\chi(\upsilon))),\mathcal{H}_2(\mathfrak{H}_2(\chi(\upsilon)))\right)\right|\\ &\times\frac{d\upsilon}{(\upsilon-c_0)^{1-\mu}}\\ &-\frac{1}{\Gamma(\theta)}\int_{c_0}^{u_1}\left(\frac{(u_1-c_0)^{\mu}-(\upsilon-c_0)^{\mu}}{\mu}\right)^{\theta-1}\left|\mathfrak{O}\left(\upsilon,\chi(\upsilon),\mathcal{H}_1(\mathfrak{H}_1(\chi(\upsilon))),\mathcal{H}_2(\mathfrak{H}_2(\chi(\upsilon)))\right)\right|\\ &\times\frac{d\upsilon}{(\upsilon-c_0)^{1-\mu}}\\ &+\frac{\left|(u_2-c_0)^{\mu}-(u_1-c_0)^{\mu}\right|}{|\tilde{\Psi}_1|}\left[|\varsigma_1|\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\left|\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\right|_{u=C}\right.\\ &\left.+|\varsigma_2|\,{}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta+\delta}\left|\mathfrak{O}\left(u,\chi_n(u),\mathcal{H}_1(\mathfrak{H}_1(\chi_n(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi_n(u)))\right)\right|_{u=\xi}\right]\\ \le& \frac{|\tilde{\Psi}_2|\left|(u_2-c_0)^{\mu}-(u_1-c_0)^{\mu}\right|}{|\tilde{\Psi}_1|}+\left\{\frac{\left(u_2-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}-\frac{\left(u_1-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}\right\}\left(M_1+M_2\|\chi\|^{\beta}\right)\\ &+\frac{\left|(u_2-c_0)^{\mu}-(u_1-c_0)^{\mu}\right|}{|\tilde{\Psi}_1|}\left(M_1+M_2\|\chi\|^{\beta}\right)\left[|\varsigma_1|\frac{\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}+|\varsigma_2|\frac{\left(\xi-c_0\right)^{\mu\theta+\delta}}{\mu^{\theta+\delta}\Gamma(\theta+\delta+1)}\right]\\ \le& \frac{|\tilde{\Psi}_2|\left|(u_2-c_0)^{\mu}-(u_1-c_0)^{\mu}\right|}{|\tilde{\Psi}_1|}+\left\{\frac{\left(u_2-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}-\frac{\left(u_1-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}\right\}\left(M_1+M_2\varepsilon^{\beta}\right)\\ &+\frac{\left|(u_2-c_0)^{\mu}-(u_1-c_0)^{\mu}\right|}{|\tilde{\Psi}_1|}\left(M_1+M_2\varepsilon^{\beta}\right)\left[|\varsigma_1|\frac{\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}+|\varsigma_2|\frac{\left(\xi-c_0\right)^{\mu\theta+\delta}}{\mu^{\theta+\delta}\Gamma(\theta+\delta+1)}\right].\end{aligned}$$

If $u_2\to u_1$, then the right-side of the latest inequality tends to zero. Hence, Ω is equicontinuous operator. As a consequences, in the light of Arzelá–Ascoli theorem, with the above three steps, we deduce that Ω is continuous and completely continuous operator.

Step 4. A priori bounds.

Now, we prove that the set $\mathfrak{E}(\Omega)=\{\chi\in\mathfrak{N}:\chi=\varpi\Omega\chi,\varpi\in[0,1]\}$ is bounded. Let $\chi\in\mathfrak{E}(\Omega)$, then $\chi=\varpi\Omega\chi$ for some $\varpi\in[0,1]$. Thus, for each $u\in J$ and by using (H_2), we have

$$\begin{aligned}|\chi(u)|=|\varpi(\Omega\chi)(u)|\le|(\Omega\chi)(u)|\le&|\rho_1|+\left|\frac{\tilde{\Psi}_2(u-c_0)^{\mu}}{\tilde{\Psi}_1}\right|\\ &+\frac{1}{\Gamma(\theta)}\int_{c_0}^{u}\left(\frac{(u-c_0)^{\mu}-(\upsilon-c_0)^{\mu}}{\mu}\right)^{\theta-1}\left(M_1+M_2|\chi(u)|^{\beta}\right)\times\frac{d\upsilon}{(\upsilon-c_0)^{1-\mu}}\\ &+\left|\frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\right|\left[\frac{|\varsigma_1|}{\Gamma(\theta)}\int_{c_0}^{C}\left(\frac{(C-c_0)^{\mu}-(\upsilon-c_0)^{\mu}}{\mu}\right)^{\theta-1}\right.\\ &\left.+\frac{|\varsigma_2|}{\Gamma(\theta+\delta)}\int_{c_0}^{\xi}\left(\frac{(\xi-c_0)^{\mu}-(\upsilon-c_0)^{\mu}}{\mu}\right)^{\theta+\delta-1}\right]\times\left(M_1+M_2|\chi(u)|^{\beta}\right)\frac{d\upsilon}{(\upsilon-c_0)^{1-\mu}}\end{aligned}$$

$$\begin{aligned}|\chi(u)|=|\varpi(\Omega\chi)(u)|\le&|\rho_1|+\left|\frac{\tilde{\Psi}_2(C-c_0)^{\mu}}{\tilde{\Psi}_1}\right|+\frac{M_1\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}\\ &+\frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\frac{M_1\ |\varsigma_1|\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}+\frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\frac{M_1\ |\varsigma_2|\left(\xi-c_0\right)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma(\theta+\delta+1)}\\ &+\frac{M_2}{\Gamma(\theta)}\int_{c_0}^{u}\left(\frac{(u-c_0)^{\mu}-(\upsilon-c_0)^{\mu}}{\mu}\right)^{\theta-1}|\chi(u)|^{\beta}\frac{d\upsilon}{(\upsilon-c_0)^{1-\mu}}\end{aligned}$$

$$+\left|\frac{(C-c_0)^{\mu}}{\tilde{\Psi}_1}\right|\left[\frac{M_2\ |\varsigma_1|}{\Gamma(\theta)}\int_{c_0}^{C}\left(\frac{(C-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{\theta-1}\right.$$
$$\left.+\frac{M_2\ |\varsigma_2|}{\Gamma(\theta+\delta)}\int_{c_0}^{\xi}\left(\frac{(\xi-c_0)^{\mu}-(v-c_0)^{\mu}}{\mu}\right)^{\theta+\delta-1}\right]\times|\chi(u)|^{\beta}\frac{\mathrm{d}v}{(v-c_0)^{1-\mu}}.$$

Hence, using the result in Lemma 3.5, there exist a constant $\mathcal{L}^*>0$ such that

$$|\chi(u)|\le\mathcal{L}^*,$$

which demonstrates that the set $\mathfrak{E}(\Omega)$ is bounded. Hence, from the Schaefer's fixed point, Theorem 2.7, we conclude that the operator Ω has a fixed point which is a solution to the Caputo-conformable FVFIDEs (1). □

Theorem 3.11. *Suppose that the function $\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)$ is continuous, and there exist the constants $l_1,l_2,l_3,l_4,l_5>0$, such that:*

(i) $|\mathfrak{O}\left(u,\chi_1,\chi_2,\chi_3\right)-\mathfrak{O}\left(u,\tilde{\chi}_1,\tilde{\chi}_2,\tilde{\chi}_3\right)|\le l_1|\chi_1-\tilde{\chi}_1|+l_2|\chi_2-\tilde{\chi}_2|+l_3|\chi_3-\tilde{\chi}_3|$,
(ii) $|\mathfrak{H}_1(\chi_1)-\mathfrak{H}_1(\tilde{\chi}_1)|\le l_4|\chi_1-\tilde{\chi}_1|$
(iii) $|\mathfrak{H}_2(\chi_1)-\mathfrak{H}_2(\tilde{\chi}_1)|\le l_5|\chi_1-\tilde{\chi}_1|$,

are hold for all $u\in[c_0,C]$ and each $\chi_1,\chi_2,\chi_3,\tilde{\chi}_1,\tilde{\chi}_2,\tilde{\chi}_3\in\mathfrak{N}$. Furthermore,

$$Q=\left(\frac{\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma\left(\theta+1\right)}+\frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma\left(\theta+1\right)}+\frac{|\varsigma_2|\left(\xi-c_0\right)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma\left(\theta+\delta+1\right)}\right]\right)\times\left(l_1+l_2\alpha_1l_4+l_3\alpha_2l_5\right)<1. \tag{23}$$

Then, the Caputo-conformable FVFIDEs (1) has a unique solution.

Proof. Under sufficient smoothness on the data, the uniqueness of the solution of problem (1) is equivalent to the uniqueness of the fixed point of operator Ω, which is defined in Eq. (22). In view of the Lipschitz continuity hypothesis and for $\chi,\chi^*\in\mathfrak{N},u\in J$, we have

$$\begin{aligned}
&\left|\Omega(\chi(u))-\Omega(\chi^*(u))\right|\le{}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\left|\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\right.\\
&\quad\left.-\mathfrak{O}\left(u,\chi^*(u),\mathcal{H}_1(\mathfrak{H}_1(\chi^*(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi^*(u)))\right)\right|\\
&+\left|\frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\right|\times\left[|\varsigma_1|\,{}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta}\left|\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\right.\right.\\
&\quad\left.-\mathfrak{O}\left(u,\chi^*(u),\mathcal{H}_1(\mathfrak{H}_1(\chi^*(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi^*(u)))\right)\right|_{u=C}\\
&\quad+|\varsigma_2|\ {}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\theta+\delta}\left|\mathfrak{O}\left(u,\chi(u),\mathcal{H}_1(\mathfrak{H}_1(\chi(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)\right.\\
&\quad\left.\left.-\mathfrak{O}\left(u,\chi^*(u),\mathcal{H}_1(\mathfrak{H}_1(\chi^*(u))),\mathcal{H}_2(\mathfrak{H}_2(\chi^*(u)))\right)\right|_{u=\xi}\right]\\
&\le\frac{\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma\left(\theta+1\right)}\left(l_1+l_2\alpha_1l_4+l_3\alpha_2l_5\right)\|\chi-\chi^*\|\\
&+\frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma\left(\theta+1\right)}+\frac{|\varsigma_2|\left(\xi-c_0\right)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma\left(\theta+\delta+1\right)}\right]\times\left(l_1+l_2\alpha_1l_4+l_3\alpha_2l_5\right)\|\chi-\chi^*\|\\
&\le\frac{\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma\left(\theta+1\right)}\left(l_1+l_2\alpha_1l_4+l_3\alpha_2l_5\right)\|\chi-\chi^*\|\\
&+\frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|\left(C-c_0\right)^{\mu\theta}}{\mu^{\theta}\Gamma\left(\theta+1\right)}+\frac{|\varsigma_2|\left(\xi-c_0\right)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma\left(\theta+\delta+1\right)}\right]\times\left(l_1+l_2\alpha_1l_4+l_3\alpha_2l_5\right)\|\chi-\chi^*\|.
\end{aligned}$$

Hence, from Eq. (23) and taking supremum, we get

$$\|\Omega\chi-\Omega\chi^*\|\le Q\,\|\chi-\chi^*\|,$$

since $Q\le1$, we deduce that Ω is a contraction map. Due to Theorem 2.6, the operator Ω has a unique fixed point which is a solution of the Caputo-conformable FVFIDEs (1). □

3.3. Uniform stability of the solution

We prove the uniform stability of solution for the Caputo-conformable FVFIDEs, Eq. (1) by constructing the following definition:
Let $\bar{\chi}$ be a solution of the following equation:

$$\begin{cases}{}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta}\bar{\chi}(u)=\mathfrak{O}\left(u,\bar{\chi}(u),\mathcal{H}_1(\mathfrak{H}_1(\bar{\chi}(u))),\mathcal{H}_2(\mathfrak{H}_2(\bar{\chi}(u)))\right),\\ \bar{\chi}(c_0)=\bar{\varrho}_1,\qquad \varsigma_1\bar{\chi}(C)+\varsigma_2{}^{\mathcal{RC}}\mathfrak{I}_{c_0}^{\mu,\delta}\bar{\chi}(\xi)=\bar{\varrho}_2,\ (u\in J=[c_0,C],c_0\ge0),\end{cases} \tag{24}$$

which has an equivalent integral equation as follows:

$$\bar{\chi}(u)=\bar{\varrho}_1+\frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\left[\bar{\varrho}_2-\varsigma_1\bar{\varrho}_1-\bar{\varrho}_1\frac{\varsigma_2(\xi-c_0)^{\mu\delta}}{\mu^{\delta}\Gamma(1+\delta)}\right]$$

$$+ {}^{\mathcal{RC}}\mathfrak{I}^{\mu,\theta}_{c_0}\mathfrak{D}\left(u, \bar{\chi}(u), \mathcal{H}_1(\mathfrak{H}_1(\bar{\chi}(u))), \mathcal{H}_2(\mathfrak{H}_2(\bar{\chi}(u)))\right)$$
$$- \frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\Bigg[\varsigma_1 \, {}^{\mathcal{RC}}\mathfrak{I}^{\mu,\theta}_{c_0}\mathfrak{D}\left(u, \bar{\chi}(u), \mathcal{H}_1(\mathfrak{H}_1(\bar{\chi}(u))), \mathcal{H}_2(\mathfrak{H}_2(\bar{\chi}(u)))\right)|_{u=C}$$
$$+ \varsigma_2 \, {}^{\mathcal{RC}}\mathfrak{I}^{\mu,\theta+\delta}_{c_0}\mathfrak{D}\left(u, \bar{\chi}(u), \mathcal{H}_1(\mathfrak{H}_1(\bar{\chi}(u))), \mathcal{H}_2(\mathfrak{H}_2(\bar{\chi}(u)))\right)|_{u=\xi}\Bigg], \quad (25)$$

where $\tilde{\Psi}_1$ is nonzero constant given in Eq. (11).

Definition 3.12 ([47,48]). Let $\chi(u)$ and $\bar{\chi}(u)$ be the solutions of Eqs. (1) and (24) respectively. The solution of Caputo-conformable FVFIDEs Eq. (1) is called uniformly stable, if for any $\epsilon > 0$, there exists $\psi(\epsilon) > 0$ such that $|(\varrho_1 + \varrho_2) - (\bar{\varrho}_1 + \bar{\varrho}_2)| < \psi(\epsilon)$, then $|\chi(u) - \bar{\chi}(u)| < \epsilon$, for all $u \in J$.

Theorem 3.13. *Suppose that the hypotheses in Theorems 3.10 and 3.11 hold. Then, the solution of the Caputo-conformable FVFIDEs Eq. (1) is uniformly stable.*

Proof. Consider $\chi(u)$ be a solution of Eq. (1) and $\bar{\chi}(u)$ is a solution of Eq. (24). By taking $K := \max\left\{\left|\frac{(C-c_0)^{\mu}}{\tilde{\Psi}_1}\right|, 1 + \left|\frac{(C-c_0)^{\mu}}{\tilde{\Psi}_1}\right|\left|\varsigma_1 + \frac{\varsigma_2(\xi-c_0)^{\mu\delta}}{\mu^{\delta}\Gamma(1+\delta)}\right|\right\}$, and by using Eqs. (15) and (25), for each $u \in J$, we have

$$|\chi(u) - \bar{\chi}(u)|$$
$$\leq |\varrho_1 - \bar{\varrho}_1| + \left|\frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\right|\left|\varrho_2 - \varsigma_1\varrho_1 - \varrho_1\frac{\varsigma_2(\xi-c_0)^{\mu\delta}}{\mu^{\delta}\Gamma(1+\delta)} - \bar{\varrho}_2 + \varsigma_1\bar{\varrho}_1 + \bar{\varrho}_1\frac{\varsigma_2(\xi-c_0)^{\mu\delta}}{\mu^{\delta}\Gamma(1+\delta)}\right|$$
$$+ {}^{\mathcal{RC}}\mathfrak{I}^{\mu,\theta}_{c_0}\Big|\mathfrak{D}\left(u, \chi(u), \mathcal{H}_1(\mathfrak{H}_1(\chi(u))), \mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)$$
$$-\mathfrak{D}\left(u, \bar{\chi}(u), \mathcal{H}_1(\mathfrak{H}_1(\bar{\chi}(u))), \mathcal{H}_2(\mathfrak{H}_2(\bar{\chi}(u)))\right)\Big|$$
$$+ \left|\frac{(u-c_0)^{\mu}}{\tilde{\Psi}_1}\right| \times \Bigg[|\varsigma_1| \, {}^{\mathcal{RC}}\mathfrak{I}^{\mu,\theta}_{c_0}\Big|\mathfrak{D}\left(u, \chi(u), \mathcal{H}_1(\mathfrak{H}_1(\chi(u))), \mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)$$
$$-\mathfrak{D}\left(u, \bar{\chi}(u), \mathcal{H}_1(\mathfrak{H}_1(\bar{\chi}(u))), \mathcal{H}_2(\mathfrak{H}_2(\bar{\chi}(u)))\right)\Big|_{u=C}$$
$$+ |\varsigma_2| \, {}^{\mathcal{RC}}\mathfrak{I}^{\mu,\theta+\delta}_{c_0}\Big|\mathfrak{D}\left(u, \chi(u), \mathcal{H}_1(\mathfrak{H}_1(\chi(u))), \mathcal{H}_2(\mathfrak{H}_2(\chi(u)))\right)$$
$$-\mathfrak{D}\left(u, \bar{\chi}(u), \mathcal{H}_1(\mathfrak{H}_1(\bar{\chi}(u))), \mathcal{H}_2(\mathfrak{H}_2(\bar{\chi}(u)))\right)\Big|_{u=\xi}\Bigg]$$
$$\leq |\varrho_1 - \bar{\varrho}_1| + \left|\frac{(C-c_0)^{\mu}}{\tilde{\Psi}_1}\right||\varrho_2 - \bar{\varrho}_2| + \left|\frac{(C-c_0)^{\mu}}{\tilde{\Psi}_1}\right|\left|\varsigma_1 + \frac{\varsigma_2(\xi-c_0)^{\mu\delta}}{\mu^{\delta}\Gamma(1+\delta)}\right||\varrho_1 - \bar{\varrho}_1|$$
$$+ \frac{(C-c_0)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)}\left(l_1 + l_2\alpha_1 l_4 + l_3\alpha_2 l_5\right)\|\chi - \bar{\chi}\|$$
$$+ \frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|(C-c_0)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)} + \frac{|\varsigma_2|(\xi-c_0)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma(\theta+\delta+1)}\right]\left(l_1 + l_2\alpha_1 l_4 + l_3\alpha_2 l_5\right)\|\chi - \bar{\chi}\|$$
$$\leq K\left(|\varrho_1 - \bar{\varrho}_1| + |\varrho_2 - \bar{\varrho}_2|\right)$$
$$+ \left\{\frac{(C-c_0)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)} + \frac{(C-c_0)^{\mu}}{|\tilde{\Psi}_1|}\left[\frac{|\varsigma_1|(C-c_0)^{\mu\theta}}{\mu^{\theta}\Gamma(\theta+1)} + \frac{|\varsigma_2|(\xi-c_0)^{\mu(\theta+\delta)}}{\mu^{(\theta+\delta)}\Gamma(\theta+\delta+1)}\right]\right\}$$
$$\times\left(l_1 + l_2\alpha_1 l_4 + l_3\alpha_2 l_5\right)\|\chi - \bar{\chi}\|.$$

Substitute the value of Q, Eq. (23), into the above expression, we get

$$\|\chi - \bar{\chi}\| \leq K\left(|\varrho_1 - \bar{\varrho}_1| + |\varrho_2 - \bar{\varrho}_2|\right) + Q\|\chi - \bar{\chi}\|.$$

where Q is defined in Eq. (23).

Combining the terms that have the common norm sign, we have

$$\|\chi - \bar{\chi}\| - Q\|\chi - \bar{\chi}\| \leq K\left(|\varrho_1 - \bar{\varrho}_1| + |\varrho_2 - \bar{\varrho}_2|\right),$$

$$\|\chi - \bar{\chi}\| \leq \frac{K}{1-Q}\left(|\varrho_1 - \bar{\varrho}_1| + |\varrho_2 - \bar{\varrho}_2|\right).$$

Therefore, for all $\epsilon > 0$, take $\psi(\epsilon) = \frac{1-Q}{K}\epsilon$, such that $\|\chi - \bar{\chi}\| < \epsilon$ whenever

$$|\varrho_1 - \bar{\varrho}_1| + |\varrho_2 - \bar{\varrho}_2| < \psi(\epsilon).$$

Thus, Eq. (1) is uniformly stable and the proof is complete. □

4. Polynomial neural network method

4.1. Construction of polynomials neural network

In this section, we represent the polynomials neural network (PNN) method for solving FVFIDEs. We consider two types Chebyshev polynomials (ChebyshevNN) of the first kind and Bernstein polynomials (BernsteinNN).

Now, we explain the pre-processing and the role of u in the network: The input u's, which are chosen randomly in [0,1] in the input layer, represent the independent variable in the fractional integro-differential equation. Then, these inputs are transformed into vectors in the hidden layer using a polynomial basis expansion as $P_n(u) = [P_0(u), P_1(u), \dots, P_{n-1}(u)]$ before being passed and fed into the network, which is called the feed-forward process. Next, the neural network processes the polynomial-transformed inputs $P_n(u)$ through a series of weighted layers $w = \{w_j\}, j = 0, 1, \dots, n-1$, producing a neural network output $N(u, w, P_n(u))$ (simplify $N(u, P_n(u))$) in the form:

$$N(u, P_n(u)) = \sum_{j=0}^{n-1} w_j P_j(u), \tag{26}$$

where the parameters w_j are chosen randomly from [0,1] and will be learned during the training process. The output will be $N(P_n(u)) = w_0 P_0(u) + w_1 P_1(u) + w_2 P_2(u) + \cdots + w_{n-1} P_{n-1}(u)$.

We define the general trial neural network approximate solution $\hat{\chi}(P_n(u))$ as

$$\hat{\chi}(P_n(u)) = \sigma(u) + \phi(u, N(P_n(u))), \tag{27}$$

where the first part $\sigma(u)$ satisfies only initial or boundary conditions, and the second part $\phi(u, N(u, P_n(u)))$ is a smooth function contains a single output $N(u, P_n(u))$ of PNN which defined in Eq. (26). The network then learns the approximation solution through trainable weights using Adam optimization method which will be explained in the next subsection.

4.2. Mathematical model

To construct the model, we apply a PNN solution mentioned in Eq. (27) into Eq. (1), yields

$$^{CC}\mathfrak{D}^{\mu,\theta}_{c_0} \hat{\chi}(u, P_n(u)) = \mathfrak{O}\left(u, \hat{\chi}(u, P_n(u)), \mathcal{H}_1(\mathfrak{H}_1(\hat{\chi}(u, P_n(u)))), \mathcal{H}_2(\mathfrak{H}_2(\hat{\chi}(u, P_n(u))))\right), \tag{28}$$

with conditions $\hat{\chi}(c_0, P_n(u)) = \varrho_1$ and $\varsigma_1 \hat{\chi}(C, P_n(u)) + \varsigma_2 {}^{\mathcal{R}C}\mathfrak{I}^{\mu,\delta}_{c_0} \hat{\chi}(\xi, P_n(u)) = \varrho_2$.

We need to minimize the above equation with its boundary conditions, i.e.

$$\begin{aligned} &Minimize \\ &\left({}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0} \hat{\chi}(u, P_n(u)) - \mathfrak{O}\left(u, \hat{\chi}(u, P_n(u)), \mathcal{H}_1(\mathfrak{H}_1(\hat{\chi}(u, P_n(u)))), \mathcal{H}_2(\mathfrak{H}_2(\hat{\chi}(u, P_n(u))))\right)\right). \end{aligned} \tag{29}$$

We define the loss function as

$$\begin{aligned} E(w) = \frac{1}{N} \Bigg[\sum_{i=1}^{N} \Bigg(&{}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0} \hat{\chi}(u, P_i(u)) - \\ &\mathfrak{O}\left(u, \hat{\chi}(u, P_i(u)), \mathcal{H}_1(\mathfrak{H}_1(\hat{\chi}(u, P_i(u)))), \mathcal{H}_2(\mathfrak{H}_2(\hat{\chi}(u, P_i(u))))\right) \Bigg)^2 + \Bigg(\hat{\chi}(c_0, P_n(u)) - \varrho_1 \Bigg)^2 \\ &+ \Bigg(\varsigma_1 \hat{\chi}(C, P_n(u)) + \varsigma_2 {}^{\mathcal{R}C}\mathfrak{I}^{\mu,\delta}_{c_0} \hat{\chi}(\xi, P_n(u)) - \varrho_2 \Bigg)^2 \Bigg] + \lambda L2, \end{aligned} \tag{30}$$

where $L2 = \|w\|_2^2 = \sum_{i=1}^{n} \sum_{j=1}^{m_i} w_{ij}^2$, n is the number of layers in the model, m_i is the number of parameters (weights) in the ith layer, w_{ij} represents the jth weight in the ith layer, w_{ij}^2 represents the square of the weight, N is the number of training sample.

We incorporate the $L2$ regularization for parameter terms (weights updated during training) to ensure the stability of training process and to overcome the fitting problems. In addition, we add the hyper-parameter λ to control the regularization strength.

Since, the boundary conditions are including in the neural network solution, the above equation reduces to

$$\begin{aligned} E(w) = &\frac{1}{N} \sum_{i=1}^{N} \Bigg[{}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0} \hat{\chi}(u, P_i(u)) \\ &- \mathfrak{O}\left(u, \hat{\chi}(u, P_i(u)), \mathcal{H}_1(\mathfrak{H}_1(\hat{\chi}(u, P_i(u)))), \mathcal{H}_2(\mathfrak{H}_2(\hat{\chi}(u, P_i(u))))\right) \Bigg]^2 \\ &+ \lambda \|w\|_2^2, \\ &= \text{Loss}(w) + \lambda \|w\|_2^2, \end{aligned} \tag{31}$$

where N is the number of training sample, and the first term computes the mean squared error of the residual on the fractional equation, and hence evaluates the minimizing for this term. While the second term $\lambda \|\mathbf{w}\|_2^2$ (with $\lambda > 0$) amends the large weights, such that the L_2 regularization stabilizes the training process and helps to prevent the over-fitting. Noted that the $\text{Loss}(w)$ mentioned above satisfies the smoothness and strong convexity properties, for more details [49–51].

Computing the gradient for the loss function in Eq. (31) using chain rule, we obtain

$$\begin{aligned} \frac{\partial E}{\partial w_i} = \frac{2}{N} \sum_{i=1}^{N} &\left\{ {}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0} \hat{\chi}\big(P_i(u)\big) - \mathfrak{O}\Big(u_i, \hat{\chi}\big(P_i(u)\big), \mathcal{H}_1\big(\mathfrak{H}_1\big(\hat{\chi}\big(P_i(u)\big)\big)\big), \mathcal{H}_2\big(\mathfrak{H}_2\big(\hat{\chi}\big(P_i(u)\big)\big)\big)\Big) \right\} \\ &\times \Delta_i B'\Big(N\big(P_n(u_i)\big)\Big) P_i\big(u_i\big) + 2\lambda w_i, \end{aligned} \tag{32}$$

where Δ_i denotes the derivative of the residual with respect to $\hat{\chi}$ defined as

$$\Delta_i = \frac{d}{d\hat{\chi}_i} \left\{ {}^{CC}\mathfrak{D}^{\mu,\theta}_{c_0} \hat{\chi}_i - \mathfrak{O}\Big(u_i, \hat{\chi}_i, \mathcal{H}_1\big(\mathfrak{H}_1(\hat{\chi}_i)\big), \mathcal{H}_2\big(\mathfrak{H}_2(\hat{\chi}_i)\big)\Big) \right\}. \tag{33}$$

$B'\left(N\left(P_n\left(u_i\right)\right)\right)$ is the derivative of B evaluated at $N\left(P_n\left(u_i\right)\right)$, and $P_k\left(u_i\right)$ is the k th basis function evaluated at u_i.

Using the back propagation learning algorithm to update weights and to minimize the error function through the Adam optimization method:

$$\text{First Moment Estimate:} \quad m_{i,t} = \beta_1 m_{i,t-1} + \left(1-\beta_1\right)\frac{\partial E}{\partial w_i}$$

$$\text{Second Moment Estimate:} \quad v_{i,t} = \beta_2 v_{i,t-1} + \left(1-\beta_2\right)\left(\frac{\partial E}{\partial w_i}\right)^2$$

where w_i are the weight that need to be updated, and $m_{i,t}$ is the estimate of the first moment at iteration t and $v_{i,t}$ is the estimate of the second moment of the gradients at iteration t, which are initialized at the initial time steps as zero vectors. β_1 and β_2 are the decay rates close to 1 (the default values for $\beta_1 = 0.9$ and $\beta_2 = 0.999$). $\frac{\partial E}{\partial w_i}$ is the gradients.

First, we calculate the bias-corrected first and second moment estimates as

$$\text{Corrected First Moment Estimate:} \quad \hat{m}_{i,t} = \frac{m_{i,t}}{1-\beta_1^t}$$

$$\text{Corrected Second Moment Estimate:} \quad \hat{v}_{i,t} = \frac{v_{i,t}}{1-\beta_2^t}.$$

Then, $\hat{m}_{i,t}$ and $\hat{v}_{i,t}$ are used to update the parameters as follows:

$$w_{i,t+1} = w_{i,t} - \frac{\eta}{\sqrt{\hat{v}_{i,t}}+\epsilon}\hat{m}_{i,t}, \tag{34}$$

where η is the learning rate and ϵ is a small constant added for numerical stability.

This update procedure allows each weight in the Bernstein neural network to be updated efficiently and stably, and ensuring that the overall method converges and the training process is robust against issues like too small or excessively large gradients.

The PNN method is explained in the Algorithm 1, while the fractional operator of Caputo-conformable derivative, the Riemann–Liouville-conformable integral and the left hand-side of the training equation are computed numerically using the quadrature rule with Python program.

Algorithm 1 PNN technique for solving FVFIDEs

Require: Degree of polynomial n, Number of hidden units N, Learning rate $lr = 0.0001$, Number of epochs $e = 1000$, the fractional degree θ, error tolerance parameter ε.
Ensure: Predicted solutions, MAE, and MSE.

```
Initialize model with polynomials.
Initialize optimizer and learning rate
for epoch = 1 to e do
    Zero the parameter gradients
    Compute the loss:
        Define trial solution incorporating boundary conditions
        Compute: the integer order derivative, the fractional derivative using Caputo-conformable derivative, the fractional conformable integral by quadrature rule
        Define the Eq. (29), and use quadrature rule.
        Compute the loss function including boundary conditions and L2 regularization
    if loss is NaN then
        Stop training (If E(w) < ε stop, otherwise go next step)
    end if
    if (epoch + 1) mod 500 == 0 then
        Print the current epoch and loss
    end if
    Back-propagate the loss and update model parameters using Eq. (34)
    Step the learning rate
end for
for each value in θ do
    Initialize model and train as described above
    Test the model at specific points u
    Compute predicted solutions and exact solution
    Calculate Mean Absolute Error (MAE) and Mean Squared Error (MSE)
    Print results and plot solutions
end for
Compute Errors:
    MAE, MSE, then print and plot the results
```

As the degree of polynomial $P_n(u)$ increases, the approximation error decreases, and the solution converges to the true solution. The MSE provides a quantitative measure of this approximation error, and by minimizing the MSE, we ensure that the polynomial approximation converges to the true solution.

5. Numerical analysis

In this section, we provide the numerical analysis for using the polynomial neural network in solving the fractional Volterra–Fredholm integro-differential equation. The analysis includes investigating the error bound, convergence of the method, computational sensitivity and stability, and the computational cost.

5.1. Convergence analysis and error estimate

Saini et al. [52,53] and Kumar and Das [54] used the Taylor expansion to compute the error bound and studied the convergence analysis, Santraet al. [3] used the non-uniform meshes on multi-term time fractional convection–diffusion reaction problems for Volterra-type integro-differential equations with weakly singular kernel. In our approach, we apply the Taylor expansion in the neural network framework.

First, we need the following notations in our work.

$\{w_j\}_{j=0}^{n}$ is the actual weights used in trail solution with regularization, $\{w_j^*\}_{j=0}^{n}$ is the weights without regularization. Let $\chi \in C^2\left(\left[c_0, C\right]\right)$ be the sufficiently smooth exact solution, $\hat{\chi}$ is the trail solution (with regularization) and $\tilde{\chi}$ is the trail solution (without regularization).

Also, let $B_n^*(\chi)(u)$ be the Bernstein polynomial approximation of $\chi(u)$, obtained from the Bernstein basis (without regularization) and $B_n(\chi(u))$ is the actual output of the Bernstein neural network (with regularization). Let $E_{\text{app}}(u)$ is the approximation error, $E_{\text{opt}}(u)$ be the optimization error, and $E_{\text{reg}}(u)$ be the regularization error used to improving the stability. Define $\|\chi\|$ is the supremum norm over $[c_0, C]$, and assume that the fractional operators ${}^{CC}\mathfrak{D}_{c_0}^{\mu,\alpha}$ and $\mathcal{R}C_{\mathfrak{I}_{c_0}}^{\mu,\delta}$ are Lipschitz continuous on the appropriate function spaces.

Using the smoothness and uniformly bounded properties of Bernstein polynomials on $[0,1]$ (and hence on $[c_0, C]$ with some transformation on the interval), we show that the function $\phi(u, N(u, P_n(u)))$ in Eq. (27) satisfied the Lipschitz continuity.

Lemma 5.1. *Let $\phi : [c_0, C] \times \mathbb{R} \to \mathbb{R}$ be a function defined in Eq. (27). Assume that the function $\phi(u,s)$ is differentiable and uniformly bounded with respect to its second argument. For all $u \in [c_0, C]$ and $s \in \mathbb{R}$, there exist a constant $L_\phi > 0$ such that*

$$\left|\frac{\partial \phi}{\partial u}(u, N(u, B_n(u)))\right| \leq L_\phi.$$

Then, $\phi(u, N(u, B_n(u)))$ is Lipschitz continuous.

Proof. For simplicity, we use the notation $s := N(u, B_n(u))$ which is defined exactly in Eq. (26).

Let $s_1, s_2 \in \mathbb{R}$ be the arbitrary values in the range of the weighted sum. Then, for each fixed $u \in [c_0, C]$ and by applying the mean value theorem, there exists some $c \in \left[s_1, s_2\right]$ such that

$$\phi(u, N(u, s_1)) - \phi(u, N(u, s_2)) = \frac{\partial \phi}{\partial s}(u, c)\left(s_1 - s_2\right)$$

Taking the absolute value of both sides yields

$$\left|\phi(u, N(u, s_1)) - \phi(u, N(u, s_2))\right| = \left|\frac{\partial \phi}{\partial s}(u, c)\right|\left|s_1 - s_2\right|$$

Since ϕ is uniformly bounded, it follows that

$$\left|\phi(u, N(u, s_1)) - \phi(u, N(u, s_2))\right| \leq L_\phi \left|s_1 - s_2\right|$$

Thus, for every fixed u, the function $\phi(u, N(u, B_n(u)))$ is Lipschitz continuous with constant L_ϕ, which completes the proof. □

Lemma 5.2. *Let $\chi \in C^2\left(\left[c_0, C\right]\right)$ be the exact solution of Eq. (1), and $B_n(\chi)$ be the Bernstein polynomials for the neural network. Then, we have*

$$\|\chi - B_n(\chi)\| \leq \frac{\gamma_1}{n},$$

where $\gamma_1 = \frac{1}{4}\sup_{u \in [c_0, C]}\left|\chi''(u)\right|$. Moreover, the error is of the following order:

$$\|\chi - B_n(\chi)\| = O\left(\frac{1}{n}\right). \tag{35}$$

Proof. First, set $F = \chi - \sigma$. Then, using Eq. (26), the output neural network for Bernstein polynomials $N(u, B_n(u))$ can be written as

$$N(u, B_n(u)) = \sum_{j=0}^{n-1} w_j B_j(u). \tag{36}$$

By comparing this notation with the notation used in Eq. (5), we noted that $w_j = F\left(\frac{j}{n}\right)$. Since F is sufficiently smooth such as $F \in C^2[c_0, C]$, then by using Taylor expansion, we have

$$F\left(\frac{j}{n}\right) = F(u) + F'(u)\left(\frac{j}{n} - u\right) + \frac{F''\left(\xi_j\right)}{2}\left(\frac{j}{n} - u\right)^2, \tag{37}$$

where $\xi_j \in (u, \frac{j}{n})$.

Substitute Eq. (37) into Eq. (5), we obtain

$$\begin{aligned}
B_n(F)(u) &= \sum_{j=0}^{n}\left[F(u) + F'(u)\left(\frac{j}{n} - u\right) + \frac{F''\left(\xi_j\right)}{2}\left(\frac{j}{n} - u\right)^2\right]B_{n,j}(u) \\
&= F(u)\sum_{j=0}^{n} B_{n,j}(x) + F'(u)\sum_{j=0}^{n}\left(\frac{j}{n} - u\right)B_{n,j}(u) + \frac{1}{2}\sum_{j=0}^{n} F''\left(\xi_j\right)\left(\frac{j}{n} - u\right)^2 B_{n,j}(u), \\
&= F(u) + \frac{1}{2}\sum_{j=0}^{n} F''\left(\xi_j\right)\left(\frac{j}{n} - u\right)^2 B_{n,j}(u).
\end{aligned}$$

Note that the error is entirely determined by the third term, since the second term vanishes because the Bernstein basis is constructed such that the mean is u. Thus,

$$F(u) - B_n(F)(u) = -\frac{1}{2}\sum_{j=0}^{n} F''\left(\xi_j\right)\left(\frac{j}{n} - u\right)^2 B_{n,j}(u).$$

Taking the absolute value, we obtain

$$\left|F(u) - B_n(F)(u)\right| \leq \frac{1}{2}\sup_{u\in[c_0,C]}\left|F''(u)\right|\sum_{j=0}^{n}\left(\frac{j}{n} - u\right)^2 B_{n,j}(u).$$

Note that the term $\sum_{j=0}^{n}\left(\frac{j}{n} - u\right)^2 B_{n,j}(u)$ is the variance of a binomial random variable with parameters n and u, which equals to $\frac{u(1-u)}{n}$. Thus,

$$\left|F(u) - B_n(F)(u)\right| \leq \frac{1}{2}\sup_{u\in[c_0,C]}\left|F''(u)\right|\frac{u(1-u)}{n}.$$

Take the supremum, we have

$$\|F - B_n(F)\| \leq \frac{\gamma_1}{n},$$

where $\gamma_1 = \frac{1}{4}\sup_{u\in[c_0,C]}\left|F''(u)\right|$. Substitute χ, we have

$$\|\chi(u) - \sigma(u) - B_n(\chi)\| \leq \frac{\gamma_1}{n},$$

where $\gamma_1 = \frac{1}{4}\sup_{u\in[c_0,C]}\left|\chi''(u)\right|$. This is exactly

$$\|\chi(u) - \tilde{\chi}(u, B_n(u))\| \leq \frac{\gamma_1}{n}$$

where $\tilde{\chi}(u, B_n(u)) = \sigma(u) + \phi(u, N(u, B_n(u)))$. Hence, the error is uniform over the domain with order

$$\|\chi - B_n(\chi)\| = O\left(\frac{1}{n}\right), \tag{38}$$

which completes the proof. □

In the case that the neural network is working without adding regularization parameter with generating weights w_j^*, the actual neural network solution $\tilde{\chi}(u)$ can be expressed as

$$\tilde{\chi}(u, B_n(u)) = \sigma(u) + \phi(u, N(u, B_n(u))), \tag{39}$$

such that the weighted sum $\sum_{j=0}^{n} w_j^* B_{n,j}(u)$ perfectly approximates $\hat{\chi}(u) - \sigma(u)$ after being processed by $\phi(u, N(u, B_n(u)))$ then ideally, we would have $\tilde{\chi}(u) \approx \hat{\chi}(u)$. From Lemma 5.2, we conclude that

$$\|\chi(u) - \tilde{\chi}(u, B_n(u))\| \leq \frac{\gamma_1}{n}.$$

Due to finite iterations, numerical precision, and possible non-convexity of the loss function in the numerical method, we only achieve an approximate minimize. In Eq. (31), set

$$\begin{aligned}\mathcal{R}(u, w_j) = & {}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta}\hat{\chi}(u, P_j(u)) \\ & - \mathfrak{O}\left(u, \hat{\chi}(u, P_j(u)), \mathcal{H}_1(\mathfrak{H}_1(\hat{\chi}(u, P_j(u)))), \mathcal{H}_2(\mathfrak{H}_2(\hat{\chi}(u, P_j(u))))\right)\end{aligned} \tag{40}$$

be the residual for a given set of weights w.

In our training algorithm, we used Adam optimizer to find the weights w_j which differs from these ideal weights $w^*{}_j$. The difference $w^*{}_j - w_j$ gives rise to an additional error in the neural network output. Define the optimization error term in residual form by $E_{opt} := \mathcal{R}(u, w_j^*) - \mathcal{R}(u, w_j)$. The following result gives the residual error. Also, we use the smoothness and strong convexity properties of $\mathrm{Loss}(w)$ as known in the neural network methods.

Lemma 5.3. *Let $\hat{\chi}(u)$ and $\tilde{\chi}(u)$ are the trail neural network solutions defined in Eqs. (27) and (39), respectively. Then, the optimization error $E_{opt}\,(u) := \mathcal{R}(u, w_j^*) - \mathcal{R}(u, w_j)$ is bounded with maximum error given by*

$$\varepsilon_{opt} = \sup_{u\in[c_0,C]} E_{opt}\,(u) \leq \gamma_2(\lambda), \tag{41}$$

where $\mathcal{R}(u, w_j)$ is define in Eq. (40), and $\gamma_2(\lambda)$ be the estimated function depends on the regularization parameter λ. Moreover, if $\varepsilon_{opt} \to 0$, then the training converges.

Proof. Set $\Delta\chi := \hat{\chi}(u; w^*{}_j) - \tilde{\chi}(u; w_j)$, such that $B_n(u; w^*{}_j)$ and $B_n(u; w_j)$ be the Bernstein polynomials with respect to wights $w^*{}_j$ and w_j, respectively. Then $\Delta\chi$ can be written as

$$\Delta\chi = \phi(u, N(u; B_n(u; w^*{}_j))) - \phi(u; N(u, B_n(u; w_j))) \tag{42}$$

Using Lemma 5.1, the function $\phi(u, N(u, B_n(u)))$ is Lipschitz continuous with constant L_ϕ, then for any two inputs $\sum_{j=0}^{n} w^*{}_j B_{n,j}(u)$ and $\sum_{j=0}^{n} w_j B_{n,j}(u)$, we have

$$\|\Delta\chi\| \leq L_\phi \|\sum_{j=0}^{n} w^*{}_j B_{n,j}(u) - \sum_{j=0}^{n} w_j B_{n,j}(u)\| \leq L_\phi \|\sum_{j=0}^{n}(w^*{}_j - w_j)B_{n,j}(u)\|. \tag{43}$$

Since $E_{\text{opt}}(u) := \mathcal{R}(u, w_j^*) - \mathcal{R}(u, w_j)$, then the error optimization can be simplified as

$$\begin{aligned} E_{\text{opt}}(u) :=& \mathcal{R}(u, w_j^*) - \mathcal{R}(u, w_j) \\ =& {}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta} \hat{\chi}(u, B_j(u; w_j^*)) - {}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta} \tilde{\chi}(u, B_j(u; w_j)) \\ &- \mathfrak{D}\left(u, \hat{\chi}(u, B_j(u; w^*{}_j)), \mathcal{H}_1(\mathfrak{H}_1(\hat{\chi}(u, B_j(u; w_j^*)))), \mathcal{H}_2(\mathfrak{H}_2(\hat{\chi}(u, B_j(u; w_j^*))))\right) \\ &+ \mathfrak{D}\left(u, \tilde{\chi}(u, B_j(u; w_j)), \mathcal{H}_1(\mathfrak{H}_1(\tilde{\chi}(u, B_j(u; w_j)))), \mathcal{H}_2(\mathfrak{H}_2(\tilde{\chi}(u, B_j(u; w_j))))\right). \end{aligned} \tag{44}$$

Using the linearity and bounded properties of Caputo-conformable fractional derivatives Lemma 3.8, we have

$$\left|{}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta} \hat{\chi}(u, B_j(u; w_j^*)) - {}^{CC}\mathfrak{D}_{c_0}^{\mu,\theta} \tilde{\chi}(u, B_j(u; w_j))\right| \le \frac{(u - c_0)^{\mu(n-\theta)}}{\mu^{n-\theta}\Gamma(n-\theta+1)} \left\|\mathfrak{D}_{c_0}^{\mu,n} \Delta\chi\right\|. \tag{45}$$

Employing the hypotheses in Theorem 3.11 and the result of Eq. (45) into Eq. (44), then take the supremum over $[c_0, C]$,we obtain

$$|E_{\text{opt}}(u)| \le \frac{(C - c_0)^{\mu(n-\theta)}}{\Gamma(n-\theta+1)\mu^{n-\theta}}(\tilde{M} - \hat{M}) + \hat{K}\|\Delta\chi\|, \tag{46}$$

where $\tilde{M}$ and $\hat{M}$ are the bounded for the nth conformable derivatives for $\hat{\chi}(u, B_j(u; w^*{}_j))$ and $\tilde{\chi}(u, B_j(u; w_j))$, respectively, on $\left[c_0, C\right]$, and $\hat{K} = (\ell_1 + \ell_2 + \ell_3 + \ell_4\alpha_1 + \ell_5\alpha_2)$. Substitute Eq. (43) into Eq. (46), we obtain

$$\varepsilon_{opt} := \|E_{\text{opt}}(u)\| \le \frac{(C - c_0)^{\mu(n-\theta)}}{\Gamma(n-\theta+1)\mu^{n-\theta}}(\tilde{M} - \hat{M}) + \hat{K}L_\phi \|\sum_{j=0}^{n}(w^*{}_j - w_j)B_{n,j}(u)\|. \tag{47}$$

Using the convex combination propriety of the Bernstein basis functions and the L_2-norm for the weights and its estimated by the regularization parameter λ, we have

$$\sup_{u\in[c_0,C]} \|\sum_{j=0}^{n}\left(w_j^* - w_j\right)B_{n,j}(u)\| \le \|w^* - w\|_2 \le C\lambda \tag{48}$$

Thus, the overall optimization error in Eq. (47) is bounded by

$$\varepsilon_{opt} \le \frac{(C - c_0)^{\mu(n-\theta)}}{\Gamma(n-\theta+1)\mu^{n-\theta}}(\tilde{M} - \hat{M}) + \hat{K}L_\phi C\lambda. \tag{49}$$

Note that this estimation depends on λ, for simplicity, let

$$\gamma_2(\lambda) = \frac{(C - c_0)^{\mu(n-\theta)}}{\Gamma(n-\theta+1)\mu^{n-\theta}}(\tilde{M} - \hat{M}) + \hat{K}L_\phi C\lambda, \tag{50}$$

and hence

$$\varepsilon_{\text{opt}} \le \gamma_2(\lambda). \tag{51}$$

If $\varepsilon_{\text{opt}} \to 0$, then the training converges, and the proof is completed. □

The last error sources for our method is the regularization error denoted by E_{reg}. The L_2 regularization term $\lambda\|w\|_2^2$ biases the weight values to avoid over-fitting and ensure stability. This introduces a bias between the ideally optimized network (without regularization) and the actual network output. We denote

$$E_{\text{reg}} = \|\tilde{\chi}(u) - \hat{\chi}(u)\| = O(\lambda)$$

meaning that the error from the regularization is of the order λ.

Recall the definitions of these notations, $\tilde{\chi}(u, B_n(u))$ defined in Eq. (39) which gives the trial neural network solution without using regularization parameter λL_2, and let w_k^λ denotes the weight comes from the neural network solution with regularization parameter in the trial solution $\hat{\chi}(u, B_n(u))$ which is defined in Eq. (27).

Then, the regularization error is the difference between the training solution (without regularization) and the actual training (with regularization), which is written as

$$E_{\text{reg}} = \|\tilde{\chi}(u, B_n(u)) - \hat{\chi}(u, B_n(u))\|.$$

We have the following result:

Lemma 5.4. *Let $\phi(u, N(u, B_n(u)))$ satisfies the continuity of Lipschitz property, and consider the smoothness and strong convexity properties of the loss function* Loss(w). *Let $\hat{\chi}$ and $\tilde{\chi}$ be the neural network solutions as defined in Eqs. (27) and (39), respectively. Then the regularization error E_{reg} is bounded with maximum error*

$$E_{reg} = \|\tilde{\chi}(u, B_n(u)) - \hat{\chi}(u, B_n(u))\| = O(\lambda). \tag{52}$$

Proof. Using the continuity Lipschitz property Lemma 5.1 with the same process as in Lemma 5.3, we have

$$E_{\text{reg}} = \|\sum_{j=0}^{n}\left(w_j^* - w_j^\lambda\right)B_{n,j}(u)\| \le L_\phi\|s^* - s^\lambda\|, \tag{53}$$

where $s^* = \sum_{j=0}^{n} w_j^* B_{n,j}(u)$ and $s^\lambda = \sum_{j=0}^{n} w_j^\lambda B_{n,j}(u)$.

Using property of the bounded of Bernstein basis Eq. (7), then the norm of the weighted difference is controlled by the difference between the weights. Hence, the difference between the non-regularized and regularized minimizing is of order λ. To justify this, let $L(w)$ be the non-regularized loss function with its minimizer w^*, then the first-order optimality condition gives $\nabla L(w^*) = 0$. Also, for the regularized loss function $J(w) = L(w) + \lambda\|w\|_2^2$, where $\lambda > 0$ is the regularization parameter. If w^λ minimizes $J(w)$, then its first-order optimality condition is $\nabla J\left(w^\lambda\right) = \nabla L\left(w^\lambda\right) + 2\lambda w^\lambda = 0$.

Subtract both optimality conditions, we obtain

$$\nabla L\left(w^\lambda\right) + 2\lambda w^\lambda - \nabla L\left(w^*\right) = 0.$$

Since $\nabla L(w^*) = 0$, this simplifies to:

$$\nabla L\left(w^\lambda\right) + 2\lambda w^\lambda = 0 \tag{54}$$

Assume that $L(w)$ is twice continuously differentiable and that its gradient is Lipschitz continuous (which follows from the smoothness) near w^*. Then we can apply a first-order Taylor expansion around w^* for the gradient:

$$\nabla L\left(w^\lambda\right) = \nabla L\left(w^*\right) + \nabla^2 L\left(w^*\right)\left(w^\lambda - w^*\right) + R,$$

where R is the remainder of order $O\left(\|w^\lambda - w^*\|_2^2\right)$. Since $\nabla L(w^*) = 0$, we have

$$\nabla L\left(w^\lambda\right) = \nabla^2 L\left(w^*\right)\left(w^\lambda - w^*\right) + O\left(\left\|w^\lambda - w^*\right\|_2^2\right).$$

For sufficiently small λ (and hence w^λ close to w^*), the quadratic remainder can be neglected in a first order analysis. Thus, we approximate:

$$\nabla L\left(w^\lambda\right) \approx \nabla^2 L\left(w^*\right)\left(w^\lambda - w^*\right). \tag{55}$$

Substitute Eq. (55) into Eq. (54), we obtain

$$\nabla^2 L\left(w^*\right)\left(w^\lambda - w^*\right) + 2\lambda w^\lambda \approx 0.$$

Assuming that $L(w)$ is strongly convex in a neighborhood of w^* (i.e. its Hessian $\nabla^2 L(w^*)$ is positive definite), the matrix $\nabla^2 L(w^*)$ is invertible. We then obtain

$$w^\lambda - w^* \approx -2\lambda\left(\nabla^2 L\left(w^*\right)\right)^{-1} w^\lambda.$$

Taking the L_2 norm on both sides, we get:

$$\left\|w^\lambda - w^*\right\|_2 \leq 2\lambda\left\|\left(\nabla^2 L\left(w^*\right)\right)^{-1}\right\|_2\left\|w^\lambda\right\|_2.$$

Under the assumption that $\|w^\lambda\|_2$ remains bounded and that the smallest eigenvalue of $\nabla^2 L(w^*)$ is at least $\mu > 0$ (i.e. $\left\|\left(\nabla^2 L(w^*)\right)^{-1}\right\|_2 \leq 1/\mu$), we deduce:

$$\left\|w^\lambda - w^*\right\|_2 = O(\lambda).$$

Let $\Delta w = w^\lambda - w^*$ be the difference between the above weights, then $\|\Delta w\|_2 = O(\lambda)$. Returning to our error term Eq. (53), we obtain

$$\|\sum_{j=0}^{n}\left(w_j^* - w_j^\lambda\right) B_{n,j}(u)\| \leq \|\Delta w\|_2 \cdot \max_u \|\left\{B_{n,j}(x)\right\}_{j=0}^n\|.$$

Note that $\|\Delta w\|_2$ depends on λ, then there exist a constant $Z > 0$ such that $\|\Delta w\|_2 = Z\lambda$. Moreover, the Bernstein basis functions on a compact interval are uniformly bounded (with the maximum L_2 norm of the vector being at most 1), take $L_B = \max_u \|\left\{B_{n,j}(x)\right\}_{j=0}^n\|$, we obtain

$$\|\sum_{j=0}^{n}\left(w_j^* - w_j^\lambda\right) B_{n,j}(u)\| \leq Z\lambda L_B. \tag{56}$$

Set

$$\gamma_3(\lambda) = Z L_B \lambda > 0, \tag{57}$$

$$\|\sum_{j=0}^{n}\left(w_j^* - w_j^\lambda\right) B_{n,j}(u)\| \leq \gamma_3(\lambda). \tag{58}$$

Hence,

$$E_{\text{reg}} = \|\tilde{\chi}(u, B_n(u)) - \hat{\chi}(u, B_n(u))\| = O(\lambda) \tag{59}$$

and the proof is completed. □

Note that E_{opt} in Lemma 5.3 and E_{reg} in Lemma 5.4 are related but not the same, since they come from different aspects of error. Hence, we can combine them in the final error later.

The following theorem, with its assumptions and result, forms the rigorous foundation for both the error bound and the convergence of the Bernstein neural network method for solving the given fractional integro-differential equation with boundary conditions.

Now, we present the main theorem as follows:

Theorem 5.5 (*Error Bound and Convergence Analysis*). *Let $\chi \in C^2\left(\left[c_0, C\right]\right)$ be the exact solution of Eq. (1), where $\alpha \in (1,2]$, $\mu \in (0,1]$ and $\delta > 0$ with the trail neural network solution defined in Eq. (27). Then, the total error of the Bernstein neural network approximation satisfies*

$$\|\chi(u) - \hat{\chi}(u)\| \leq \frac{\gamma_1}{n} + \gamma_2(\lambda) + \gamma_3(\lambda),$$

where $\gamma_1 > 0$, $\gamma_2(\lambda)$ and $\gamma_3(\lambda)$ are defined in Lemma 5.2, Eqs. (50) and (57), respectively.

Moreover, for n large enough and λ is sufficiently small such that

$$\lim_{n\to\infty,\lambda\to 0} \|\chi(u) - \hat{\chi}(u)\| = 0,$$

and the residual

$$\|{}^{CC}\mathfrak{D}^{\mu,\alpha}_{c_0}\left(\chi(u) - \hat{\chi}(u)\right)\| \to 0,$$

then the approximation method is convergent.

Proof. We decompose the total error between the true solution and the network approximation as

$$\chi(u) - \hat{\chi}(u) = \chi(u) - \tilde{\chi}(u, B_n(u)) + \tilde{\chi}(u, B_n(u)) - \hat{\chi}(u)$$

which can be divided into

$$\chi(u) - \hat{\chi}(u) = \underbrace{\chi(u) - \tilde{\chi}(u, B_n(u))}_{E_{app}} + \underbrace{\tilde{\chi}(u, B_n(u)) - \hat{\chi}(u)}_{E_{opt}+E_{reg}}$$

Taking norm and applying the triangle inequality gives:

$$\|\chi(u) - \hat{\chi}(u)\| \leq \|\chi(u) - \tilde{\chi}(u, B_n(u))\| + \|\tilde{\chi}(u, B_n(u)) - \hat{\chi}(u)\|.$$

Using Lemmas 5.2–5.4, we conclude that the overall error is bounded by

$$\|\chi(u) - \hat{\chi}(u)\| \leq \frac{\gamma_1}{n} + \gamma_2(\lambda) + \gamma_3(\lambda). \tag{60}$$

Since

$$\lim_{n\to\infty,\lambda\to 0} \|\chi(u) - \hat{\chi}(u)\| = 0,$$

the Lipschitz continuity property of Caputo-conformable operator, there exist a constant $L_\theta > 0$ such that

$$\|{}^{CC}D^{\mu,\theta}_{c_0}(\chi)(u) - {}^{CC}D^{\mu,\theta}_{c_0}\left(\hat{\chi}(u)\right)\| \leq L_\theta \|\chi(u) - \hat{\chi}(u)\|.$$

Using the linearity property of the fractional operator, we have

$${}^{CC}D^{\mu,\theta}_{c_0}(\chi) - {}^{CC}D^{\mu,\theta}_{c_0}\left(\hat{\chi}\right) = {}^{CC}D^{\mu,\theta}_{c_0}\left(\chi - \hat{\chi}\right).$$

Therefore,

$$\|{}^{CC}D^{\mu,\theta}_{c_0}(\chi) - {}^{CC}D^{\mu,\theta}_{c_0}\left(\hat{\chi}\right)\| \leq L_\theta \|\chi - \hat{\chi}\|. \tag{61}$$

Substitute Eq. (60) into Eq. (61), we obtain

$$\|{}^{CC}D^{\mu,\theta}_{c_0}(\chi) - {}^{CC}D^{\mu,\theta}_{c_0}\left(\hat{\chi}\right)\| \leq L_\theta\left(\frac{\gamma_1}{n} + \gamma_2(\lambda) + \gamma_3(\lambda)\right). \tag{62}$$

By choosing n large enough and λ is sufficiently small, then

$$\lim_{n\to\infty,\lambda\to 0} \|{}^{CC}D^{\mu,\theta}_{c_0}(\chi) - {}^{CC}D^{\mu,\theta}_{c_0}\left(\hat{\chi}\right)\| = 0.$$

Hence, we conclude that $\hat{\chi}(u)$ converges to $\chi(u)$ and the proposed method is convergent. The proof is completed. □

Now, we extended this work to the Chebyshev neural network, use Eq. (10) and set the fixed constant

$$c_{i,n} = (2n-1)!\frac{(-1)^{n-i}}{(2i-1)!(2n-2i-1)!}, \tag{63}$$

which converts the Bernstein basis into Chebyshev polynomials. If these coefficients are bounded, then any error bound obtained for a Bernstein approximation can be scaled by an appropriate constant to yield an error bound for the Chebyshev approximation. For this purpose, let the trail neural network solution of Chebyshev neural network method be defined by

$$\hat{\chi}(T_n(u)) = \sigma(u) + \phi(u, N(T_n(u))), \tag{64}$$

where the output neural network is

$$N(T_n(u)) = \sum_{j=0}^{n-1} w_j T_j(u). \tag{65}$$

Then, the following result investigates the error bound and the convergence analysis of the Chebyshev neural network.

Theorem 5.6. *Assume that $\chi \in C^2\left(\left[c_0, C\right]\right)$ be the exact solution of Eq. (1), where $\alpha \in (1,2]$, $\mu \in (0,1]$, $\delta > 0$, and $M > 0$ be an upper bound on the norm of the coefficient vector $\left\{c_{i,n}\right\}_{i=0}^{n}$. For $\gamma_1 > 0$, $\gamma_2(\lambda)$ and $\gamma_3(\lambda)$ are defined in Lemma 5.2, Eqs. (50) and (57), respectively, the total error for the Chebyshev neural network approximation satisfies*

$$\|\chi(u) - \hat{\chi}(u)\| \leq M(\frac{\gamma_1}{n} + \gamma_2(\lambda) + \gamma_3(\lambda)),$$

If n large enough and λ is sufficiently small such that

$$\lim_{n\to\infty,\lambda\to 0} \|\chi(u) - \hat{\chi}(u)\| = 0,$$

and the residual

$$\|{}^{CC}\mathfrak{D}_{c_0}^{\mu,\alpha}\left(\chi(u) - \hat{\chi}(u)\right)\| \to 0.$$

Then, the Chebyshev neural network method is convergent.

Proof. By following the construction proof of Theorem 5.5, the proof of this theorem can easily be established.

Since $T_n(u)$ is a linear combination of $B_{n,i}(u)$, the error in approximating $\chi(u)$ by a Chebyshev expansion constructed using the same sample points can be estimated by

$$\|\chi(u) - \hat{\chi}(u)\| = \left|\left|\chi(u) - \sum_{i=0}^{n} \chi\left(\frac{i}{n}\right) T_{n,i}(u)\right|\right| \leq M\frac{\gamma_1}{n}$$

where M is an upper bound on the norm of the coefficient vector $\left\{c_{i,n}\right\}_{i=0}^{n}$.

Using results in Theorem 5.5 and Eq. (10), we conclude that the Chebyshev neural network has the error bound

$$\|\chi(u) - \hat{\chi}(u)\| \leq M(\frac{\gamma_1}{n} + \gamma_2(\lambda) + \gamma_3(\lambda)).$$

Choosing n large enough and λ sufficiently small, then

$$\lim_{n\to\infty,\lambda\to 0} \|\chi(u) - \hat{\chi}(u)\| = 0.$$

Therefore, the Chebyshev neural network is bounded and convergent. Also, by the Lipschitz continuity of the fractional operators, the residual

$$\|{}^{CC}\mathfrak{D}_{c_0}^{\mu,\alpha}\left(\chi(u) - \hat{\chi}(u)\right)\| \to 0.$$

Hence, $\hat{\chi}(u)$ converges to $\chi(u)$ and the method is convergent. □

5.2. Computational complexity

In a practical situation, the running time of an algorithm naturally depends on the speed with which the underlying computations are performed. Therefore computational complexity, is usually measured in terms of number of operations (additions, multiplications, and storage) [55,56].

Time Complexity:

The forward evaluation which comes from different sources such as Bernstein basis evaluation in Eqs. (26) and (6) which implies binomial coefficients, multiplications, and power computations. Summing over n terms makes the forward evaluation cost is of order $O(n)$ for each u. Additionally, back-propagation process which is computed the gradient with respect to each weight w_k. We have n weights with trial solution $\hat{\chi}(u, P_n(u))$, the cost per training example is also of order $O(n)$. Moreover, the fractional operator where Caputo-conformable derivative or the Riemann–Liouville-conformable integral are computed numerically using quadrature rule. Suppose we use r quadrature nodes, then the cost per evaluation is of order $O(r)$. With m training points, every iteration of training requires computing these fractional derivatives or integrals for each training point, yielding the cost of order $O(m \times r)$. Hence, combining the costs over all times per iteration gives the total of order $O(m \times n \times r)$. In implementation code, we choose r is fixed constant, and then the main factors effect on the time complexity is the number of training points m and Bernstein basis functions n such that the cost is of order $O(m \times n)$.

Space Complexity:

The computation cost comes from different resources, such as network parameters like the n weights which sorted by neural network with space cost of order $O(n)$. The computation during the forward propagation and back-propagation such as the gradients are computed with storage of order $O(n)$, while the quadrature nodes q during the computation of the fractional operators does not affect on the computational space since q is relative small as compared to m and n. Hence, the overall space complexity is dominated by the weights of order $O(n)$. Note that the boundary conditions does not affect the computational complexity because $\sigma(u)$ is typically a known, simple function.

The above analysis shows that the Bernstein neural network method measures linearly with respect to the parameters, making it computationally efficient and scalable for solving the fractional differential equation with boundary conditions.

5.3. Computational stability and sensitivity

In this subsection, we introduce a brief discussion related to the affect of using some parameters on the stability of the method. One of these parameter is the value of computation in the regularization term L_2: $\lambda\|\mathbf{w}\|_2^2$ which controls the magnitude of the weights. It eliminates the over-fitting and ensuring that the large gradients or weight updates do not grow unbounded in the training process, and hence improves the numerical stability of the backward propagation procedure. Another parameter is the adaptive learning rate in Adam optimizer method. This parameter automatically scales the gradients and prevents overshooting problem, and reserves the convergence through training process in the case of the presence of noisy gradient estimates.

Moreover, the method's sensitivity is determined by the choice of λ. If the value of λ is too large, it may bias the solution, where as if it is too small it may allow over-fitting. Additionally, the number of Bernstein basis functions n is very sensitive since choosing a larger n will improve the approximation but the computational cost increases. The accuracy of numerical integration for the fractional derivative and integral terms also play main role since choosing the quadrature methods with enough nodes are required to ensure the operator evaluations are accurate.

Table 1
The MSE and MAE for ChebyshevNN and BernsteinNN computed on different degree of the polynomials, with 1000 epochs, Example 6.1.

PNN	MSE/MAE	$n = 4$	$n = 6$	$n = 8$	$n = 30$
BernsteinNN	MSE	1.2×10^{-2}	3.0×10^{-5}	6.4×10^{-3}	9.9×10^{-4}
	MAE	9.2×10^{-2}	3.9×10^{-3}	6.6×10^{-2}	2.6×10^{-2}
ChebyshevNN	MSE	9.2×10^{-3}	8.9×10^{-3}	2.6×10^{-2}	2.3×10^{-3}
	MAE	7.7×10^{-2}	7.6×10^{-2}	1.3×10^{-1}	3.6×10^{-2}

Table 2
The comparison between ChebyshevNN and BernsteinNN approximate solutions for Example 6.1.

u	Exact	Approximate solution		Error	
		BernsteinNN	ChebyshevNN	BernsteinNN	ChebyshevNN
0.1	0.12589	0.12447	0.15942	0.00143	0.03353
0.2	0.23492	0.23526	0.31800	0.00033	0.08308
0.3	0.33838	0.33725	0.47102	0.00113	0.13263
0.4	0.43838	0.43255	0.60548	0.00583	0.16709
0.5	0.53589	0.52530	0.70673	0.01059	0.17085
0.6	0.63145	0.62039	0.76896	0.01106	0.13752
0.7	0.72542	0.71950	0.80442	0.00591	0.07900
0.8	0.81805	0.81865	0.84376	0.00060	0.02571
0.9	0.90953	0.91164	0.91728	0.00210	0.00775
1.0	1.00000	1.00435	1.00435	0.00435	0.00435

6. Examples

Example 6.1. Consider the following Caputo-conformable FVFIDEs:

$$ {}^{CC}\mathfrak{D}_0^{0.9,1.5}\chi(u) = \frac{10}{29} - \frac{10u}{19} - \frac{5u^{14/5}}{14} + \int_0^1 (u-v)\ \chi(v)\mathrm{d}v + \int_0^u \chi^2(v)\mathrm{d}v, $$

with boundary conditions,

$$ \chi(0) = 0, \qquad \chi(1) + {}^{\mathcal{RC}}\mathfrak{I}_0^{0.9,1.5}\chi\left(\frac{1}{3}\right) = \frac{16\sqrt{\frac{10}{\pi}}}{729\sqrt[4]{3}} + 1,\ u \in [0,1]. \tag{66} $$

The exact solution is $\chi(u) = u^{0.9}$.

To construct a trial solution that satisfies these conditions by design, we express $\chi(u)$ in the form

$$ \hat{\chi}(u) = \hat{\chi}(u, P_n(u)) = Cu + u(1-u)\Big(u - \frac{1}{3}\Big)N(u,\mathbf{w}), \tag{67} $$

where: $C = 0.832$, and the neural network component is given by

$$ N(u,\mathbf{w}) = \sum_{k=0}^{n} w_k\, B_{k,n}(u), \quad \text{with} \quad B_{k,n}(u) = \binom{n}{k} u^k (1-u)^{n-k}. $$

The neural network trail solution $\hat{\chi}(u, P_n(u)) = u(1-u)\ N(u, P_n(u)) + u\ C$ where $C = \frac{16\sqrt{\frac{10}{\pi}}}{729\sqrt[4]{3}} + 1$, and $N(u, P_n(u)) = \sum_{j=1}^{n} \alpha_j P_{j-1}(u)$.

Table 1 shows the values of MSE and MAE for ChebyshevNN and BernsteinNN, which are computed on different degree of the polynomials, with 1000 epochs, learning rate = 0.0001, and $\lambda = 10^{-4}$. It clearly shows that BernsteinNN gives smaller error than ChebyshevNN.

Table 2 demonstrates the comparison between ChebyshevNN and BernsteinNN approximate solutions for Example 6.1 at $\theta = 1.5$ and $\mu = 0.9$, 1000 epochs and learning rate = 0.0001, degree = 6. Fig. 1 represents the approximate solutions and the error functions for both methods. The results show the better approximation for BernsteinNN than ChebyshevNN.

Example 6.2. Consider the following Caputo-conformable FVFIDEs:

$$ {}^{CC}\mathfrak{D}_0^{0.5,1.8}\chi(u) = \frac{-u}{3} - \frac{\Gamma(3)\ u^{\frac{7}{4}}}{(\frac{1}{2})^{\frac{7}{2}}\,\Gamma(\frac{9}{2})} + \int_0^1 u\ \big(\chi(v) - 1\big)\mathrm{d}v + \int_0^u \frac{\sqrt{\frac{2}{v}}}{\Gamma(\frac{3}{2})}\Big(\sqrt{u} - \sqrt{v}\Big)^{\frac{1}{2}} \chi^2(v)\mathrm{d}v, $$

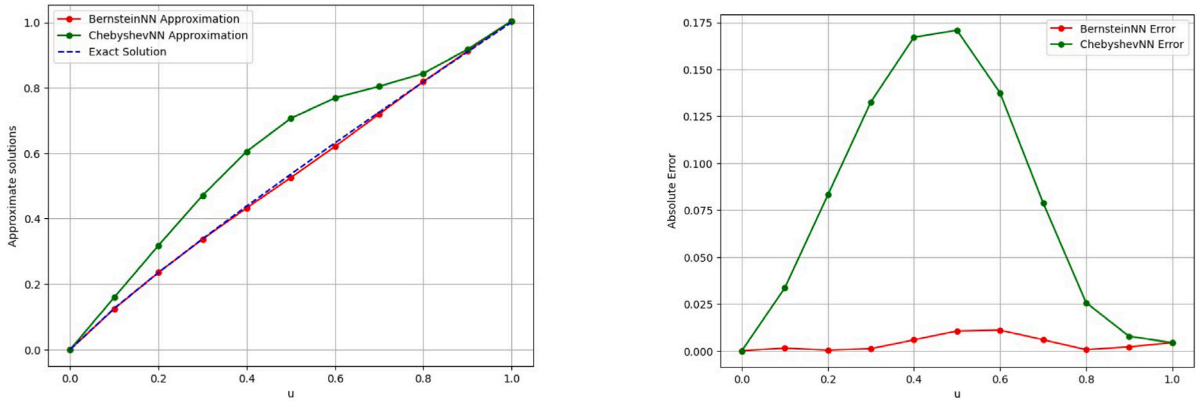
$$ \chi(0) = 0,\ \ \chi(1) + \varsigma_2\, {}^{\mathcal{RC}}\mathfrak{I}_0^{0.5,1.5}\chi(0.5) = 2,\ \ u \in [0,1]. \tag{68} $$

The exact solution is $\chi(u) = 2\sqrt{u}$.

The neural network trail solution $\hat{\chi}(u, P_n(u)) = u(1-u)\ N(u, P_n(u)) + 2\ u$, where $N(u, P_n(u)) = \sum_{j=1}^{n} \alpha_j P_{j-1}(u)$.

Table 3 shows the values of MSE and MAE for ChebyshevNN and BernsteinNN which are computed on different degree of the polynomials, with 1000 epochs, learning rate = 0.0001, and $\lambda = 10^{-4}$. For $n = 8$, the execution time for ChebyshevNN is 70.45 s and for BernsteinNN is 66.32 s. Note that BernsteinNN gives smaller error than ChebyshevNN with less computation times.

Table 4 demonstrates the comparison between ChebyshevNN and BernsteinNN approximate solutions for Example 6.2 at $\theta = 1.8$, $\mu = 0.5$, 1000 epochs and learning rate = 0.0001, degree of polynomial is 4. Fig. 2 represents the approximate solutions and the error functions for both methods. The results show the better approximation for BernsteinNN than ChebyshevNN.



(a) Approximate solutions via exact solution (b) Errors

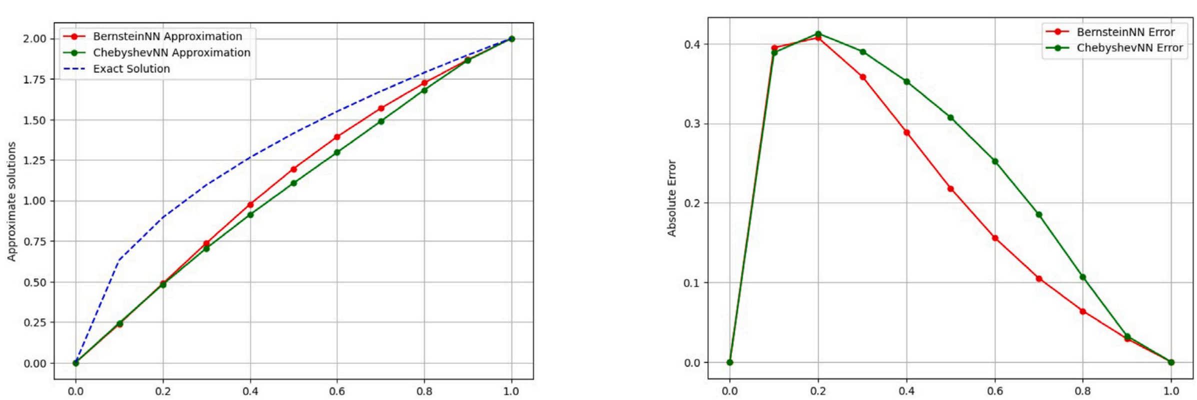
Fig. 1. A comparison between the approximate solutions using ChebyshevNN and BernsteinNN for Example 6.1.

Table 3
The MSE and MAE for ChebyshevNN and BernsteinNN computed on different degree of the polynomials, Example 6.2.

PNN	MSE/MAE	$n = 4$	$n = 6$	$n = 8$
BernsteinNN	MSE	5.6×10^{-2}	9.9×10^{-2}	7.8×10^{-2}
	MAE	1.8×10^{-1}	2.6×10^{-1}	2.3×10^{-1}
ChebyshevNN	MSE	7.3×10^{-2}	1.2×10^{-1}	7.4×10^{-2}
	MAE	2.2×10^{-1}	3.1×10^{-1}	2.3×10^{-1}

Table 4
The comparison between ChebyshevNN and BernsteinNN approximate solutions for Example 6.2.

u	Exact	Approximate solution		Error	
		BernsteinNN	ChebyshevNN	BernsteinNN	ChebyshevNN
0.1	0.63246	0.23761	0.24371	0.39484	0.38874
0.2	0.89443	0.48700	0.48158	0.40743	0.41284
0.3	1.09545	0.73679	0.70505	0.35865	0.39040
0.4	1.26491	0.97580	0.91212	0.28911	0.35279
0.5	1.41421	1.19581	1.10674	0.21840	0.30747
0.6	1.54919	1.39298	1.29639	0.15621	0.25281
0.7	1.67332	1.56801	1.48796	0.10531	0.18536
0.8	1.78885	1.72489	1.68196	0.06397	0.10690
0.9	1.89737	1.86835	1.86489	0.02902	0.03247
1.00	2.00000	2.00000	2.00000	0.00000	0.00000

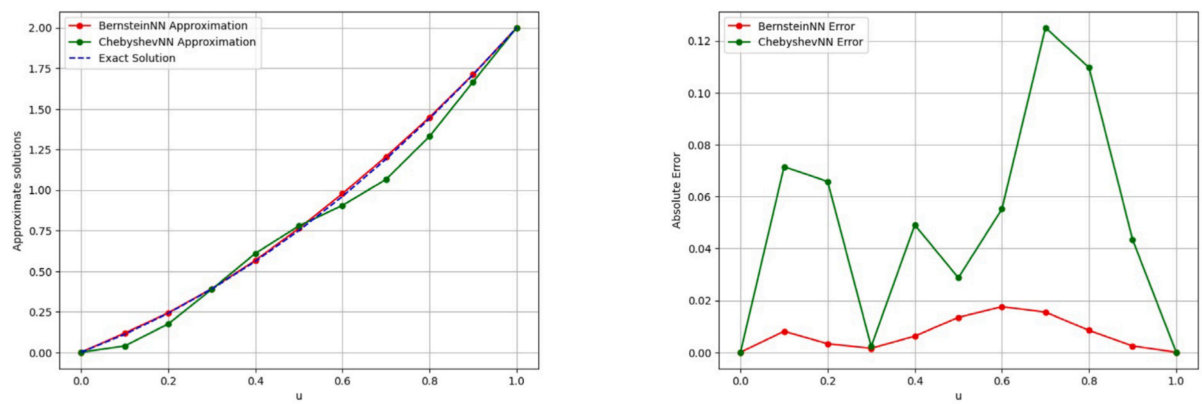


(a) Approximate solutions via exact solution (b) Errors

Fig. 2. A comparison between ChebyshevNN and BernsteinNN for Example 6.2.

Table 5
The comparison between ChebyshevNN and BernsteinNN approximate solutions for Example 6.3 with polynomials degree $n = 8$.

u	Exact	Approximate solution		Error	
		BernsteinNN	ChebyshevNN	BernsteinNN	ChebyshevNN
0.1	0.11000	0.11808	0.03852	0.00808	0.07148
0.2	0.24000	0.24327	0.17420	0.00327	0.06580
0.3	0.39000	0.39151	0.38777	0.00151	0.00223
0.4	0.56000	0.56620	0.60907	0.00620	0.04907
0.5	0.75000	0.76344	0.77869	0.01344	0.02869
0.6	0.96000	0.97756	0.90482	0.01756	0.05518
0.7	1.19000	1.20547	1.06494	0.01547	0.12506
0.8	1.44000	1.44841	1.33030	0.00841	0.10970
0.9	1.71000	1.71241	1.66663	0.00241	0.04337
1.0	2.00000	2.00000	2.00000	0.00000	0.00000



(a) Approximate solutions via exact solution

(b) Errors

Fig. 3. A comparison between ChebyshevNN and BernsteinNN for Example 6.3 at $n = 8$.

Example 6.3. Consider the following Caputo-conformable FVFIDEs:

$$ {}^{CC}\mathfrak{D}_0^{0.1,\frac{\sqrt{7}}{2}}\chi(u) = 1 - e^2 - \log\left(1 + u + u^2\right) - \frac{4u^{2+\frac{\sqrt{7}}{2}}}{(\sqrt{7}-4)\Gamma\left(2 - \frac{\sqrt{7}}{2}\right)} + \int_0^1 (1+2v)e^{\chi(v)}dv + \int_0^u \frac{(1+2v)}{1+\chi(v)}dv, $$

with boundary conditions

$$ \chi(0) = 1, \quad \chi(1) = e, \tag{69} $$

and the exact solution is $\chi(u) = u^2 + u$.

The neural network trail solution $\hat{\chi}(u, P_n(u)) = -\left(1 - \frac{u}{\pi}\right) + \frac{u}{\pi} + u(\pi - u)\ N(u, P_n(u))$, with $N(u, P_n(u)) = \sum_{j=1}^{n} \alpha_j P_{j-1}(u)$.

Table 5 demonstrates the comparison between ChebyshevNN and BernsteinNN approximate solutions for Example 6.3 at $\theta = \frac{\sqrt{7}}{2}$ and $\mu = 0.1$, 1000 epochs and learning rate = 0.0001, degree of polynomial is $n = 8$. Fig. 3 represents the approximate solutions and the error functions for both methods computed at polynomial degree $n = 8$, with execution time for ChebyshevNN is 79.80 s and for BernsteinNN is 70.29 s. The results show the better approximation for BernsteinNN than ChebyshevNN.

7. Conclusion

In this article, the new Caputo-conformable FVFIDEs with three-point non-local Riemann–Liouville conformable integral boundary conditions was considered. The existence and uniqueness of solutions were investigated by employing the Banach contraction principle, Scheafer's fixed point theorem with the help of the newly constructed generalized Gronwall inequality. A new approach was developed using ChebyshevNN and BernsteinNN with extreme learning machine algorithm and without activation functions. The error function was developed with new terms $L2$ regularization and the hyper-parameter λ related to the boundary conditions, to improve the stability and control the regularization strength. Moreover, Adam optimization method is employed to minimize the error, and the MSE and MAE were computed for the comparison purpose between the proposed polynomials. Numerical results showed the efficiency of the suggested method to minimize the error in solving this type of equation. In future works, the developed algorithm could be extended to solve the multi-dimensional and higher order integro-differential equations and partial differential equations. Also, it could deal with the singularity and permutation problems.

CRediT authorship contribution statement

Kawthar Alsa'di: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Nik Mohd Asri B. Nik Long:** Writing – review & editing, Writing – original draft, Validation, Supervision, Investigation, Funding acquisition, Conceptualization. **Norazak Senu:** Writing – review & editing, Conceptualization. **Z.K. Eshkuvatov:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Funding: This work was supported by the Universiti Putra Malaysia through Putra Grant [GP/2023/9752700].

Appendix

To evaluate the value of the following integral in terms of Beta function:

$$I_1 = \int_{c_0}^{u} \left(\frac{\left(u-c_0\right)^{\mu} - \left(v-c_0\right)^{\mu}}{\mu} \right)^{q(\theta-1)} \frac{1}{\left(v-c_0\right)^{q(1-\mu)}} dv.$$

Substituting $t = \frac{v-c_0}{u-c_0}$, and change of variable:

$$t = \frac{v-c_0}{u-c_0}, \quad \text{which gives} \quad v = c_0 + t\left(u-c_0\right), \quad \text{and} \quad dv = \left(u-c_0\right) dt.$$

Rewriting terms in the integral: $\left(v-c_0\right)^{q(1-\mu)} = \left(t\left(u-c_0\right)\right)^{q(1-\mu)} = t^{q(1-\mu)}\left(u-c_0\right)^{q(1-\mu)}$. Also, using the substitution for the first term: $\left(u-c_0\right)^{\mu} - \left(v-c_0\right)^{\mu} = \left(u-c_0\right)^{\mu}(1-t^{\mu})$. Thus, the integral transforms into:

$$I_1 = \int_0^1 \left(\frac{\left(u-c_0\right)^{\mu}(1-t^{\mu})}{\mu} \right)^{q(\theta-1)} \frac{\left(u-c_0\right) dt}{t^{q(1-\mu)}\left(u-c_0\right)^{q(1-\mu)}}$$

Next, expanding the power terms of $u-c_0$ as:

$$\left(\frac{\left(u-c_0\right)^{\mu}(1-t^{\mu})}{\mu} \right)^{q(\theta-1)} = \left(u-c_0\right)^{\mu q(\theta-1)} \frac{(1-t^{\mu})^{q(\theta-1)}}{\mu^{q(\theta-1)}}$$

Then, the integral can be expressed as:

$$I_1 = \left(u-c_0\right)^{1+\mu q(\theta-1)-q(1-\mu)} \frac{1}{\mu^{q(\theta-1)}} \int_0^1 (1-t^{\mu})^{q(\theta-1)} t^{-q(1-\mu)} dt$$

By changing the variable $x = t^{\mu}$ in the integral term, such that $dx = \mu t^{\mu-1} dt$, $t = x^{1/\mu}$, and $dt = \frac{dx}{\mu} x^{(1/\mu)-1}$, we obtain

$$\int_0^1 (1-x)^{q(\theta-1)} x^{\frac{-q(1-\mu)}{\mu}} \frac{dx}{\mu} x^{(1/\mu)-1}$$

Simplifying the exponent of x :

$$\int_0^1 (1-x)^{q(\theta-1)} x^{\frac{-q(1-\mu)}{\mu} + \frac{1}{\mu} - 1} \frac{dx}{\mu}.$$

The integral term can be expressed in terms of Beta function as

$$\mathbb{B}(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx,$$

where

$$p = q(\theta-1)+1, \quad q = \frac{1-q(1-\mu)}{\mu}.$$

Thus,

$$I_1 = \left(u-c_0\right)^{1+\mu q(\theta-1)-q(1-\mu)} \frac{\mathbb{B}\left(q(\theta-1)+1, \frac{1-q(1-\mu)}{\mu}\right)}{\mu^{q(\theta-1)+1}}.$$

Simplify,

$$I_1 = \frac{(u-c_0)^{\mu\theta q-q+1}}{\mu^{q(\theta-1)+1}} \mathbb{B}\Big(q\theta - q + 1, \frac{1}{\mu}(1-q) + q\Big).$$

Similarly, this technique can be used to evaluate the following integral terms for the coefficients a_3 and a_4

$$I_2 = \int_{c_0}^{C} \left(\frac{(C-c_0)^{\mu} - (v-c_0)^{\mu}}{\mu} \right)^{q(\theta-1)} \frac{1}{(v-c_0)^{q(1-\mu)}} \mathrm{d}v,$$

after simplify, we have

$$I_2 = \frac{(C - c_0)^{\mu\theta q - q + 1}}{\mu^{q(\theta-1)+1}} \mathbb{B}\Big(q\theta - q + 1, \frac{1}{\mu}(1-q) + q\Big),$$

and for the integral

$$I_3 = \int_{c_0}^{\xi} \left(\frac{(\xi - c_0)^\mu - (\upsilon - c_0)^\mu}{\mu}\right)^{q(\theta+\delta-1)} \frac{1}{(\upsilon - c_0)^{q(1-\mu)}} \mathrm{d}\upsilon,$$

after simplify, we have

$$I_3 = \frac{(\xi - c_0)^{q\mu(\theta+\delta) - q + 1}}{\mu^{q(\theta+\delta-1)+1}} \mathbb{B}\Big(q(\theta+\delta) - q + 1, \frac{1}{\mu}(1-q) + q\Big).$$

References

[1] A.M. Yang, Y. Han, Y.Z. Zhang, L.T. Wang, D. Zhang, X.J. Yang, On local fractional Volterra integro–differential equations in fractal steady heat transfer, Therm. Sci. 20 (suppl. 3) (2016) 789–793.
[2] Ş. Yüzbaşı, A numerical approximation for Volterra's population growth model with fractional order, Appl. Math. Model. 37 (5) (2013) 3216–3227.
[3] S. Santra, J. Mohapatra, P. Das, D. Choudhuri, Higher order approximations for fractional order integro–parabolic partial differential equations on an adaptive mesh with error analysis, Comput. Math. Appl. 150 (2023) 87–101.
[4] P. Das, S. Rana, H. Ramos, A perturbation-based approach for solving fractional-order Volterra–Fredholm integro–differential equations and its convergence analysis, Int. J. Comput. Math. 97 (10) (2020) 1994–2014.
[5] P. Das, S. Rana, H. Ramos, On the approximate solutions of a class of fractional order nonlinear Volterra integro–differential initial value problems and boundary value problems of first kind and their convergence analysis, J. Comput. Appl. Math. 404 (2022) 113116.
[6] P. Das, S. Rana, H. Ramos, Homotopy perturbation method for solving Caputo-type fractional-order Volterra-Fredholm integro–differential equations, Comput. Math. Methods 1 (5) (2019) e1047.
[7] P. Das, S. Rana, Theoretical prospects of fractional order weakly singular Volterra integro–differential equations and their approximations with convergence analysis, Math. Methods Appl. Sci. 44 (11) (2021) 9419–9440.
[8] V. Saw, P. Das, H.M. Srivastava, Investigations of a class of Liouville-Caputo fractional order Pennes bioheat flow partial differential equations through orthogonal polynomials on collocation points, Bull. Sci. Math. 201 (2025) 103637.
[9] S.K. Ntouyas, J. Tariboon, Fractional integral problems for Hadamard–Caputo fractional Langevin differential inclusions, J. Appl. Math. Comput. 51 (2016) 13–33.
[10] J. Tariboon, S.K. Ntouyas, W. Sudsutad, Nonlocal Hadamard fractional integral conditions for nonlinear Riemann-Liouville fractional differential equations, Bound. Value Probl. 2014 (2014) 1–16.
[11] K. Karthikeyan, A. Debbouche, D.F. Torres, Analysis of Hilfer fractional integro-differential equations with almost sectorial operators, Fractal Fract. 5 (1) (2021) 22.
[12] B. Lupinska, Properties of the Katugampola fractional operators, Tatra Mt. Math. Publ. 79 (2021) 135–148.
[13] A.B. Makhlouf, J. Sallay, Some results on proportional Caputo neutral fractional stochastic differential equations, Discret. Contin. Dyn. Syst.-S 17 (10) (2024) 3102–3115.
[14] A.B. Makhlouf, L. Mchiri, H.M. Srivastava, Some existence and uniqueness results for a class of proportional Liouville-Caputo fractional stochastic differential equations, Bull. Sci. Math. 189 (2023) 103349.
[15] R. Khalil, M. Al Horani, A. Yousef, M. Sababheh, A new definition of fractional derivative, J. Comput. Appl. Math. 264 (2014) 65–70.
[16] T. Abdeljawad, On conformable fractional calculus, J. Comput. Appl. Math. 279 (2015) 57–66.
[17] O. Naifar, A. Jmal, A. Ben Makhlouf, Non-fragile H_∞ observer for Lipschitz conformable fractional–order systems, Asian J. Control 24 (5) (2022) 2202–2212.
[18] E. Piotrowska, Ł. Sajewski, Analysis of an electrical circuit using two-parameter conformable operator in the Caputo sense, Symmetry 13 (5) (2021) 771.
[19] A. Aphithana, S.K. Ntouyas, J. Tariboon, Existence and Ulam–Hyers stability for Caputo conformable differential equations with four-point integral conditions, Adv. Difference Equ. 2019 (2019) 1–17.
[20] D. Baleanu, S. Etemad, S. Rezapour, On a Caputo conformable inclusion problem with mixed Riemann–Liouville conformable integro-derivative conditions, Adv. Difference Equ. 2020 (1) (2020) 473.
[21] B. Samet, C. Vetro, P. Vetro, Fixed point theorems for α–ψ-contractive type mappings, Nonlinear Anal. 75 (4) (2012) 2154–2165.
[22] S.T. Thabet, S. Etemad, S. Rezapour, On a new structure of the pantograph inclusion problem in the Caputo conformable setting, Bound. Value Probl. 2020 (1) (2020) 171.
[23] L. Rabhi, M. Al Horani, R. Khalil, Existence results of mild solutions for nonlocal fractional delay integro-differential evolution equations via Caputo conformable fractional derivative, AIMS Math. 7 (7) (2022) 11614–11634.
[24] A. Amara, S. Etemad, S. Rezapour, Approximate solutions for a fractional hybrid initial value problem via the Caputo conformable derivative, Adv. Difference Equ. 2020 (1) (2020) 608.
[25] Z. Peng, Y. Li, Q. Zhang, Y. Xue, Extremal solutions for Caputo conformable differential equations with p-Laplacian operator and integral boundary condition, Complexity 2021 (1) (2021) 1097505.
[26] S.S. Tantawy, Solving linear systems of fractional integro–differential equations by Haar and Legendre wavelets techniques, Part. Differ. Equ. Appl. Math. 10 (2024) 100683.
[27] M. Mandal, et al., Convergence analysis and numerical implementation of projection methods for solving classical and fractional Volterra integro-differential equations, Math. Comput. Simulation 225 (2024) 889–913.
[28] K. Sadri, D. Amilo, E. Hinçal, K. Hosseini, S. Salahshour, A generalized Chebyshev operational method for Volterra integro-partial differential equations with weakly singular kernels, Heliyon 10 (5) (2024).
[29] H. Li, P. Shi, X. Li, Machine learning for nonlinear integro-differential equations with degenerate kernel scheme, Commun. Nonlinear Sci. Numer. Simul. 138 (2024) 108242.
[30] L. Yuan, Y.-Q. Ni, X.-Y. Deng, S. Hao, A-PINN: Auxiliary physics informed neural networks for forward and inverse problems of nonlinear integro–differential equations, J. Comput. Phys. 462 (2022) 111260.
[31] T. Allahviranloo, A. Jafarian, R. Saneifard, N. Ghalami, S. Measoomy Nia, F. Kiani, U. Fernandez-Gamiz, S. Noeiaghdam, An application of artificial neural networks for solving fractional higher-order linear integro-differential equations, Bound. Value Probl. 2023 (1) (2023) 74.
[32] R. Saneifard, A. Jafarian, N. Ghalami, S.M. Nia, Extended artificial neural networks approach for solving two-dimensional fractional-order Volterra-type integro-differential equations, Inf. Sci. 612 (2022) 887–897.
[33] Z. Hajimohammadi, F. Baharifard, A. Ghodsi, K. Parand, Fractional Chebyshev deep neural network (FCDNN) for solving differential models, Chaos, Solitons Fractals 153 (2021) 111530.
[34] X. Shao, L. Yang, A. Guo, A feedforward neural network based on Legendre polynomial for solving linear Fredholm integro–differential equations, Int. J. Comput. Math. 100 (7) (2023) 1480–1499.
[35] H. Liu, H. Liu, J. Xu, L. Li, J. Song, Jacobi neural network method for solving linear differential-algebraic equations with variable coefficients, Neural Process. Lett. 53 (5) (2021) 3357–3374.
[36] H. Sun, M. Hou, Y. Yang, T. Zhang, F. Weng, F. Han, Solving partial differential equation based on Bernstein neural network and extreme learning machine algorithm, Neural Process. Lett. 50 (2019) 1153–1172.
[37] S. Chakraverty, S. Mall, Single layer Chebyshev neural network model with regression-based weights for solving nonlinear ordinary differential equations, Evol. Intell. 13 (4) (2020) 687–694.
[38] K. Shah, B. Abdalla, T. Abdeljawad, et al., An efficient matrix method for coupled systems of variable fractional order differential equations, Therm. Sci. 27 (2023) S195–S210.

[39] F.A. Shah, K. Shah, T. Abdeljawad, et al., Numerical modelling of advection diffusion equation using Chebyshev spectral collocation method and Laplace transform, Results Appl. Math. 21 (2024) 100420.
[40] I. Podlubny, Fractional Differential Equations: an Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications, Elsevier, 1998.
[41] S.G. Samko, Fractional integrals and derivatives, Theory Appl. (1993).
[42] F. Jarad, E. Uğurlu, T. Abdeljawad, D. Baleanu, On a new class of fractional operators, Adv. Difference Equ. 2017 (2017) 1–16.
[43] Y. Zhou, Basic Theory of Fractional Differential Equations, World scientific, 2023.
[44] A. Rababah, Transformation of Chebyshev–Bernstein polynomial basis, Comput. Methods Appl. Math. 3 (4) (2003) 608–622.
[45] J.R. Wang, X. Xiang, W. Wei, Q. Chen, The generalized Grönwall inequality and its application to periodic solutions of integro–differential impulsive periodic system on Banach space, J. Inequal. Appl. 2008 (2008) 1–22.
[46] W. Wei, X. Xiang, Y. Peng, Nonlinear impulsive integro–differential equations of mixed type and optimal controls, Optimization 55 (1–2) (2006) 141–156.
[47] C. Corduneanu, Principles of Differential and Integral Equations, Vol. 295, American Mathematical Soc., 1977.
[48] Z. Gao, L. Yang, Z. Luo, Stability of the solutions for nonlinear fractional differential equations with delays and integral boundary conditions, Adv. Difference Equ. 2013 (1) (2013) 43.
[49] A. Jadbabaie, A. Makur, D. Shah, Federated optimization of smooth loss functions, IEEE Trans. Inform. Theory 69 (12) (2023) 7836–7866.
[50] B. Qi, W. Gong, L. Li, Extended convexity and smoothness and their applications in deep learning, 2024, arXiv preprint arXiv:2410.05807.
[51] M. Akhtar, M. Tanveer, M. Arshad, RoBoSS: A robust, bounded, sparse, and smooth loss function for supervised learning, IEEE Trans. Pattern Anal. Mach. Intell. (2024).
[52] S. Saini, P. Das, S. Kumar, Parameter uniform higher order numerical treatment for singularly perturbed Robin type parabolic reaction diffusion multiple scale problems with large delay in time, Appl. Numer. Math. 196 (2024) 1–21.
[53] S. Saini, P. Das, S. Kumar, Computational cost reduction for coupled system of multiple scale reaction diffusion problems with mixed type boundary conditions having boundary layers, Rev. Real Acad. Cienc. Exactas Físicas Y Nat. Ser. A. Mat. 117 (2) (2023) 66.
[54] S. Kumar, P. Das, A uniformly convergent analysis for multiple scale parabolic singularly perturbed convection-diffusion coupled systems: Optimal accuracy with less computational time, Appl. Numer. Math. 207 (2025) 534–557.
[55] S. Haykin, Neural Networks: a Comprehensive Foundation, Prentice Hall PTR, 1994.
[56] A. Afzal Aghaei, K. Parand, Hyperparameter optimization of orthogonal functions in the numerical solution of differential equations, Math. Methods Appl. Sci. 48 (1) (2025) 605–625.