Ferdowsi University of Mashhad

Iranian Journal of Numerical Analysis and Optimization

Volume 14, Number 4

December 2024

Serial Number: 31

Ferdowsi University of Mashhad, Iran

In the Name of God

Iranian Journal of Numerical Analysis and Optimization (IJNAO)

This journal is authorized under the registration No. 174/853 dated 1386/2/26 (2007/05/16), by the Ministry of Culture and Islamic Guidance.

Volume 14, Number 4, December 2024

ISSN-Print: 2423-6977, **ISSN-Online:** 2423-6969

Publisher: Faculty of Mathematical Sciences, Ferdowsi University of Mashhad

Published by: Ferdowsi University of Mashhad Press

Printing Method: Electronic

Address: Iranian Journal of Numerical Analysis and Optimization

Faculty of Mathematical Sciences, Ferdowsi University of Mashhad

P.O. Box 1159, Mashhad 91775, Iran.

Tel. : +98-51-38806222 , **Fax:** +98-51-38807358

E-mail: ijnao@um.ac.ir

Website: http://ijnao.um.ac.ir

This journal is indexed by:

- SCOPUS
- ZbMATH Open
- ISC
- DOAJ
- SID
- Civilica
- Magiran
- Mendeley
- Academia.edu
- Linkedin

• The Journal granted the International degree by the Iranian Ministry of Science, Research, and Technology.

Iranian Journal of Numerical Analysis and Optimization

Volume 14, Number 4, December 2024

Ferdowsi University of Mashhad - Iran

Iranian Journal of Numerical Analysis and Optimization

This journal is published under the auspices of Ferdowsi University of Mashhad

* Full Professor

** Associate Professor

We would like to acknowledge the help of Miss Narjes khatoon Zohorian in the preparation of this issue.

Letter from the Editor–in–Chief

I would like to welcome you to the Iranian Journal of Numerical Analysis and Optimization (IJNAO). This journal has been published two issues per year and supported by the Faculty of Mathematical Sciences at the Ferdowsi University of Mashhad. The faculty of Mathematical Sciences with the centers of excellence and the research centers is well-known in mathematical communities in Iran.

The main aim of the journal is to facilitate discussions and collaborations between specialists in applied mathematics, especially in the fields of numerical analysis and optimization, in the region and worldwide. Our vision is that scholars from different applied mathematical research disciplines pool their insight, knowledge, and efforts by communicating via this international journal. In order to assure the high quality of the journal, each article is reviewed by subject-qualified referees. Our expectations for IJNAO are as high as any well-known applied mathematical journal in the world. We trust that by publishing quality research and creative work, the possibility of more collaborations between researchers would be provided. We invite all applied mathematicians especially in the fields of numerical analysis and optimization to join us by submitting their original work to the Iranian Journal of Numerical Analysis and Optimization.

We would like to inform all readers that the Iranian Journal of Numerical Analysis and Optimization (IJNAO), has changed its publishing frequency from ”Semiannual” to a ”Quarterly” journal since January 2023. The four journal issues per year will be published in the months of March, June, September, and December. One of our goals is to continue to improve the speed of both the review and publication processes, while try continuing to publish the best available international research in numerical analysis and optimization, with the high scientific and publication standards that the journal is known for.

Ali R. Soheili

Editor-in-Chief

Contents

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 991–1015
https://doi.org/10.22067/ijnao.2024.86954.1395
https://ijnao.um.ac.ir/

Research Article

A numerical solution of parabolic quasi-variational inequality nonlinear using Newton-multigrid method

M. Bahi*, M. Beggas, N. Nesba and A. Imtiaz

Abstract

In this article, we apply three numerical methods to study the $\mathbb{L}^\infty$-convergence of the Newton-multigrid method for parabolic quasi-variational inequalities with a nonlinear right-hand side. To discretize the problem, we

*Corresponding author

Received 20 February 2024; revised 04 May 2024; accepted 13 May 2024

Mostafa Bahi
Department of Mathematics, Faculty of exact Sciences, University of EL-OUED, Algeria.
e-mail: bahi-mostafa@univ-eloued.dz

Mohammed Beggas
Department of Mathematics, Faculty of exact Sciences, University of EL-OUED, Algeria.
e-mail: mohammed-beggas@univ-eloued.dz

Nourelhouda Nesba
Department of Mathematics, Faculty of exact Sciences, University of EL-OUED, Algeria.
e-mail: nesba-nourelhouda@univ-eloued.dz

Ahmad Imtiaz
Institute of Informatics and Computing in Energy (IICE), Universiti Tenaga Nasional, Kajang, Selangor, Malaysia. e-mail: imtiazkakakhil@gmail.com

How to cite this article
Bahi, M., Beggas, M., Nesba, N. and Imtiaz, A., A numerical solution of parabolic quasi-variational inequality nonlinear using Newton-multigrid method. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 991-1015. https://doi.org/10.22067/ijnao.2024.86954.1395

utilize a finite element method for the operator and Euler scheme for the time. To obtain the system discretization of the problem, we reformulate the parabolic quasi-variational inequality as a Hamilton–Jacobi–Bellman equation. For linearizing the problem on the coarse grid, we employ Newton's method as an external interior iteration of the Jacobian system. On the smooth grid, we apply the multigrid method as an interior iteration on the Jacobian system. Finally, we provide a proof for the $\mathbb{L}^\infty$-convergence of the Newton-multigrid method for parabolic quasi-variational inequalities with a nonlinear right-hand, by giving a numerical example for this problem.

AMS subject classifications (2020): Primary 65M55, 65M60; Secondary, 65F10.

Keywords: Newton's method; Multigrid method; Parabolic variational inequality; Finite element method; Hamilton–Jacobi–Bellman equation.

1 Introduction

Before studying the $\mathbb{L}^\infty$-convergence of the Newton-multigrid method for parabolic quasi-variational inequalities (PQVIs) with a nonlinear right-hand side, we will define our problem. Let $\mathbb{A}$ be an elliptic operator, and let Ω be an open domain in $\mathbb{R}^d$, where $d \geq 1$, with sufficiently regular bounds Γ. We provide the definition of our proposed problem:
Find $\mathtt{u} \in L^2(H_0^1(\Omega); [0,T])$ solution of

$$\begin{cases} \dfrac{\partial \mathtt{u}}{\partial t} + \mathbb{A}\mathtt{u} \leq f(\mathtt{u}); \qquad \mathtt{u} - \psi \leq 0, \\ \left(\dfrac{\partial \mathtt{u}}{\partial t} + \mathbb{A}\mathtt{u} - f(\mathtt{u})\right)(\mathtt{u} - \psi) = 0 \ \text{ in } \mathbf{Q}_t, \\ \mathtt{u}(x,t) = \mathtt{u}_0 \qquad in \quad \Omega_h, \\ \mathtt{u}(x,t) = 0 \qquad\quad in \quad \sum_h, \end{cases} \tag{1}$$

where $\mathbf{Q}_T$ is defined as $\mathbf{Q}_T = \Omega \times [0,T]$, $\sum = \Gamma \times [0,T]$, $T < \infty$, and the obstacle $\psi \in W^{2,\infty}$ such that $\psi \geq 0$.

We assumed that the function $f(\cdot)$ is nonlinear, continuous, and c-Lipschitz; that is,

$$|f(\mathtt{y})-f(\mathtt{z})|\leq c\,|\mathtt{y}-\mathtt{z}| \quad \text{for all } \mathtt{y},\mathtt{z}\in\mathbb{R}, \quad c\ >0.$$

We have divided this research into six parts as follows:

Part one: We formulate our problem and provide the necessary and sufficient conditions for studying the $\mathbb{L}^\infty$-convergence of Newton-multigrid methods.

Part two: To facilitate understanding, we introduce important symbols and hypotheses that will be used in our research. Subsequently, we present a theorem that proves the existence and uniqueness of the solution for the continuous problem.

Part three: To study the discrete problem, we define the space $\mathbb{V}_h$ that

$$\mathbb{V}_h=\{\mathtt{v}_h\in L^2(H_0^1(\Omega);[0,T])\,\cap\, C^1(H_0^1(\overline{\Omega});[0,T]),\ \mathtt{v}_h(\cdot,t)\in P_1\}$$

with P_1 is a Lagrange polynomial with a degree ≤ 1.

To discretize, we employ the finite element approximation on the operator $\mathbb{A}$ and use the Euler scheme for time discretization, resulting in the following discrete problem. Find $\mathtt{u}_h^n\in\mathbb{V}_h$ as a solution of

$$\begin{cases}\left\langle \mathbb{B}_h\mathtt{u}_h^n,\mathtt{v}_h-\mathtt{u}_h^n\right\rangle \geq \left\langle F^n(\mathtt{u}_h^n),\mathtt{v}_h-\mathtt{u}_h^n\right\rangle, \\ \\ \mathtt{u}_h^n\leq r_h\psi, \\ \\ \mathtt{u}_h(x)=\mathtt{u}_h(x,0) \quad in \quad \Omega_h, \\ \\ \mathtt{u}(x,t)=0 \quad in \quad \Sigma_h.\end{cases} \tag{2}$$

Part four: To solve the PQVI using Newton-multigrid methods, we discretize the domain through the finite element method and employ the Euler scheme for time discretization. Subsequently, we reformulate the PQVI as the Hamilton–Jacobi–Bellman (HJB) equation, expressed as follows:

$$\max_{1\leq i\leq N}\left(\mathbb{B}_{k,i}\mathtt{u}_{k,i}^v-F(\mathtt{u}_{k,i}^v);\ \mathtt{u}_{k,i}^v-r_h\psi\right)=0.$$

We employ the iterative procedure introduced by Hoppe (see [22, 23, 24]) to solve the obtained system. Subsequently, we apply Newton's method to address the nonlinear system by linearizing it to obtain the Jacobian system. Following this, we employ the multigrid method on the Jacobian system in all iterations.

Part five: To establish the $\mathbb{L}^\infty$-convergence of the Newton-multigrid method, we rely on the approximation and smoothing suggestions proposed by Huckbach (see[19, 20]). In this research, we demonstrate the $\mathbb{L}^\infty$-convergence in the addressed problem based on the techniques presented in Huckbach's work.

Part six: Here, we give a numerical example that confirms the results obtained in the theoretical part.

2 Newton-multigrid method

2.1 Assumptions and symbols

Consider $\mathbb{A}$ as an elliptic operator defined by

$$\mathbb{A} = -\sum_{1\le i,j\le d}^{d} \mathtt{a}_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^{d} \mathtt{a}_j(x)\frac{\partial}{\partial x_j} + \mathtt{a}_0(x). \tag{3}$$

The coefficients $\mathtt{a}_{ij}(x)$, $\mathtt{a}_j(x)$, $\mathtt{a}_0(x) \in L^\infty(\Omega) \cap C^2(\bar{\Omega})$ meet the following conditions:

$$\mathtt{a}_{ij}(x) = \mathtt{a}_{ji}(x), \quad \mathtt{a}_0(x) \ge \alpha > 0, \quad \alpha \textit{ is a constant}, \quad x \in \bar{\Omega}, \tag{4}$$

$$\sum_{1\le i,j\le d}^{d} \mathtt{a}_{ij}(x)\zeta_i\zeta_j \ge \gamma|\zeta|^2; \quad \zeta \in \mathbb{R}^d \quad \gamma > 0,$$

and $\mathtt{a}(\cdot,\cdot)$ is associated with the elliptic operator $\mathbb{A}$, where $\mathtt{a}(\cdot,\cdot)$ is a continuous and coercive bilinear form, as given by

$$\mathtt{a}(\mathtt{u},\mathtt{v}) = \int_\Omega \left[\sum_{1\le i,j\le d}^{d} \mathtt{a}_{ij}(x)\frac{\partial \mathtt{u}}{\partial x_i}\frac{\partial \mathtt{v}}{d x_j} + \sum_{j=1}^{d} \mathtt{a}_j(x)\frac{\partial \mathtt{u}}{\partial x_j} + \mathtt{a}_0(x)\mathtt{u}\mathtt{v} \right] dx, \tag{5}$$

$$\mathtt{a}(\mathtt{v},\mathtt{v})+\omega\|\mathtt{v}\|^2_{L^2(\Omega)} \geq \gamma\|\mathtt{v}\|^2_{H^1_0(\Omega)}, \quad \gamma>0, \qquad \omega>0, \quad \text{for all } \mathtt{v}\in H^1_0(\Omega). \tag{6}$$

Let $f(\mathtt{u})$ be a nonlinear right-hand function and f is c-Lipschitzian given by

$$f(\mathtt{u}) \in L^\infty(H^1_0(\Omega);[0,T]\cap C^1(H^1_0(\Omega);[0,T])), \quad f>0, \quad and \quad c<\alpha.$$

2.2 Continuous problem

After applying Green's formulation, we obtain the weak formulation of our proposed problem, $\mathtt{v}$ in $H^1_0(\Omega)$ and $\mathtt{v}\leqslant\psi$ and find $\mathtt{u}$ in $L^2(H^1_0(\Omega);[0,T])$ a solution to the following problem:

$$\begin{cases} \left(\dfrac{\partial \mathtt{u}}{\partial t},\mathtt{v}-\mathtt{u}\right)+\mathtt{a}(\mathtt{u},\mathtt{v}-\mathtt{u}) \geq (f(\mathtt{u}),\mathtt{v}-\mathtt{u}), \\ \mathtt{u}\leq\psi, \\ \mathtt{u}(x,t)=\mathtt{u}_0 \qquad in \quad \Omega, \\ \mathtt{u}(x,t)=0 \quad in \quad \Sigma. \end{cases} \tag{7}$$

Certainly, the problem (7) has a unique solution, as supported by fixed-point theory and in accordance with the previously established assumptions; see [7, 10, 12].

3 Discretization

3.1 Discretize

To establish a multigrid loop, consider a decreasing sequence $\{h_k\}_{k=0}^{l}$ such that

$$0<h_k<h_{k+1}<1, \quad 0\leq k\leq m-1,$$

that $\{T_h\}$ is a family of uniform triangulations, where $\Omega_h = \underset{T\in T_h}{\cup} T_h$, and that for all T_h, we have

$$\Omega_h\subset\Omega_{h+1}\subset\Omega, \quad dist(\Gamma_h,\Gamma)\leq ch_k^2, \quad h_{k+1}h_k\leq c_1.$$

We associate each h_k with the following symbols to facilitate our work:

$$\Omega_{h_k} = \Omega_h, \qquad \mathbb{V}_{h_k} = \mathbb{V}_h, \qquad \mathbb{A}_{h_k} = \mathbb{A}_h.$$

We consider ϕ_h^i, $i = 1, 2, \ldots, m(h)$ as the conventional basis for affine functions defined as follows:

$$\phi_h^i(K_h^j) = \delta_{ij},$$

where K_h^j denotes a vertex of the triangulation T_h. Let $\mathbb{U}_h = R^{mh}$. We consider the restriction operator defined as

$$r_h : \mathbb{U}_h \longrightarrow \mathbb{V}_h,$$

$$r_h \mathtt{v} = \sum_{i=1}^{m(h)} \mathtt{v}(K_h^i)\phi_h^i(x). \tag{8}$$

We provide a definition for the following adjoint operator $r_h^* : \mathbb{V}_h \longrightarrow \mathbb{U}_h$ satisfying

$$\langle r_h \mathtt{u}, \mathtt{u} \rangle_{L^2} = \langle \mathtt{u}, r_h^* \mathtt{v} \rangle, \qquad \text{for all } \mathtt{u} \in \mathbb{U}_h, \mathtt{v} \in \mathbb{V}_h.$$

The norm $\|\cdot\|_\infty$ (on $\mathbb{U}_k$) and the norm $\|\cdot\|_{L^\infty}$ (on $\mathbb{V}_k$) are equivalent, which are indicated by $\|\cdot\|_\infty$.

Lemma 1. [3] There exist λ_1 and λ_2 independent of h such that

1. $\|r_h(\mathtt{u})\|_\infty = \|\mathtt{u}\|_\infty$, for all $\mathtt{u} \in \mathbb{U}_h$,

2. $\lambda_1 \|\mathtt{v}\|_\infty \leq \|r^*(\mathtt{v})\|_\infty \leq \lambda_2 \|\mathtt{v}\|_\infty$, for all $\mathtt{v} \in \mathbb{V}_h$.

3.2 Discrete problem

We apply the semi-implicit Euler scheme to discretize problem (7) with respect to time. Consequently, we are in seek of a sequence of elements $\mathtt{u}_h^n \in \mathbb{V}_h$ that approaches $\mathtt{u}(t_n)$, $t_n = n\delta t$, with initial data $\mathtt{u}_0$ and $\dfrac{\partial \mathtt{u}_h}{\partial t} = \frac{\mathtt{u}_h^n - \mathtt{u}_h^{n-1}}{\delta t}$. Thus, we have for $n = 1, \ldots, N$, and for all $\mathtt{v}_h \in \mathbb{V}_h$,

$$\begin{cases} (\dfrac{\mathtt{u}_h^n - \mathtt{u}_h^{n-1}}{\delta t}, \mathtt{v}_h - \mathtt{u}_h^n) + \mathtt{a}(\mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n) \geq (f(\mathtt{u}_h^n), \mathtt{v}_h - \mathtt{u}_h^n), \\ \mathtt{u}_h^n \leq r_h\psi; \quad \mathtt{v}_h^n \leq r_h\psi. \end{cases} \tag{9}$$

We can write the problem (9) as for all $\mathtt{v}_h \in \mathbb{V}_h$,

$$\begin{cases} \mathtt{a}(\mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n) + \dfrac{1}{\delta t}(\mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n) \geq (f(\mathtt{u}_h^n) + \dfrac{\mathtt{u}_h^{n-1}}{\delta t}, \mathtt{v}_h - \mathtt{u}_h^n), \\ \mathtt{u}_h^n \leq r_h\psi; \quad \mathtt{v}_h^n \leq r_h\psi. \end{cases} \tag{10}$$

When we put $\omega = \dfrac{1}{\delta t}$ and $F(u_h^n) = f(\mathtt{u}_h^n) + \omega \mathtt{u}_h^{n-1}$, we obtain

$$\begin{cases} \omega(\mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n) + \mathtt{a}(\mathtt{u}_h^n, \mathtt{v}_h - u_h^n) \geq (F_h^n(\mathtt{u}_h^n), \mathtt{v}_h - \mathtt{u}_h^n), \\ \mathtt{u}_h^n \leq r_h\psi. \end{cases} \tag{11}$$

We can express the problem (11) in the following form:
Find $\mathtt{u}_h^n \in \mathbb{V}_h$, for all $\mathtt{v}_h \in \mathbb{V}_h$

$$\begin{cases} \mathtt{b}(\mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n) \geq (F_h^n(\mathtt{u}_h^n), \mathtt{v}_h - \mathtt{u}_h^n), \\ \mathtt{u}_h^n \leq r_h\psi, \\ \mathtt{u}(x,t) = \mathtt{u}_0 \qquad in \quad \Omega_h, \\ \mathtt{u}(x,t) = 0 \quad in \quad \Sigma_h, \end{cases} \tag{12}$$

where

$$\mathtt{b}(\mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n) = \omega(\mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n) + \mathtt{a}(\mathtt{u}_h^n, \mathtt{u}_h - \mathtt{u}_h^n).$$

The bilinear form $\mathtt{b}(\cdot,\cdot)$ exhibits strong coerciveness. Consequently, the formulation (12) denotes a coercive and continuous issue concerning elliptic quasi-variational inequalities (see[8]).

Denote by $\mathbb{B}_h$ the outcome obtained by addressing problem (12) through a finite element method, resulting in the solution to the subsequent problem ascertain $\mathtt{u}_h^n \in \mathbb{V}_h$ such that

$$\begin{cases} \langle \mathbb{B}_h \mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n \rangle \ \geq \ \langle F^n(\mathtt{u}_h^n), \mathtt{v}_h - \mathtt{u}_h^n \rangle, \\ \mathtt{u}_h^n \leq r_h\psi, \\ \mathtt{u}_h(x) = \mathtt{u}_h(x,t) \qquad in \quad \Omega_h, \\ \mathtt{u}(x,t) = 0. \quad in \quad \Sigma_h, \end{cases} \tag{13}$$

where $(\mathbb{B}_h)_{i,j}$ denotes the finite element matrix defined by

$$(\mathbb{B}_h)_{i,j} = \mathtt{b}(\phi_i, \phi_j) = \mathtt{a}(\phi_i, \phi_j) + \omega(\phi_i, \phi_j),$$

and $\{\phi_i\}$ is basis of $\mathbb{V}_h,\ 1 \leq i, 1 \leq j$.

The discretization matrices $\mathbb{B}_h$ and the generic coefficient matrices $\mathtt{a}(\phi_h^i, \phi_h^j)$ are introduced in a natural progression, where the customary basic functions are denoted by $\phi_h = 1, 2, \ldots, m(h_h)$. Now that these definitions have been established, we may state the discrete problem as follows: Determine $\mathtt{u}_h^n \in \mathbb{V}_h$, representing the solution for

$$\begin{cases} \langle \mathbb{B}_h \mathtt{u}_h^n, \mathtt{v}_h - \mathtt{u}_h^n \rangle \ \geq \ \langle F^n(\mathtt{u}_h^n), \mathtt{v}_h - \mathtt{u}_h^n \rangle \,, \\ \mathtt{u}_h^n \leq r_h \psi, \\ \mathtt{u}_h(x) = \mathtt{u}_h(x,t) \qquad in \quad \Omega_h, \\ \mathtt{u}(x,t) = 0 \quad in \quad \Sigma_h. \end{cases} \tag{14}$$

The matrix $(\mathbb{B}_h)$ with coefficients $\mathtt{b}(\phi_i, \phi_j)$ is an M-matrix (see([16, 17, 18]). The problem (14) has a unique solution (as is well known), and we obtain the following regularity result.

Theorem 1. [10] Let $\mathtt{u}^\infty$ and $\mathtt{u}_h^n$ be the solutions of (7) and (14), respectively. Then

$$\|\mathtt{u}_h^n - \mathtt{u}^\infty\|_\infty \leq C \left[h_k^2 |\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha} \right)^N \right], \tag{15}$$

where $\mathtt{u}^\infty$ is the asymptotic solution of the problem (12), and $\mathtt{u}_h^n$ is the discrete solution calculated at the moment $T = n\delta t, C > \ 0$.

3.3 HJB-formulation of discrete problem

Before commencing work in this section, we will use the following symbols:

$$\mathbb{B}_h = \mathbb{B}_k^v, \quad \mathtt{u}_h^n = \mathtt{u}_k^*, \quad \mathbb{U}_h = \mathbb{U}_k, \quad \mathbb{V}_h = \mathbb{V}_k.$$

The problem (14) can be expressed as the following HJB equation. Let $\mathtt{u}_k^v$ denote the unique solution of the discrete HJB equation,

$$\max_{1\leq i\leq N}\left(\mathbb{B}_{k,i}\mathtt{u}^v_{k,i}-F(\mathtt{u}^v_{k,i});\ \mathtt{u}^v_{k,i}-r_k\psi\right)=0.$$

We put

$$\mathtt{u}^v_{k,i}-r_k\psi=\lambda^v_{k,i}.$$

Iterative steps are as follows:

Step 1: We select an initial vector $\mathtt{u}^0_k\in\mathbb{R}^{n_k}$.

Step 2: Let $\mathtt{u}^v_k\in\mathbb{R}^{n_k}, v\geq 0$, and we compute $\mathtt{u}^{(v+1)}_k$, where $\mathtt{u}^{(v+1)}_k$ is a solution of the following equation:

$$\mathbb{B}^v_k\mathtt{u}^{v+1}_k-F^v_k(\mathtt{u}^v_k)=0, \tag{16}$$

where

$$\mathbb{B}^v_k=\begin{cases}\mathbb{B}_{k,i}(\mathtt{u}^v_k)\ if\ \mathbb{B}_{k,i}\mathtt{u}^v_{k,i}-F(\mathtt{u}^v_{k,i})\ >\ \lambda^v_{k,i}\\ \quad I_{k,i}\quad if\ \mathbb{B}_{k,i}\mathtt{u}^v_{k,i}-F(\mathtt{u}^v_{k,i})\ \leq\ \lambda^v_{k,i}\end{cases} \tag{17}$$

$$F^v_k(\mathtt{u}^v_k)=\begin{cases}F^v_k(\mathtt{u}^v_k)\ if\ \mathbb{B}_{k,i}\mathtt{u}^v_{k,i}-F(u^v_{k,i})\ >\ \lambda^v_{k,i}\\ \quad\psi_k\quad if\ \mathbb{B}_{k,i}\mathtt{u}^v_{k,i}-F(\mathtt{u}^v_{k,i})\ \leqslant\lambda^v_{k,i}\end{cases} \tag{18}$$

Consider $\mathtt{u}^*_k$ as the unique solution to the discrete HJB equation,

$$\max_{1\leq i\leq N}\left(\mathbb{B}_{k,i}\mathtt{u}^*_{k,i}-F(\mathtt{u}^*_{k,i});\ \mathtt{u}^*_{k,i}-\psi_k\right)=0.$$

The following theorem represents the equivalence between the HJB equation and the PQVI by using the finite elements method, and we will rely on its proof of the work done by Hoppe on the elliptic quasi-variational inequality using the finite differences method.

Theorem 2. Let $\mathtt{u}^v_k$ be the iterate obtained by the previous iterative scheme, satisfying the HJB-equation above. Additionally, we suppose that $\mathbb{B}_k$ is monotone. Then, the sequence $\mathtt{u}^v_k$, where $(v\geq 0)$, exhibits monotonically decreasing convergence toward the unique solution $\mathtt{u}^*_k$ of (13).

Proof. Let $\mathtt{u}^v_k$ represent an iteration, and let $\mathtt{u}^*_k$ denote a solution of the HJB equation. To prove that $\mathtt{u}^v_k$ converges towards $\mathtt{u}^*_k$, it is adequate to show that $(\mathtt{u}^v_k,\mathtt{v}\geq 0)$ consistently decreases towards $\mathtt{u}^*_k$, as outlined below:

$$\mathtt{u}^*_k\leq\mathtt{u}^{v+1}_k\leq\mathtt{u}^v_k.$$

To begin, let us consider $\mathtt{u}^0_k$ satisfying

$$\max(\mathbb{B}_k^0 \mathtt{u}_k^0 - F_k, \mathtt{u}_k^0 - \psi_k) \geq 0,$$

for all $v \geq 0$,

$$\max(\mathbb{B}_k^v \mathtt{u}_k^v - F_k, \mathtt{u}_k^v - \psi_k \geq (\mathbb{B}_k^v \mathtt{u}_k^v - F_k) \geq 0.$$

Using (16), we get

$$(\mathbb{B}_k^v \mathtt{u}_k^v - \mathbb{B}_k^v u_k^{v+1}) \geq 0.$$

We have $\mathbb{B}_k^v$, which is linear and monotone,

$$\mathbb{B}_k^v (\mathtt{u}_k^v - \mathtt{u}_k^{v+1}) \geq 0.$$

Then

$$(\mathtt{u}_k^v - \mathtt{u}_k^{v+1}) \geq 0,$$

$$\mathtt{u}_k^{v+1} \leq \mathtt{u}_k^v. \tag{19}$$

We know that $\mathtt{u}_k^*$ is a solution of (14). We have

$$\max(\mathbb{B}_k^v \mathtt{u}_k^* - F_k, \mathtt{u}_k^* - \psi_k) = 0,$$

again, using (16)

$$\mathbb{B}_k^v \mathtt{u}_k^* - \mathbb{B}_k^v \mathtt{u}_k^{v+1} \leq 0,$$

$$\mathbb{B}_k^v \mathtt{u}_k^* \leq \mathbb{B}_k^v \mathtt{u}_k^{v+1}.$$

We have $\mathbb{B}_k^v$ is a monotone matrix; that is, $\mathbb{B}_k^v$ is an invertible matrix ($(\mathbb{B}_k^v)^{-1}$ exists). Thus

$$(\mathbb{B}_k^v)^{-1} \mathbb{B}_k^v \mathtt{u}_k^* \leq (\mathbb{B}_k^v)^{-1} \mathbb{B}_k^v \mathtt{u}_k^{v+1}.$$

Then,

$$\mathtt{u}_k^* \leq \mathtt{u}_k^{v+1}. \tag{20}$$

From (19) and (20), we obtain

$$\mathtt{u}_k^* \leq \mathtt{u}_k^{v+1} \leq \mathtt{u}_k^v.$$

The uniqueness: To demonstrate uniqueness, we suppose that there are two solutions $\mathtt{u}_1^*$ and $\mathtt{u}_2^*$ of (14).

First, we put $\mathtt{u}_1^* = \mathtt{u}_k^*$ and $\mathtt{u}_2^* = \mathtt{u}_k^{v+1}$,

$$\max(\mathbb{B}_k^v \mathtt{u}_1^* - F_k, \mathtt{u}_1^* - \psi_k) = 0 \geq (\mathbb{B}_k^v u_1^* - \mathbb{B}_k^v u_k^{v+1}),$$

$$\mathbb{B}_k^v \mathtt{u}_1^* - \mathbb{B}_k^v \mathtt{u}_2^* \leq 0, \qquad \mathbb{B}_k^v \mathtt{u}_1^* \leq \mathbb{B}_k^v \mathtt{u}_2^*.$$

Therefore

$$\mathtt{u}_1^* - \mathtt{u}_2^* \leq 0. \tag{21}$$

Secondly, we put $u_2^* = u_k^*$ and $u_1^* = u_k^{v+1}$,

$$\max(\mathbb{B}_k^v \mathtt{u}_2^* - F_k, \mathtt{u}_2^* - \psi_k) = 0 \geq (\mathbb{B}_k^v \mathtt{u}_2^* - \mathbb{B}_k^v \mathtt{u}_k^{v+1}).$$

$$\mathbb{B}_k^v \mathtt{u}_2^* - \mathbb{B}_k^v \mathtt{u}_1^* \leq 0 \quad \textit{then} \quad \mathbb{B}_k^v \mathtt{u}_2^* \leq \mathbb{B}_k^v \mathtt{u}_1^*.$$

We have

$$\mathtt{u}_1^* - \mathtt{u}_2^* \geq 0. \tag{22}$$

From (21) and (22), we get

$$\mathtt{u}_1^* = \mathtt{u}_2^*.$$

Then, the solution of (14) ($\mathtt{u}_k^v$), approximates towards the solution of problem (13), $\mathtt{u}_k^v \mapsto \mathtt{u}_k^*$. □

3.4 Description of the Newton-multigrid method for PQVIs

In addressing the nonlinear system (16), we utilize the Newton-multigrid method, a hybrid technique that merges Newton's method for nonlinear systems with the multigrid method for linear systems. The application of the Newton-multigrid approach involves the following procedural steps:

First step: By utilizing Newton's method, we obtain the nonlinear system, recognizing that f is a Lipschitz function.

Second step: In every linear step of the system (16), we seek the Jacobian system to derive a linear system connected with the Jacobian matrix.

Third step: We use the multigrid method to solve the linear system resulting from the Jacobian matrix.

Let $\mathtt{u}_k^v$ be the exact solution of the system (16), and let $\mathtt{w}_k^v$ be an approximation to $\mathtt{u}_k^v$. We define the residual as follows:

$$\mathfrak{R}_k(\mathtt{u}_k^v) = F_k^v(\mathtt{u}_k^v) - \mathbb{B}_k^v \mathtt{w}_k^v. \tag{23}$$

Substituting (16) in (23), we find

$$\mathfrak{R}_k(\mathtt{u}_k^v) = \mathbb{B}_k^v \mathtt{u}_k^v - \mathbb{B}_k^v \mathtt{w}_k^v, \tag{24}$$

$$\mathfrak{R}_k(\mathtt{u}_k^v) = \mathbb{B}_k^v \ (\mathtt{u}_k^v - \mathtt{w}_k^v) = \mathbb{B}_k^v (e_k^v), \tag{25}$$

where $\mathtt{e}_k^v$ is the error with $\mathtt{e}_k^v = \mathtt{u}_k^v - \mathtt{w}_k^v$. In the context of (25), determining the error $\mathtt{e}_k^v$ for the linear equation (25) on the coarse grid, as commonly done in the multigrid method, is not feasible. Consequently, we adopt (25) as a residual equation, given the nonlinear characteristics of f.

To facilitate access to the solution, we choose $\mathcal{H}$ as a nonlinear operator, where

$$\mathcal{H}_k(\mathtt{u}_k^v) = \mathbb{B}_k^v \mathtt{u}_k^v - F_k^v(\mathtt{u}_k^v) = 0. \tag{26}$$

The residual in the fine grid can be rewritten as follows:

$$\mathfrak{R}_k(\mathtt{u}_k^v) = -\mathcal{H}_k(\mathtt{u}_k^v). \tag{27}$$

To solve (26), we employ the Newton iteration method as follows:

$$\mathtt{u}_k^{v+1} = \mathtt{u}_k^v + \frac{\mathfrak{R}_k(\mathtt{u}_k^v)}{\mathcal{J}_k(\mathtt{u}_k^v)}, \tag{28}$$

where $\mathcal{J}_k(\mathtt{u}_k^v)$ represents the Jacobian matrix of the nonlinear system, and $\mathcal{J}_k(\mathtt{u}_k^v) = \mathcal{H}_k^{'}(\mathtt{u}_k^v)$.

From (28), we can derive the following Jacobian linear system for $\mathtt{e}_k^v$:

$$\mathcal{J}_k(\mathtt{u}_k^v)\mathtt{e}_k^v = \mathfrak{R}_k(\mathtt{u}_k^v), \tag{29}$$

where $\mathtt{u}_k^{v+1} - \mathtt{u}_k^v = \mathtt{e}_k^v$.

The solution to the linear system (29) is utilized as an approach to solve the nonlinear system (26). Therefore, to find the solution for the nonlinear system (26), we employ the multigrid method to solve the associated linear system (29).

3.5 Multigrid technique

In addressing the linear system (29), we employ the multigrid technique. By selecting an iteration $\mathtt{e}_k^v$, where $v > 0$, within the multigrid method, we derive $\overline{\mathtt{e}}_k^v$ through the application of an iterative method. This iterative process is utilized to solve the system (29), utilizing α as the coefficient expression as follows:

$$\overline{\mathtt{e}}_k^v = \mathcal{S}_k^v(\mathtt{e}_k^v). \tag{30}$$

Here, $\mathcal{S}_k^v$ stands for the iteration or smoothing operator, and α represents the number of iterations executed. The solution to (29) is denoted as $\mathtt{e}^\star$.

The error is defined as $\mathtt{E}_k^v = \overline{\mathtt{e}}_k^v - \mathtt{e}_k^\star$, and the residual is also considered

$$\mathtt{d}_k^v = \mathfrak{R}_k(\mathtt{u}_k^v) - \mathcal{J}_k(\mathtt{u}_k^v)\overline{\mathtt{e}}_k^v.$$

We can write (29) as

$$\mathcal{J}_k(\mathtt{u}_k^v)(\overline{\mathtt{e}}_k^v + \mathtt{E}_k^v) = \mathfrak{R}_k(\mathtt{u}_k^v). \tag{31}$$

We derive the subsequent residual equation,

$$\begin{aligned}
&\mathcal{J}_k(\mathtt{u}_k^v)\mathtt{E}_k^v = \mathfrak{R}_k(\mathtt{u}_k^v) - \mathcal{J}_k(\mathtt{u}_k^v)\overline{\mathtt{e}}_k^v = \mathtt{d}_k^v,\\
&\mathcal{J}_k(\mathtt{u}_k^v)\mathtt{E}_k^v = \mathtt{d}_k^v.
\end{aligned}$$

To fully determine $\mathtt{E}_k^v$, it is necessary to compute $\mathtt{E}_k^v$ at the level $(k-1)$ as the solution to the coarse grid system,

$$\mathcal{J}_{k-1}(\mathtt{e}_{k-1}^v)\mathtt{E}_{k-1}^v = \mathtt{d}_{k-1}^v. \tag{32}$$

In the context, $\mathtt{E}_{k-1}^v$, $\mathcal{J}_{k-1}(\mathtt{e}_{k-1}^v)$, $\mathtt{d}_{k-1}^v$ represent approximations at the $(k-1)$ of $\mathtt{E}_k^v$, $\mathcal{J}_k(\mathtt{e}_k^v)$, $\mathtt{d}_k^v$, respectively.

We have

$$\begin{aligned}
\mathtt{E}_{k-1}^v &= \mathtt{R}_k\mathtt{E}_k^v,\\
\mathtt{d}_{k-1}^v &= \mathtt{R}_k\mathtt{d}_k^v,\\
\mathcal{J}_{k-1}^v(\mathtt{e}_{k-1}^v) &= \mathcal{R}_k\mathcal{J}_k^v(e_k^v)\mathcal{P}_k,
\end{aligned}$$

where $\mathcal{R}_k$ is the restriction matrix and $\mathcal{P}_k$ is the prolongation matrix. Owing to the nested structure, we employ the clearly defined identity operator,

$$\Psi : \mathbb{V}_{k-1} \longrightarrow \mathbb{V}_k,$$

for defining the prolongation and restriction operators, that is,

$$\begin{aligned} \mathcal{R}_k &= \mathcal{P}_k^t \\ \mathcal{P}_k &= r_k r_{k-1}^{-1}. \end{aligned} \tag{33}$$

Note: To solve the linear system (29) for the two sequences Ω_k and Ω_{k-1}, we employ the multigrid technique as an internal iteration. Additionally, to address system (26), we utilize Newton's method as an external iteration. This results in the iterative solution of the system (32), achieved by applying a two-grid iteration repeatedly to the sequence $\Omega_k\{k = 0, \ldots, m_k\}$ until the iteration process is halted.

4 $\mathbb{L}^\infty$-convergence of multigrid method

This paragraph explores the assessment of uniform convergence for the multigrid algorithm (inner iteration) and the convergence characteristics of Newton's method (outer iteration). The assumptions employed in these convergence analyses closely resemble those utilized in multigrid methods specifically crafted for addressing nonlinear equations. We now present the main hypotheses:

Hypothesis 1:

$$\text{There exists } \mathtt{u}_k^\star \in V_k \quad \textit{such that} \quad \mathcal{H}_k(\mathtt{u}_k^\star)^{-1} = 0.$$

Hypothesis 2 (inverse):

$$\mathcal{H}_k'(\mathtt{u}_k^\star)^{-1} \textit{ is exist and } \|\mathcal{H}_k'(\mathtt{u}_k^\star)^{-1}\|_\infty \leq k \quad \textit{with } k > 0.$$

Hypothesis 3 (continuous):

$$\text{for all } \epsilon > 0, \quad \text{there exists } \eta > 0, \quad \|I - \mathcal{H}_k'(\mathtt{u}_k^\star)\mathcal{H}_k'(\mathtt{u}_k)^{-1}\|_\infty \leq \epsilon,$$

whenever

$$\|(\mathtt{u}_k - \mathtt{u}_k^\star)\|_\infty \leq \eta.$$

Hypothesis 4: For any $\mathtt{u}_k$ in the neighborhood of $\mathtt{u}_k^\star$, there is a linear mapping denoted as $\mathcal{H}_k^{'}(\mathtt{u}_k)$ such that for all $\epsilon > 0$ there exists $\eta > 0$ such that

$$\|\mathcal{H}_k(\mathtt{u}_k) - \mathcal{H}_k(\mathtt{u}_k^\star) - \mathcal{H}_k^{'}(\mathtt{u}_k^\star)\|(\|(\mathtt{u}_k - \mathtt{u}_k^\star)\|_\infty \leq \epsilon\|(\mathtt{u}_k - \mathtt{u}_k^\star)\|_\infty$$

hold, whenever

$$\|(\mathtt{u}_k - \mathtt{u}_k^\star)\|_\infty \leq \eta.$$

4.1 Matrix associated with the MGHJB algorithm

The two-grid method's iteration matrix, which includes α_1 presmoothing and α_2 postsmoothing at level $k-1$, can be represented as follows:

$$\mathbb{TG}_k(\alpha_1, \alpha_2) = \mathcal{S}_k^{\alpha_2}\,(\mathcal{J}_k^{-1} - \mathcal{P}_k\,\mathcal{J}_{k-1}^{-1}\mathcal{R}_k)\mathcal{J}_k\,\mathcal{S}_k^{\alpha_1}, \qquad k = 1, 2, \ldots. \tag{34}$$

Theorem 3. [26, 27] The multigrid approach is a linear iterative technique characterized by the iteration matrix $\mathbb{MG}_k$, which is expressed as follows:

$$\mathbb{MG}_0 = 0 \quad if \quad k = 0,$$

$$\mathbb{MG}_k = \mathcal{S}_k^{\alpha_2}\left(\mathbb{I}_k - \mathcal{P}_k\,(\mathbb{I}_k - \mathbb{MG}_{k-1})\,(\mathcal{J}_{k-1})^{-1}\mathcal{R}_k\right)\mathcal{J}_k\,\mathcal{S}_k^{\alpha_1} \quad if \quad k = 1, 2, \ldots, \tag{35}$$

$$\begin{aligned}\mathbb{MG}_k &= \mathcal{S}_k^{\alpha_2}\left(\mathbb{I}_k - \mathcal{P}_k\,(\mathbb{I}_k - \mathbb{MG}_{k-1})\,(\mathcal{J}_k^{-1}\mathcal{R}_k\right)\mathcal{J}_k\,\mathcal{S}_k^{\alpha_1} \\ &= \mathbb{TG}_k + \mathcal{S}_k^{\alpha_2}\mathcal{P}_k(\mathbb{MG}_{k-1})\mathcal{J}_k^{-1}\mathcal{R}_k\mathcal{J}_k\mathcal{S}_k^{\alpha_1} \quad if \quad k = 1, 2, \ldots.\end{aligned}$$

4.2 Approximation property

The following approximation property is based on the inequality (15) in Theorem 1 and in Lemma 1. The demonstration of the theorem is adapted from the one in [21], originally presented for the problem of variational inequality.

4.3 The main result

The following theorems give us the convergence of the multigrid matrix of the parabolic quasi variational, in its proof we will rely on the work done by Haiour [21] in the elliptic quasi-variational inequality.

Theorem 4. $\mathbb{X}$ is the iteration matrix of the multigrid, given by

$$\mathbb{X} = \left[\mathcal{J}_k^{-1} - \mathcal{P}_k\, \mathcal{J}_{k-1}^{-1}\mathcal{R}_k\right]. \tag{36}$$

Under the previous assumptions, the matrix $\mathbb{X}$ satisfies the following approximation properties:

$$\|\mathbb{X}\|_\infty \leq C\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]. \tag{37}$$

Proof. According to Theorem 1, let $\mathtt{u} \in H_0^1(\Omega)$. We have

$$\|\mathtt{u}_k^n - \mathtt{u}^\infty\|_\infty \leq C\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\nabla F\|_\infty.$$

Then

$$\|\mathtt{u}_k^\star - \mathtt{u}^\infty\|_\infty \leq C\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\nabla F\|_\infty.$$

Let $\mathtt{u}_k^\star$ and $\mathtt{u}_{k-1}^\star$ be solutions to the problem (13). Then by applying Theorem 1,

$$\begin{aligned}
\|\mathtt{u}_k^* - \mathtt{u}_{k-1}^*\|_\infty &\leq \|\mathtt{u}_k^* - \mathtt{u}^\infty\|_\infty + \|\mathtt{u}_{k-1}^* - \mathtt{u}^\infty\|_\infty \\
&\leq C_1\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\nabla F\|_\infty \\
&\quad + C_2\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\nabla F\|_\infty \\
&\leq (C_1+C_2)\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\nabla F\|_\infty.
\end{aligned} \tag{38}$$

We obtain

$$\|\mathtt{u}_k^* - \mathtt{u}_{k-1}^*\|_\infty \leq C\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\nabla F\|_\infty, \tag{39}$$

where $C = (C_1 + C_2)$.

We use the Galerkin discretization to obtain

$$\mathtt{b}(r_k\mathtt{u},\ r_k\mathtt{v}) = \langle(\mathbb{B}_k),\ \mathtt{v}\rangle_{L^2(\Omega)}\,;\ \text{for all } \mathtt{u},\mathtt{v}\in\mathbb{U}_k.$$

So

$$\left(\mathtt{b}(\mathtt{u},\ \mathtt{v})\right)^{'} = \langle(\mathcal{J}_k),\ \mathtt{v}_k\rangle_{L^2(\Omega)}\,;\ \text{for all } \mathtt{u},\mathtt{v}\in\mathbb{U}_k.$$

Also, we have for all $\mathtt{v}\in\mathbb{V}$

$$\left(\mathtt{b}((r_k)^{-1}\,(\mathcal{J}_k)^{-1}(F_k),\ \mathtt{v})\right)^{'} = \left\langle(r^*_{k-1})^{-1}\,\nabla F\ \ \mathtt{v}\right\rangle_{L^2(\Omega)}.$$

Let $\mathtt{u}_k\in\mathbb{V}_k$ and $\mathtt{u}_{k-1}\in\mathbb{V}_{k-1}$. Then

$$\mathtt{b}(\mathtt{u}_k,\ \mathtt{v}) = \left\langle\left(r^*_k\right)^{-1}F,\ \mathtt{v}\right\rangle_{L^2(\Omega)},$$
$$\mathtt{b}(\mathtt{u}_{k-1},\ \mathtt{v}) = \left\langle\left(r^*_{k-1}\right)^{-1}F,\ \mathtt{v}\right\rangle_{L^2(\Omega)},$$

which implies $\mathtt{u}^\star_k = r_k^{-1}\mathcal{J}_k^{-1}\nabla(F)$ and $\mathtt{u}^\star_{k-1} = r_{k-1}^{-1}\mathcal{J}_{k-1}^{-1}\mathcal{R}_k\nabla(F)$. Using (39) and (33) and Lemma 1, we get

$$\|r_k^{-1}\mathcal{J}_k^{-1}\nabla(F) - r_{k-1}^{-1}\mathcal{J}_{k-1}^{-1}\mathcal{R}_k\nabla(F)\|_\infty \le C\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\nabla F\|_\infty.$$

Then

$$\|\mathcal{J}_k^{-1} - r_k r_{k-1}^{-1}\mathcal{J}_{k-1}^{-1}\mathcal{R}_k\|_\infty\|\nabla F\|_\infty \le C\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\nabla F\|_\infty.$$

The proof is now concluded from

$$\|\mathcal{J}_k^{-1} - \mathcal{P}_k\mathcal{J}_{k-1}^{-1}\mathcal{R}_k\|_\infty \le C\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right],$$
$$\|\mathbb{X}\|_\infty \le C\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right].$$

□

4.4 Smoothing property

To establish a smoothing property, we form the matrix $\mathcal{J}_k^v = \mathcal{L}_k - \mathcal{N}_k$, and rely on the subsequent assumptions:
$\mathcal{L}_k$ is considered regular, where

$$\|\mathcal{L}_k^{-1}\,\mathcal{N}_k\|_\infty \leq 1 \qquad \text{for all } k, \tag{40}$$

$$\|\mathcal{L}_k^{-1}\|_\infty \leq \frac{C}{h^2}, \quad \text{for all } k, \text{ with } C \text{ independent } k. \tag{41}$$

We utilized a relaxation method with an iteration matrix as a smoother,

$$\mathcal{S}_k = I_k - \omega \mathcal{J}_k^{-1}\mathcal{N}_k, \qquad \omega \in [0;1].$$

Theorem 5. [26, 27] Given that the preceding assumptions and notations are met, there exists a constant C, independent of k, such that the following smoothing property holds:

$$\|\mathcal{J}_k\,\mathcal{S}_k^\alpha\|_\infty \leq \frac{C}{\sqrt{\alpha}\,h_k^2}. \tag{42}$$

Following the approximation and smoothing properties, it is essential to establish the subsequent stability bound:

$$\text{There exists } C_s \quad \|\mathcal{S}_k^\alpha\|_\infty \leq C_s, \qquad \text{for all k and } \alpha. \tag{43}$$

The convergence analysis relies on the following decomposition of the iteration matrix for the two-grid method, where $\alpha_2 = 0$:

$$\begin{aligned}\|\mathbb{TG}_k(\alpha_1,0)\|_\infty &= \|\left[(\mathcal{J}_k)^{-1} - \mathcal{P}_k\,(\mathcal{J}_{k-1})^{-1}\,\mathcal{R}_k\right]\;\mathcal{J}_k\,\mathcal{S}_k^{\alpha_1}\,\|_\infty \\ &\leq \|(\mathcal{J}_k)^{-1} - \mathcal{P}_k\,(\mathcal{J}_{k-1})^{-1}\,\mathcal{R}_k\|_\infty\,\|\mathcal{J}_k\,\mathcal{S}_k^{\alpha_1}\|_\infty.\end{aligned}$$

Now, choosing more than two grids for a hierarchy comprising, we can formulate the matrix $\mathbb{X}$ in (36) by recursively utilizing the matrix $\mathbb{MG}_k$ in (35) for all levels. Assuming the validity of condition (36), convergence outcomes can be easily inferred from the preceding findings.

Theorem 6. [25] Examine a multigrid approach for a given iterative matrix (35). Then, based on the earlier assumption, for the specified parameter value $\alpha_1 = \alpha,\ \alpha_2 = 0,$ for all $\varepsilon \in [0,1],$ there exists $\alpha^* \leq \alpha :$

$$\|\mathbb{MG}_k\|_\infty \leq \varepsilon.$$

Combining the approximation and smoothness properties with (3.5) allows us to utilize identical parameters, as outlined in the subsequent theorem. This theorem constitutes the primary outcome of our study.

Theorem 7. Given the preceding assumptions and notations, the iterated u_k^v, $v > 0$, for two meshes k and $k-1$, adhere to the following relationship: There exists $C > 0$:

$$\|\mathtt{u}_k^{v+1} - \mathtt{u}_k^*\|_\infty \leq \frac{C}{\sqrt{\alpha}\, h_k^2}\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\mathtt{u}_k^v - \mathtt{u}_k^*\|_\infty.$$

Proof. We have

$$\begin{aligned}
\|\mathtt{u}_k^{v+1} - \mathtt{u}_k^*\|_\infty &= \|(\mathcal{J}_k\, \mathcal{S}_k^\alpha)(I_k - \mathcal{P}_k(I_k - \mathbb{MG}_{k-1})(\mathcal{J}_{k-1}^{-1})\mathcal{R}_k(\mathtt{u}_k^v - \mathtt{u}_k^*)\|_\infty \\
&\leq \|(\mathcal{J}_k\, \mathcal{S}_k^\alpha)\|_\infty \|(I_k - \mathcal{P}_k(I_k - \mathbb{MG}_{k-1})(\mathcal{J}_{k-1}^{-1})\mathcal{R}_k\|_\infty \|\mathtt{u}_k^v - \mathtt{u}_k^*\|_\infty \\
&\leq \frac{C_1 C_2}{\sqrt{\alpha}\, h_k^2}\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\mathtt{u}_k^v - \mathtt{u}_k^*\|_\infty \\
&\leq \frac{C}{\sqrt{\alpha}\, h_k^2}\left[h_k^2|\log(h_k)|^3 + \left(\frac{1+nc}{1+n\alpha}\right)^N\right]\|\mathtt{u}_k^v - \mathtt{u}_k^*\|_\infty,
\end{aligned}$$

where $C = C_1 \times C_2$. □

5 Numerical example

Example 1. In this numerical example, we find the solution of the following problem (44) by using multigrid method and Gauss–Siedel method:

$$\begin{cases} \dfrac{\partial u}{\partial t} + Au \leq f(u) \ in \ [0,1] \times \Omega, \\ \\ \left\langle \dfrac{\partial u}{\partial t} + Au - f(u) \, ; \mathtt{u} - \psi \right\rangle = 0, \\ \\ u(x,t) = 0, \ \text{ in } [0,1] \times \Gamma, \\ \\ u(x,0) = 0, \ \ \text{in } \Omega, \end{cases} \tag{44}$$

where

$$\Omega = \{(x_1, x_2 | x_1 + x_2 \leq 1)\}, \quad \Delta t = 0.01, \quad A\mathtt{u} = -\Delta \mathtt{u}, \quad f(u) = \cos 2u, \quad \psi = 0.$$

For discretization in space, we have used PDE toolbox in MATLAB (r2018) to generate the mesh, and semi-discretization on time, and start-iterate $u_k^0 = (0, \ldots, 0)^t \in R^{1024}$.

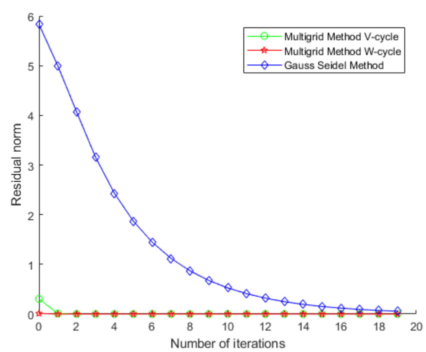


Figure 1: Comparison between the convergence of multigrid method and Gauss–Siedel methods.

We note from Figure 1 that the multigrid method (V-cycle, W-cycle) gives us the solution with the least number of iterations, that is, after one iteration, while the Gauss–Siedel method gives after more than 20 iterations. From here, we conclude that the multigrid method gives us the solution to the sets of large linear equations with the least number of iterations.

Figures 2 and 3 represent the solution to the problem using MATLAB 2018 using the Gauss–Siedel method and the multigrid method (V-cycle and W-cycle).

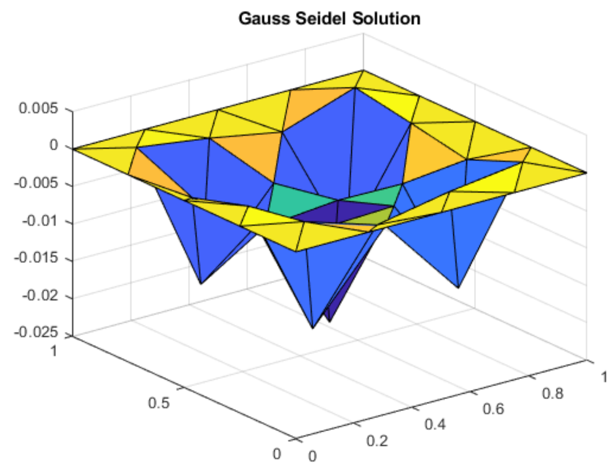
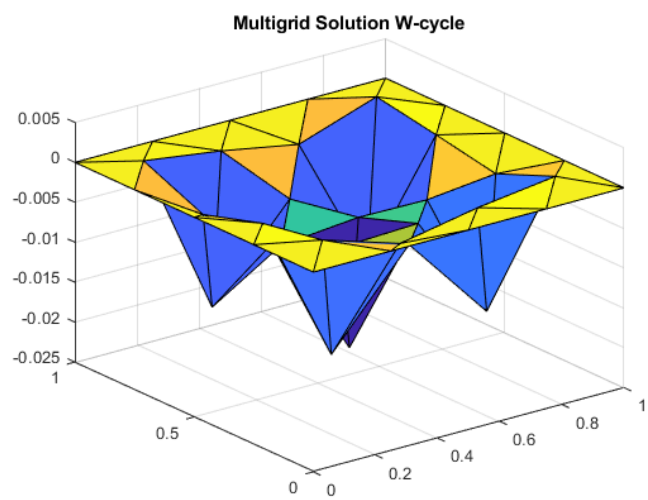


Figure 2: Solution of P.Q.V.I using Gauss–Siedel method (after 20 iterations).



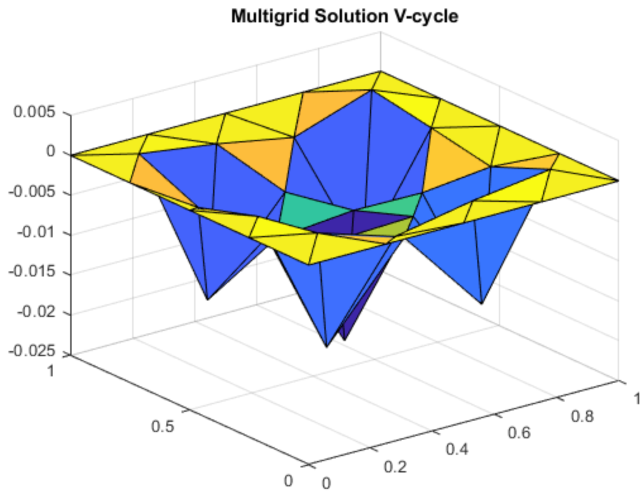


Figure 3: Solution of parabolic quasi-variational inequality after two iterations of multigrid method.

Acknowledgments

Greetings and appreciation are extended to the members of the journal. We also welcome all directions and inquiries submitted by reviewers.

References

[1] Ahmad, I., Seadawy, I., Ahmad, H., Thounthong, P. and Wang, F. *Numerical study of multi-dimensional hyperbolic telegraph equations arising in nuclear material science via an efficient local meshless method*, Int. J. Nonlinear Sci. Numer. Simul. 23(1) (2022), 115–122.

[2] Beggas, M. and Haiour, M. *The maximum norm analysis of schwarz method for elliptic quasi-variationnal inequalities*, Kragujevac Journal of Mathematics 45(5) (2021), 635–645.

[3] Belouafi, M-S. and Beggas, M. *Maximum norm convergence of Newton-multigrid methods for elliptic quasi-variational inequalities with nonlinear source terms*, Adv. Math., Sci. J. 11(10) (2022), 969–983.

[4] Bencheikh, M.A. and Haiour, M. *L_∞-error analysis for parabolic quasi-variational inequalities related to impulse control problems*, Comput. Math. Model. 28 (2017), 89–108.

[5] Bencheikh, M.A., Boulaaras, S. and Haiour, M. *An optimal L_∞-error estimate for an approximation of a parabolic variational inequality*, Numer. Func. Anal. Opt. 37(1) (2016), 1–18.

[6] Bensoussan, A. and Lions, J. *Impulse control and quasi-variational inequalities*, Gauthier Villars, Paris, 1984.

[7] Boulaaras, S. and Haiour, M. *A new approach to asymptotic behavior for a finite element approximation in parabolic variational inequalities*, Int. Sch. Res. Notices 2011 (2011).

[8] Boulaaras, S. and Haiour, M. *L_∞-asymptotic behavior for a finite element approximation in parabolic quasi-variational inequalities related to impulse control problem*, Appl. Math. Comput. 217 (2011), 6443–6450.

[9] Boulaaras, S. and Haiour, M. *The finite element approximation in parabolic quasivariational inequalities related to impulse control problem with mixed boundary conditions*, J. Taibah Univ. Sci. 7(3) (2013), 105–113.

[10] Boulaaras, S. and Haiour, M. *The finite element approximation of evolutionary Hamilton–Jacobi–Bellman equations with nonlinear source terms*, Indag. Math. 24 (2013), 161–173.

[11] Boulaaras, S. and Haiour, M. *The theta time scheme combined with a finite-element spatial, approximation in the evolutionary Hamilton–Jacobi–Bellman equation with linear source terms*, Comput. Math. Model. 25(3) (2014), 423–438.

[12] Boulaaras, S. and Haiour, M. *A new proof for the existence and uniqueness of the discrete evolutionary HJB equation*, Appl. Math. Comput. 262 (2015), 42–55.

[13] Boulbrachene, M. and Chentouf, B. *The finite element approximation of Hamilton–Jacobi–Bellman equations: The noncoercive case*, J. Appl. Math. Comput. 158(2) (2004), 585–592.

[14] Brandt, A. *Algebraic multigrid theory: The symmetric case*, Appl. Math. Comput. 19 (1986), 23–56.

[15] Brown, P.N., Vassilevski, P.S. and Woodward, C.S. *On mesh-independent convergence of an inexact Newton–multigrid algorithm*, SIAM J. Comput. 25(2) (2003), 570–590.

[16] Ciarlet, P.G. and Raviart, P-A. *Maximum principle and uniform convergence for the finite element method*, Comput. Methods Appl. Mech. Eng. 2(1) (1973), 17–31.

[17] Cortey-Dumont, P. *Approximation numérique d'une inéquation quasi variationnelle liée à des problèmes de gestion de stock*, RAIRO. Anal. numér. 14(4) (1980), 335–346.

[18] Cortey-Dumont, P. *On finite element approximation in the L_∞ norm of variational inequalities*, Numer. Math. (Heidelb.), 47(1) (1985), 45–57.

[19] Hackbusch, W. *Multi-grid methods and applications*, Springer- Verlag (Springer Series in Computational Mathematics, Berlin and New York, 1985.

[20] Hackbusch, W. and Mittelmann, H-D. *On multi-grid methods for variational inequalities*, Numer. Math. (Heidelb.) 42(1) (1983), 65–76.

[21] Haiour, M. *Etude de la convergence uniforme de la methode multigrilles appliquees aux problemes frontieres libres*, PhD thesis, Universite de Annaba-Badji Mokhtar, 2004.

[22] Hoppe, H.W. *Multi-grid methods for Hamilton-Jacobi-Bellman equations*, Numer. Math. (Heidelb.) 49(2) (1986), 239–254.

[23] Hoppe, H.W. *Multigrid algorithms for variational inequalities*, SIAM J. Numer. Anal. 24(5) (1987), 1046–1065.

[24] Hoppe, H.W. *Une méthode multigrille pour la solution des problèmes d'obstacle*, ESAIM: Math. Model. Numer. Anal. 24(6) (1990), 711–735.

[25] Nesba, N-H. Beggas, M. Belouafi, M-S. Imtiaz, A. and Hijaz, A. *Multigrid methods for the solution of nonlinear variational inequalities*, Eur. J. Pure Appl. 16(3) (2023), 1956–1969.

[26] Reusken, A. *On maximum norm convergence of multigrid methods for elliptic boundary value problems*, SIAM J. Numer. Anal. 31(2) (1994), 378–392.

[27] Reusken, A. *Introduction to multigrid methods for elliptic boundary value problems*, Inst. Fur. Geometrie und Praktische Mathematik, 2008.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1016–1036
https://doi.org/10.22067/ijnao.2024.86494.1380
https://ijnao.um.ac.ir/

Research Article

Designing the sinc neural networks to solve the fractional optimal control problem

R. Heydari Dastjerdi*, and G. Ahmadi

Abstract

Sinc numerical methods are essential approaches for solving nonlinear problems. In this work, based on this method, the sinc neural networks (SNNs) are designed and applied to solve the fractional optimal control problem (FOCP) in the sense of the Riemann–Liouville (RL) derivative. To solve the FOCP, we first approximate the RL derivative using Grunwald–Letnikov operators. Then, according to Pontryagin's minimum principle for FOCP and using an error function, we construct an unconstrained minimization problem. We approximate the solution of the ordinary differential equation obtained from the Hamiltonian condition using the SNN. Simulation results show the efficiencies of the proposed approach.

*Corresponding author

Received 20 January 2024; revised 22 Aprail 2024; accepted 13 May 2024

Rasoul Heydari Dastjerdi
Department of Mathematics, Payame Noor University, Tehran, Iran. e-mail: rasoul59heydari@pnu.ac.ir

Ghasem Ahmadi
Department of Mathematics, Payame Noor University, Tehran, Iran. e-mail: g.ahmadi@pnu.ac.ir

How to cite this article
Heydari Dastjerdi, R. and Ahmadi, G., Designing the sinc neural networks to solve the fractional optimal control problem. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1016-1036. https://doi.org/10.22067/ijnao.2024.86494.1380

AMS subject classifications (2020): Primary 49M25; Secondary 68T07, 26A33.

Keywords: Sinc numerical method; Neural network; Sinc neural network; Pontryagin's minimum principle; Fractional optimal control problem.

1 Introduction

In the last decades, based on the sinc approximation, there have been developed a variety of numerical methods, which are now called the sinc numerical methods [30, 31]. Since 1974, Stenger [29] has studied the sinc approximation methods for the numerical problems. These methods are used as useful tools to solve linear and nonlinear problems arising from scientific and engineering applications. In general, the error in the sinc numerical method for the single exponential is $O(\exp(-c\sqrt{n}))$ with some positive c and n is the number of nodes used in the method [30]. In 2002, the sinc function was formed by double exponential transformation by Sugihara. He [33, 32] discovered that the error of the new method with a positive value of c is equal to $O(\exp(-\frac{cN}{\ln N}))$.

By expanding the normal integer calculus to noninteger calculus, fractional calculus is created. Today, fractional calculus is used in many branches of science and engineering, which shows its importance and application [4, 5, 14, 17, 22, 24, 28, 34, 37, 38]. An optimal control problem that includes at least one fractional derivative term in the performance index or differential equation dominating the system's dynamics is called the fraction optimal control problem. The fractional optimal control problem (FOCP) can be introduced concerning different definitions of fractional derivatives, where Riemann–Liouville (RL), Caputo, and Grunwald–Letnikov (GL) fractional derivatives are the most essential types of fractional derivatives. Because of the importance of this category of problems, research has been done to solve the FOCP. In [6], a solution of multidimensional FOCPs with inequality constraint by multiwavelets is presented. SIR and SEIR epidemic models are considered by a formulation for optimal control problems for a class of fuzzy fractional differential systems [9]. Nemati, Lima, and Torres [21] used modified hat functions for solving the FOCPs. The authors in [18] gave an extended modal series method and linear programming strategy to solve the FOCPs. In [35], for solving FOCPs with vector components, the authors used a Bernoulli polynomial method. It is hard to find the exact solution for the Hamiltonian system of equations because there exist both right and left in the Hamiltonian system. In [2, 3], the authors, to find the optimal solution of the FOCPs with the RL and Caputo fractional derivative sense, solved the Hamiltonian system of equations, which provides the necessary optimization conditions. To transform the Caputo-type FOCP into a system of algebraic equations, the authors of [7] have employed the GL approximation.

In this work, we consider the FOCP as follows:

$$\begin{aligned} &\min J\Big(x(t), u(t)\Big) = \textstyle\int_{t_0}^{t_f} F\Big(t, x(t), u(t)\Big)dt, \\ &\text{s.t. } {}_{t_0}D_{t_f}^{\alpha} x(t) = G\Big(t, x(t), u(t)\Big), \\ &\qquad x(t_0) = x_0, \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^p$ and $u(t) \in \mathbb{R}^q$ are the state and control variables, respectively. We assume that the integrand F, for all its arguments, has continuous first and second partial derivatives, and G is Lipschitz continuous on a set $\Omega \subset \mathbb{R}^p$. In addition, $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_p]^T, i = 1, 2, \ldots, p$ and $n = [\alpha_i] + 1$, where $[\alpha_i]$ is the integer part of α_i.

The left and right RL fractional derivatives are, respectively, defined in the following form:

$$\begin{aligned} {}_aD_t^{\alpha} f(t) &= \frac{1}{\boldsymbol{\Gamma(n-\alpha)}}\left(\frac{d}{dt}\right)^n \int_a^t (t-z)^{n-\alpha-1} f(z)dz, \\ {}_tD_b^{\alpha} f(t) &= \frac{1}{\boldsymbol{\Gamma(n-\alpha)}}\left(-\frac{d}{dt}\right)^n \int_t^b (z-t)^{n-\alpha-1} f(z)dz. \end{aligned}$$

In the last decade, artificial neural networks (ANNs) have been applied to solving different problems, such as the FOCPs. The results show that the ANNs are accurate and efficient for many problems, and this method is comparable with other methods obtained by mathematical algorithms. The authors provided a method to solve the continuous-time direct adaptive optimal control for partially unknown nonlinear systems in [36]. Sabouri, Effati, and Pakdaman [25] used a neural network approach for solving a class of the FOCPs. Ghasemi and Nazemi [12] have designed a neural network to solve the FOCPs by using Mittag-Leffler. The multilayer perceptron (MLP) with a back propagation learning algorithm is employed for neonatal disease diagnosis [8]. Lagaris and Likas [16] solved the ordinary differential equations and partial differential equations for boundary and initial value problems by ANNs. The fuzzy neural networks algorithm is used for detecting cardiovascular diseases [26]. Recently, Ghasemi, Nazemi, and Hosseinpour [13] investigated the nonlinear FOCPs using ANNs. There are many references in theory and applications of neural networks, such as mathematical programming [19, 20] and optimal control problems [10].

In this work, after introducing the sinc numerical method, we design the sinc neural networks (SNNs). To solve the FOCP using this network, we first approximate the RL derivative used in the FOCP by the GL operator. Then, using Pontryagin's minimum principle (PMP) and the optimality conditions in the Hamiltonian function, we construct the unrestricted optimality problem. Finally, the adjustable parameters in the trial solution related to this problem are determined by the SNN. The main reason for using the SNN in solving FOCPs can be found in the simplicity of the network. Also, SNN has no bias and increases the efficiency of the network due to the linearity of the unknowns of the problem with respect to the output.

The article is arranged as follows: Section 2 introduces the sinc numerical method. In Section 3, the SNN is presented. In Section 4, the proposed technique is utilized for solving FOCPs. Numerical examples are presented in Section 5, and Section 6 contains concluding remarks.

2 Sinc numerical method

In recent decades, sinc numerical methods have been used to solve various problems due to their exponential convergence rate. In [30], sinc function is fully introduced. In this section, we briefly state some of its definitions and features.

The sinc function is defined for each $x \in (-\infty, \infty)$ as follows:

$$Sinc(x) = \begin{cases} \dfrac{\sin \pi x}{\pi x}, & x \neq 0, \\ 1, & x = 0. \end{cases}$$

For $h > 0$, and any integer k, the translated sinc functions with evenly spaced nodes are defined as

$$S(k, h)(x) = Sinc\left(\frac{x}{h} - k\right).$$

By these functions, a set of interpolation functions can be formed in the following form:

$$S(k, h)(jh) = \delta_{j,k} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases} \tag{2}$$

Assume that f is a function defined on real line. Then for $h > 0$, the cardinal function corresponding to f is introduced by

$$C(f, h)(x) = \sum_{k=-\infty}^{\infty} f(kh) S(k, h)(x), \tag{3}$$

while the series in (3) converges [30]. From the relations (2) and (3), it can be concluded that the function f is interpolated by the cardinal function in points $\{nh\}_{n=-\infty}^{+\infty}$. Suppose that in the complex w -plane, D_d shows the infinite strip region of width $2d$ as

$$D_d = \{w = t + is : |s| < d\}.$$

We use a conformal map ϕ as $\phi(\Gamma) = R$ for problems on a subinterval $\Gamma \subseteq R$, such that ϕ has the inverse ψ and is a conformal map of the simply-connected domain D, where $(0, 1) \subseteq D$, onto D_d, then on a subinterval $\Gamma = (0, 1) = \psi(R)$ with $\phi(0) = -\infty$, and $\phi(1) = +\infty$. The following interpolation can be defined as

$$f(x) \cong \sum_{k=-N}^{N} f(x_k) S_k(x),$$

where $x_k = \psi(kh)$ is defined as the sinc grid points and

$$S_k(x) = S(k,h)o\phi(x) = Sinc\left(\frac{\phi(x)}{h} - k\right)$$

is the translated sinc basic function. As a result of these discussions, double exponential and single exponential can be introduced, respectively, as follows:

$$z = \psi_1(w) = \frac{1}{2} + \frac{1}{2}\tanh\left(\frac{\pi}{2}\sinh(w)\right),$$
$$z = \psi_2(w) = \frac{1}{2} + \frac{1}{2}\tanh\left(\frac{w}{2}\right),$$

where the interval $(-\infty, +\infty)$ is mapped onto $(0,1)$.

3 Sinc neural network (SNN)

In this part, to solve FOCP, we introduce the structure of the SNN. There are few works in the field of neural networks with the sinc activation function. SNN with a single input and a single output has been utilized for the approximation of functions with one variable by Elwasif and Fausett [11] in 1996. Suppose that $T = [t_1, t_2, \dots, t_n]^T$, and $O = [o_1, o_2, \dots, o_q]^T$ are the input and output vectors of the SNN, respectively. Then, the SNN can be structured as follows:

$$O = \sum_{i=1}^{m} W_i S_i(T), \tag{4}$$

where $W = [W_1, W_2, \dots, W_m]$ is a $q \times n$ matrix of weights between the nodes of $S_i(T)$ and outputs, and $S_i(T), 1 \leq i \leq m$, is the translated sinc basic function with input vector T. The structure of SNN can be observed in Figure 1.

The structure of SNN has been motivated from the sinc numerical method as a fascinating approach in solving nonlinear problems numerically. Sinc function is a famous and beneficial function in science and engineering. It is a smooth function with positive and negative amounts. It oscillates, and its output approaches zero when its input tends to infinity.

Unlike the MLP neural network, the structure of SNN is simple with a few number of trainable parameters. In fact, it contains only one layer of

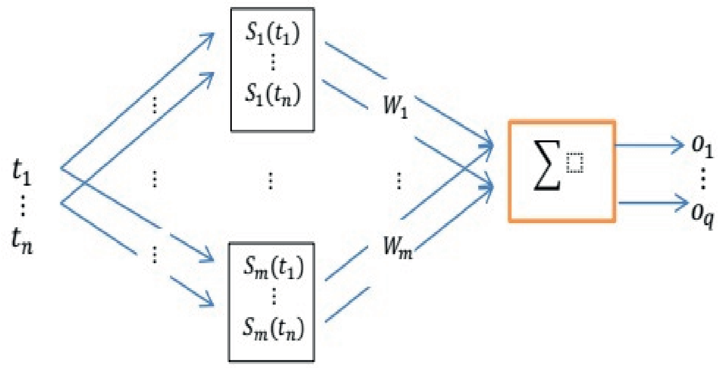


Figure 1: The structure of SNN.

trainable parameters. In SNN, the activation functions are the sinc basis functions, and therefore, the accuracy of modeling results increases when the number of sinc basis functions (nodes $S_i(T)$) increases in its structure.

To train the SNN, we can use various existing methods such as gradient methods, methods based on Lyapunov's stability theory, and so on. Since in these neural networks we only have adjustable parameters in one layer, the output of the neural network is linear with respect to the parameters, and this makes it easier to train it, and it becomes possible to use a variety of learning algorithms. In this article, based on the error functions obtained from the existing mathematical relations, an optimization problem is created, and by solving it, we get the optimal parameters of the neural network.

4 Applying SNN to solve FOCP

Before we resolve the problem, we approximate the RL fractional derivative with the GL operators. Assume $y \in c^l([a, b]), l \in N$. Then the l-order derivative of y at $t \in (a, b]$ can be defined as

$$ {}_aD_t^l y(t) = \lim_{h \to 0} \frac{1}{h^l} \sum_{k=0}^{l}, (-1)^k \binom{l}{k} y(t - kh), \tag{5}$$

where $\binom{l}{k} = \frac{l!}{k!(l-k)!}$ are the binomial coefficients. By replacing the real number α instead of l, binomial coefficients can be developed using the gamma function as

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt, \quad z > 0,$$

where $\Gamma(1) = 1$ and $\Gamma(z+1) = z\Gamma(z)$, for any $z > 0$. Now, the binomial coefficients are defined for the real number α as $\binom{\alpha}{k} = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha-k+1)}$. So, (5) can be extended to the fractional order $\alpha > 0$ by means of

$$_aD_t^\alpha y(t) = \lim_{N\to\infty} \frac{1}{h^\alpha} \sum_{k=0}^{N} w_k^{(\alpha)} y(t-kh), \tag{6}$$

where $h = \frac{t-a}{N}$ and $t \in (a, b]$. Besides, we have $w_k^{(\alpha)} = (-1)^k \binom{\alpha}{k}$, which can be evaluated by using the recurrence formula as follows:

$$w_0^{(\alpha)} = 1, \quad w_k^{(\alpha)} = \big(1 - \frac{\alpha+1}{k}\big) w_{k-1}^{(\alpha)}.$$

Note that for $h > 0$, relation (6) is left GL fractional derivative, and for $h < 0$, it is called the right GL fractional derivative.

Now, to solve problem (1), we define the Hamiltonian function as

$$H\big(x(t), u(t), p(t), t\big) = F\big(x(t), u(t), t\big) + p(t).G\big(x(t), u(t), t\big),$$

where $p(t) \in \mathbb{R}^q$ is the co-state vector. Assume that $x^*(t), u^*(t)$, and $p^*(t)$ are the optimal state, control, and co-state functions, respectively. Then, to minimize the objective function (1), a necessary condition for $u^*(t)$ is as follows:

$$H\big(x^*(t), u^*(t), p^*(t), t\big) \le H\big(x^*(t), u(t), p^*(t), t\big), \tag{7}$$

where $t \in [t_0, t_f]$, for all impossible controls. Equation (7) is called the PMP, where an optimal control must minimize the Hamiltonian function [15]. According to the PMP, if $x(t), u(t)$, and $p(t)$ are the optimal solutions in (7), then they must satisfy the following conditions:

$$\begin{cases} \dfrac{\partial H(x,u,p,t)}{\partial x} = {}_tD_{t_f}^\alpha p(t), \\ \dfrac{\partial H(x,u,t,p)}{\partial p} = {}_{t_0}D_t^\alpha x(t), \\ \dfrac{\partial H(x,u,p,t)}{\partial u} = 0. \end{cases} \tag{8}$$

We try to solve the system of equations (8) using the SNN. It can define the main trial solutions using the SNN (4), so that the initial or boundary conditions are satisfied, so they are constructed in the following form:

$$\begin{cases} x_T = x_0 + (t - t_0)o_x, \\ p_T = (t - t_f)o_p, \\ u_T = o_u, \end{cases} \tag{9}$$

where

$$o_x = \sum_{i=1}^{m} W_i^x S_i^x(T), \quad o_p = \sum_{i=1}^{m} W_i^p S_i^p(T), \quad o_x = \sum_{i=1}^{m} W_i^u S_i^u(T),$$

are state, co-state, and control neural networks, respectively. Note that since $x(t_f)$ is free, then $p(t_f) = 0$. By placing the trial solutions (9) in relation (8), we have

$$\begin{cases} \dfrac{\partial H_T}{\partial x_T} = {}_tD_{t_f}^{\alpha} p_T(t), \\ \dfrac{\partial H_T}{\partial p_T} = {}_{t_0}D_t^{\alpha} x_T(t), \\ \dfrac{\partial H_T}{\partial u_T} = 0, \end{cases} \tag{10}$$

where $H_T = H\big(x_T(t), u_T(t), p_T(t), t\big)$. To solve the system of equations (10) as an unconstrained minimization problem, we put them on the $m+1$ points of the interval $[t_0, t_f]$ as $t_k = t_0 + \frac{t_f - t_0}{m}k, k = 0, 1, \ldots, m$. According to the relation (5), the left and right RL fractional derivatives can be calculated as follows:

$$\begin{aligned} {}_{t_0}D_{t_k}^{\alpha} x_T &\simeq \frac{1}{h^\alpha} \sum_{j=0}^{k} w_j^{(\alpha)} x_T(t_{k-j}), \quad k = 1, 2, \ldots, m, \\ {}_{t_k}D_{t_f}^{\alpha} p_T &\simeq \frac{1}{h^\alpha} \sum_{j=0}^{m-k} w_j^{(\alpha)} p_T(t_{k+j}), \quad k = 0, 1, \ldots, m-1. \end{aligned}$$

Now, we construct the optimization problem as

$$\min E(\eta) = \frac{1}{2} \sum_{k=0}^{m} \{E_1(t_k, \eta) + E_2(t_k, \eta) + E_3(t_k, \eta)\}, \tag{11}$$

where $\eta = (w_x, w_p, w_u)$ and

$$\begin{cases} E_1(t_k,\eta) = \left[\dfrac{\partial H_T}{\partial x_T} - \frac{1}{h^\alpha}\sum_{j=0}^{m-k} w_j^{(\alpha)} p_T(t_{k+j})\right]^2, & k=0,1,\dots,m-1,\\ \\ E_2(t_k,\eta) = \left[\dfrac{\partial H_T}{\partial p_T} - \frac{1}{h^\alpha}\sum_{j=0}^{k} w_j^{(\alpha)} x_T(t_{k-j})\right]^2, & k=1,2,\dots,m,\\ \\ E_3(t_k,\eta) = \left[\dfrac{\partial H_T}{\partial u_T}\right]^2, & k=1,2,\dots,m. \end{cases}$$

Based on [13, Lemma 1], if η satisfies the system of equations (11), then η is an optimal solution of (1). So, one can verify that the minimization problem (11) is equivalent to the following problem:

$$\min_{\eta} E(\eta) = \frac{1}{2}||\phi(\eta)||^2. \tag{12}$$

Optimization algorithms such as steepest descent, Newton, quasi-Newton, and conjugate gradient can be used to solve the problem (12). The main advantage of the mentioned method is that it is not very complicated, and the accuracy of the approximation solution can be increased using more training points in the interval $[t_0, t_f]$.

Remark 1. In the present work, on the basis of sinc numerical methods and using the sinc basis functions, we propose the SNNs for the optimal control of fractional linear and nonlinear dynamic systems. Then, the results have been compared with MLP. In [13], the MLP is used for the optimal control of fractional dynamic systems. However, aside from the utilized neural structure, the main approach to solving FOCP is similar to [13].

In this work, we introduce the SNN as a new neural network with a simple structure for solving the FOCPs. Due to the beneficial properties of sinc function, as described in Section 3, the SNN is powerful in solving the problems. The simulation results show the capabilities of this neural network in solving FOCPs.

Remark 2. Generally, utilizing the neural networks to solve the fractional optimal control follows the presented approach in [13], and therefore, we can spread the presented theory for the convergence and stability of the method in that article to the present work.

5 Numerical examples

In order to show the efficiency and application of the proposed technique, three problems are solved in this section. In these numerical calculations, the fractional derivatives are approximated using the GL operators. In all the examples, five nodes are applied in the SNN to solve FOCP. To show the efficiency of the proposed method, we have compared this method with MLP and shown the error graph of each of these methods in the figures.

In this work, all numerical computations have been coded in MATLAB R2017b with 16GB RAM. In the proposed approach, to adjust the parameters of neural structures, we are faced with some unconstrained nonlinear optimization problems. These optimization problems have been solved using the "fminunc" function, and the optimization algorithm has been selected as "quasi-newton".

Example 1. As the first example, consider the FOCP as follows:

$$\begin{aligned}
&\min\ J = \tfrac{1}{2}\textstyle\int_0^1 \big(3x^2(t) + u^2(t)\big)dt,\\
&{}_0D_t^{\alpha} x(t) = -x(t) + u(t),\\
&x(0) = 0, \quad x(1) = 2.
\end{aligned}$$

The exact solution of this problem [13] in the case of $\alpha = 1$ is equal to

$$\begin{cases}
x(t) = \frac{2}{\sinh 2}\sinh(2t),\\
u(t) = \frac{2}{\sinh 2}\big(2\cosh(2t) + \sinh(2t)\big).
\end{cases}$$

The Hamiltonian function of this example is as follows:

$$H\big(x, u, p, t\big) = \frac{1}{2}\big(3x^2(t) + u^2(t)\big) + p(t)\big(-x(t) + u(t)\big).$$

According to relation (10), we have

$$\begin{cases}
{}_tD_1^{\alpha} p(t) = 3x(t) - p(t),\\
{}_0D_t^{\alpha} x(t) = -x(t) + u(t),\\
u(t) + p(t) = 0,\\
x(0) = 0, \quad x(1) = 2.
\end{cases}$$

Due to the initial conditions, the trial solutions are selected as follows:

$$\begin{cases} x_T = 2t + t(t-1)o_x, \\ p_T = o_p, \\ u_T = o_u. \end{cases}$$

Figure 2 shows the exact and approximate diagram of the state and control function. Also, the error graph of the proposed method and the MLP are compared.

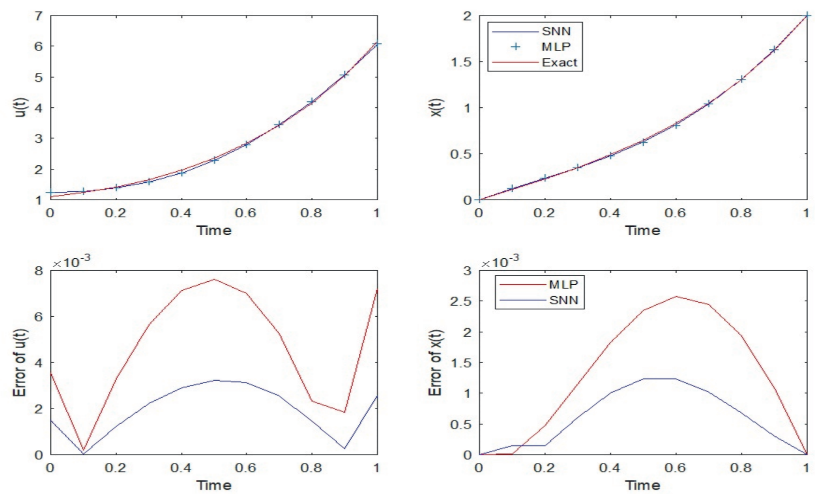


Figure 2: Exact and approximated state and control functions and their errors for Example 1.

In Table 1, we have expressed the error related to approximated state and control functions in the number of points and the process time in the SNN is 2.4 seconds, while it is 2.51 seconds in the MLP.

Example 2. Consider the following FOCP [23]:

$$\begin{aligned} &\min\ J = \tfrac{1}{2}\int_0^1 \left(x^2(t) + u^2(t)\right)dt, \\ &{}_0D_t^{\alpha} x(t) = -x(t) + u(t), \\ &x(0) = 1. \end{aligned}$$

The exact solution of this example with $\alpha = 1$ is

Table 1: Parameter values for Example 1

t	Error of $x(t)$ (SNN)	Error of $x(t)$ (MLP)	Error of $u(t)$ (SNN)	Error of $u(t)$ (MLP)
0.0	0	0	0.0023	0.0043
0.1	10^{-5}	10^{-5}	0.0004	0.0002
0.2	0.0001	0.0004	0.0012	0.0032
0.3	0.0004	0.0010	0.0025	0.0058
0.4	0.0006	0.0016	0.0034	0.0075
0.5	0.0009	0.0022	0.0039	0.0081
0.6	0.0011	0.0025	0.0037	0.0075
0.7	0.0011	0.0024	0.0029	0.0056
0.8	0.0010	0.0021	0.0012	0.0023
0.9	0.0006	0.0012	0.0014	0.0024
1.0	0	0	0.0050	0.0087

$$\begin{cases} x(t) &= \cosh(\sqrt{2}t) + \beta \sinh(\sqrt{2}t), \\ u(t) &= (1+\sqrt{2}\beta)\cosh(\sqrt{2}t) + (\sqrt{2}+\beta)\sinh(\sqrt{2}t), \end{cases}$$

where $\beta = -\frac{\cosh(\sqrt{2})+(\sqrt{2}\sinh(\sqrt{2})}{(\sqrt{2}\cosh(\sqrt{2})+\sinh(\sqrt{2})} \approx -0.98$.

Since $x(1)$ is free, we conclude $p(1) = 0$. So, due to relation (10), we have

$$\begin{cases} {}_tD_1^\alpha p(t) = x(t) - p(t), \\ {}_0D_t^\alpha x(t) = -x(t) + u(t), \\ u(t) + p(t) = 0, \\ x(0) = 1, \quad p(1) = 0. \end{cases}$$

From the initial conditions, the trial solution is written as

$$\begin{cases} x_T = 1 + to_x, \\ p_T = (t-1)o_p, \\ u_T = o_u. \end{cases}$$

Figure 3 shows the exact and approximate diagram of the state and control function, as well as the error diagram of the proposed method and MLP.

In Table 2, we have expressed the error related to approximated state and control functions in the number of points and the process time in the SNN is 3.14 seconds, while it is 3.3 seconds in the MLP.

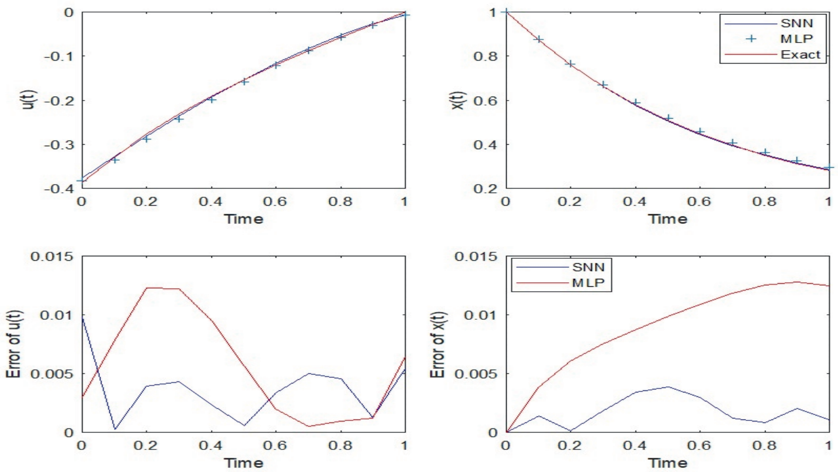


Figure 3: Exact and approximated state and control functions and their errors for Example 2.

Table 2: Parameter values for Example 2

t	Error of $x(t)$ (SNN)	Error of $x(t)$ (MLP)	Error of $u(t)$ (SNN)	Error of $u(t)$ (MLP)
0.0	0	0	0.0098	0.0030
0.1	0.0014	0.0038	0.0002	0.0078
0.2	0.0002	0.0061	0.0039	0.0123
0.3	0.0019	0.0075	0.0043	0.0122
0.4	0.0034	0.0087	0.0024	0.0095
0.5	0.0038	0.0098	0.0006	0.0056
0.6	0.0030	0.0109	0.0034	0.0020
0.7	0.0012	0.0119	0.0050	0.0005
0.8	0.0008	0.0125	0.0046	0.0009
0.9	0.0020	0.0128	0.0013	0.0012
1.0	0.0011	0.0125	0.0055	0.0065

Example 3. In this example, we consider the FOCP as [1]

$$\begin{aligned}
&\min J = \tfrac{1}{2}\int_0^1 \big(x^2(t) + u^2(t)\big)dt,\\
&{}_0D_t^{\alpha} x(t) = tx(t) + u(t),\\
&x(0) = 1.
\end{aligned}$$

According to (10), we have

$$\begin{cases} {}_tD_1^{\alpha}p(t) = x(t) + tp(t), \\ {}_0D_t^{\alpha}x(t) = tx(t) + u(t), \\ u(t) + p(t) = 0, \\ x(0) = 1, \quad p(1) = 0. \end{cases}$$

It is clear that $p(1) = 0$, because $x(1)$ is free. Also, we can select the trial solutions as

$$\begin{cases} x_T = 1 + to_x, \\ p_T = (t-1)o_p, \\ u_T = o_u. \end{cases}$$

This example does not have an exact solution; therefore, only approximate solutions for the control and state functions are shown in Figure 4. Also, approximate the state and control functions with different values of α are shown. The operation time is 3.5 seconds.

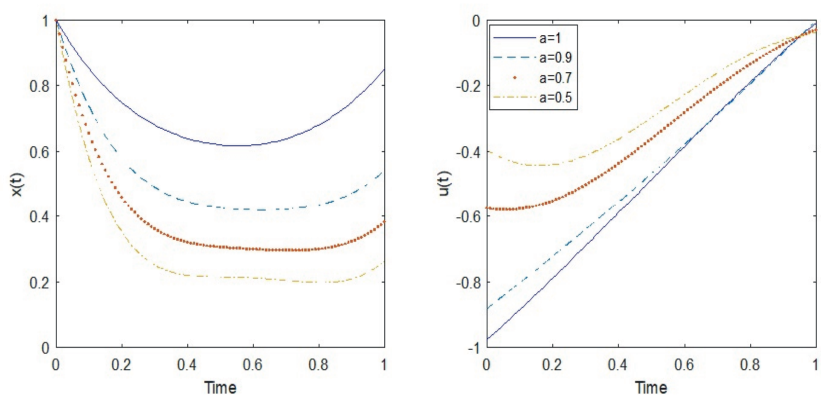


Figure 4: Approximated of the state and control functions by assuming different values of α for Example 3.

Example 4. Consider the nonlinear FOCP as follows [27]:

$$\begin{aligned} \min\ J(x,u) &= \tfrac{1}{2}\int_0^1 \big(0.625x^2(t) + 0.5x(t)u(t) + 0.5u^2(t)\big)dt, \\ &{}_0D_t^{\alpha}x(t) = 0.5x(t) + u(t), \\ &x(0) = 1. \end{aligned}$$

The exact solution of this problem in the case of $\alpha = 1$ is equal to

$$\begin{cases} x(t) = 0.99999 - 0.761547t + 0.499522t^2 - 0.124986t^3, \\ \qquad\quad +0.0379117t^4 - 0.00284627t^5, \\ u(t) = \frac{-\left(\tanh(1-t)+0.5\right)\cosh(1-t)}{\cosh(1)}. \end{cases}$$

Figure 5 shows the exact and approximate diagram of the state and control function. In Table 3, we have expressed the error related to approximated state and control functions in the number of points and the process time in the SNN is 2.71 seconds, while it is 2.98 seconds in the MLP.

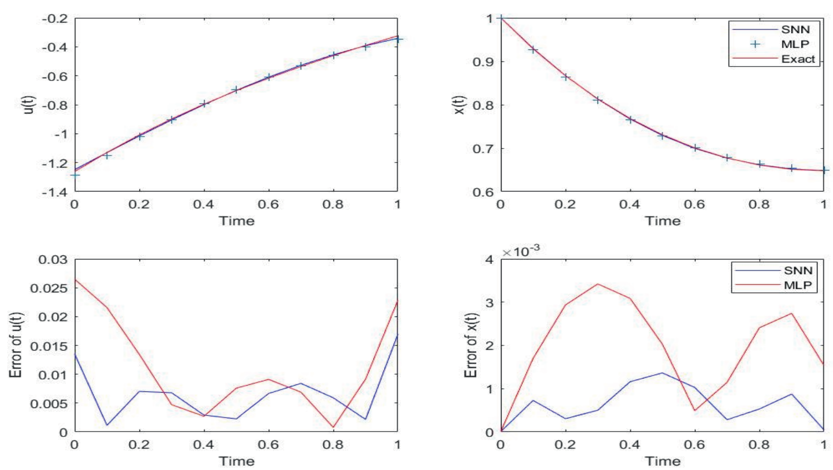


Figure 5: Exact and approximated state and control functions and their errors for Example 4.

6 Conclusion

In this work, based on the sinc numerical method, we presented a novel structure for the SNN and used this network to solve the FOCPs. Compared to MLP, this network has a simpler structure and fewer variables. The results obtained from the simulations showed that the method used is more efficient than the MLP. In general, ANNs have lower computational complexity. We can increase the accuracy of the trial solution by increasing the number of training points or different optimization algorithms. One of the important

Table 3: Parameter values for Example 4

t	Error of $x(t)$ (SNN)	Error of $x(t)$ (MLP)	Error of $u(t)$ (SNN)	Error of $u(t)$ (MLP)
0.0	0	0	0.0135	0.0265
0.1	0.0007	0.0017	0.0011	0.0216
0.2	0.0003	0.0029	0.0070	0.0134
0.3	0.0005	0.0034	0.0068	0.0047
0.4	0.0012	0.0031	0.0029	0.0027
0.5	0.0014	0.0020	0.0023	0.0076
0.6	0.0010	0.0005	0.0067	0.0091
0.7	0.0003	0.0011	0.0084	0.0069
0.8	0.0005	0.0024	0.0059	0.0008
0.9	0.0009	0.0027	0.0022	0.0092
1.0	0	0.0015	0.0169	0.0228

advantages of ANNs is that we can train the neural network at some points, but the final approximation solution can be calculated at each arbitrary point in the training interval. As a future work, the proposed method can be utilized to solve delay FOCPs.

Acknowledgements

Authors are grateful to the anonymous referees and editor for their constructive comments.

References

[1] Agrawal, O.M.P. *A general formulation and solution scheme for fractional optimal control problem*, Nonlinear Dyn. 38 (2004), 323–337.

[2] Agrawal, O.M.P. *A formulation and numerical scheme for fractional optimal control problems*, J. Vib. Control. 14 (2008), 1291–1299.

[3] Agrawal, O.M.P. and Baleanu, D. *A Hamiltonian formulation and a direct numerical scheme for fractional optimal control problems*, J. Vib. Control. 13 (2007), 1269–1281.

[4] Ahmad, H., Khan, M., Ahmad, I., Omri, M. and Alotaibi, M. *A meshless method for numerical solutions of linear and nonlinear time-fractional Black–Scholes models*, AIMS Math. 8 (2023), 19677–19698.

[5] Alquran, M., Sulaiman, T., Yusuf, A., Alshomrani, A. and Baleanu, D. *Nonautonomous lump-periodic and analytical solutions to the (3+1)-dimensional generalized Kadomtsev–Petviashvili equation*, Nonlinear Dyn. 111 (2023), 11429–11436.

[6] Ashpazzadeh, E., Lakestani, M. and Yildirim, A. *Biorthogonal multiwavelets on the interval for solving multidimensional fractional optimal control problems with inequality constraint*, Optim. Control Appl. Math. 41(5) (2020), 1477–1494.

[7] Biswas, R.K. and Sen, S. *Fractional optimal control problems with specified final time*, J. Comput. Nonlinear Dyn. 6(2) (2010), 021009.

[8] Chowdhury, D.R., Chatterjee, M. and Samanta, R.K. *An artificial neural network model for neonatal disease diagnosis*, International Journal of Artificial Intelligence and Expert Systems (IJAE), 2(3) (2011), 96–106.

[9] Dong, N.P., Long, H. and Khastan, A. *Optimal control of a fractional order model for granular SEIR epidemic with uncertainty*, Commun. Nonlinear Sci. Numer. Simul. 88 (2020), 105312.

[10] Effati, S. and Pakdaman, M. *Optimal control problem via neural networks*, Neural Comput. Appl. 23 (2013), 2093–2100.

[11] Elwasif, W. and Fausett, L.V. *Function approximation using a sinc neural network*, In: Proceedings of the SPIE, Volume 2760 (1996), 690–701.

[12] Ghasemi, S. and Nazemi, A.R. *A neural network method based on Mittag-Leffler function for solving a class of fractional optimal control problems*, AUT J. Model. Simul. 50(2) (2018), 211–218.

[13] Ghasemi, S., Nazemi, A. and Hosseinpour, S. *Nonlinear fractional optimal control problems with neural network and dynamic optimization schemes*, Nonlinear Dyn. 89 (2017), 2669–2682.

[14] Hashemi, M., Mirzazadeh, M. and Ahmad, H. *A reduction technique to solve the (2+1)-dimensional KdV equations with time local fractional derivatives*, Opt. Quantum Electron, 55(8) (2023), 721.

[15] Kirk, D.E. *Optimal control theory: An introduction*, Donver Publication, Inc. Mineola, New York, 2004.

[16] Lagaris, I.E. and Likas, A. *Hamilton–Jacobi theory over time scales and applications to linear-quadratic problems*, IEEE Trans Neural Netw. 9(5) (2012), 987–1000.

[17] Latif, S., Sabir, Z., Raja, M., Altamirano, G., Núñez, R., Gago, D., Sadat, R. and Ali, M. *IoT technology enabled stochastic computing paradigm for numerical simulation of heterogeneous mosquito model*, Multimed Tools Appl. 82 (2023), 18851–18866.

[18] Mohammadzadeh, E., Pariz, N., Hosseini sani, S.K. and Jajarmi, A. *Optimal Control for a Class of Nonlinear Fractional-Order Systems Using an Extended Modal Series Method and Linear Programming Strategy*, J. Control 10(1) (2016), 51–64.

[19] Nazemi, A. and Effati, S. *An application of a merit function for solving convex programming problems*, Comput. Ind. Eng. 66(2) (2013), 212–221.

[20] Nazemi, A. and Omidi, F. *A capable neural network model for solving the maximum flow problem*, J. Comput. Appl. Math. 236(14) (2012), 3498–3513.

[21] Nemati, S., Lima, P.M. and Torres, D.F.M. *A numerical approach for solving fractional optimal control problems using modified hat functions*, Commun. Nonlinear Sci. Numer. Simul. 78 (2019), 104849.

[22] Qayyum, M., Ahmad, E., Tauseef Saeed, S., Ahmad, H. and Askar, S. *Homotopy perturbation method-based soliton solutions of the time-fractional (2+1)-dimensionalWu–Zhang system describing long dispersive gravity water waves in the ocean*, Front Phys. 11 (2023), 1178154.

[23] Saberi Nik, H., Effati, S. and Yildirim, A. *Solution of linear optimal control systems by differential transform method*, Neural Comput. Appl. 23 (2013), 1311–1317.

[24] Sabir, Z. and Guirao, J. *A soft computing scaled conjugate gradient procedure for the fractional order Majnun and Layla romantic story*, Mathematics 11 (2023), 835.

[25] Sabouri, J., Effati, S. and Pakdaman, M. *A neural network approach for solving a class of fractional optimal control problems*,Neural Process. Lett. 45 (2017), 59–74.

[26] Shi, J., Sekar, B.D., Dong, M.C. and Hu, X.Y. *Extract knowledge from site-sampled data sets and fused hierarchical neural networks for detecting cardiovascular diseases*, 2012 International Conference on Biomedical Engineering and Biotechnology, (2012), 275–279.

[27] Singh, H., Pandey, R.K. and Kumar, D. *A reliable numerical approach for nonlinear fractional optimal control problems*, Int. J. Nonlinear Sci. Numer. Simul. 22(5) (2021), 495–507.

[28] Souayeh, B. and Sabir, Z. *Designing Hyperbolic Tangent Sigmoid Function for Solving theWilliamson Nanofluid Model*, Fractal Fract. 7 (2023), 350.

[29] Stenger, F. *A "Sinc-Galerkin" method of solution of boundary value problems*, J. Math. Comp. 33(145) (1979), 85–109.

[30] Stenger, F. *Sinc-related methods, Numerical Methods Based on Sinc and Analytic Functions*, Springer, New York, 1993.

[31] Stenger, F. *Summary of Sinc numerical methods*, J. Comput. Appl. Math. 121 (2000), 379–420.

[32] Sugihara, M. *Optimality of the double exponential formula–functional analysis approach–*, Numer. Math. (Heidelb.) 75 (1997), 379–395.

[33] Sugihara, M. *Near optimality of the sinc approximation*, Math. Comput. 72 (2003), 767–786.

[34] Sulaiman, T.A., Yusuf, A., Alshomrani, A.S. and Baleanu, D. *Wave solutions to the more general (2+1)-dimensional Boussinesq equation arising in ocean engineering*, Int. J. Mod. Phys. 2350214 (2023), 2350214.

[35] Taherpour, V., Nazari, M. and Nemati, A. *A new numerical Bernoulli polynomial method for solving fractional optimal control problems with vector components*, Comput. Method. Differ. Equ. 9(2) (2021), 446–466.

[36] Vrabie, D. and Lewis, F. *Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems*, Neural Netw. 22(3) (2009), 237–246.

[37] Zarin, R., Khan, M., Khan, A. and Yusuf, A. *Deterministic and fractional analysis of a newly developed dengue epidemic model*, Waves Random Complex Media, (2023), 1–34.

[38] Zil-E-Huma, Butt, A., Raza, N., Ahmad, H., Ozsahin, D. and Tchier, F. *Different solitary wave solutions and bilinear form for modified mixed-KDV equation*, Optik, 287 (2023), 171031.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1037–1068
https://doi.org/10.22067/ijnao.2024.86883.1392
https://ijnao.um.ac.ir/

Research Article

A space-time least-squares support vector regression scheme for inverse source problem of the time-fractional wave equation

A. Mohammadi and A. Tari Marzabad*

Abstract

The inverse problems in various fields of applied sciences and industrial design are concerned with the estimation of parameters that cannot be directly measured. In this work, we present a novel numerical approach for addressing the fractional inverse source problem by a machine learning algorithm and considering the ideas behind the spectral methods. The introduced algorithm utilizes a space-time Galerkin type of least-squares support vector regression to approximate the unknown source in a finite-dimensional space, providing a stable and efficient solution. With the proposed machine learning method, we overcome the limitations of classical

*Corresponding author

Received 15 February 2024; revised 8 June 2024; accepted 11 June 2024

Abumoslem Mohammadi
Department of Mathematics, Shahed University, Tehran, Iran. e-mail: abumoslem.mohammadi@shahed.ac.ir

Abolfazl Tari Marzabad
Department of Mathematics, Shahed University, Tehran, Iran. e-mail: tari@shahed.ac.ir

How to cite this article
Mohammadi, A. and Tari Marzabad, A., A space-time least-squares support vector regression scheme for inverse source problem of the time-fractional wave equation. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1037-1068. https://doi.org/10.22067/ijnao.2024.86883.1392

numerical methods and offer a promising alternative for tackling inverse source problems while avoiding overfitting by carefully selecting regularization parameters. To validate the effectiveness of our approach and illustrate an exponential convergence, we present some test problems along with the corresponding numerical results. The proposed method's superior accuracy compared to the existing methods is also illustrated.

AMS subject classifications (2020): Primary 35R11; Secondary 65M70, 65M22.

Keywords: Machine learning; Support vector machines; Inverse Source problem; Time fractional wave equation; Space-time Galerkin.

1 Introduction

Inverse problems are a significant area of research in various fields, including physics, engineering, geophysics, medical imaging, and many others. The fundamental nature of an inverse problem lies in determining the cause or source of an observed effect or measurement.

In the process of designing industrial system components, there are often instances where the desired future state of the system is known as a set of observations, and the aim is to calculate the underlying causes that result from those observations. As an illustration, consider certain diffusion processes where the final temperature, together with the initial and boundary conditions, is known. However, the value of the diffusion coefficient is yet to be determined. Likewise, in other scenarios, information regarding the vibrations of a rod or the voltage at a specific future time is provided, necessitating the determination of the associated source term. These two situations can be classified as the coefficient inverse problem and the inverse source problem, respectively.

Within the medical domain, the coefficient inverse problem plays a notable role in drug delivery applications. In this context, the proper selection of ingredients is vital to achieve a targeted final density of medicine within a predetermined time frame within the blood vessels. Inverse source problems appear in diverse domains, including crack analysis in nondestructive material evaluation [32], solving inverse heat equations to determine boundary conditions on inaccessible surfaces of a scramjet combustor [6], cell detection [7], models of light propagation in optical tomography across various scales [2], as well as determining absorption and diffusion coefficients in photo-acoustic imaging [27]. Additional instances include the retrieval of the adsorption-desorption source density rate, the exploration of elasticity problems [36], and the investigation of functionally graded materials [11].

Two important types of inverse source problems concern the wave and diffusion equations. In this work, we focus on the inverse damped wave problems and present a novel numerical simulation method based on Galerkin-type least-squares support vector regression (LS-SVR), which is a supervised machine learning algorithm.

The numerical solution of inverse problems that involve time-fractional wave equations has attracted considerable attention due to its wide range of applications in various scientific and engineering domains [3, 16, 25, 28]. Table 1 provides a short literature review of the recent works, focusing on the applications, methods, and key references in the field of inverse problems.

In recent years, support vector machines (SVM), support vector regression (SVR), and LS-SVR have emerged as powerful and versatile tools for solving differential and integral equations [18, 22, 23, 35], among others. These machine learning methods offer advantages such as high accuracy, robustness, and the ability to handle large datasets.

In this research article, our primary objective is to present a space-time Galerkin method that implements LS-SVR as a numerical solution for the inverse time-fractional wave equations. In the proposed algorithms, we aim to improve the efficiency and effectiveness of solving inverse problems in the context of time-fractional wave models. Through extensive numerical experiments and comparative analyses, we will demonstrate the superior performance of these methods in terms of accuracy, stability, and computational efficiency.

The time-fractional wave equations concerning the inverse source problems arise in reconstructing the source functions. They have found many other applications in engineering [3, 13, 25, 28, 17]. Here we focus on such equations. We denote the space domain by $\Lambda \subset \mathbb{R}^d$, which is assumed to be

Table 1: A literature review for the numerical methods for inverse problems. Let ICs and BCs stand for initial and boundary conditions.

inverse type	ref	method	applications
all types	[37]	*review work*	vehicle–bridge interaction dynamics
n-pixel image	[20]	machine learning	image reconstruction
source	[1]	deep learning	seismic inverse problem
ICs, BCs	[4]	adjoint method	heat transfer problem
source	[16]	finite element method	wave source reconstruction
source	[19]	real-time reconstruction	locations, and magnitudes of wave sources

a bounded Lipschitz domain, the time domain by $I = [0, T]$, the space-time domain by $\Omega = \Lambda \times I$, and the boundary of the space domain by $\partial\Lambda = \Gamma$. We consider the following inverse wave equation with damping:

$$tial_t^2 u + \tilde{g}(\partial_t^\alpha u) - \nabla \cdot (a\nabla u) + cu + K * u = f + \tilde{s}, \qquad \text{for all } (x,t) \in\Omega, \quad (1)$$

with the unknown source f. Here the operator $*$ denotes the convolution integral given by

$$(k * u)(t) = \int_0^t K(t-s)u(s)ds.$$

In this problem, $u,\ a,\ c$, and $\tilde{s}$ are spatiotemporal functions, $K = K(t)$, and $f = f(x)$. The damping term $\tilde{g}$ may be a nonlinear monotonic function, and the kernel K is assumed as a smooth function [12]. The linear and nonlinear cases of $\tilde{g}(x) = x$ or x^2 are discussed exclusively for the integer order $\alpha = 1$ in [29].

The problem (1) is equipped with the boundary and initial conditions

$$u(x,0) = h_0(x), \quad \text{for } x \in \Lambda, \quad (2)$$
$$u_t(x,0) = h_1(x), \quad \text{for } x \in \Lambda, \quad (3)$$
$$u|_\Gamma = 0, \quad \text{for } t \in I. \quad (4)$$

In this paper, the final observation is assumed as

$$u(x,T) = \tilde{\psi}_T(x). \quad (5)$$

The goal is to recover an unknown spatially distributed source $f = f(x) \in L^2(\Lambda)$, as well as the velocity $u \in H^1(\Omega)$.

Wave propagation in some complex media has special non-local and memory-dependent properties that are better captured in modeling with fractional derivatives. This approach allows for the incorporation of memory effects into wave equations, which leads to more accurate models [17]. For the inverse source problem (1)–(5), the uniqueness of the solution is discussed in [29].

Stability issues and challenges in inverse problems arise due to their ill-posed nature, sensitivity to noise, computational complexity, and uncertainties in the models. However, by employing regularization techniques, Bayesian inference, data filtering and preprocessing, careful model selection and validation, and proper regularization parameter selection, it is possible to mitigate these issues and obtain stable and reliable solutions to inverse problems. The proposed machine learning method in this work utilizes a Tikhonov regularization to overcome the stability issues. These aspects are discussed in the numerical results.

This paper introduces a space-time Galerkin LS-SVR approach with the orthogonal polynomial kernel that offers a unified framework for discretization, implementation, and regularization. It also demonstrates rapid convergence in both the spatial and temporal dimensions for solving the inverse time-fractional wave equation (1)–(4). The proposed method effectively addresses stability issues and achieves exponential convergence. During the training process, the solution is sought within a finite-dimensional space using simple orthogonal polynomial functions. Polynomial kernels in LS-SVR offer advantages in inverse source problems due to their ability to capture nonlinear relationships, computational efficiency, simplicity, interpretability, robustness to noise, and flexibility in model customization.

The main contribution of this work is to propose a space-time LS-SVR method, utilizing an orthogonal polynomial kernel for solving the inverse source problem of the time-fractional wave equation.

The paper is structured as follows. In Section 2, a concise overview of fractional derivatives, SVM for classification and regression, as well as approximation by orthogonal kernels is presented. Section 3 focuses on introducing our LS-SVR numerical method for simulating damped wave equations with memory. In Section 4, some test problems are provided to demonstrate the efficacy of the method. Finally, the paper concludes with some closing remarks.

2 Preliminaries

In this section, we present some preliminaries required in the next sections. The following subsections include the basic definitions of fractional calculus, function approximation with orthogonal polynomials, weighted residual methods, and SVM for classification and regression.

2.1 Modeling with fractional derivatives

Fractional derivatives provide a flexible tool for modeling certain phenomena in science and engineering offering greater compatibility with the experimental results compared to classical derivatives. It is now well established that the fractional diffusion and wave equations exhibit more realistic results and support various applications ranging from population dynamics to mathematical finance (e.g., as Black–Scholes equations). Due to the lack of analytical methods for fractional partial differential equations, developing numerical techniques is an ongoing research field.

Let $n-1 \leq \alpha < n,\ n \in \mathbb{N}$. The Caputo and Riemann–Liouville definitions for fractional derivative of order $\alpha > 0$, denoted by ${}_a^cD_t^{(\alpha)}f$ and ${}_a^{RL}D_t^{(\alpha)}f$, are, respectively, given by [8, 24]

$$ {}_a^cD_t^{(\alpha)}f(t) = \frac{1}{\Gamma(n-\alpha)} \int_a^t \frac{f^{(n)}(\tau)}{(t-\tau)^{\alpha-n+1}} d\tau, \tag{6} $$

$$ {}_a^{RL}D_t^{(\alpha)}f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{\partial}{\partial t} \int_a^t \frac{f(\tau)}{(t-\tau)^{\alpha-n+1}} d\tau, \tag{7} $$

where $n-1 \leq \alpha < n$ and $n \in \mathbb{N}$.

The right Riemann–Liouville derivative of order α is given by

$$ {}_{t}^{RL}D_{T}^{(\alpha)}f(t) = \frac{(-1)^n}{\Gamma(n-\alpha)}\frac{\partial}{\partial t}\int_t^T \frac{f(\tau)}{(\tau-t)^{\alpha-n+1}}d\tau. \tag{8} $$

2.2 LS-SVR and orthogonal kernels

Machine learning techniques, such as SVM and SVR, have gained widespread recognition within the research community and industry for their ability to classify data and predict patterns in large datasets. SVM addresses classification tasks, while SVR focuses on regression problems. Both algorithms formulate an optimization problem with inequality constraints to achieve their objectives. Similarly, the least squares variants, LS-SVM, and LS-SVR, convert these inequalities into equalities, simplifying the quadratic programming into a system of linear equations. This transformation allows for efficient training and easier solution computation. For more details, see, for example, [5, 9].

For a known data set (x_i, y_i), $i = 1, \ldots, N$, where the real numbers y_i's are the target value for the independent data $x_i \in R^{n_d}$, LS-SVR finds the wights w_i's and the bias $b \in \mathbb{R}$ in $y(x) = w^T\phi(x) + b$ with known functions ϕ_i, by solving the following optimization problem:

$$ \begin{aligned} &\min \ \frac{1}{2}w^T w + \frac{\gamma}{2}e^T e \\ &\text{s.t.} \quad y_i = w^T\phi(x_i) + b + e_i, \quad i = 1, \ldots, N. \end{aligned} \tag{9} $$

Here $\gamma \in \mathbb{R}^+$ denotes the tuning parameter $, w = [w_1, \ldots, w_M]^T, \phi = [\phi_1, \ldots, \phi_M]^T$. The hyperparameter γ controls the trade-off between the model complexity and the accuracy of the regression. It is used to penalize the deviation of the predictions from the actual target values. A smaller value of γ allows for more complex models that are able to capture fine details and noise in the data but may suffer from overfitting. In our work, in which an inverse fractional partial differential equation is considered without noise, we follow the same values as the recent publications [18, 22].

Utilizing the Lagrangian, which combines the objective function and a linear combination of the constraints using the dual variables α_i, this quadratic programming problem is transformed into a dual problem given by an equivalent linear system. The following theorem establishes the equivalence between the primal problem, represented as a quadratic programming formulation (9), and a linear system [18, 21, 33, 34].

Theorem 1. An equivalent form of the quadratic programming (9) is given by

$$\begin{bmatrix} W+\frac{1}{\gamma}I_N & \mathbf{1}_N \\ \mathbf{1}_N^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ b \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} \\ 0 \end{bmatrix}, \tag{10}$$

in which W is a positive definite matrix with the entries $W_{i,j} = \phi^T(x_i)\phi(x_j)$, I_N is the identity matrix, $\mathbf{1}_N^T = [1,\dots,1] \in \mathbb{R}^N$, $\boldsymbol{y} = [y_1,\dots,y_N]^T$, and the dual variables $\boldsymbol{\alpha} = [\alpha_1,\dots,\alpha_N]^T$. Then with the kernel $K(x,x_j) = \phi^T(x)\phi(x_j)$, we have

$$y(x) = \sum_{j=1}^{N} \alpha_j K(x,x_j) + b. \tag{11}$$

The kernel function computes similarities between pairs of data points in the input space. The purpose of the kernel in SVM is to transform the data into a higher-dimensional feature space, where it becomes easier to separate or make accurate predictions. The Mercer theorem states that a symmetric positive definite kernel function is necessary and sufficient for SVM to guarantee the existence of a corresponding feature space mapping [15].

To use orthogonal polynomials as kernels in SVR, we can define the kernel function as follows: $K(x,x') = \phi^T(x)\cdot\phi(x')$, where $\phi(x)$ is the feature map that transforms the input space into the space of orthogonal polynomials. We denote the set of polynomials with a degree less than or equal to M on the unit interval $[0,1]$ by $\mathcal{P}_M$. A basis for this set is given by Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ with $\alpha,\beta > -1$, which are orthogonal on $[-1,1]$, [10, 31]

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(x) P_m^{(\alpha,\beta)}(x) \chi^{(\alpha,\beta)}(x)dx = 0, \quad n \neq m, \tag{12}$$

with $\chi^{(\alpha,\beta)}(x) = (1-x)^{\alpha}(1+x)^{\beta}$. These functions result in fewer computations due to the sparsity of the involved matrices in the variational formulation of the method. The endpoint values are also given by

$$P_n^{(\alpha,\beta)}(-1) = (-1)^n \binom{n+\beta}{n}, P_n^{(\alpha,\beta)}(1) = \binom{n+\alpha}{n}, \tag{13}$$

which are used in handling boundary conditions for an efficient formulation of the proposed method [30].

A special case when $\alpha = 0, \beta = 0$, the Legendre polynomials are easily obtained by the recurrence formula

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad n \geq 1, \tag{14}$$

starting with $P_0(x) = 1, P_1(x) = x$. By (13), they have the boundary values $P_n(-1) = (-1)^n, P_n(1) = 1$.

We use the shifted Legendre polynomials that are orthogonal on the desired domain $[0, T]$ as

$$L_n(t) = P_n(\frac{2t}{T} - 1), \quad n = 0, 1, 2, \dots, \tag{15}$$

with

$$\int_I L_m(t)L_n(t)dt = \frac{T}{2n+1}\delta_{m,n}. \tag{16}$$

3 LS-SVR for the inverse time-fractional wave equation

In this section, we first recall the Sobolev spaces and the specific spaces required in our formulation and the associated norms [14, 26, 31]. Then, we present the variational formulation and a space-time Galerkin LS-SVR for (1)–(5).

3.1 Functional spaces

The Lebesgue spaces L^p, $p \geq 1$ on the domain Λ, denoted as $L^p(\Lambda)$, refers to the space of measurable functions u with pth power Lebesgue integrable on Λ, with the inner product. For $p = 2$, it is a Hilbert space with the inner product $(u, v) = \int_\Lambda uvd\Lambda$ endowed with the norm $\|u\|_2 = (u, u)^{1/2}$. On the other hand, the Sobolev spaces $W^{k,p}(\Lambda)$ consist of functions with weak derivatives up to order k that are also in $L^p(\Lambda)$. It can be mathematically expressed as

$$W^{k,p}(\Lambda) = \{u : D^s u \in L^p(\Lambda), |s| \leq k\},$$

where $D^s u$ represents the weak derivative of u with respect to the multi-index s. The associated Sobolev norm $\|u\|_{k,p}$ is given by $\|u\|_{k,p} = (\sum_{i=0}^{k} \|u^{(i)}\|_p^p)^{1/p}$. For the special case, where $p = 2$, $W^{k,2}(\Lambda)$ is a Hilbert space, denoted by $H^k(\Lambda)$. These functional spaces are defined on the spatial domain Λ. Also, these spaces on the temporal domain I are defined similarly. For a functional space X, the $x - t$ spaces for $s > 0$, are defined as

$$H^s(I;X) = \{v : \|v(\cdot,t)\|_X \in H^s(I)\},$$
$$H_0^s(I;X) = \{v : \|v(\cdot,t)\|_X \in H_0^s(I)\},$$

with the subscript zero as the Sobolev space with compact support.

Also, we define the Banach space $B^s(\Omega) = H^s(I; L^2(\Lambda)) \cap L^2(I; H_0^1(\Lambda))$ equipped with the norm

$$\|v\|_{B^s(\Omega)} = \|v\|_{H^s(I;L^2(\Lambda))} + \|v\|_{L^2(I;H_0^1(\Lambda))}.$$

Also, $C_0^\infty(I)$ refers to the smooth functions on I with compact support with the closure in terms of $\|\cdot\|_2$, $H_0^2(I)$.

Note that we first use a change of variable for the damped wave equation (1)–(5) as

$$v(x,t) = u(x,t) + h_0(x) + t h_1(x),$$

to make the initial conditions homogeneous. So we have

$$\partial_t^2 u + g(\partial_t^\alpha u) - \nabla \cdot (a \nabla u) + \beta u + K * u = f + s, \qquad \text{for all } (x,t) \in \Omega, \quad (17)$$

with the initial conditions

$$u(x,0) = 0, u_t(x,0) = 0, \qquad \text{for all } x \in \Lambda, \quad (18)$$

the BCs

$$u|_\Gamma = 0, \qquad \text{for all } t \in I, \quad (19)$$

and the final condition

$$u(x,T) = \psi_T(x) := \tilde{\psi}_T(x) - h_0(x) - th_1(x). \tag{20}$$

Now, we introduce the variational formulation of the problem (17)–(20), an easy-to-compute associated bilinear form as well as a space-time Galerkin LS-SVR for the numerical simulation of the problem.

3.2 LS-SVR with variational formulation of the inverse time-fractional wave equation

Here, we first present the LS-SVR for $\mathcal{L}u = f$ with a linear differential operator $\mathcal{L}$ and with some appropriate homogeneous boundary conditions. The approximate solution in dual variables by using a kernel function is assumed as

$$u(x) = \sum_{j=0}^{M} \alpha_j K(x, x_j), \tag{21}$$

with a kernel $K(\cdot,\cdot)$ in which the weights are given by the following problem:

$$\begin{aligned} &\min \frac{1}{2} w^T w + \frac{\gamma}{2} e^T e \\ &\text{s.t. } (w^T \mathcal{L}\phi(x), \psi_i) - (f(x), \psi_i) = e_i, \quad i = 0, \ldots, M, \end{aligned} \tag{22}$$

with the assumption that ϕ_i's satisfy the homogeneous boundary conditions so do u (21).

The quadratic programming (22) may be written as an equivalent linear system. This is given by the following theorem.

Theorem 2. The dual variables α_i's in (21) are given by

$$(W + \frac{1}{\gamma} I_M)\boldsymbol{\alpha} = \boldsymbol{b}, \tag{23}$$

with

$$\begin{aligned} &b_i = (f, \psi_i), \quad i = 0, \ldots, M, \\ &W_{k,i} = \sum_{j=0}^{M} (\mathcal{L}\phi_j, \psi_k)(\mathcal{L}\phi_j, \psi_i), \qquad k, i = 0, \ldots, M. \end{aligned} \tag{24}$$

Proof. We consider the Lagrangian as

$$L = \frac{1}{2}w^T w + \frac{\gamma}{2}e^T e - \sum_{i=0}^{M} \alpha_i((w^T \mathcal{L}\phi, \psi_i) - (f, \psi_i) - e_i).$$

Then, we have the following optimality conditions:

$$\frac{\partial L}{\partial w_k} = 0 \to w_k = \sum_{i=0}^{M} \alpha_i(\mathcal{L}\phi_k, \psi_i), \tag{25}$$

$$\frac{\partial L}{\partial e_k} = 0 \to \gamma e_k + \alpha_k = 0, \tag{26}$$

$$\frac{\partial L}{\partial \alpha_k} = 0 \to (w^T \mathcal{L}\phi, \psi_k) - (f, \psi_k) - e_k = 0, \tag{27}$$

for $k = 0, \ldots, M$. So, substituting e_k from (26) into (27), we have

$$\sum_{j=0}^{M} w_j(\mathcal{L}\phi_j, \psi_k) + \frac{1}{\gamma}\alpha_k = (f, \psi_k).$$

Now, by (25), we get

$$\sum_{j=0}^{M}\sum_{i=0}^{M} \alpha_i(\mathcal{L}\phi_j, \psi_i)(\mathcal{L}\phi_j, \psi_k) + \frac{1}{\gamma}\alpha_k = (f, \psi_k).$$

Using the Kronecker delta, this can be written as

$$\sum_{i=0}^{M} \alpha_i(\sum_{j=0}^{M}(\mathcal{L}\phi_j, \psi_k)(\mathcal{L}\phi_j, \psi_i) + \frac{1}{\gamma}\alpha_k\delta_{ki}) = (f, \psi_k),$$

which is the desired result. □

Note that the system (23) has a simpler form than (10) for known datasets since the constant function is involved in the basis functions.

Now, we present an algorithm for the numerical simulation of the fractional inverse source problem in the LS-SVR using orthogonal Legendre polynomials as the kernel. For the approximation in space $[0, \pi]$, and time interval $[0, T]$, we use combinations of $\phi_m(x) = P_m(\frac{2}{\pi}x - 1)$ and $\psi_n(t) = P_n(2t/T - 1)$, respectively, which satisfy the boundary conditions.

We consider the weak formulation of (17)–(20), on the space-time domain $\Omega = \Lambda \times I$:

$$(\partial_t^2 u, v)_\Omega + (g(\partial_t^\alpha u), v)_\Omega - (\nabla \cdot (a\nabla u), v)_\Omega + c(u, v)_\Omega + (K * u, v)_\Omega$$
$$= (f, v)_\Omega + (s, v)_\Omega, \qquad (28)$$

for all $v \in B^{\alpha/2}(\Omega)$. For $f \in L^2(\Omega)$, the problem is to find $u \in B^{\alpha/2}(\Omega)$ such that

$$\mathcal{A}(u, v) = \mathcal{F}(v), \qquad \text{for all } v \in B^{\alpha/2}(\Omega), \qquad (29)$$

in which the bilinear form $\mathcal{A}$ is a manipulation of (28) and the the functional $\mathcal{F}$ are given by

$$\begin{aligned}&\mathcal{A}(u, v) = -(\partial_t u, \partial_t v)_\Omega + (g(\partial_t^\alpha u), v)_\Omega + ((a\nabla u), \nabla v)_\Omega + c(u, v)_\Omega + (K * u, v)_\Omega,\\ &\mathcal{F}(v) = (f, v)_\Omega + (s, v)_\Omega. \end{aligned} \qquad (30)$$

We later consider a flexible margin with a tolerance for errors in the LS-SVR for the constraints (29). The bilinear map in the weak form characterizes the interactions between the unknown function and the derivatives using suitable test functions. The variational formulation, based on the bilinear form, leads to an algebraic system of equations enabling us to solve using numerical methods. On the other hand, the Lax–Milgram lemma ensures the existence and uniqueness of the solution as well as the continuous dependence of the solution on the given data when $\mathcal{A}$ in (30) is a continuous, bounded, and weakly coercive function, that is,

$$\sup_{\|v\|=1} |\mathcal{A}(u, v)| \geq c\|u\| \qquad \text{for all } u,$$
$$\sup_{\|u\|=1} |\mathcal{A}(u, v)| > 0, \qquad \text{for all } v \neq 0.$$

Note that the term $(\tilde{g}(\partial_t^\alpha u), v)_\Omega$ may be written in terms of the derivative of order $\alpha/2$ by using integration by parts, the definitions of the left and right Riemann–Liouville derivatives and some algebraic manipulations. We do this by following the work of [14, Lemmas 2.1 and 2.6].

Lemma 1. Let $s \in (0, 1), u(t) \in H^s(I), v(t) \in C_0^\infty(I)$. Then, the inner product of a fractional derivative with a function changes the order as

$$({}^{RL}\partial_t^\alpha u, v)_I = (u, {}_T\partial_t^\alpha v)_I. \qquad (31)$$

Proof. By (7) and integration by parts, we have

$$\begin{aligned}
({}^{RL}\partial_t^{\alpha} u, v)_I &= \int_0^T v\, {}^{RL}\partial_t^{\alpha} u dt \\
&= \frac{1}{\Gamma(1-\alpha)} \int_0^T \frac{d}{dt} \int_0^t \frac{u(x,\tau)}{(t-\tau)^{\alpha}} d\tau\, v(x,t) dt \\
&= \frac{v(x,t)}{\Gamma(1-\alpha)} \int_0^t \frac{u(x,\tau)}{(t-\tau)^{\alpha}} d\tau |_0^T - \frac{1}{\Gamma(1-\alpha)} \int_0^T \int_0^t \frac{u(x,\tau)}{(t-\tau)^{\alpha}} d\tau \partial_t v dt \\
&= -\frac{1}{\Gamma(1-\alpha)} \int_0^T \int_0^t \frac{u(x,\tau)}{(t-\tau)^{\alpha}} d\tau \partial_t v dt \\
&= -\frac{1}{\Gamma(1-\alpha)} \int_0^T \int_\tau^T \frac{\partial_t v}{(t-\tau)^{\alpha}} dt\, u(x,\tau) d\tau \\
&= -\frac{1}{\Gamma(1-\alpha)} \int_0^T (\frac{d}{d\tau} \int_\tau^T \frac{v}{(t-\tau)^{\alpha}} dt)\, u(x,\tau) d\tau \\
&= (u, {}_T\partial_t^{\alpha} v)_I .
\end{aligned}$$

□

We use (31) to get the following result.

Lemma 2. Let $0 < \alpha < 1$, $u \in {}_0H^1(I), v \in {}_0H^{\alpha/2}(I)$. Then,

$$(\partial_t^{\alpha} u, v)_I = (\partial_t^{\alpha/2} u, {}_t\partial_T^{\alpha/2} v)_I . \tag{32}$$

Proof. First, we consider a sequence $v_n \in C_0^{\infty}(I)$ such that

$$\|v_n - v\|_{H^{\alpha/2}(I)} \to 0, \quad n \to \infty.$$

By (31), we have

$$\begin{aligned}
(\partial_t^{\alpha} u, v_n)_I &= (\partial_t^{\alpha/2} \partial_t^{\alpha/2} u, v_n)_I \\
&= (\partial_t^{\alpha/2} u, {}_t\partial_T^{\alpha/2} v_n)_I .
\end{aligned}$$

By using the Schwarz inequality, we get

$$|(\partial_t^{\alpha} u, v_n)_I - (\partial_t^{\alpha} u, v)_I| \le \|\partial_t^{\alpha} u\| \|v - v_n\|.$$

This tends to zero as $n \to \infty$, which leads to (32).

Now, for a constant $a \in \mathbb{R}$, we have the standard weak formulation for the problem (17)–(20) as follows. Find $u \in B^{\alpha/2}(\Omega)$ with

$$-(\partial_t u, \partial_t v)_\Omega + \zeta(\partial_t^{\alpha/2} u, {}_t\partial_T^{\alpha/2} v)_\Omega + a(\partial_x u, \partial_x v)_\Omega + c(u,v)_\Omega + (K * u, v)_\Omega$$

$$= (f, v)_\Omega + (s, v)_\Omega, \tag{33}$$

for all v in the infinite-dimensional space $B^{\alpha/2}(\Omega)$. To be practical, we present an LS-SVR formulation by orthogonally projecting this form into trial and test spaces with finite dimensions. To discretize the problem (17)–(20) in the space-time domain, the approximate solution in space is sought in $P_M^0(\Lambda)$, and the space of polynomials on Λ with degree at most M having compact support on Λ. The approximate space in time includes the functions $v \in P_N$ with $v(t = 0) = 0$. The space-time trial space is written as

$$S_{M,N} = P_M^0(\Lambda) \otimes P_N^E(I). \tag{34}$$

For simplicity, let $S_L = S_{M,N}$ be the finite-dimensional space in both directions. $\square$

The LS-SVR for the problem (17)–(20) is as follows. By letting a tolerance in the bilinear form, we seek $u_L \in S_L$ by considering

$$\begin{aligned} &\min \ \frac{1}{2} w^T w + \frac{\gamma}{2} e^T e \\ &\text{s.t.} \quad \mathcal{A}(u_L, v_L) - \mathcal{F}(v_L) = e_i, \quad i = 0, \cdots, M, \quad \text{for all } v_L \in S_L, \end{aligned} \tag{35}$$

in which

$$u_L(x,t) = \sum_{m=0}^{M} \sum_{n=0}^{N} u_{m,n} \phi_m(x) \psi_n(t), \tag{36}$$

$$f_L(x) = \sum_{m=0}^{M} f_m \phi_m(x), \tag{37}$$

with the trial functions. To facilitate the computations in inner products and make use of the orthogonality in (35), the test functions are chosen as $v_L = \phi_i(x)\psi_j(t) \in S_L$. So, we have a sparse system of linear equations as constraints in (35), and we call the optimization problem (35) with $u_L \in S_L$ as (36), $f_L \in P_M^0(\Lambda)$ as (37), the space-time Galerkin LS-SVR for the inverse problem (17)–(20). The constraints in (35) are written as follows:

$$\begin{aligned} &-(\partial_t u_L, \partial_t v_L)_\Omega + \zeta(\partial_t^{\alpha/2} u_{L,t}\, \partial_T^{\alpha/2} v_L)_\Omega + a(\partial_x u_L, \partial_x v_L)_\Omega + c(u_L, v_L)_\Omega \\ &+(K * u_L, v_L)_\Omega = (f_L, v_L)_\Omega + (s, v_L)_\Omega + e. \end{aligned} \tag{38}$$

Note that there are alternatives such as collocating at some points and/or using other test functions. However, from a computational point of view, Galerkin LS-SVR (35) is preferred. We present analytical relations for computing the inner products in (38). The relation (38) is written in the form

$$\begin{aligned} -P^{0,0}UQ^{1,1} + \zeta P^{0,0}UQ^{\alpha/2,\alpha/2} + P^{0,0}UQ^{\alpha/2,\alpha/2} + aP^{1,1}UQ^{0,0} \\ + cP^{0,0}UQ^{0,0} + P^{0,0}UQ_c^{0,0} = F. \end{aligned}$$

In this relation, the matrices P, Q, U, and F are given by

$$\begin{aligned} P^{r,s} &= (\partial_x^r \phi, \partial_x^s \phi)_\Lambda, \\ Q^{r,s} &= (\partial_x^r \psi, \partial_x^s \psi)_I, \\ Q_c^{0,0} &= (\partial_x^r \psi, K * \partial_x^s \psi)_I, \\ U &= [u_{m,n} : 0 \le m, n \le M], \\ F &= (f_L + \tilde{s}, \phi_i(x)\psi_j(t))_\Omega. \end{aligned}$$

We treat the inner products in the first term analytically and the second term numerically for the sake of efficiency. The analytical part provides the use of orthogonality, which leads to sparse matrices for an efficient computational cost.

4 Numerical experiments

In this section, we consider some linear and nonlinear test problems to illustrate the convergence of the proposed space-time Galerkin LS-SVR algorithm for the inverse source problems. The purpose is to check the convergence behavior of the numerical solutions for increasing degrees of polynomials M and N and some fractional derivatives α. We plot the errors with three measures L^∞, L^2, and H^1-errors in semi-log scale to illustrate the exponential convergence. The figures include the convergence behavior in both spatial and temporal dimensions with a fixed M and N, respectively. The exponential convergence is confirmed since the logarithm of errors behaves as linear functions versus the polynomial degree. The L^∞, L^2, and H^1-errors are computed by

$$L^\infty(u_L) \approx \max_{0\le i,j\le \mathcal{M}} |U(x_i,t_j) - u_L(x_i,t_j)|, \tag{39}$$

$$L^2(u_L) \approx (\frac{1}{\mathcal{M}^2}\sum_{i,j=1}^{\mathcal{M}} |U(x_i,t_j) - u_L(x_i,t_j)|^2)^{1/2}, \tag{40}$$

$$H^1(u_L) \approx (L^2(u_L) + L^2(\partial_x u_L))^{\frac{1}{2}}, \tag{41}$$

with $x_i = \frac{i}{\mathcal{M}}$, $i = 0,\dots,\mathcal{M}$. The errors are computed on a grid of 101×101 equidistant points on the $x-t$ domain for all numerical examples. That is, we set the number of test points, $\mathcal{M} = 100$. For the errors of the source function f, we use similar measures but on space only. We use the experimental order of convergence (EOC) in time as

$$EOC_i = \frac{\ln(e_{i+1}/e_i)}{\ln(N_i/N_{i+1})},$$

in which e_i is the error with N_i as the number of training points, $\Delta t = 1/N_i$, $i = 1,2,\dots$ and a fixed number of spatial basis functions M. The EOC in space is computed in a similar manner.

Example 1. As the first example, we consider the linear case, $g(u) = u$, with the following data:

$$h_0 = 0, h_1 = \sin(x), u|_\Gamma = 0,\ a = 1,\ c = 2,\ K = -1,$$
$$\psi_T = \sin(x), s = 2\sin(x)\cos(t) + 2\sin(t)\cos(x),$$

on the domain $(x,t) \in [0,\pi] \times [0,\pi/2]$ with the exact solution $u(x,t) = \sin(t)\sin(x)$ and $f = -\sin(x)$. The problem (1)–(5) with these data was also considered for $\alpha = 1$ in [29]. With same parameters, the Landweber-type algorithm discussed in [29], gives the following errors with seven iterations

$$\|u_7(x,T) - \psi_T(x)\| \le 10^{-4},$$
$$\|f_7 - f\| \le 0.018.$$

The proposed scheme, that is, the space-time Galerkin LS-SVR algorithm, gives the following results after seven iterations

$$\|u_7(x,T) - \psi_T(x)\| \le 1.31e-05,$$
$$\|f_7 - f\| \le 2.27e-03.$$

We report the numerical errors as well as the convergence rates in terms of L^∞, L^2, and H^1 measures versus M with $N = 16$ and $\alpha = 1$, to demonstrate the spatial convergence of our proposed scheme for the wave and source functions in Tables 2 and 3, respectively. Running time is also reported for these results. It is worth noting that the results of Tables 2 and 3 are obtained at once, giving wave and source functions, respectively. Therefore, the time is reported only for one of these tables.

Table 2: L^∞, L^2, and H^1 errors and spatial convergence rates of the wave function versus M with $N = 16$ and $\alpha = 1$ for Example 1.

M	L^∞	EOC	L^2	EOC	H^1	EOC	time (s)
2	5.22e-02		2.33e-02		1.01e-01		2.01
4	1.44e-03	5.1799	7.16e-04	5.0242	5.29e-03	4.2549	10.61
6	1.98e-05	10.5723	9.12e-06	10.7610	1.34e-04	9.0655	29.53
8	1.48e-07	17.0196	6.99e-08	16.9324	2.68e-06	13.5984	63.94
10	1.34e-09	21.0830	5.10e-10	22.0504	3.07e-08	20.0289	113.37
12	1.86e-11	23.4600	8.21e-12	22.6471	2.92e-10	25.5333	194.41

The spectral convergence of the method is seen in Figures 1 and 2 in which a linear semi-log plot verifies the exponential convergence for the wave and source functions, respectively, for $\alpha = 1$.

Table 3: L^∞, L^2, and H^1 errors and spatial convergence rates for the source function f versus M with $N = 16$ and $\alpha = 1$ for Example 1 time is the same as 2.

M	$\|f - f_N\|_\infty$	EOC	$\|f - f_N\|_2$	EOC	$\|f - f_N\|_{H_1}$	EOC
2	5.04e-01		2.38e-01		6.06e-01	
4	5.78e-02	3.1243	2.72e-02	3.1293	1.35e-01	2.1664
6	2.64e-03	7.6115	1.18e-03	7.7385	9.95e-03	6.4314
8	6.95e-05	12.6432	3.11e-05	12.6392	3.34e-04	11.7984
10	1.06e-06	18.7460	4.73e-07	18.7586	6.07e-06	17.9606
12	1.08e-08	25.1560	4.83e-09	25.1438	4.59e-08	26.7914

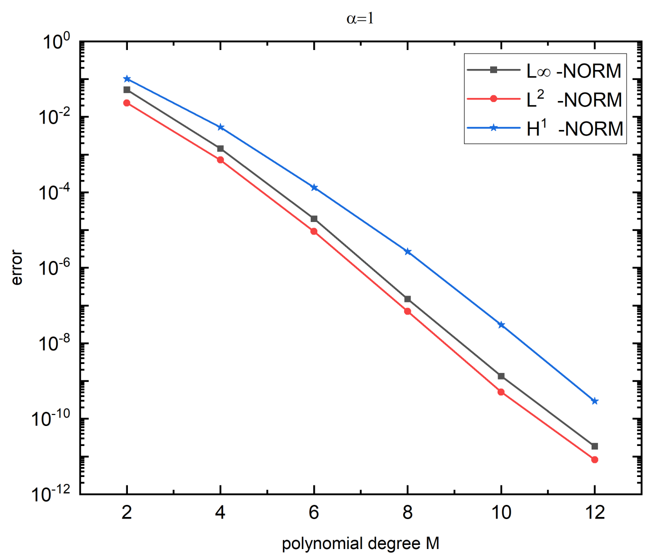


Figure 1: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 1$ for Example 1.

Tables 4, 5, and 6 demonstrate the spatial convergence of our proposed scheme as well as the running time for the wave function versus M with $N = 16$ for some fractional orders.

Table 4: L^∞, L^2, and H^1 errors and spatial convergence rates of the wave function versus M with $N = 16$ and $\alpha = 0.1$ for Example 1.

M	L_∞	EOC	L_2	EOC	H_1	EOC	time (s)
2	5.76e-02		3.03e-02		9.76e-02		2.06
4	1.61e-03	5.1609	7.61e-04	5.3153	5.15e-03	4.2442	10.98
6	7.34e-05	7.6161	3.48e-05	7.6086	1.41e-04	8.8737	26.01
8	4.61e-07	17.6246	2.10e-07	17.7636	2.71e-06	13.7367	55.78
10	3.03e-09	22.5184	1.24e-09	22.9986	3.12e-08	20.0063	162.14
12	2.87e-11	25.5561	1.46e-11	24.3627	2.75e-10	25.9509	257.39

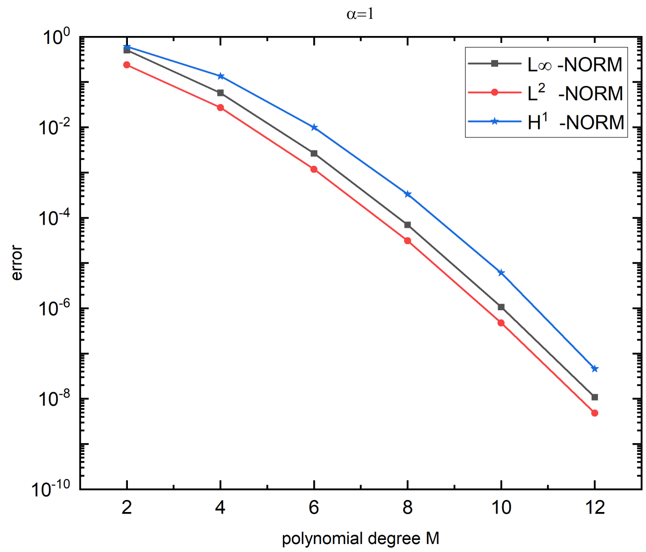


Figure 2: L^{∞}, L^2, and H^1 errors for the source function versus M with $N = 16$ and $\alpha = 1$ for Example 1.

The spectral convergence of the method is seen in Figures 3, 4, and 5, where a linear semi-log plot verifies the exponential convergence for the wave function for some fractional derivatives.

Table 5: L^{∞}, L^2, and H^1 errors and spatial convergence rates of the wave function versus M with $N = 16$ and $\alpha = 0.5$ for Example 1.

M	L^{∞}	EOC	L^2	EOC	H^1	EOC
2	5.22e-02		2.26e-02		9.72e-02	
4	1.55e-03	5.0737	7.48e-04	4.9171	5.21e-03	4.2216
6	4.23e-05	8.8817	1.95e-05	8.9946	1.30e-04	9.1026
8	3.00e-07	17.2022	1.36e-07	17.2604	2.67e-06	13.5061
10	2.25e-09	21.9269	8.78e-10	22.5987	3.09e-08	19.9830
12	2.47e-11	24.7468	1.27e-11	23.2339	2.80e-10	25.7990

Now, we investigate the effect of increasing the number of training points in the time direction. To do so, we fix the number of used polynomials in the

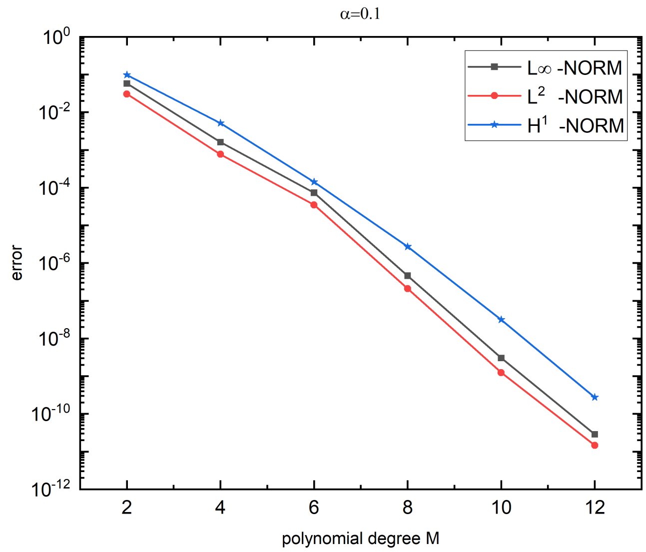


Figure 3: L^{∞}, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.1$ for Example 1.

Table 6: L^{∞}, L^2, and H^1 errors and spatial convergence rates of the wave function versus M with $N = 16$ and $\alpha = 0.9$ for Example 1.

M	L_∞	EOC	L_2	EOC	H_1	EOC
2	5.22e-02		2.27e-02		1.01e-01	
4	1.46e-03	5.1600	7.24e-04	4.9706	5.28e-03	4.2577
6	2.24e-05	10.3021	1.06e-05	10.4175	1.33e-04	9.0793
8	1.72e-07	16.9261	7.97e-08	16.9991	2.67e-06	13.5854
10	1.51e-09	21.2212	5.69e-10	22.1478	3.07e-08	20.0121
12	2.47e-11	24.7468	1.27e-11	23.2339	2.80e-10	25.7990

space dimension $M = 15$. We report the L^{∞}, L^2, and H^1 errors to show the temporal convergence of our proposed scheme for the wave function versus N and $\alpha = 1$, $\alpha = 0.5$ in Tables 7 and 8, respectively.

The temporal convergence of our proposed scheme for the wave function versus N with $M = 15$ and $\alpha = 1$, $\alpha = 0.5$ in Figures 6 and 7, respectively.

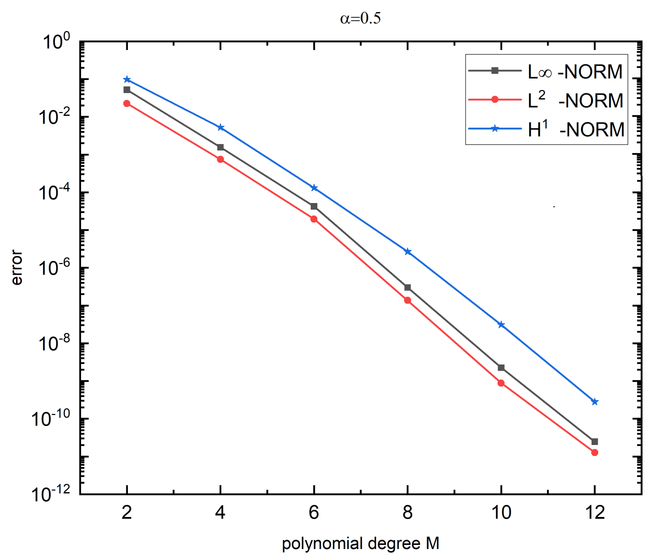


Figure 4: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.5$ for Example 1.

Table 7: L^∞, L^2, and H^1 errors and temporal convergence rates of the wave function versus N with $M = 15$ and $\alpha = 1$ for Example 1.

N	L_∞	EOC	L_2	EOC	H_1	EOC
2	6.85e-02		3.41e-02		4.83e-02	
4	1.94e-03	5.1420	8.92e-04	5.2566	1.60e-03	4.9159
6	7.69e-06	13.6400	3.24e-06	13.8554	5.08e-06	14.1873
8	2.26e-08	20.2645	1.21e-08	19.4316	1.63e-08	19.9592
10	6.32e-11	26.3481	3.28e-11	26.4876	4.55e-11	26.3562
12	2.26e-13	30.8988	1.06e-13	31.4540	1.41e-12	19.0549

The next example illustrates the convergence of the proposed method in space and time for the wave and source functions in a nonlinear case.

Example 2. Consider the problem (1) with the same parameters but the nonlinear function $g(u) = u^2$,

$$s = \sin^2(x)\cos^2(t) + 2\sin(t)\sin(x) + \cos(t)\sin(x),$$

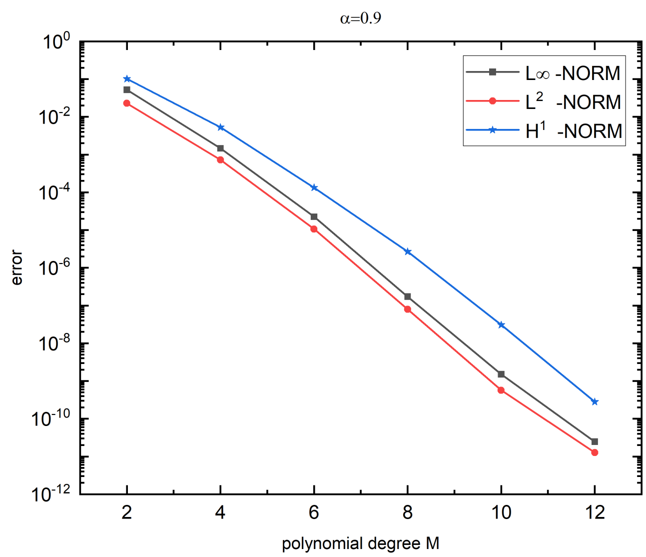


Figure 5: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.9$ for Example 1.

Table 8: L^∞, L^2, and H^1 errors and temporal convergence rates of the wave function versus N with $M = 15$ and $\alpha = 0.5$ for Example 1.

N	L^∞	EOC	L^2	EOC	H^1	EOC
2	6.85e-02		3.41e-02		4.83e-02	
4	2.38e-03	4.8471	1.11e-03	4.9411	1.72e-03	4.8115
6	9.07e-06	13.7370	3.86e-06	13.9628	5.89e-06	14.0008
8	2.94e-08	19.9238	1.53e-08	19.2246	2.16e-08	19.4948
10	8.05e-11	26.4426	4.25e-11	26.3781	5.83e-11	26.5069
12	4.18e-13	28.8530	1.88e-13	29.7322	1.49e-12	20.1119

on the same domain $[0, \pi] \times [0, \pi/2]$ as for Example 1 with the exact solution $u(x, t) = \sin(t)\sin(x)$ and $f(x) = -\sin(x)$. The problem (1)–(5) with these data was also considered for $\alpha = 1$ in [29]. The results (errors) obtained by the Landweber-type algorithm discussed in [29] with 13 iterations are

$$\|u_{13}(x, T) - \psi_T(x)\| \le 0.001,$$

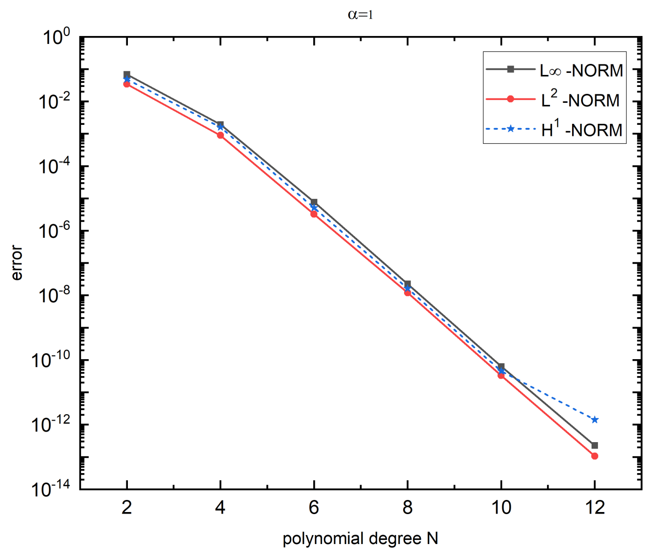


Figure 6: L^∞, L^2, and H^1 errors for the wave function versus N with $M = 15$ and $\alpha = 1$ for Example 1.

$$\|f_{13} - f\| \leq 0.02.$$

By the proposed scheme of this paper for this example with 13 iterations, the following results are obtained

$$\|u_{13}(x,T) - \psi_T(x)\| \leq 1.90e - 11,$$
$$\|f_{13} - f\| \leq 8.66e - 09.$$

We report the numerical errors as well as the convergence rates in terms of L^∞, L^2, and H^1 measures versus M with $N = 16$ and $\alpha = 1$ to demonstrate the spatial convergence of our proposed scheme for the wave and source functions in Tables 9 and 10, respectively.

The spectral convergence of the method is seen in Figures 8 and 9 in which linear semi-log plot verifies the exponential convergence for the wave and source functions, respectively, for $\alpha = 1$.

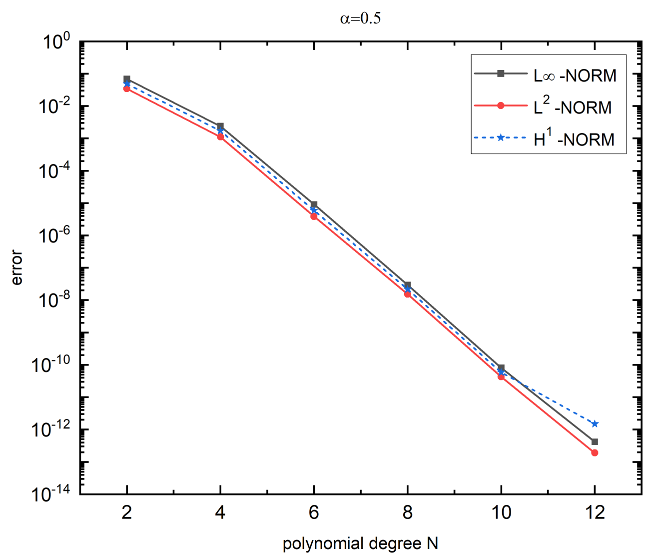


Figure 7: L^∞, L^2, and H^1 errors for the wave function versus N with $M = 15$ and $\alpha = 0.5$ for Example 1.

Table 9: L^∞, L^2, and H^1 norm errors of the wave function versus M with $N = 16$ and $\alpha = 1$ for the time-fractional wave equation using space-time Galerkin LS-SVR.

M	L^∞	EOC	L^2	EOC	H^1	EOC
2	5.84e-02		3.08e-02		9.77e-02	
4	1.62e-03	5.1719	7.56e-04	5.3484	5.11e-03	4.2570
6	7.13e-05	7.7030	3.48e-05	7.5923	1.41e-04	8.8545
8	3.92e-07	18.0873	1.77e-07	18.3579	2.62e-06	13.8541
10	2.40e-09	22.8364	9.66e-10	23.3515	3.07e-08	19.9274
12	2.45e-11	25.1454	1.24e-11	23.8889	2.74e-10	25.8822

Table 11 demonstrates the convergence in time of the proposed scheme for the wave function, versus N with $M = 15$ for $\alpha = 0.5$.

The spectral convergence in time of the method is seen in Figure 10 in which a linear semi-log plot verifies the exponential convergence for the wave function for $\alpha = 0.5$.

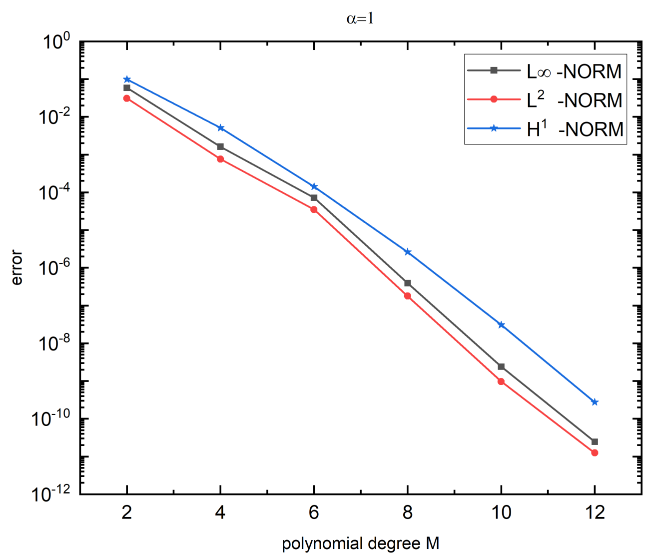


Figure 8: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 1$ for Example 2.

Table 10: L^∞, L^2, and H^1 errors and spatial convergence rates of the source function versus M with $N = 16$ and $\alpha = 1$ for Example 2.

M	$\|f - f_N\|_\infty$	EOC	$\|f - f_N\|_2$	EOC	$\|f - f_N\|_{H_1}$	EOC
2	4.76e-01		2.23e-01		5.12e-01	
4	5.48e-02	3.1187	2.55e-02	3.1285	1.24e-01	2.0458
6	2.63e-03	7.4894	1.20e-03	7.5379	1.07e-02	6.0425
8	6.95e-05	12.6300	3.12e-05	12.6864	3.51e-04	11.8784
10	1.06e-06	18.7460	4.73e-07	18.7730	6.08e-06	18.1757
12	1.08e-08	25.1560	4.83e-09	25.1438	5.05e-08	26.2766

As observed from the numerical results, the proposed LS-SVR, which employs orthogonal kernels with Legendre polynomials, yields approximate models with the desired accuracy and demonstrates favorable convergence behavior as the number of training points in both dimensions increases.

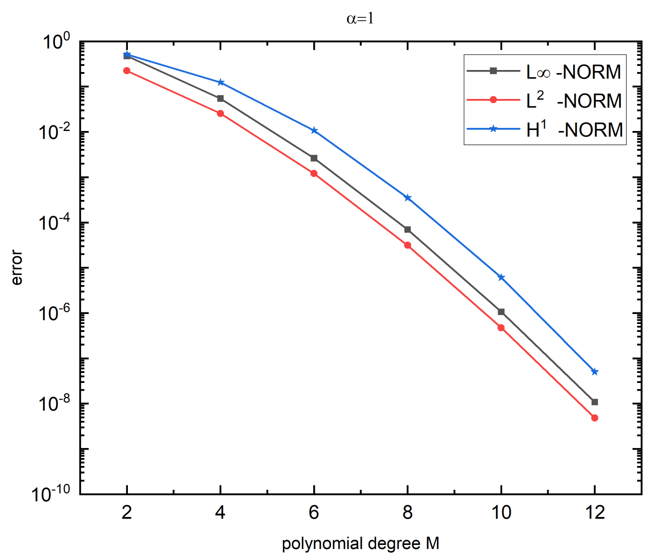


Figure 9: L^{∞}, L^2, and H^1 errors for the source function versus M with $N = 16$ and $\alpha = 1$ for Example 2.

Table 11: L^{∞}, L^2, and H^1 errors of the source function and temporal convergence rates versus N with $M = 15$ and $\alpha = 1$ for Example 2.

N	L^{∞}	EOC	L^2	EOC	H^1	EOC
2	6.85e-02		3.41e-02		4.83e-02	
4	2.40e-03	4.8350	1.13e-03	4.9154	1.67e-03	4.8541
6	9.55e-06	13.6305	4.23e-06	13.7811	6.45e-06	13.7040
8	3.25e-08	19.7546	1.65e-08	19.2803	2.34e-08	19.5323
10	8.37e-11	26.7171	4.59e-11	26.3716	6.30e-11	26.5182
12	5.87e-13	27.2045	2.83e-13	27.9110	1.66e-12	19.9445

5 Conclusion

In this study, we addressed an inverse source problem concerning the time-fractional wave equation. We presented a space-time Galerkin LS-SVR for the numerical simulation of the problem. We employed an orthogonal polynomial

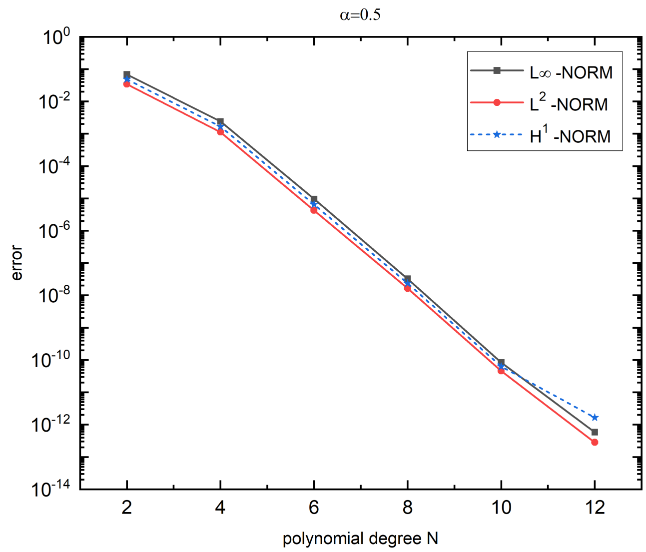


Figure 10: L^{∞}, L^2, and H^1 errors for the wave function versus N with $M = 15$ and $\alpha = 0.5$ for Example 2.

kernel in both space and time variables simultaneously. We presented the method as a quadratic programming problem and subsequently transformed it into a system of linear algebraic equations by introducing dual variables. The structure of the resulting system was also discussed, and some numerical test problems were given to demonstrate the efficacy of the proposed method. The method can be further developed for the reconstruction of sources in the presence of noisy data in future works as well as for other cases of inverse problems.

Acknowledgments

The authors are very grateful to the anonymous referees and the editor of this journal for carefully reading the paper and for their constructive comments and suggestions to improve the paper.

Funding: The authors did not receive support from any organization for the

submitted work.

Financial interests: The authors declare that they have no conflict of interest.

Availability of data and material: Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Code availability: Not applicable.

Authors' contributions: The order of the author list reflects contributions to the paper. All authors read and approved the final manuscript.

References

[1] Adler, A., Araya-Polo, M., and Poggio T., *Deep learning for seismic inverse problems: Toward the acceleration of geophysical analysis workflows,* IEEE Signal Process. Mag. 38(2) (2021), 89–119.

[2] Arridge, S.R. and Schotland, J.C., *Optical tomography: forward and inverse problems,* Inverse Probl. 25(12) (2009), 123010.

[3] Ashurov, R.R. and Faiziev, Y.É., *Inverse problem for finding the order of the fractional derivative in the wave equation,* Math.l Notes, 110(5-6) (2021), 842–852.

[4] Bourquin, F. and Nassiopoulos, A. *Inverse reconstruction of initial and boundary conditions of a heat transfer problem with accurate final state,* Int. J. Heat Mass Transf. 54(15-16) (2011), 3749–3760.

[5] Campbell, C. and Ying, Y. *Learning with support vector machines,* Springer Nature, 2022.

[6] Cui, M., Mei, J., Zhang, B-W, Xu, B.-B., Zhou, L., and Zhang, Y., *Inverse identification of boundary conditions in a scramjet combustor with a regenerative cooling system,* Appl. Therm. Eng. 134 (2018), 555–563.

[7] Del Aguila Pla, P. and Jaldén, J. *Cell detection by functional inverse diffusion and non-negative group sparsity—part ii: Proximal optimiza-*

tion and performance evaluation, IEEE Trans. Signal Process. 66(20) (2018), 5422–5437.

[8] Diethelm, K., *The analysis of fractional dierential equations: An application-oriented exposition using dierential operators of Caputo type,* Springer, 2010.

[9] Ernst, F. and Schweikard, A. *Fundamentals of Machine Learning: Support Vector Machines Made Easy,* Utb GmbH, 2020.

[10] Guo, B. *Spectral methods and their applications,* World Scientific, 1998.

[11] Hu, W., Gu, Y., and Fan, C.M., *A meshless collocation scheme for inverse heat conduction problem in three-dimensional functionally graded materials,* Eng. Anal. Bound. Elem. 114 (2020), 1–7.

[12] Jiang, D., Liu, Y., and Yamamoto, M., *Inverse source problem for a wave equation with final observation data,* Mathematical Analysis of Continuum Mechanics and Industrial Applications, 26 (2017), 153–164.

[13] Kinash, N. and Janno, J., *An inverse problem for a generalized fractional derivative with an application in reconstruction of time-and space-dependent sources in fractional diffusion and wave equations,* Mathematics, 7(12) (2019), 1138.

[14] Li, X. and Xu, C., *A space-time spectral method for the time-fractional diffusion equation,* SIAM J. Numer. Anal. 47(3) (2009), 2108–2131.

[15] Liu, Y. and Xie, M., *Rebooting data-driven soft-sensors in process industries: A review of kernel methods,* J. Process Control, 89 (2020), 58–73.

[16] Lloyd, S., Schaal, C. and Jeong, C., *Inverse modeling and experimental validation for reconstructing wave sources on a 2D solid from surficial measurement,* Ultrasonics, 128 (2023), 106880.

[17] Mainardi, F., *Fractional calculus in wave propagation problems,* In Forum der Berliner Mathematischer Gesellschaft, vol. 19, pp. 20-52. 2011.

[18] Mehrkanoon, S. and Suykens, J.A., *Learning solutions to partial differential equations using LS-SVM,* Neurocomputing, 159 (2015), 105–116.

[19] Ohe, T., *Real-time reconstruction of moving point/dipole wave sources from boundary measurements,* Inverse Probl. Sci. Eng. 28(8) (2020), 1057–1102.

[20] Ongie, G., Jalal, A., Metzler, C.A., Baraniuk, R.G., Dimakis, A.G., and Willett, R., *Deep learning techniques for inverse problems in imaging,* IEEE J. Sel. Areas Inf. Theory, 1(1) (2020), 39–56.

[21] Pakniyat, A., Parand, K., and Jani, M., *Least squares support vector regression for differential equations on unbounded domains,* Chaos Solit. Fractals. 151 (2021), 111232.

[22] Parand, K., Aghaei, A.A., Jani, M. and Sahleh, R., *Solving integral equations by ls-svr. In Learning with Fractional Orthogonal Kernel Classifiers in Support Vector Machines: Theory, Algorithms and Applications,* Singapore: Springer Nature Singapore, 2023.

[23] Parand, K., Aghaei, A.A., Kiani, S., Zadeh, T.I. and Khosravi, Z., *A neural network approach for solving nonlinear differential equations of Lane–Emden type,* Eng. Comput. (2023), 1–7.

[24] Podlubny, I., *Fractional Differential Equations,* Academic Press, New York, 1999.

[25] Qiu, L., Hu, C. and Qin, Q.H., *A novel homogenization function method for inverse source problem of nonlinear time-fractional wave equation,* Appl. Math. Lett.109 (2020), 106554.

[26] Reddy, B.D., *Introductory functional analysis: with applications to boundary value problems and finite elements,* Springer Science and Business Media, 2013.

[27] Ren, K. and Triki, F., *A Global stability estimate for the photo-acoustic inverse problem in layered media,* Eur. J. Appl. Math. 30(3) (2019), 505–528.

[28] Sakamoto, K. and Yamamoto, M., *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems,* J. Math. Anal. Appl. 382(1) (2011), 426–447.

[29] Seliga, L. and Slodicka, M., *An inverse source problem for a damped wave equation with memory,* J. Inverse Ill-Posed Probl. 24(2) (2016), 111–122.

[30] Shen, J. and Sheng, C.-T., *An efficient space–time method for time-fractional diffusion equation,* J. Sci. Comput. 81 (2019), 1088–1110.

[31] Shen J., Tang T., and Wang L.L., *Spectral methods: algorithms, analysis and applications,* Springer, 2011.

[32] Smith, R.C. and Demetriou, M.A., *Research directions in distributed parameter systems,* SIAM, 2003

[33] Suykens, J.A.K, Gestel, T.V., De Brabanter, J., De Moor, B., and Vandewalle, J., *Least squares support vector machines,* Singapore: World Scientific Publishing Company, 2002.

[34] Suykens, J.A.K. and Vandewalle, J., *Least squares support vector machine classifiers,* Neural Process. Lett. 9(3) (1999), 293–300.

[35] Taheri, T., Aghaei, A.A., and Parand, K., *Bridging machine learning and weighted residual methods for delay differential equations of fractional order,* Appl. Soft Comput. (2023) 110936.

[36] Yang, J.P. and Hsin, W.-C., *Weighted reproducing kernel collocation method based on error analysis for solving inverse elasticity problems,* Acta Mech. 230 (2019), 3477–3497.

[37] Zhu, X., and Law, S.-S., *Recent developments in inverse problems of vehicle–bridge interaction dynamics,* J. Civ. Struct. Health Monit. 6 (2016), 107–128.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1069–1105
https://doi.org/10.22067/ijnao.2024.85120.1331
https://ijnao.um.ac.ir/

Research Article

A semi-analytic and numerical approach to the fractional differential equations

K.N. Sachin, K. Suguntha Devi* and S. Kumbinarasaiah

Abstract

A class of linear and nonlinear fractional differential equations (FDEs) in the Caputo sense is considered and studied through two novel techniques called the Homotopy analysis method (HAM). A reliable approach is proposed for solving fractional order nonlinear ordinary differential equations, and the Haar wavelet technique (HWT) is a numerical approach for both integer and noninteger orders. Perturbation techniques are widely applied to gain analytic approximations of nonlinear equations. However, perturbation methods are essentially based on small physical parameters (called

*Corresponding author

Received 31 October 2023; revised 09 Juan 2024; accepted 18 Juan 2024

K.N. Sachin
Department of Mathematics, Bangalore University, Bengaluru-560056, India. e-mail: sachinnarayana05@gmail.com

K. Suguntha Devi
Department of Mathematics, Bangalore University, Bengaluru-560056, India. e-mail: suganthadevik77@gmail.com

S. Kumbinarasaiah
Department of Mathematics, Bangalore University, Bengaluru-560056, India. e-mail: kumbinarasaiah@gmail.com

How to cite this article
Sachin, K.N., Suguntha Devi, K. and Kumbinarasaiah, S., A semi-analytic and numerical approach to the fractional differential equations. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1069-1105. https://doi.org/10.22067/ijnao.2024.85120.1331

perturbation quantity), but unfortunately, many nonlinear problems have no such kind of small physical parameters at all. HAM overcomes this, and HWT does not require any parameters. Due to this, we opt for HAM and HWT to study FDEs. We have drawn a semi-analytical solution in terms of a series of polynomials and numerical solutions for FDEs. First, we solve the models by HAM by choosing the preferred control parameter. Second, HWT is considered. Through this technique, the operational matrix of integration is used to convert the given FDEs into a set of algebraic equation systems. Four problems are discussed using both techniques. Obtained results are expressed in graphs and tables. Results on convergence have been discussed in terms of theorems.

AMS subject classifications (2020): Primary 26A33; Secondary 34A08, 65H20, 90C30, 65T60.

Keywords: Homotopy analysis method; Haar wavelet; Convergence; Collocation method.

1 Introduction

Many engineering and scientific areas, including mathematical modeling in chemistry, physics, electrodynamics, and aerodynamics, involve fractional differential equations (FDEs). For instance, fractional derivatives can be used to describe nonlinear seismic oscillation, and they can also be used to solve the flaw in the continuum traffic flow assumption in the fluid-dynamic traffic model. Recent research has shown that FDEs are helpful tools for modeling various physical phenomena [29, 34]. This is due to the accurate representation of physical phenomena requiring the knowledge of both the present and the past, which can also be accomplished through the application of fractional calculus. Additionally, as the model contains memory terms, fractional derivatives offer an effective tool for explaining memory and hereditary qualities of different materials and actions [12]. This memory concept covers the past and how it affects the present and the future. Compared to integer-order, the fractional-order models are much more precise; this phenomenon has drawn the interest of numerous academicians [27, 17].

In the literature, we have yet to find an approachable analytical method for the general form of FDEs. So, we need to switch over to the semi-analytical or numerical methods.

Consider the general FDE of the following form:

$$aD^{\alpha}y(t) + bg(y(t)) + cy(t) = f(t) \tag{1}$$

with initial conditions

$$y(\beta_1) = \alpha_1, \quad y'(\beta_1) = \alpha_2,$$

or boundary conditions

$$y(\beta_2) = \alpha_3, \quad y'(\beta_2) = \alpha_4,$$

where $\alpha \in \mathbb{R}$ and a, b, c, β_1, β_2, α_1, α_2, α_3, and α_4 are real constants, $g(y(t))$ is nonlinear term, and $f(t)$ is continuous function, for all $t \geq 0$.

A method that yields the analytic solution after some iterations is called a semi-analytical method. Here, we considered one of the semi-analytic methods to solve FDEs called the Homotopy analysis method (HAM). The HAM creates a convergent series solution for nonlinear mathematical models by using the idea of the Homotopy in the topology. It is possible to use a Homotopy-Maclaurin series to handle the system's nonlinearities. In 1992, HAM was created for the first time in Shanghai Jiaotong University by Liao Shijun for his Ph.D. thesis [20]. Then, in 1997, it was modified by adding an auxiliary parameter [19] C_0 ($\neq 0$), known as the convergence-control parameter [21]. A nonphysical variable called the convergence-control parameter offers a straightforward approach to confirming and enforcing the convergence of a solution series. In analytical and semi-analytic techniques to nonlinear differential equations, it is unusual for the HAM to demonstrate the convergence naturally. First, unlike other series expansion techniques, the HAM does not rely on either small or large physical factors directly, which allows it to apply to both strongly and weakly nonlinear problems, overcoming some inherent limitations of the standard perturbation methods. Second, the HAM unifies the Adomian decomposition method (ADM), the delta expansion method, the Homotopy perturbation method, and the Lyapunov artificial small parameter approach [18, 30]. Strong convergence of the solution over broader geographical and parameter domains is frequently possible due to the method's enhanced generality. Third, the HAM offers super flexibility in the solution's representation and in the method by which it is expressly achieved. The basis functions of the intended solution and the related auxiliary linear operator of the Homotopy can both be chosen with a significant deal of freedom. Finally, the HAM offers a straightforward method to guarantee the convergence of the solution series, in contrast to the other analytic approximation methods. The Homotopy analysis approach can be used with other nonlinear differential equations methods, including spectral methods [25] and Padé approximants. To enable the linear method to handle nonlinear systems, it may also be supplemented with computational techniques like the boundary element method. The HAM, in contrast to the discrete computational approach of Homotopy continuation, is an analytic approximation method. In contrast, the HAM uses the Homotopy parameter only theoretically to show that a nonlinear system may be split into an infinite set of linear systems that are solved analytically. The HAM was created as an analytical approximation technique for the computer epoch with the intention of "computing with functions instead of numbers." Using the HAM in conjunction with a computer algebra system like Maple or Mathematica, we can quickly obtain analytic approximations of a higher-order strongly nonlinear problem. BVPh is a Mathematica package based on the HAM that has been made available online for solving nonlinear boundary-value problems as a result of the HAM's recent successes in several disciplines [1].

Solutions for the highly complex mathematical model by semi-analytical techniques require much time for each iteration, and sometimes even an infinite series of terms also cannot be possible to put in the analytical format. For such models, numerical techniques are better. In this study, we considered the Haar wavelet technique (HWT) to solve FDEs. When representing data or other functions, wavelets are mathematical functions that meet specific criteria. Since Joseph Fourier realized that sines and cosines could be superposed to describe other functions in the early 1800s, approximation utilizing the superposition of functions has been used. In the 1980s and 1990s, wavelets were created as an alternative to Fourier analysis of signals. Jean Morlet, Baroness Ingrid Daubechies, Alex Grossman, Palle Jorgensen, Yves Meyer, Ronald Coifman, Alfred Haar, and Stephane Mallat were a few key players in this invention. Yet, the scale at which we examine the data has a specific significance in wavelet analysis. Different scales or resolutions of data are processed by wavelet algorithms. Wavelet transforms are extremely helpful for signal analysis, compression, and denoising. Fourier analysis is impoverished at approximating sharp spikes when investigating its solutions; however, we can employ approximation functions that are tidily contained in finite domains appreciations to wavelet analysis. For estimating data with sharp discontinuities, wavelets work well. Compared to other methods [14, 13], the wavelet approach gives better outcomes. Numerical methods based on wavelets are effective for solving FDEs [36]. The Haar wavelets are compactly supported, and orthogonal with multi-resolution analysis [15]. Here we found several methods on FDEs, for instance, the reduced differential transform method [35], Sawi transform combined with the homotopy perturbation method [9], ADM [2], Legendre wavelet operational matrix method [3], fractional reduced differential transform method [37], fractional differential transform method [8] and positive solutions method [4], Adomian decomposition Shehu transform method [10], homotopy technique [28, 7], Adomian decomposition general transform method [5].

The paper is arranged as described as follows. Preliminaries of the Haar wavelets and their operational integration matrix, fractional derivative, and HAM are covered in Section 2. HAM and HWT have been explained in Section 3. Convergence analysis of HAM and HWT is drawn in Section 4. The implementation of test problems is in Section 5. The conclusion is enclosed with concluding remarks in Section 6.

2 Preliminaries

Definition 1 (Haar Wavelet)**.** A wavelet can be expressed as a real valued function $\Psi(t)$ that satisfies the following conditions [33]:

$$\int_{-\infty}^{\infty} \Psi(t)dt = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} |\Psi(t)|^2 dt = 1.$$

This means that $\Psi(t)$ is an oscillatory function with zero mean, and the wavelet function has unit energy. more precisely, wavelets are defined as

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi(\frac{t-b}{a}), \quad a \neq 0, b \in \mathbb{R}.$$

Here a and b represent, respectively, the dilation and translation. Consider an interval $[A, B] \subset \mathbb{R}$, which is divided into m equal subintervals, each of width $\Delta t = \frac{B-A}{m}$. The i^{th} orthogonal set of Haar functions on the interval $[A, B]$ is defined as

$$h_i(t) = \begin{cases} 1, & \zeta_1(i) \leq t < \zeta_2(i), \\ -1, & \zeta_2(i) \leq t < \zeta_3(i), \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where

$$\zeta_1(i) = A + \frac{k-1}{2^j} m\Delta t,$$
$$\zeta_2(i) = A + \frac{k-(\frac{1}{2})}{2^j} m\Delta t,$$
$$\zeta_3(i) = A + \frac{k}{2^j} m\Delta t,$$

for $i = 1, 2, \ldots, m, m = 2^J$ and J is a positive integer, which is called the maximum level of resolution. Here j and k represent the integer decomposition of the index i; that is, $i = k + 2^j - 1$, $0 \leq j < 1$, and $1 \leq k < 2^j + 1$.

Equation (2) is valid for $i \geq 2$, and for $i = 1$, we have

$$h_i(t) = \begin{cases} 1, & \text{for} \quad x \in [A, B], \\ 0, & \text{otherwise}. \end{cases} \tag{3}$$

The Haar wavelet operational matrix Q^α for integration of the general order α is given by

$$Q^\alpha H_m(t) = J^\alpha H_m(t) = [J^\alpha h_0(t), J^\alpha h_1(t), J^\alpha h_2(t), \ldots, J^\alpha h_{m-1}(t)],$$
$$Q^\alpha H_m(t) = [Qh_0(t), Qh_1(t), Qh_2(t), \ldots, Qh_{m-1}(t)], \tag{4}$$

where

$$Qh_i(t) = \begin{cases} 0 & A \leq t < \zeta_1(i), \\ \Phi_1 & \zeta_1(i) \leq t < \zeta_2(i), \\ \Phi_2 & \zeta_2(i) \leq t < \zeta_3(i), \\ \Phi_3 & \zeta_2(i) \leq t < B, \end{cases} \tag{5}$$

in which

$$\begin{aligned} \Phi_1 &= \frac{(t-\zeta_1(i))^\alpha}{\Gamma(\alpha+1)}, \\ \Phi_2 &= \frac{(t-\zeta_1(i))^\alpha}{\Gamma(\alpha+1)} - 2\frac{(t-\zeta_2(i))^\alpha}{\Gamma(\alpha+1)}, \\ \phi_3 &= \frac{(t-\zeta_1(i))^\alpha}{\Gamma(\alpha+1)} - 2\frac{(t-\zeta_2(i))^\alpha}{\Gamma(\alpha+1)} + \frac{(t-\zeta_3(i))^\alpha}{\Gamma(\alpha+1)}. \end{aligned}$$

Equation (5) is valid for $i \geq 1$. For $i = 0$, we have

$$Qh_0(t) = \begin{cases} \frac{t^\alpha}{\Gamma(\alpha+1)} & t \in [A, B], \\ 0 & \text{otherwise.} \end{cases}$$

For instance, if $\alpha \in \mathbb{R}$, then we have the following cases:

Case 1: For $\alpha = 1, J = 3$

$$Q^1 H_m(t) = \begin{bmatrix} 0.0625 & 0.1875 & 0.3125 & 0.4375 & 0.5625 & 0.6875 & 0.8125 & 0.9375 \\ 0.0625 & 0.1875 & 0.3125 & 0..4375 & 0.4375 & 0.3125 & 0.1875 & 0.0625 \\ 0.0625 & 0.1875 & 0.1875 & 0.0625 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0625 & 0.1875 & 0.1875 & 0.0625 \\ 0.0625 & 0.0625 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0625 & 0.0625 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0625 & 0.0625 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.0625 & 0.0625 \end{bmatrix}.$$

Case 2: For $\alpha = 2, J = 3$

$Q^2 H_m(t)$

$$
= \begin{bmatrix}
0.00195313 & 0.0175781 & 0.0488281 & 0.0957031 & 0.158203 & 0.236328 & 0.330078 & 0.439453 \\
0.00195313 & 0.0175781 & 0.0488281 & 0.0957031 & 0.154297 & 0.201172 & 0.232422 & 0.248047 \\
0.00195313 & 0.0175781 & 0.0449219 & 0.0605469 & 0.0625 & 0.0625 & 0.0625 & 0.0625 \\
0 & 0 & 0 & 0 & 0.00195313 & 0.0175781 & 0.0449219 & 0.0605469 \\
0.00195313 & 0.0136719 & 0.015625 & 0.015625 & 0.015625 & 0.015625 & 0.015625 & 0.015625 \\
0 & 0 & 0.00195313 & 0.0136719 & 0.015625 & 0.015625 & 0.015625 & 0.015625 \\
0 & 0 & 0 & 0 & 0.00195313 & 0.0136719 & 0.015625 & 0.015625 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.00195313 & 0.0136719
\end{bmatrix}.
$$

Case 3: Similarly, we obtain the operational matrix for $\alpha = 1.5, J = 3$

$$
Q^{1.5}H_m(t)
$$

$$
= \begin{bmatrix}
0.0117539 & 0.0610753 & 0.131413 & 0.217686 & 0.317357 & 0.428818 & 0.550933 & 0.682843 \\
0.0117539 & 0.0610753 & 0.131413 & 0.217686 & 0.293849 & 0.306667 & 0.288107 & 0.24747 \\
0.0117539 & 0.0610753 & 0.107905 & 0.0955356 & 0.0662843 & 0.0545208 & 0.047633 & 0.0428933 \\
0 & 0 & 0 & 0 & 0.0117539 & 0.0610753 & 0.107905 & 0.0955356 \\
0.117539 & 0.0375674 & 0.0210165 & 0.0159352 & 0.0133974 & 0.0117908 & 0.010654 & 0.00979445 \\
0 & 0 & 0.117539 & 0.0375674 & 0.0210165 & 0.0159352 & 0.0133974 & 0.0117908 \\
0 & 0 & 0 & 0 & 0.117539 & 0.0375674 & 0.0210165 & 0.0159352 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.117539 & 0.0375674
\end{bmatrix}.
$$

In the same way, we can generate the operational matrix of Haar wavelets for different values of α as per our requirements.

Definition 2. [32] A real function $f(x)$, for all $x > 0$, is said to be in the space $C_\mu, \mu \in \mathbb{R}$, if there exists a real number $p > \mu$, such that $f(x) = x^p f_1(x)$, where $f_1(x) \in C[0, \infty)$ and it is said to be in the space C_μ^n if and only if $f^{(n)}(x) \in C_\mu, n \in \mathbb{N}$.

Definition 3. [32] The fractional derivative of $f(x)$ in the Caputo sense is defined as

$$
D^\alpha f(x) = I^{n-\alpha} D^n f(x) = \frac{1}{\Gamma(n-\alpha)} \int_0^x (x-t)^{x-\alpha-1} f^n(t) dt, \tag{6}
$$

for $n - 1 < \alpha \leq n, n \in \mathbb{N}, x > 0, f \in C_{-1}^n$.

Definition 4. The fractional integration of x^n in the Caputo sense is given by

$$
I^\alpha x^n = \frac{x^{n+\alpha}\Gamma(1+n)}{\Gamma(1+n+\alpha)}, \tag{7}
$$

for $n - 1 < \alpha \leq n, n \in \mathbb{N}, x > 0$.

3 Method of solution

3.1 Homotopy analysis method

Consider the nonlinear FDEs with different physical conditions,

$$\mathscr{N}[y(t)] = 0, \quad t \geq 0, \tag{8}$$

where $\mathscr{N}$ represents the nonlinear differential operator, and $y(t)$ is the function to be determined.

Zeroth order deformation equation:
Let $y_0(t)$ be the initial approximation to the actual solution of (8). Liao constructed zeroth deformation equations taking the auxiliary function $\mathscr{H}(t)$ $(\neq 0)$ and auxiliary parameter $\hbar$ $(\neq 0)$ as [16],

$$(1-q)\mathscr{L}[\phi(t;q) - y_0(t)] = q\hbar\mathscr{H}(t)\mathscr{N}[\phi(t;q)], \tag{9}$$

where, $\phi(t;q)$ is a unknown function and $\mathscr{L}$ is a Linear operator.

When $q = 0$, (9) becomes, $\phi(t;0) = y_0\ (t)$ and at $q = 1$, (9) becomes, $\phi(t;1) = y(t)$. So as the q varies from 0 to 1, the function $\phi(t;q)$ varies from initial approximation $y_0\ (t)$ to the actual solution $y(t)$. Defining the mth order deformation derivatives,

$$y_m(t) = \frac{1}{m!}\frac{\partial^m \phi(t;q)}{\partial q^m}, \tag{10}$$

expanding $\phi(t;q)$, and using the Taylor series with respect to q, we get

$$\phi(t;q) = y_0(t) + \sum_{m=1}^{\infty} y_m(t)q^m. \tag{11}$$

As we know at $q = 1$ $\phi(t;q)$ becomes the required solution, (11) at $q = 1$ becomes

$$\phi(t;1) = y(t) = y_0(t) + \sum_{m=1}^{\infty} y_m(t). \tag{12}$$

Similarly, mth order deformation equation is given by

$$\mathscr{L}[y_m(t) - \chi_m y_{m-1}(t)] = \hbar\mathscr{H}(t)\mathscr{R}_m(y_{m-1}(t)), \tag{13}$$

where

$$\chi_m = \begin{cases} 0 & \text{if} \quad m \leq 1, \\ 1 & \text{otherwise}, \end{cases} \tag{14}$$

$$\mathscr{R}_m(y_{m-1}(t)) = \frac{1}{(m-1)!}\frac{\partial^{m-1}[\mathscr{N}[\phi(t;q)]]}{\partial q^{m-1}}. \tag{15}$$

Thus $y_1(t)$, $y_2(t)$, $y_3(t)$, $\ldots$ can be obtained from solving (13). The mth order approximation of $y(t)$ [21, 22, 23] is given by

$$y(t) = \sum_{m=0}^{m} y_m(t). \tag{16}$$

Equation (16) is the semi-analytical solution of (8).

NOTE: The above method is the same for solving fractional order differential equations. The inverse of the linear operator will be integration in a nonfractional order differential equation, whereas, for fractional order, it will be a fractional integration.

3.2 Haar wavelet method

Consider the general FDEs of the form

$$aD^{\alpha}y(t) + bg(y(t)) + cy(t) = f(t), \tag{17}$$

with initial conditions

$$y(\beta_1) = \alpha_1, \quad y'(\beta_1) = \alpha_2.$$

The Haar wavelet approximation is given as

$$D^{''}y(t) = \sum_{i=1}^{m} a_i H_m(t), \tag{18}$$

$$D^{'}y(t) = y^{'}(0) + \sum_{i=1}^{m} a_i Q^1 H_m(t),$$

$$D^{'}y(t) = \beta_2 + \sum_{i=1}^{m} a_i Q^1 H_m(t). \tag{19}$$

Integrating the above equation with respect to “t” from 0 to t, we get

$$y(t) = y(0) + \beta_2 t + \sum_{i=1}^{m} a_i Q^2 H_m(t),$$
$$y(t) = \beta_1 + \beta_2 t + \sum_{i=1}^{m} a_i Q^2 H_m(t). \tag{20}$$

Fractionally differentiate the above equation of order α, we obtain

$$D^\alpha y(t) = D^\alpha\{\beta_1 + \beta_2 t\} + \sum_{i=1}^{m} a_i Q^{2-\alpha} H_m(t), \tag{21}$$

where $Q^{2-\alpha}H_m(t)$ is the $(2-\alpha)$th order operational matrix of integration.

Substituting (18), (20), and (21) in (17), we get

$$\begin{aligned} &aD^\alpha\{\beta_1 + \beta_2 t\} + \sum_{i=1}^{m} a_i Q^{2-\alpha} H_m(t) + bg(\beta_1 + \beta_2 t + \sum_{i=1}^{m} a_i Q^2 H_m(t)) \\ &\quad + c(\beta_1 + \beta_2 t + \sum_{i=1}^{m} a_i Q^2 H_m(t)) = f(t). \end{aligned} \tag{22}$$

Collocate this equation by $t_l = A + (l - 0.5)(B - A)/m$, which yields the following equation:

$$\begin{aligned} &aD^\alpha\{\beta_1 + \beta_2 t_l\} + \sum_{i=1}^{m} a_i Q^{2-\alpha} H_m(t_l) + bg(\beta_1 + \beta_2 t_l + \sum_{i=1}^{m} a_i Q^2 H_m(t_l)) \\ &\quad + c(\beta_1 + \beta_2 t_l + \sum_{i=1}^{m} a_i Q^2 H_m(t_l)) = f(t_l). \end{aligned} \tag{23}$$

Solving the system of algebraic system equation (23) by the Newton–Raphson method gives the unknown coefficient values. Substituting these obtained unknown values in (20) will provide the wavelet numerical solution of (17).

4 Convergence analysis

Theorem 1. [21] As long as the series $y_0(t) + \sum_{m=1}^{\infty} y_m(t)$ converges, where $y_m(t)$ is governed by the higher order deformation equation number the χ_m given by (14), it must be the exact solution.

Theorem 2. [26] Let $\phi_0, \phi_1, \phi_2, \dots$ be the solution components of a given equation. The series solution $\sum_{k=0}^{\infty} \phi_k(t)$ converges if there exists $0 < \gamma < 1$ such that $||\phi_{k+1}|| \leq \gamma \|\phi_k\|$, for all $k \geq k_0$ for some $k_0 \in \mathbb{N}$.

Theorem 3. [26] Assume that the series solution $\sum_{k=0}^{\infty} \phi_k(t)$ is convergent to the solution $y(t)$. If the truncation series $\sum_{k=0}^{m} \phi_k(t)$ is used as an approximation to the solution $y(t)$, then the maximum absolute truncation error is estimated as, $\|y(t) - \sum_{k=0}^{m} \phi_k(t)\| \leq \frac{1}{1-\gamma} \gamma^{m+1} \|\phi_0(t)\|$.

Theorem 4. [6] Suppose that the functions $D_*^{\alpha} u_k(t)$ obtained by using Haar wavelets are the approximation of $D_*^{\alpha} u(t)$. Then we have an exact upper bound as follows:

$$\|D_*^{\alpha} u(t) - D_*^{\alpha} u_k(t)\|_E \leq \frac{M}{\Gamma(m-\alpha).(m-\alpha)} \frac{1}{[1-2^{2(\alpha-m)}]^{\frac{1}{2}}} \frac{1}{k^{m-\alpha}},$$

where $\|u(t)\|_E = (\int_0^1 u^2(t) dt)^{\frac{1}{2}}$.

5 Applications

Example 1. Consider the linear fractional model [24],

$$D^{\alpha} y(t) + y(t) = f(t), \quad t \geq 0, \quad and \quad 1 \leq \alpha \leq 2, \tag{24}$$

with the initial conditions

$$y^{(k)}(0) = 0, \quad k = 0, 1, \tag{25}$$

where $f(t) = te^{-t}$. The exact solution is $y(t) = -\frac{1}{2} e^{-t} (-1 - t + e^t \cos(t))$.

Method of implementation: Applying the HAM method to solve (24) and concerning (9), the zeroth order deformation is given by

$$(1-q)\mathscr{L}[\phi(t;q) - y_0(t)] = q\hbar \mathscr{H}(t)[D^{\alpha}\phi(t;q) + \phi(t;q) - f(t)]. \tag{26}$$

According to the condition (25), we choose the initial approximation as $y_0(t) = t^2$ and take the linear operator as $L = D^{\alpha}$ with $L(C_1 + C_2 t) = 0$, where C_1 and C_2 are integral constants with $\mathscr{H}(t) = 1$. Therefore, mth order deformation is given by

$$D^\alpha[y_m(t) - \chi_m y_{m-1}(t)] = \hbar \mathscr{R}_m(y_{m-1}(t)), \tag{27}$$

where

$$\mathscr{R}_m(y_{m-1}(t)) = \frac{1}{(m-1)!} \frac{\partial^{m-1}[D^\alpha \phi(t;q) + \phi(t;q) - f(t)]}{\partial q^{m-1}}, \tag{28}$$

subject to

$$y_m(0) = 0, \qquad y'_m(0) = 0. \tag{29}$$

Applying I^α, the inverse of D^α on either sides of (27), we get

$$y_m(t) = \chi_m y_{m-1}(t) + \hbar I^\alpha[\mathscr{R}_m(y_{m-1}(t))] + C_1 + C_2 t, \quad m \geq 1,$$

where C_1 and C_2 are calculated by using (29).

The HAM series solution upto first seven terms when $\alpha = 1.5$ with $\hbar = -1$ is given by

$$\begin{aligned} y(t) = {} & 0.306625t^{2.5} - 0.178204t^{3.5} - 0.0424592t^4 + 0.0607691t^{4.5} + 0.0172735t^5 \\ & - 0.0116399t^{5.5} - 0.00441781t^6 + 0.00192798t^{6.5} + 0.000664872t^7 \\ & - 0.000284779t^{7.5} - 0.0000894778t^8 + 0.0000471467t^{8.5} + 0.0000110138t^9 \\ & - 6.78818 \times 10^{-6}t^{9.5} - 1.54989 \times 10^{-6}t^{10} + 8.99097 \times 10^{-7}t^{10.5} \\ & + 2.42827 \times 10^{-7}t^{11} - 1.19467 \times 10^{-7}t^{11.5} - 2.23355 \times 10^{-8}t^{12} \\ & + 2.2627 \times 10^{-8}t^{12.5} + 2.62537 \times 10^{-9}t^{13} + 3.25559 \times 10^{-10}t^{13.5} \\ & - 4.43969 \times 10^{-10}t^{14} - 3.57732 \times 10^{-11}t^{14.5} - 5.74907 \times 10^{-12}t^{15} + \cdots . \end{aligned}$$

The HAM series solution upto first seven terms when $\alpha = 2$ with $\hbar = -1$ is given by

$$\begin{aligned} y(t) = {} & \frac{t^3}{6} - \frac{t^4}{12} + \frac{t^5}{60} - \frac{t^6}{360} + \frac{t^7}{1680} - \frac{t^8}{10080} + \frac{t^9}{90720} - \frac{t^{10}}{907200} + \frac{t^{11}}{7983360} \\ & - \frac{t^{12}}{79833600} + \frac{23t^{13}}{1037836800} + \frac{t^{14}}{10897286400} - \frac{139t^{15}}{1307674368000} - \cdots . \end{aligned}$$

Also, the above problem was solved through HWT and ND solvers. The results obtained are explained through the tables and graphs. Figure 1 shows the geometrical comparison of exact, HAM, and HTM solutions. Figure 2 reflects the graphical presentation of the HAM solution with different α values. The error analysis of HAM, HWT, and ND solver solutions with the exact solution is shown in Figure 3. Numerical values for integer values and noninteger values of α have been given in Tables 1 and 2, respectively.

Table 1: Comparison between HAM and HWT solutions for different fractional values of α in Example 1

t	$\alpha = 1.2$		$\alpha = 1.5$		$\alpha = 1.7$	
	HAM	HWT	HAM	HWT	HAM	HWT
0.0	0	0	0	0	0	0
0.1	0.002456	0.002359	0.000911	0.000872	0.000457	0.00047
0.2	0.010388	0.010175	0.004826	0.000872	0.002802	0.002739
0.3	0.023282	0.0228757	0.012430	0.012118	0.007889	0.007729
0.4	0.040204	0.039558	0.023801	0.023309	0.016138	0.015862
0.5	0.060161	0.059252	0.038723	0.0380337	0.027696	0.027294
0.6	0.082198	0.081026	0.056803	0.0559103	0.042517	0.041987
0.7	0.105453	0.104031	0.077547	0.076456	0.060412	0.059758
0.8	0.129169	0.127521	0.100408	0.099131	0.081087	0.080318
0.9	0.152700	0.150860	0.124818	0.0123377	0.104172	0.103302

Table 2: Comparison of solutions obtained from ND Solver, HAM, HWT and their absolute errors (AE) with exact solution for integer order $\alpha = 2$ of Example 1

t	Exact	ND Solver	HAM	HWT	ND Solver Error	HAM Error	HWT Error
0.0	0	0	0	0	0	0	0
0.1	0.000159	0.000159	0.000159	0.000189	4.72×10^{-9}	4.7×10^{-9}	3.09×10^{-5}
0.2	0.001205	0.001205	0.001205	0.001264	1.61×10^{-8}	1.61×10^{-8}	5.88×10^{-5}
0.3	0.003863	0.003864	0.003864	0.003947	6.12×10^{-9}	6.1×10^{-9}	8.32×10^{-5}
0.4	0.008694	0.008694	0.008694	0.008798	6.28×10^{-9}	6.3×10^{-9}	1.04×10^{-4}
0.5	0.016107	0.016107	0.016107	0.016229	7.26×10^{-9}	7.3×10^{-9}	1.22×10^{-4}
0.6	0.026382	0.026382	0.026382	0.026518	8.18×10^{-9}	8.2×10^{-9}	1.37×10^{-4}
0.7	0.039676	0.039677	0.039676	0.039825	1.30×10^{-8}	1.28×10^{-8}	1.49×10^{-4}
0.8	0.056043	0.056043	0.056043	0.056200	1.54×10^{-8}	1.42×10^{-8}	1.57×10^{-4}
0.9	0.075436	0.075436	0.075436	0.075599	1.70×10^{-8}	1.16×10^{-8}	1.63×10^{-4}

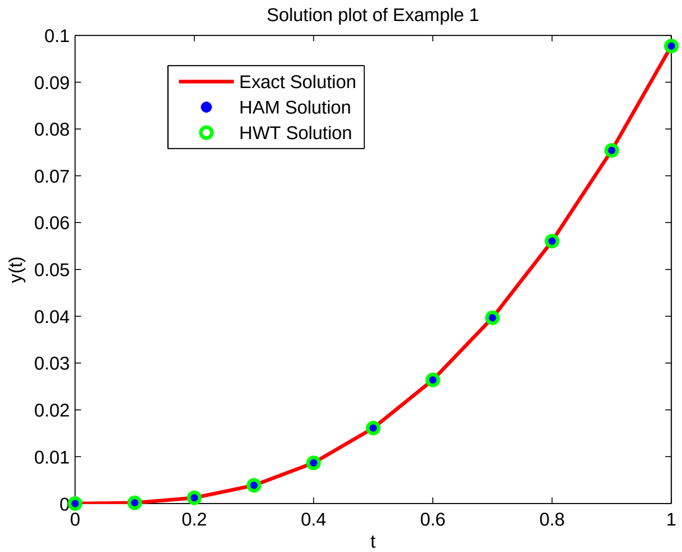


Figure 1: Comparison of the exact solution with HAM and HWT solutions, for Example 1.

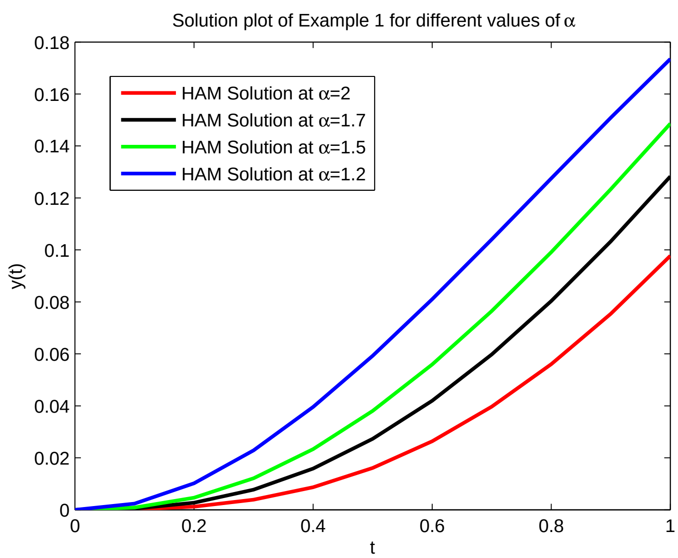


Figure 2: Graphical interpretation of the HAM solutions at different α values for Example 1.

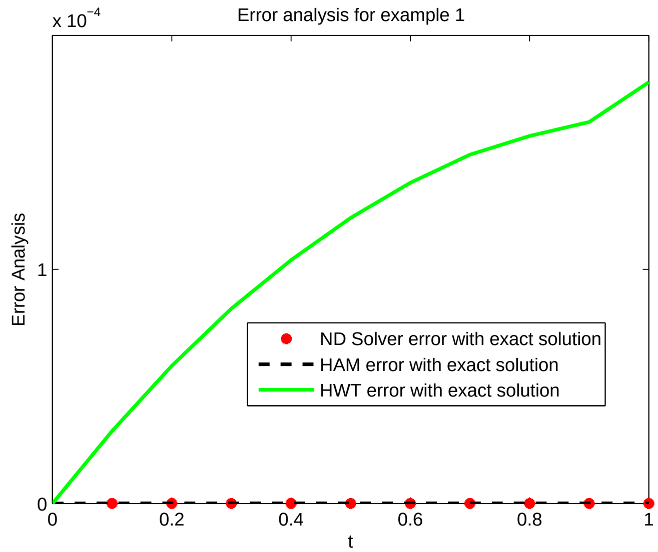


Figure 3: Graphical interpretation of the error analysis, for Example 1.

Example 2. Consider the nonlinear fractional model [24],

$$D^{\alpha}y(t) + y^2(t) = f(t), \quad 1 \leq \alpha \leq 2, \tag{30}$$

subject to the conditions

$$y(0) = 0, \quad y'(0) = 0, \tag{31}$$

where $f(t) = \frac{\Gamma(6)}{\Gamma(6-\alpha)}t^{5-\alpha} - \frac{3\Gamma(5)}{\Gamma(5-\alpha)}t^{4-\alpha} + \frac{2\Gamma(4)}{\Gamma(4-\alpha)}t^{3-\alpha} + (t^5 - 3t^4 + 2t^3)^2$. The exact solution is [24] $y(t) = t^5 - 3t^4 + 2t^3$.

Method of implementation: Applying HAM method to solve (30) and concerning (9), the zeroth order deformation is given by

$$(1-q)\mathscr{L}[\phi(t;q) - y_0(t)] = q\hbar\mathscr{H}(t)[D^{\alpha}\phi(t;q) + \phi^2(t;q) - f(t)]. \tag{32}$$

According to the condition (31), choosing the initial approximation $y_0\,(t) = \frac{t^{\alpha}}{\Gamma(\alpha+1)}$ and the linear operator as $\mathscr{L} = D^{\alpha}$ with $\mathscr{L}(C_1 + C_2t) = 0$, where C_1 and C_2 are integral constants with $\mathscr{H}(t) = 1$, therefore, the mth order deformation is given by

$$D^{\alpha}[y_m(t) - \chi_m y_{m-1}(t)] = \hbar \mathscr{R}_m(y_{m-1}(t)), \tag{33}$$

where

$$\mathscr{R}_m(y_{m-1}(t)) = \frac{1}{(m-1)!} \frac{\partial^{m-1}[D^{\alpha}\phi(t;q) + \phi^2(t;q) - f(t)]}{\partial q^{m-1}}, \tag{34}$$

subject to

$$y_m(0) = 0, \qquad y'_m(0) = 0. \tag{35}$$

Applying I^{α}, the inverse of D^{α} on either sides of (33), we get

$$y_m(t) = \chi_m y_{m-1}(t) + \hbar I^{\alpha}[\mathscr{R}_m(y_{m-1}(t))] + C_1 + C_2 t, \quad m \geq 1,$$

where C_1 and C_2 are calculated by using (35).

The HAM series upto first six terms when $\alpha = 1.5$ and taking $\hbar = -1$ is given by

$$\begin{aligned} y(t) = {} & 2t^3 - 3t^4 - 5.55112 \times 10^{-17} t^{4.5} + 1t^5 + 1.11022 \times 10^{-16} t^{8.5} - 1.66533 \\ & \times 10^{-16} t^{9.5} - 1.38778 \times 10^{-17} t^{10} - 0.00107409 t^{10.5} + 3.46945 \times 10^{-18} t^{11} \\ & + 1.38778 \times 10^{-17} t^{11.5} + 0.00410658 t^{12} - 0.00447957 t^{13} - 0.00465867 t^{13.5} \\ & + 0.00114198 t^{14} + 0.010066 t^{14.5} + 0.000170673 t^{15} - 0.00817004 t^{15.5} \\ & - 0.000182659 t^{16} + 0.0049056 t^{16.5} + 0.0000458036 t^{17} - 0.00927567 t^{17.5} \\ & - 0.000316896 t^{18} + 0.0173363 t^{18.5} + 0.00101742 t^{19} - 0.0193873 t^{19.5} \\ & - 0.00135484 t^{20} + 0.0135402 t^{20.5} + 0.00107153 t^{21} - 0.00602444 t^{21.5} \\ & - 0.00102235 t^{22} + 0.00165852 t^{22.5} + 0.00169885 t^{23}. - 0.000236628 t^{23.5} \\ & - 0.00240855 t^{24} - 0.0000267944 t^{24.5} + 0.002328 t^{25} + 0.000048839 t^{25.5} - \cdots . \end{aligned}$$

The HAM series upto first six terms when $\alpha = 2$ and taking $\hbar = -1$, is given by

$$\begin{aligned} y(t) = {} & 2t^3 - 3t^4 + t^5 - \frac{t^{14}}{374400} + \frac{79t^{15}}{4365900} - \frac{12469t^{16}}{232848000} + \frac{19801t^{17}}{257297040} \\ & - \frac{2940419t^{18}}{198486288000} - \frac{6625093t^{19}}{43135092000} + \frac{311661149t^{20}}{985944960000} - \frac{3268452721t^{21}}{9725828112000} \\ & + \frac{45173665739t^{22}}{203779255680000} - \frac{103894399937t^{23}}{1113144184152000} + \frac{455437961213t^{24}}{20778691437504000} \\ & + \frac{18932926013t^{25}}{5081745188520000} - \cdots . \end{aligned}$$

Additionally, the HWT is used to resolve the above problem. Graphs and tables are used to explain the results that were obtained. The geometrical comparison of the exact HAM and HTM solutions is shown in Figure 4. The graphical depiction of the HAM solution with various values of α is shown in Figure 5. Figure 6 displays the error analysis of the HAM and HWT answers with the precise solution. Tables 3 and 4 provide numerical values for the integer and noninteger values of α accordingly, including the results available in the literature.

Table 3: Comparison between HAM and HWT solutions for different fractional α values in Example 2

t	$\alpha = 1.2$		$\alpha = 1.5$		$\alpha = 1.7$	
	HAM	HWT	HAM	HWT	HAM	HWT
0.0	0	0	0	0	0	
0.1	0.001710	0.001799	0.001710	0.001790	0.001710	0.001799
0.2	0.011520	0.011555	0.011520	0.011565	0.011520	0.011511
0.3	0.032129	0.032988	0.032130	0.032153	0.032130	0.032125
0.4	0.061432	0.061457	0.061440	0.061477	0.061440	0.061410
0.5	0.093701	0.093365	0.093750	0.093799	0.093750	0.093702
0.6	0.120744	0.120875	0.120958	0.120921	0.120960	0.120999
0.7	0.132991	0.132822	0.133760	0.133796	0.133770	0.133765
0.8	0.120401	0.120897	0.122840	0.122856	0.122878	0.122825
0.9	0.072946	0.072365	0.080041	0.080035	0.080182	0.080147
1.0	-0.019915	-0.019896	-0.000521	-0.000988	-0.000034	-0.000988

Table 4: Comparison of solutions obtained from HAM, HOFLMSM, HWT, and their AE with exact solution for integer order $\alpha = 2$ of Example 2

t	Exact	HAM	HOFLMSM [24]	HWT	HAM Error	HOFSMLM Error [24]	HWT Error
0	0	0	0	0	0	0	0
0.1	0.001710	0.001710	-	0.002044	0	-	3.34×10^{-4}
0.2	0.011520	0.0115200	0.011517	0.012099	0	3.222×10^{-6}	5.79×10^{-4}
0.3	0.032130	0.032130	-	0.032866	2.1×10^{-9}	-	7.35×10^{-4}
0.4	0.061440	0.061440	0.061411	0.062269	3.66×10^{-8}	2.920×10^{-5}	8.29×10^{-4}
0.5	0.093750	0.093750	-	0.094641	3.268×10^{-7}	-	8.91×10^{-4}
0.6	0.120960	0.120958	0.120920	0.12187	1.991×10^{-6}	3.972×10^{-5}	9.26×10^{-4}
0.7	0.13377	0.133760	-	0.134730	9.592×10^{-6}	-	9.60×10^{-4}
0.8	0.122880	0.122840	0.122845	0.012389	3.979×10^{-5}	3.49×10^{-5}	1.01×10^{-3}
0.9	0.08019	0.080041	-	0.081295	1.494×10^{-4}	-	1.10×10^{-3}
1.0	0	-0.000521	-0.000015	-0.001106	5.212×10^{-4}	1.545×10^{-5}	1.10×10^{-3}

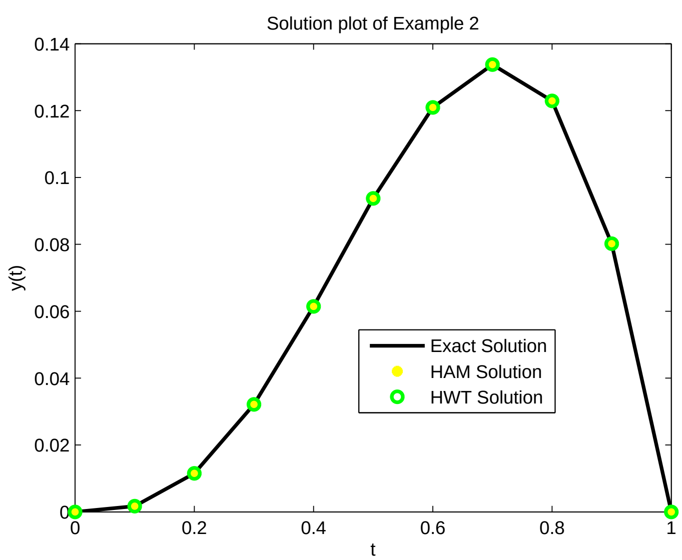


Figure 4: Comparison of the exact solution with HAM and HWT solutions, for Example 2.

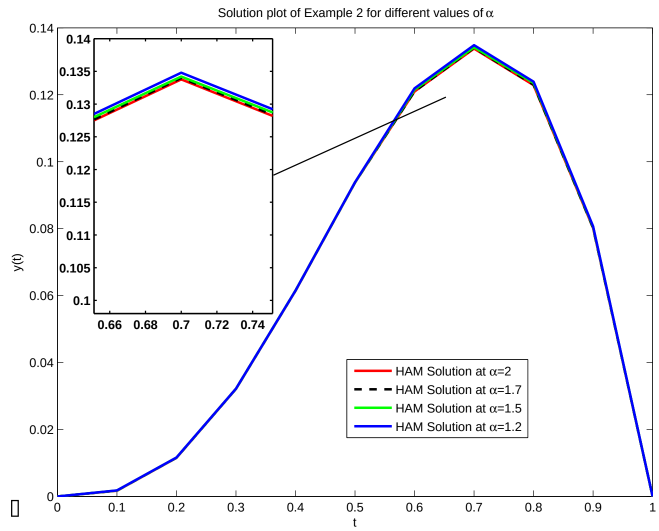


Figure 5: Graphical interpretation of the HAM solutions at different α values for Example 2.

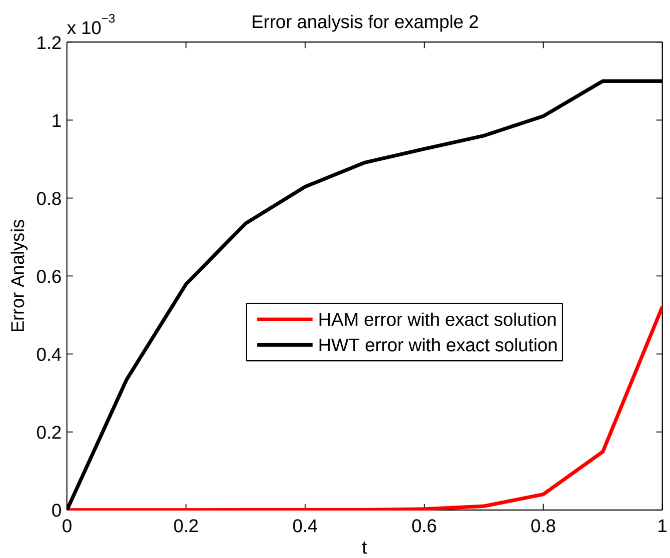


Figure 6: Graphical interpretation of the error analysis, for Example 2.

Example 3. Consider the nonlinear fraction model [31],

$$D^\alpha y(t) - t\sinh(y(t)) - 1 = 0, \quad 2 \le \alpha \le 3, \tag{36}$$

subject to the conditions

$$y(0) = 0, \quad y(0.25) = 1, \quad y(1) = 0. \tag{37}$$

Method of implementation: Applying HAM method to solve (36) and concerning (9), the zeroth order deformation is given by

$$(1-q)\mathscr{L}[\phi(t;q) - y_0(t)] = q\hbar\mathscr{H}(t)[D^\alpha\phi(t;q) - \sinh(\phi(t;q)) - 1]. \tag{38}$$

According to the condition (37) and choosing the initial approximation $y_0(t) = \frac{t^\alpha}{\Gamma(\alpha+1)}$, we take the linear operator $\mathscr{L}= D^\alpha$ with $\mathscr{L}(C_1 + C_2 t + C_3 t^2) = 0$, where C_1, C_2, and C_3 are integral constants with $\mathscr{H}(t) = 1$. Therefore, the mth order deformation is given by

$$D^\alpha[y_m(t) - \chi_m y_{m-1}(t)] = \hbar\mathscr{R}_m(y_{m-1}(t)), \tag{39}$$

where

$$\mathscr{R}_m(y_{m-1}(t)) = \frac{1}{(m-1)!}\frac{\partial^{m-1}[D^\alpha\phi(t;q) - \sinh(\phi(t;q)) - 1]}{\partial q^{m-1}}, \tag{40}$$

subject to

$$y_m(0) = 0, \quad y_m(0.25) = 0, \quad y_m(1) = 0. \tag{41}$$

Applying I^α, the inverse of D^α on either sides of (39), we get

$$y_m(t) = \chi_m y_{m-1}(t) + \hbar I^\alpha[\mathscr{R}_m(y_{m-1}(t))] + C_1 + C_2 t + C_3 t^2, \quad m \ge 1,$$

where C_1, C_2, and C_3 are calculated by using (41).

The HAM series upto first six terms when $\alpha = 2.5$ and taking $\hbar = -1$ is given by

$$\begin{aligned} y(t) = {} & 5.42172t - 5.85394t^2 + 0.300901t^{2.5} + 0.13254t^{3.5} + 0.00333693t^{4.5} \\ & - 0.0107536t^{5.5} + 0.00486111t^6 + 6.71025\times 10^{-7}t^{6.5} + 0.00133921t^7 \\ & - 8.70736\times 10^{-6}t^{7.5} + 0.0000293912t^8 + 0.0000396298t^{8.5} \\ & - 0.000104465t^9 - 0.000025573t^{9.5} + 0.000110933t^{10} - 0.000061787t^{10.5} \end{aligned}$$

$$
\begin{aligned}
&+ 0.0000292694t^{11} - 0.0000240961t^{11.5} + 9.8944 \times 10^{-6}t^{12} - 8.36776 \\
&\times 10^{-7}t^{12.5} - 7.54444 \times 10^{-7}t^{13} + 2.01592 \times 10^{-6}t^{13.5} - 1.14756 \\
&\times 10^{-6}t^{14} - 1.27511 \times 10^{-7}t^{14.5} + 3.67453 \times 10^{-7}t^{15} - 1.41232 \\
&\times 10^{-7}t^{15.5} + 1.8676 \times 10^{-8}t^{16} - 6.94883 \times 10^{-9}t^{16.5} + 1.11916 \times 10^{-8}t^{17} \\
&+ 1.86632 \times 10^{-9}t^{17.5} - 1.1398 \times 10^{-8}t^{18} + 7.4029 \times 10^{-9}t^{18.5} - 7.60145 \\
&\times 10^{-10}t^{19} - 1.21074 \times 10^{-9}t^{19.5} + 1.38057 \times 10^{-10}t^{20} + 4.59674 \\
&\times 10^{-10}t^{20.5} + 2.71044 \times 10^{-10}t^{21} - 9.18445 \times 10^{-10}t^{21.5} + 7.2516 \\
&\times 10^{-10}t^{22} - 2.61436 \times 10^{-10}t^{22.5} + 3.73559 \times 10^{-11}t^{23} - 4.92754 \\
&\times 10^{-13}t^{23.5} + 1.0376 \times 10^{-12}t^{24} - 6.09013 \times 10^{-13}t^{24.5} + 1.22102 \times 10^{-13}t^{25}.
\end{aligned}
$$

The HAM series upto first six terms when $\alpha = 3$ and taking $h = -1$ is given by

$$
\begin{aligned}
y(t) = {}& 5.39493t - 5.62542t^2 + \frac{t^3}{6} + 0.0642361t^4 + 0.00101757t^5 - 0.00241502t^6 \\
&+ 0.000793762t^7 + 0.000184995t^8 + 5.78057 \times 10^{-6}t^9 - 0.0000101977t^{10} \\
&+ 6.76308 \times 10^{-6}t^{11} - 1.72474 \times 10^{-6}t^{12} - 3.70131 \times 10^{-7}t^{13} + 2.20798 \\
&\times 10^{-7}t^{14} + 4.34937 \times 10^{-8}t^{15} - 2.3555 \times 10^{-8}t^{16} + 7.1741 \times 10^{-9}t^{17} \\
&- 1.46382 \times 10^{-9}t^{18} + 2.83531 \times 10^{-10}t^{19} - 5.93013 \times 10^{-11}t^{20} \\
&+ 2.72701 \times 10^{-11}t^{21} - 1.25142 \times 10^{-11}t^{22} + 5.48224 \times 10^{-12}t^{23} \\
&- 2.21553 \times 10^{-12}t^{24} + 7.74129 \times 10^{-13}t^{25} - 1.78755 \times 10^{-13}t^{26} \\
&+ 1.97752 \times 10^{-14}t^{27} - 1.1748 \times 10^{-16}t^{28} + 1.46076 \times 10^{-17}t^{29}.
\end{aligned}
$$

The above problem is also solved using the HWT and ND solver. The graphs and tables provide an explanation of the results. Figure 7 compares ND solver, HAM, and HTM solutions geometrically. The HAM solution is graphically depicted with various values of α in Figure 8. Figure 9 depicts the error analysis of the solutions from the HAM, HWT, and ND solvers. Tables 5 and 6 contain numerical values for the fractional and nonfractional values of α, respectively.

Table 5: Comparison between HAM and HWT solutions for different fractional values of α in Example 3

t	$\alpha = 2.2$		$\alpha = 2.5$		$\alpha = 2.7$	
	HAM	HWT	HAM	HWT	HAM	HWT
0.1	0.4844060642	0.484351	0.4838689878	0.484351	0.4836640338	0.483640
0.2	0.8558871839	0.855887	0.8555770932	0.855887	0.8554919337	0.855491
0.3	1.1166930600	1.116703	1.1170945299	1.117105	1.1171714074	1.117176
0.4	1.2687652290	1.268765	1.2701278494	1.270128	1.2702944408	1.270294
0.5	1.3139530304	1.313949	1.3162987718	1.316293	1.3164440681	1.316442
0.6	1.2540670349	1.254067	1.2571981039	1.257198	1.2572226052	1.257222
0.7	1.0909051719	1.090910	1.0944110459	1.094417	1.0942652940	1.094268
0.8	0.8262733802	0.826273	0.8295325852	0.829533	0.8292490252	0.829249
0.9	0.4620075965	0.461998	0.4641792631	0.464165	0.4638997104	0.463893
1.0	0	0	0	0	0	0

Table 6: Comparison of solutions obtained from HAM, HWT, and their AE with ND Solver solution for integer order $\alpha = 3$ of Example 3

t	ND Solver	HAM	HWT	HAM Error	HWT Error
0.0	0	0	0	0	0
0.1	0.4850443102	0.4834121287	0.483293	1.63×10^{-3}	1.75×10^{-3}
0.2	0.8564165257	0.8554060228	0.855347	1.01×10^{-3}	1.06×10^{-3}
0.3	1.1158509874	1.1172133089	1.117272	1.36×10^{-3}	1.42×10^{-3}
0.4	1.2655112703	1.2702191975	1.270362	4.70×10^{-3}	4.85×10^{-3}
0.5	1.3078943289	1.3159609254	1.316153	8.06×10^{-3}	8.25×10^{-3}
0.6	1.2456730868	1.2561260143	1.256409	1.04×10^{-2}	1.07×10^{-2}
0.7	1.0815016154	1.0925510067	1.093045	1.10×10^{-2}	1.15×10^{-2}
0.8	0.8178293569	0.8272214124	0.828006	9.39×10^{-3}	1.10×10^{-2}
0.9	0.4567647525	0.4622736694	0.463128	5.509×10^{-3}	6.3×10^{-3}
1.0	0	0	0	0	0

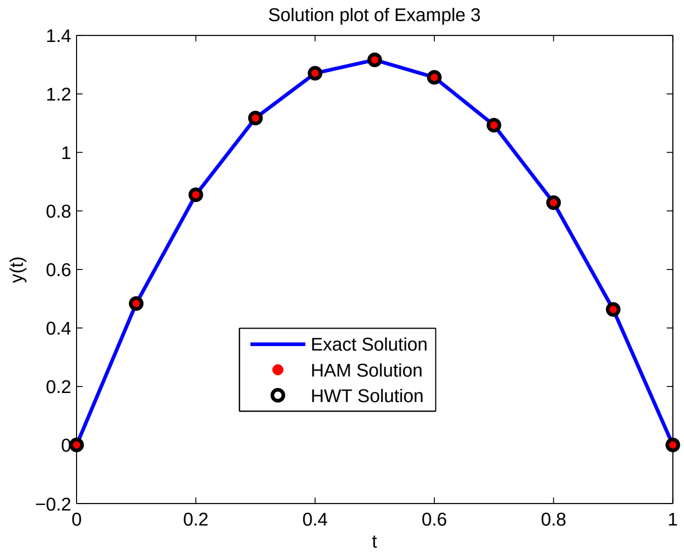


Figure 7: Comparison of the exact solution with HAM and HWT solutions, for Example 3.

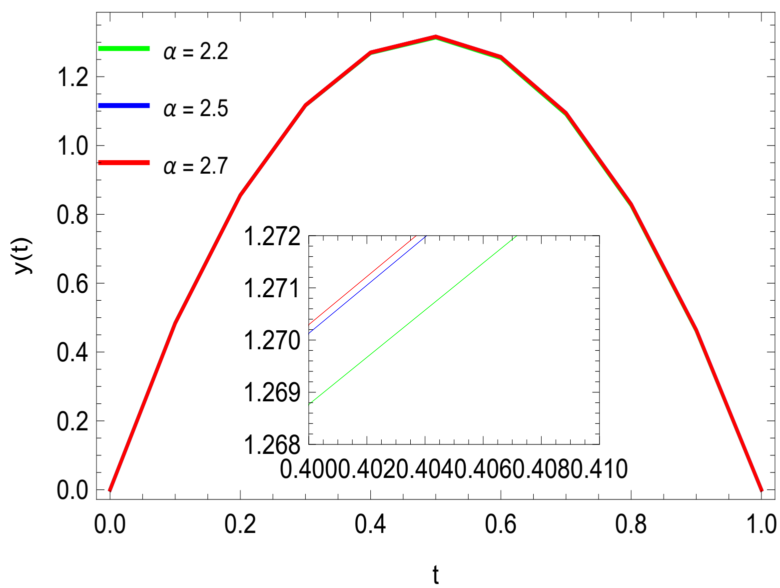


Figure 8: Graphical interpretation of the HAM solutions at different α values for Example 3.

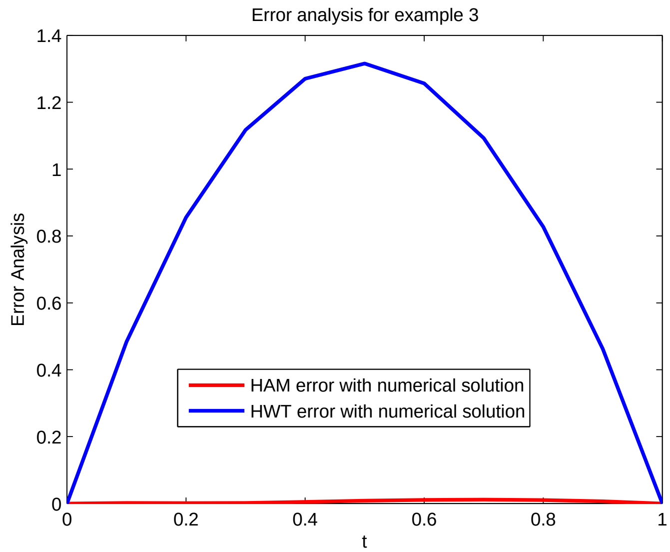


Figure 9: Graphical interpretation of the error analysis, for Example 3.

Example 4. Consider the nonlinear fractional model,

$$D^{\alpha}y(t)-\sin(y(t))=0,\quad 1\leq\alpha\leq 2, \tag{42}$$

subject to the conditions

$$y(0)=\pi,\quad y'(0)=-2. \tag{43}$$

Method of implementation: Applying HAM to solve (42) and concerning (9), the zeroth order deformation is given by

$$(1-q)\mathscr{L}[\phi(t;q)-y_0(t)]=q\hbar\mathscr{H}(t)[D^{\alpha}\phi(t;q)-\sin(\phi(t;q))]. \tag{44}$$

According to the condition (43) and choosing initial approximation $y_0(t)=\pi-2t$, we take the linear operator as $\mathscr{L}=D^{\alpha}$ with $\mathscr{L}(C_1+C_2t)=0$, where C_1 and C_2 are integral constants with $\mathscr{H}(t)=1$. Therefore, mth order deformation is given by

$$D^{\alpha}[y_m(t)-\chi_m y_{m-1}(t)]=\hbar\mathscr{R}_m(y_{m-1}(t)), \tag{45}$$

where

$$\mathscr{R}_m(y_{m-1}(t)) = \frac{1}{(m-1)!}\frac{\partial^{m-1}[D^\alpha \phi(t;q) - \sin(\phi(t;q))]}{\partial q^{m-1}} \tag{46}$$

subject to

$$y_m(0) = 0, \qquad y'_m(0) = 0. \tag{47}$$

Applying I^α, the inverse of D^α on either sides of (45), we get

$$y_m(t) = \chi_m y_{m-1}(t) + \hbar I^\alpha[\mathscr{R}_m(y_{m-1}(t))] + C_1 + C_2 t, \quad m \geq 1,$$

where C_1 and C_2 are calculated using (47).

The HAM series upto first six terms when $\alpha = 2$ and taking $h = -1$ is given by

$$\begin{aligned} y(t) = \frac{1}{16986931200}&(100t(72(642315 - 7616t^2)\cos(2t) + 52736\cos(6t) \\ &+ 405\cos(8t) + 48(9(5381 - 40t^2)\cos(4t) + 4t(45691 - 48t^2 + 4293\cos(2t) \\ &+ 50\cos(4t))\sin(2t))) - 5(-3397386240\pi + 4168250428t \\ &+ 1152t^3(-835 + 16t^2) + 1660356360\sin(2t) + 66396240\sin(4t) \\ &+ 2073250\sin(6t) + 33885\sin(8t)) - 1026\sin(10t)). \end{aligned}$$

Additionally, the above problem was solved using the HWT and ND solver. Through the use of graphs and tables, outcomes are explained. Exact HAM and HTM solutions are geometrically compared in Figure 10. Figure 11 shows a graphic representation of the HAM solution with various values for α. Figure 12 displays the exact solution and error analysis of the HAM, HWT, and ND solver solutions. In Tables 7 and 8, respectively, numerical values for integer and noninteger values of α are provided.

Table 7: Comparison between HAM and HWT solutions for different fractional values of α in Example 4

t	$\alpha = 1.2$		$\alpha = 1.5$		$\alpha = 1.7$	
	HAM	HWT	HAM	HWT	HAM	HWT
0.1	2.946783	2.94657	2.943491	2.94343	2.942547	2.94256
0.2	2.765230	2.76433	2.752249	2.75192	2.747753	2.74762
0.3	2.598414	2.59546	2.570587	2.56965	2.559792	2.55934
0.4	2.446371	2.43943	2.400060	2.39788	2.380531	2.37946
0.5	2.308211	2.29522	2.241409	2.23711	2.211257	2.20912
0.6	2.182297	2.16154	2.094669	2.08729	2.052699	2.04892
0.7	2.066382	2.03708	1.959240	1.94802	1.905048	1.89911
0.8	1.957728	1.92051	1.833943	1.81866	1.767964	1.75956
0.9	1.853242	1.8106	1.7171035	1.69842	1.640616	1.62994
1.0	1.749611	1.70617	1.606626	1.58642	1.521720	1.50973

Table 8: Comparison of solutions obtained from HAM, HWT, and their AE with ND Solver solution for integer order $\alpha = 2$ of Example 4

t	ND Solver	HAM	HWT	HAM Error	HWT Error
0.1	2.941925	2.941925	2.94216	8.8×10^{-9}	2.35×10^{-4}
0.2	2.744233	2.744233	2.74474	1.57×10^{-8}	5.17×10^{-4}
0.3	2.550495	2.550395	2.55114	5.4×10^{-9}	7.44×10^{-4}
0.4	2.362110	2.362110	2.36307	3.5×10^{-9}	9.60×10^{-4}
0.5	2.180831	2.180831	2.18198	4.87×10^{-8}	1.14×10^{-3}
0.6	2.007722	2.007721	2.00903	2.033×10^{-7}	1.30×10^{-3}
0.7	1.843649	1.843648	1.84509	7.507×10^{-7}	1.44×10^{-3}
0.8	1.689183	1.689181	1.69073	2.3361×10^{-6}	1.55×10^{-3}
0.9	1.544628	1.544622	4.54625	5.8886×10^{-6}	1.62×10^{-3}
1.0	1.410054	1.410042	1.41194	1.16915×10^{-5}	1.89×10^{-3}

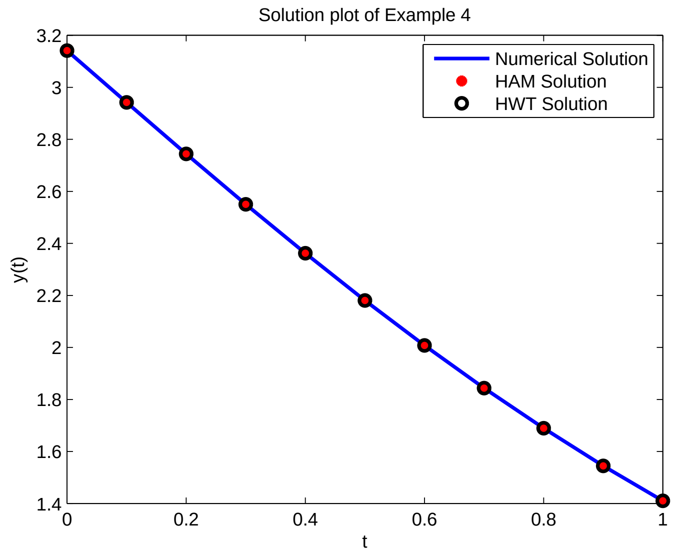


Figure 10: Comparison of the exact solution with HAM and HWT solutions, for Example 4.

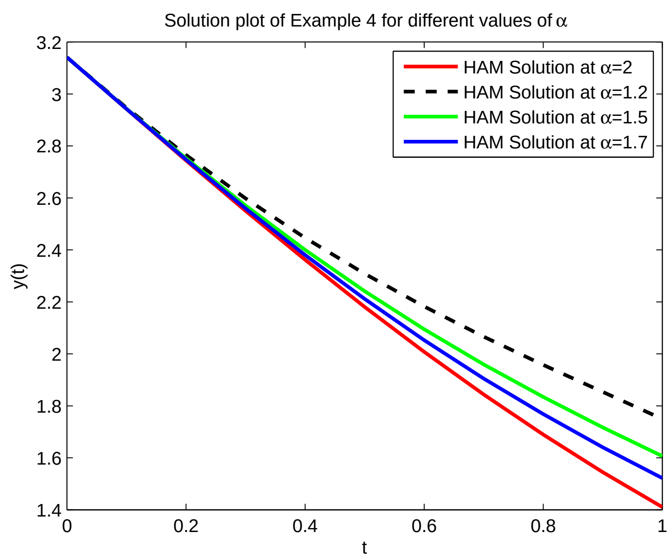


Figure 11: Graphical interpretation of the HAM solutions at different α values for Example 4.

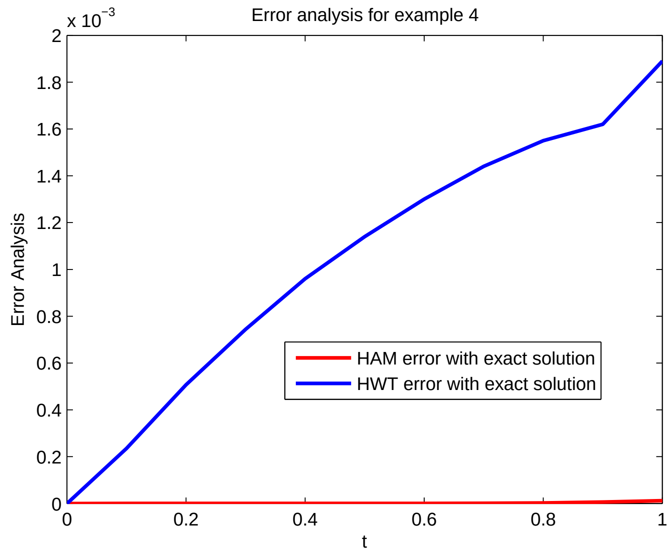


Figure 12: Graphical interpretation of the error analysis, for Example 4.

Example 5. Consider the one-dimensional nonlinear homogeneous time fractional Korteweg–de Vries (KdV) equation [11],

$$D_t^{\alpha}[v(x,t)] + 6vv_x + v_{xxx} = 0, \quad 0 \le \alpha \le 1, \tag{48}$$

with the initial condition

$$v(x,0) = \frac{1}{2} sech^2\left(\frac{x}{2}\right). \tag{49}$$

The exact solution [11] is $v(x,t) = \frac{1}{2} sech^2\left(\frac{x+t}{2}\right)$.

Method of implementation: Applying HAM method to solve (48) and concerning (9), the zeroth order deformation is given by

$$\begin{aligned}&(1-q)\mathscr{L}[\phi(x,t;q) - v_0(x,t)]\\ &\quad = q\hbar\mathscr{H}(t)[D_t^{\alpha}\phi(x,t;q) + 6\phi(x,t;q)\phi(x,t;q)_x + \phi(x,t;q)_{xxx}].\end{aligned} \tag{50}$$

According to the condition (49), choosing the initial approximation $v_0(x,t) = \frac{1}{2} sech^2\left(\frac{x}{2}\right)$, we take the linear operator $\mathscr{L} = D_t^{\alpha}$ with $\mathscr{L}(C_1) = 0$, where C_1 is a integral constant with $\mathscr{H}(t) = 1$. Therefore, mth order deformation is given by

$$D_t^{\alpha}[v_m(x,t) - \chi_m v_{m-1}(x,t)] = \hbar \mathscr{R}_m(v_{m-1}(x,t)), \tag{51}$$

where

$$\begin{aligned}&\mathscr{R}_m(v_{m-1}(x,t))\\ &= \frac{1}{(m-1)!}\frac{\partial^{m-1}[D_t^{\alpha}\phi(x,t;q) + 6\phi(x,t;q)\phi(x,t;q)_x + \phi(x,t;q)_{xxx}]}{\partial q^{m-1}},\end{aligned} \tag{52}$$

subject to

$$v_m(x,0) = 0, \tag{53}$$

Applying I_t^{α}, the inverse of D_t^{α} on either sides of (51), we get

$$v_m(x,t) = \chi_m v_{m-1}(xt) + \hbar I_t^{\alpha}[\mathscr{R}_m(v_{m-1}(x,t))] + C_1, \quad m \geq 1,$$

where C_1 is calculated by using (53).

The HAM series upto first six terms when $\alpha = 1$ and taking $\hbar = -1$ is given by

$$\begin{aligned}v(x,t) = &\frac{2}{\left(e^{\frac{-x}{2}} + e^{\frac{x}{2}}\right)^2} + \frac{2e^x t}{(1+e^x)^4} - \frac{2e^{3x} t}{(1+e^x)^4} + \frac{4e^x\left(-1+e^x\right)t}{(1+e^x)^3}\\ &+ \frac{e^x t^2}{(1+e^x)^4} - \frac{4e^{2x}t^2}{(1+e^x)^4} + \frac{e^{3x}t^2}{(1+e^x)^4} + \cdots .\end{aligned}$$

The HAM series upto first six terms when $\alpha = 0.4$ and taking $h = -1$ is given by

$$\begin{aligned}y(t) = &\frac{2}{\left(e^{\frac{-x}{2}} + e^{\frac{x}{2}}\right)^2} + \frac{2.25412e^x t^{0.4}}{(1+e^x)^6} + \frac{4.50824e^{2x} t^{0.4}}{(1+e^x)^6} - \frac{4.50824e^{4x} t^{0.4}}{(1+e^x)^6}\\ &- \frac{2.25412e^{5x} t^{0.4}}{(1+e^x)^6} + \frac{4.50824e^x(e^x - 1)t^{0.4}}{(1+e^x)^3}\\ &+ \frac{2.25412e^x t^{0.8}}{(1+e^x)^6} - \frac{4.29469e^{2x} t^{0.8}}{(1+e^x)^6} + \cdots .\end{aligned}$$

The graphs and tables provide an explanation of the results. Figure 13 compares Exact, HAM, and FRPSM solutions geometrically. The HAM solution is graphically depicted with various values of α in Figure 14. Figure 15 depicts the error analysis of the solutions from the HAM, FRSPM, and Exact.

Tables 9 and 10 contain numerical values for the fractional and nonfractional values of α, respectively.

Table 9: Comparison of solutions obtained from exact solutions and HAM and AE of HAM, FRSM, ADM solutions with exact solution for integer order $\alpha = 1$ at $x = 0$ of Example 5

t	Exact	HAM	HAM Error	FRPSM Error[11]	ADM Error[11]
0	0.500000	0.500000	0.	-	-
0.1	0.498752	0.498752	5.551×10^{-17}	2.900×10^{-9}	2.080×10^{-6}
0.2	0.495033	0.495033	5.551×10^{-17}	1.879×10^{-7}	3.31×10^{-5}
0.3	0.488917	0.488917	1.110×10^{-16}	2.216×10^{-6}	3.054×10^{-4}
0.4	0.480521	0.480521	5.551×10^{-17}	1.184×10^{-5}	2.763×10^{-4}
0.5	0.470007	0.470007	5.551×10^{-17}	4.465×10^{-5}	2.500×10^{-3}
0.6	0.457568	0.457568	2.220×10^{-16}	-	-
0.7	0.443426	0.443426	6.106×10^{-16}	-	-
0.8	0.427819	0.427819	3.608×10^{-15}	-	-
0.9	0.411001	0.411001	7.622×10^{-14}	-	-
1.0	0.393224	0.393224	1.171×10^{-12}	-	-

Table 10: Comparison between HAM and FRPSM solutions for different fractional values of α at $x = 2$ in Example 5

t	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$		$\alpha = 1$	
	HAM	FRPSM[11]	HAM	FRPSM[11]	HAM	FRPSM[11]	HAM	FRPSM[11]	HAM	FRPSM[11]
0	0.209987	-	0.209987	-	0.209987	-	0.209987	-	0.209987	-
0.1	0.354751	0.354750978	0.294966	0.29496648	0.259396	0.259395686	0.238566	0.238566145	0.226368	0.226368188
0.2	0.382232	0.382232355	0.327693	0.327692525	0.288355	0.28835476	0.261508	0.261507702	0.243526	0.243526238
0.3	0.400997	0.400996543	0.353185	0.353185341	0.31353	0.313530228	0.283442	0.283441821	0.261461	0.261461322
0.4	0.415703	0.415702913	0.375006	0.375006055	0.336761	0.336761221	0.30503	0.305029944	0.280173	0.280173439
0.5	0.427989	0.427988593	0.394504	0.394504163	0.35877	0.358769892	0.326536	0.326536438	0.299663	0.299662589
0.6	0.438641	-	0.412363	-	0.379928	-	0.348096	-	0.319929	-
0.7	0.448105	-	0.428985	-	0.400457	-	0.369786	-	0.340972	-
0.8	0.456661	-	0.44463	-	0.420499	-	0.391655	-	0.362792	-
0.9	0.464496	-	0.459478	-	0.440155	-	0.413735	-	0.385390	-
1.0	0.471742	-	0.47366	-	0.459495	-	0.436046	-	0.408764	-

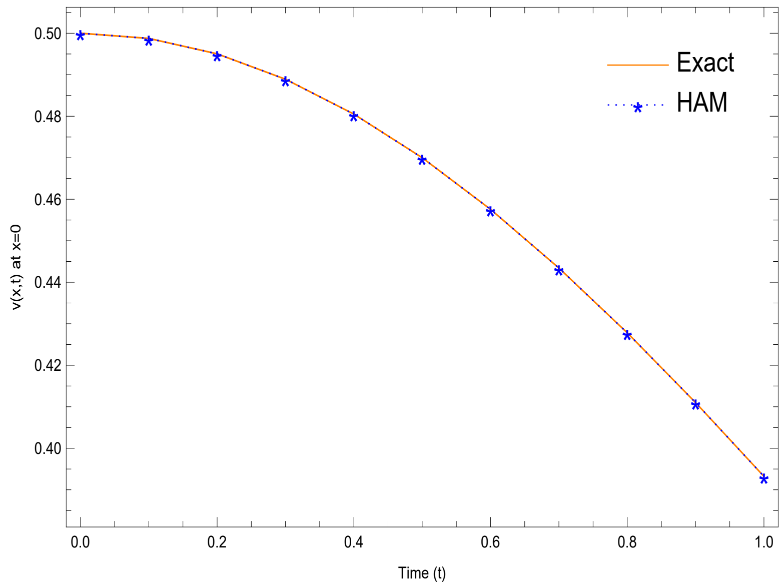


Figure 13: Comparison of the exact solution with HAM and HWT solutions, for Example 5.

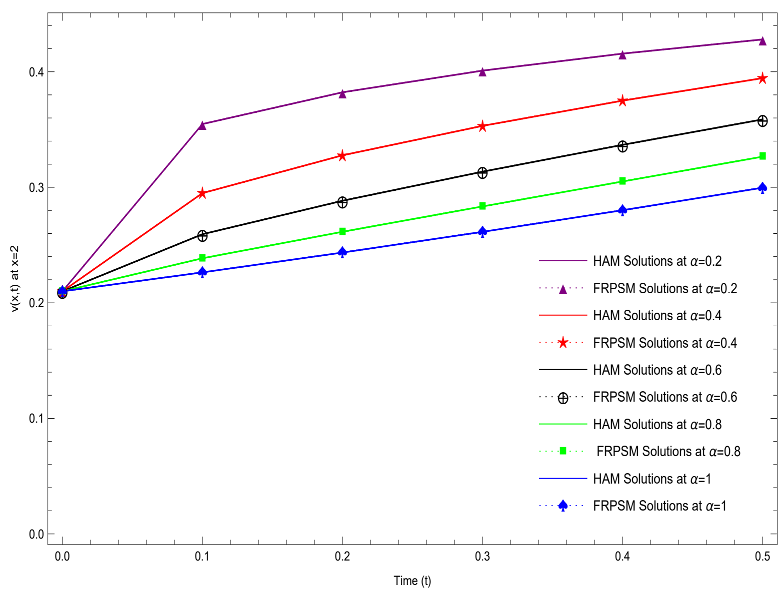


Figure 14: Graphical interpretation of the HAM solutions at different α values for Example 5.

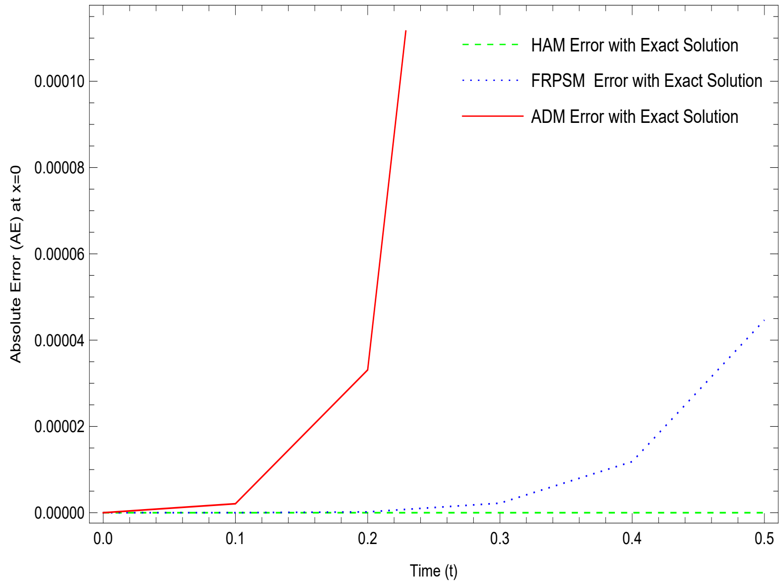


Figure 15: Graphical interpretation of the error analysis, for Example 5.

6 Conclusion

In this study, we discussed the linear and nonlinear fractional models of different orders through two different methods: the HAM (Homotopy analysis method) and the HWT (Haar wavelet technique). Here, we considered five problems to justify the two different methods. HAM is the semi-analytical method that yields the analytical solution of a given model after more deformations. The HWT is a numerical technique that solves the models numerically with the help of the Mathematica software. The obtained solutions were numerically tabulated in Tables 1–10. Figures 1–14 show the performance of the methods. Here, HAM consumes more time to yield solutions for the different models, but HWT yields the numerical results in less time. This study reveals that HAM provides solutions with high accuracy compared to HWT. From the tables and figures, we conclude that HAM is better than HWM, FRPSM [11], and HOFLMSM [24].

Acknowledgements

Authors are grateful to there anonymous referees and editor for their constructive comments.

References

[1] Adomian, G. *Solving frontier problems of physics: the decomposition method, Springer*, Dordrecht, 1994.

[2] Ahmed, H.F., Bahgat, M.S. and Zaki, M., *Numerical approaches to system of fractional partial differential equations*, J. Egypt. Math. Soc. 25(2) (2017), 141–150.

[3] Balaji, S. *Legendre wavelet operational matrix method for solution of fractional order Riccati differential equation*, J. Egypt. Math. Soc. 23(2) (2015), 263–270.

[4] Cabada, A. and Wang, G. *Positive solutions of nonlinear fractional differential equations with integral boundary value conditions*, J. Math. Anal. Appl. 389(1) (2012), 403–411.

[5] Chauhan, J.P. and Khirsariya, S.R. *A semi-analytic method to solve nonlinear differential equations with arbitrary order*, Results Control Optim. 12 (2023), 100267.

[6] Chen, Y., Yi, M. and Yu, C. *Error analysis for numerical solution of fractional differential equation by Haar wavelets method*, J. Comput. Sci. 3(5) (2012), 367–373.

[7] El-Ajou, A., Odibat, Z., Momani, S. and Alawneh, A. *construction of analytical solutions to fractional differential equations using homotopy analysis method*, IAENG Int. J. Appl. Math. 40(2) (2010).

[8] Kanth, A.R. and Aruna, K. *Solution of fractional third-order dispersive partial differential equations*, Egypt. J. Basic Appl. Sci. 2(3) (2015), 190–199.

[9] Khirsariya, S.R. and Rao, S.B. *On the semi-analytic technique to deal with nonlinear fractional differential equations*, J. Appl. Math. Comput. Mech. 22(1) (2023).

[10] Khirsariya, S.R. and Rao, S.B. *Solution of fractional Sawada–Kotera–Ito equation using Caputo and Atangana–Baleanu derivatives*, Math. Methods Appl. Scie. 46(15) (2023), 16072–16091.

[11] Khirsariya, S., Rao, S. and Chauhan, J. *Semi-analytic solution of time-fractional Korteweg-de Vries equation using fractional residual power series method*, Results Nonlinear Anal. 5(3) (2022), 222–234.

[12] Khirsariya, S.R., Rao, S.B. and Hathiwala, G.S. *Investigation of fractional diabetes model involving glucose–insulin alliance scheme*, Int. J. Dyn. Contr. 12(1) (2024), 1–14.

[13] Kumbinarasaiah, S. *A novel approach for multi dimensional fractional coupled Navier–Stokes equation*, SeMA J. 80(2) (2023), 261–282.

[14] Kumbinarasaiah, S. and Mulimani, M. *The Fibonacci wavelets approach for the fractional Rosenau–Hyman equations*, Results Control Optim. 11 (2023), 100221.

[15] Kumbinarasaiah, S. and Mundewadi, R.A. *Numerical solution of fractional-order integro-differential equations using Laguerre wavelet method*, J. Inform. Optim. Scie. 43(4) (2022), 643–662.

[16] Kumbinarasaiah, S. and Preetham, M.P. *A study on homotopy analysis method and clique polynomial method*, Comput. Methods Differ. Equ. 10(3) (2022), 774–788.

[17] Lazarević, M.P. and Spasić, A.M. *Finite-time stability analysis of fractional order time-delay systems: Gronwall's approach*, Math. Comput. Model. 49(3-4) (2009), 475–481.

[18] Liang, S. and Jeffrey, D.J. *Comparison of homotopy analysis method and homotopy perturbation method through an evolution equation*, Commun. Nonlinear Sci. Numer. Simul. 14(12) (2009), 4057–4064.

[19] Liao, S. *An explicit, totally analytical approximate solution for Blasius equation [J]*, Int. J. Non-Linear Mech. 34 (1999), 759–778.

[20] Liao, S.J. *The proposed homotopy analysis technique for the solution of nonlinear problems*, Diss. PhD thesis, Shanghai Jiao Tong University, 1992.

[21] Liao, S. *Beyond perturbation: introduction to the homotopy analysis method*, CRC press, 2003.

[22] Liao, S. *Homotopy analysis method in nonlinear differential equations*, Beijing: Higher education press, 2012.

[23] Liao, S. ed. *Advances in the homotopy analysis method*, World Scientific, 2013.

[24] Lin, R. and Liu, F. *Fractional high order methods for the nonlinear fractional ordinary differential equation*, Nonlinear Anal. Theory Methods Appl. 66(4) (2007), 856–869.

[25] Motsa, S.S., Sibanda, P., Awad, F.G. and Shateyi, S. *A new spectral-homotopy analysis method for the MHD Jeffery–Hamel problem*, Comput. Fluids 39(7) (2010) , 1219–1225.

[26] Odibat, Z.M. *A study on the convergence of homotopy analysis method*, Appl. Math. Comput. 217(2) (2010), 782–789.

[27] Podlubny, I. *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*, Elsevier, 1998.

[28] Prakash, A., Veeresha, P., Prakasha, D.G. and Goyal, M. *A homotopy technique for a fractional order multi-dimensional telegraph equation via the Laplace transform*, Eur. Phys. J. Plus 134 (2019), 1–18.

[29] Rahimy, M. *Applications of fractional differential equations*, Appl. Math. Sci 4(50) (2010), 2453–2461.

[30] Sajid, M. and Hayat, T. *Comparison of HAM and HPM methods in nonlinear heat conduction and convection equations*, Nonlinear Anal. Real World Appl. 9(5) (2008), 2296–2301.

[31] Semary, M.S. and Hany, N.H. *The homotopy analysis method for strongly nonlinear initial/boundary value problems*, Int. J. Modern Math. Sci. 9(3) (2014), 154–172.

[32] Shawagfeh, N.T. *Analytical approximate solutions for nonlinear fractional differential equations*, Appl. Math. Comput. 131 (2-3) (2002), 517–529.

[33] Shiralasetti, S. *Some results on Haar wavelets matrix through linear algebra*, Wavelet and Linear Algebra 4(2) (2017), 49–59.

[34] Srinivasa, K. and Rezazadeh, H. *Numerical solution for the fractional-order one-dimensional telegraph equation via wavelet technique*, Int. J. Nonlinear Sci. Numer. Simul. 22(6) (2021), 767–780.

[35] Srivastava, V.K., Awasthi, M.K. and Kumar, S. *Analytical approximations of two and three dimensional time-fractional telegraphic equation by reduced differential transform method*, Egypt. J. Basic Appl. Sci. 1(1) (2014), 60–66.

[36] Srinivasa, K., Baskonus, H.M. and Guerrero Sánchez, Y. *Numerical solutions of the mathematical models on the digestive system and covid-19 pandemic by hermite wavelet technique*, Symmetry 13 (12) (2021), 2428.

[37] Srivastava, V.K., Kumar, S., Awasthi, M.K. and Singh, B.K. *Two-dimensional time fractional-order biological population model and its analytical solution*, Egypt. J. Basic Appl. Sci. 1(1) (2014), 71–76.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1106–1139
https://doi.org/ 10.22067/ijnao.2024.87836.1434
https://ijnao.um.ac.ir/

Research Article

A fuzzy solution approach to multi-objective fully triangular fuzzy optimization problem

N. Swain, S. Maharana and S. Nayak*,

Abstract

Numerous optimization problems comprise uncertain data in practical circumstances and such uncertainty can be suitably addressed using the concept of fuzzy logic. This paper proposes a computationally efficient solution methodology to generate a set of fuzzy non-dominated solutions of a fully fuzzy multi-objective linear programming problem, which incorporates all its parameters and decision variables expressed in form of triangular fuzzy

*Corresponding author

Received 29 April 2024; revised 1 July 2024; accepted 23 July 2024

Neelam Swain
Department of Mathematics, School of Applied Sciences, KIIT Deemed to be University, Bhubaneswar, 751024, Odisha, India. e-mail: 2381147@kiit.ac.in

Sujit Maharana
Department of Mathematics, School of Applied Sciences, KIIT Deemed to be University, Bhubaneswar, 751024, Odisha, India. e-mail: sujitmaharana863@gmail.com

Suvasis Nayak
Department of Mathematics, School of Applied Sciences, KIIT Deemed to be University, Bhubaneswar, 751024, Odisha, India. e-mail: suvasis.nayakfma@kiit.ac.in

How to cite this article
Swain, N., Maharana, S. and Nayak, S., A fuzzy solution approach to multi-objective fully triangular fuzzy optimization problem. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1106-1139. https://doi.org/ 10.22067/ijnao.2024.87836.1434

numbers. The fuzzy parameters associated with the objective functions are transformed into interval forms by utilizing the fuzzy-cuts, which subsequently generates the equivalent interval valued objective functions. The concept of centroid of triangular fuzzy numbers derives the deterministic form of the constraints. Furthermore, the scalarization process of weighting sum approach and certain concepts of interval analysis are used to generate the fuzzy non-dominated solutions from which the compromise solution can be determined based on the corresponding real valued expressions of fuzzy optimal objective values resulted due to the ranking function. Three numerical problems and one practical problem are solved for illustration and validation of the proposed approach. The computational results are also discussed as compared to some existing methods.

AMS subject classifications (2020): 90C05; 90C29; 90C70.

Keywords: Multi-objective optimization; Fully fuzzy programming; Triangular fuzzy numbers; Interval valued functions; Weighting sum approach.

1 Introduction

Linear optimization has several applications in context of decision making real world problems but the associated data often exist in ambiguous state, which cannot be defined or categorized exactly. To manage such data with uncertainty, the concept of fuzzy logic is appropriately useful in formulating practical problems into suitable fuzzy optimization models. The concept of fuzzy was first introduced by Zadeh [31], which is rapidly implemented in numerous fields of applications including business, economics, management, engineering, health care, and many more. A crisp form of linear programming problem (LPP) usually exists in the following form:

$Max\ Z = \sum_{j=1}^{n} c_j x_j$

subject to

$\sum_{j=1}^{n} a_{ij} x_j (\leq, =, \geq) b_i, \quad i = 1, 2, \ldots, m$

$x_j \geq 0$, where $x = (x_j) \in \mathbb{R}^n, c_j, a_{ij}, b_i \in \mathbb{R}$.

The linear optimization that comprises all the technological coefficients, resource constants, and decision variables expressed in form of fuzzy numbers, is interpreted as fully fuzzy linear programming problem (FFLPP) [18, 19]. Practical problems of several fields incorporate multiple conflicting objective functions that need simultaneous optimization over some common constraints; such problems are mathematically modeled as multi-objective optimization problems (MOOP) [7]. An MOOP comprises a set of non-dominated solutions [21] from which the decision-maker (DM) determines the compromise solution. Lotfi et al. [19] proposed a method to solve FFLPP using the lexicography method where the associated symmetric triangular fuzzy numbers (TFNs) are approximated by nearest symmetric triangular numbers to derive its symmetric fuzzy solution. Das, Mandal and Edalatpanah [10] derived a method to solve FFLPP based on multi-objective LPP, and lexicographic ordering approach where all its parameters are considered as trapezoidal fuzzy numbers. Ezzati, Khorram, and Enayati [12] developed a method to solve FFLPP, which converts the fuzzy problem into a multi-objective LPP (MOLPP) and found its solution using lexicographic method. Hosseinzadeh and Edalatpanah [15] proposed a method of solution for FFLPP using lexicography method where the parameters are considered as L-R fuzzy numbers. Kumar, Kaur, and Singh [18] developed a method of solution for FFLPP with equality constraints using the ranking functions to derive the fuzzy optimal solution whereas Allahviranloo et al.[2] proposed another solution approach for FFLPP using a new linear ranking function for defuzzification and found its equivalent form based on some proposed theoretical results. Ebrahimnejad [11] developed a solution approach for FFLPP, assuming the parameters as non-negative TFNs where FFLPP is formulated into tri-objective LPP. Khan, Ahmad, and Maan [17] proposed a modified version of simplex method for solving FFLPP, which utilizes both the concepts of ranking function and gauss elimination method to derive the optimal solution. Khalili, Nasseri, and Taghi-Nezhad [16] developed a fuzzy interactive approach based on the concept of membership function to solve a fully fuzzy mixed integer LPP. Daneshrad and Jafari [8] proposed an algorithm to find out the non-dominated solutions of an FFLPP by considering its parameters as symmetric trapezoidal fuzzy numbers. Das [9] proposed a modified algorithm to solve and derive fuzzy optimal solutions of an FFLPP containing equality constraints. Arana-Jiménez [3] proposed an algorithm to determine the non-dominated solutions of an FFLPP involving TFNs without applying ranking functions but an equivalent MOLPP is formulated, which generates the corresponding fuzzy optimal solutions. Pérez-Cañedo et al. [24] developed a solution approach for fully fuzzy linear regression problem comprising LR fuzzy numbers where this problem is transformed into an equivalent fully fuzzy MOLPP and fuzzy goal programming is used to find its solution. Bas and Ozkok [6] developed an iterative approach to solve a fully fuzzy linear fractional programming problem (LFPP) i.e., objective function exists as the ratio of linear functions, by transforming it to crisp MOLPP.

Some practical problems can be designed as optimization models with multiple goals that can be mathematically expressed as the problems of multi-objective optimization by the DMs. Fully fuzzy multi-objective LPP(FFMOLPP) has multiple conflicting linear objective functions with all its parameters and variables both in the objective functions and constraints are expressed as fuzzy numbers. Jimenez [4] proposed a method to solve FFMOLPP and derived fuzzy Pareto optimal solutions without using any ranking functions. Hop [28] studied the relationship between fuzzy numbers and solved a fully fuzzy multi-objective decision making problem based on the concepts of absolute fuzzy and relative fuzzy dominant degrees. Temelcan, Gonce Kocken, and Albayrak [27] also proposed an algorithm to solve an FFMOLPP where game theory approach was used for determining proper weights and solving the weighted LPP to find a fuzzy compromise solution. Pérez-Cañedo, Verdegay, and Miranda Perez [25] proposed fuzzy epsilon constraint method to derive fuzzy Pareto optimal solutions of FFMOLPP using lexicographic ranking process. Hamadameen and Hassan [14] developed a compromise solution algorithm to solve FFMOLPP by using revised simplex and Gaussian elimination methods. Sharma and Aggarwal [26] proposed a solution algorithm for FFMOLPP involving LR flat fuzzy numbers as the coefficients and decision variables, which converts the fuzzy problem into a crisp LPP using the concepts of scalarization methods and nearest interval approximation for fuzzy numbers. Yang, Cao, and Lin [30] proposed a method based on lexicographic order relation to solve an FFMOLPP with TFNs by converting it into multi-level MOLPP. Arana-Jiménez [5] developed an algorithm to obtain the fuzzy Pareto solutions of an FFMOLPP by using fuzzy arithmetics and partial orders to formulate its associated crisp MOLPP instead of applying ranking functions. Fathy and Hassanien [13] used a fuzzy harmonic approach to solve a fully fuzzy multi-level MOLPP by converting it into crisp MOLPP and derived its compromise solution. Malik and Gupta [20] proposed a method to solve fully triangular intuitionistic fuzzy multi-objective LFPP using goal programming approach and provided an application in e-education system. Niksirat [23] used nearest interval approximation concept to solve a fully fuzzy multi-objective transportation problem considering its parameters in uncertain forms.

MOLPP often arises in numerous practical problems of several domains comprising ambiguous information. In view of this, the paper studies FF-MOLPP with the following major objectives:

- Develop a computationally efficient solution methodology to solve FF-MOLPP in triangular fuzzy environment.
- Generate a set of fuzzy Pareto optimal(fuzzy non-dominated) solutions of FFMOLPP so that DM gets a choice to determine the compromise solution.

As it is observed, most of the existing methods deal with single objective fully fuzzy LPP or solve FFMOLPP using ranking function for converting it into crisp MOLPP or derives only one solution as the compromise solution of FFMOLPP. However, this paper studies a multi-objective FFLPP with all triangular fuzzy parameters and derives multiple fuzzy non-dominated solutions from which DM can choose the compromise solution. The proposed solution approach is also simple to formulate and computationally efficient, which is verified through comparisons of the computational results of numerical problems with some existing methods. The proposed study utilizes multiple concepts to solve FFMOLPP such as fuzzy α-cuts and centroid concept of TFNs, concept of interval analysis, which converts interval valued to real valued objective functions, weighting sum approach and ranking function for TFNs. Different values of fuzzy cuts maintain different degrees of satisfaction for the fuzzy parameters involved in the objective functions whereas the centroid concept of TFNs makes a better comparison in between both sides of the fuzzy valued constraints. This is a novel approach in context of utilizing the aforesaid concepts to efficiently solve and derive fuzzy Pareto optimal solutions of FFMOLPP.

This paper is organized as follows: Including introduction in Section 1, Section 2 incorporates the discussions on some preliminary concepts. Section 3 comprises the problem formulation of FFMOLPP and a description of the proposed methodology along with the presentation of its algorithm and flowchart. Section 4 contains the solutions of three numerical and one practical problems of FFMOLPP using the proposed methodology and their comparative result analysis. Finally, Section 5 contains the concluding remarks of this paper.

2 Certain preliminary concepts

Definition 1. [32] Let U be the universal set; then the fuzzy set $\widetilde{T}$ on U is defined as follows:

$$\widetilde{T} = \{(x, \mu_{\widetilde{T}}(x)) : x \in U\},$$

where $\mu_{\widetilde{T}} : U \to [0, 1]$ and $\mu_{\widetilde{T}}(x) \in [0, 1]$ represents the degree of membership of x in $\widetilde{T}$.

Definition 2. [32] A fuzzy set $\widetilde{T}$ is characterized as a fuzzy number if it satisfies the following properties:

- $\widetilde{T}$ must be defined over the real line $\mathbb{R}$.
- $\widetilde{T}$ must be normal, that is, here exists at least one element $x_0 \in \mathbb{R}$ such that $\mu_{\widetilde{T}}(x_0)$=1.
- $\widetilde{T}$ is a convex fuzzy set, that is, for all $x_1, x_2 \in \mathbb{R}$ and $\lambda \in [0, 1]$, $\mu_{\widetilde{T}}(\lambda x_1 + (1 - \lambda)\, x_2) \geq \min\{\mu_{\widetilde{T}}(x_1), \mu_{\widetilde{T}}(x_2)\}$
- The membership function $\mu_{\widetilde{T}}$ of the fuzzy set $\widetilde{T}$ in $\mathbb{R}$ is piecewise continuous.
- Support of the fuzzy set $\widetilde{T}$ in $\mathbb{R}$, $S(\widetilde{T}) = \{x|\mu_{\widetilde{T}}(x) > 0\}$ must be bounded.

Definition 3. [32] A TFN on $\mathbb{R}^+$ remains in the form of $\widetilde{T} = (t_1, t_2, t_3)$, where $t_1 \le t_2 \le t_3$ and $t_1, t_2, t_3 \in \mathbb{R}^+$. The associated membership function $\mu_{\widetilde{T}}$ can be mathematically defined as follows with its geometrical interpretation stated in Figure 1:

$$\mu_{\widetilde{T}}(x) = \begin{cases} \dfrac{x - t_1}{t_2 - t_1}, & t_1 \le x \le t_2 \ , \\ \dfrac{t_3 - x}{t_3 - t_2}, & t_2 \le x \le t_3, \\ 0, & otherwise. \end{cases} \tag{1}$$

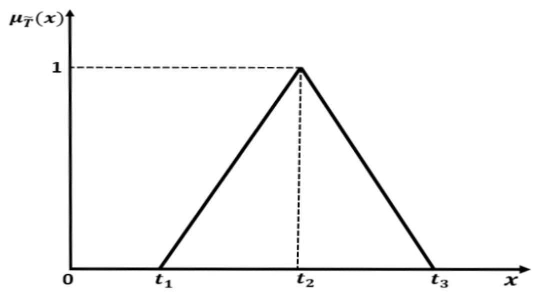


Figure 1: TFNs

Definition 4. Let $\widetilde{T}_1 = (t_1^l, t_1^m, t_1^n)$, $\widetilde{T}_2 = (t_2^l, t_2^m, t_2^n)$ be the TFNs on $\mathbb{R}$; then the arithmetic operations can be defined as follows:

- $\widetilde{T}_1 \oplus \widetilde{T}_2 = (t_1^l + t_2^l, t_1^m + t_2^m, t_1^n + t_2^n)$,
- $\widetilde{T}_1 \ominus \widetilde{T}_2 = (t_1^l - t_2^n, t_1^m - t_2^m, t_1^n - t_2^l)$,
- $\alpha \widetilde{T}_1 = \begin{cases} (\alpha t_1^l, \alpha t_1^m, \alpha t_1^n), & \alpha \ge 0, \\ (\alpha t_1^n, \alpha t_1^m, \alpha t_1^l), & \alpha < 0, \end{cases}$
- $\widetilde{T}_1 \otimes \widetilde{T}_2 = \begin{cases} (t_1^l t_2^l, t_1^m t_2^m, t_1^n t_2^n), & \widetilde{T}_1, \widetilde{T}_2 \in \mathbb{R}^+, \\ (t_1^n t_2^n, t_1^m t_2^m, t_1^l t_2^l), & \widetilde{T}_1, \widetilde{T}_2 \in \mathbb{R}^-, \\ (t_1^n t_2^l, t_1^m t_2^m, t_1^l t_2^n), & \widetilde{T}_1 \in \mathbb{R}^+, \widetilde{T}_2 \in \mathbb{R}^-, \\ (t_1^l t_2^n, t_1^m t_2^m, t_1^n t_2^l), & \widetilde{T}_1 \in \mathbb{R}^-, \widetilde{T}_2 \in \mathbb{R}^+, \end{cases}$

- $$\frac{\widetilde{T}_1}{\widetilde{T}_2} = \begin{cases} \left(\frac{t_1^l}{t_2^n}, \frac{t_1^m}{t_2^m}, \frac{t_1^n}{t_2^l}\right), & \widetilde{T}_1, \widetilde{T}_2 \in \mathbb{R}^+, \\ \left(\frac{t_1^n}{t_2^l}, \frac{t_1^m}{t_2^m}, \frac{t_1^l}{t_2^n}\right), & \widetilde{T}_1, \widetilde{T}_2 \in \mathbb{R}^-, \\ \left(\frac{t_1^n}{t_2^n}, \frac{t_1^m}{t_2^m}, \frac{t_1^l}{t_2^l}\right), & \widetilde{T}_1 \in \mathbb{R}^+, \widetilde{T}_2 \in \mathbb{R}^-, \\ \left(\frac{t_1^l}{t_2^l}, \frac{t_1^m}{t_2^m}, \frac{t_1^n}{t_2^n}\right), & \widetilde{T}_1 \in \mathbb{R}^-, \widetilde{T}_2 \in \mathbb{R}^+, \end{cases}$$

where the symbols $\oplus, \ominus$, and $\otimes$ denote addition, subtraction, and multiplication of fuzzy numbers, respectively.

Definition 5. [32] Fuzzy α-cut $\widetilde{T}_\alpha$ of the fuzzy set $\widetilde{T}$ can be defined as $\widetilde{T}_\alpha = \{x \in U : \mu_{\widetilde{T}}(x) \geq \alpha, \alpha \in [0,1]\}$. The fuzzy α-cut of a TFN $\widetilde{T} = (t_1, t_2, t_3)$ is defined as $\widetilde{T}_\alpha = [t_1 + \alpha(t_2 - t_1), t_3 - \alpha(t_3 - t_2)]$.

Definition 6. [22] Ranking function over the set of TFNs $\widetilde{T} = (t_1, t_2, t_3)$ in $\mathbb{R}$ can be defined as, $\mathcal{R}(\widetilde{T}) = \frac{t_1+2t_2+t_3}{4}$ for all $\widetilde{T} = (t_1, t_2, t_3) \in TFN(\mathbb{R})$.

Definition 7. [22] Let ABC be a triangle in xy-plane with its vertices $A(x_1, y_1)$, $B(x_2, y_2)$ and $C(x_3, y_3)$. Then its centroid is the point of intersections of the medians, which can be defined as $\left(\frac{x_1 + x_2 + x_3}{3}, \frac{y_1 + y_2 + y_3}{3}\right)$. The centroid of a TFN $\widetilde{T} = (t_1, t_2, t_3)$ is defined as $\left(\frac{t_1 + t_2 + t_3}{3}, \frac{1}{3}\right)$.

Theorem 1. [29] The optimal solution of the real valued optimization, $Max_{x \in \delta}\ F^L(x) + F^U(x)$ is a non-dominated solution of the interval valued optimization, $Max_{x \in \delta}\ [F^L(x), F^U(x)]$.

3 Problem formulation and methodology developed for FFMOLPP

FFMOLPP frequently arises in many real world problems. The use of TFNs are common and often required to address the informational ambiguity in these optimization models.

3.1 FFMOLPP formulation:

Consider the following FFMOLPP with p number of objective functions:

$$
\begin{aligned}
&Max\ \widetilde{Z}_1(\widetilde{X}) = \sum_{j=1}^{r} \widetilde{F}_{1j} \otimes \tilde{x}_j = \widetilde{F}_{11} \otimes \tilde{x}_1 \oplus \widetilde{F}_{12} \otimes \tilde{x}_2 \oplus \cdots \oplus \widetilde{F}_{1r} \otimes \tilde{x}_r \\
&Max\ \widetilde{Z}_2(\widetilde{X}) = \sum_{j=1}^{r} \widetilde{F}_{2j} \otimes \tilde{x}_j = \widetilde{F}_{21} \otimes \tilde{x}_1 \oplus \widetilde{F}_{22} \otimes \tilde{x}_2 \oplus \cdots \oplus \widetilde{F}_{2r} \otimes \tilde{x}_r \\
&\quad \vdots \\
&Max\ \widetilde{Z}_p(\widetilde{X}) = \sum_{j=1}^{r} \widetilde{F}_{pj} \otimes \tilde{x}_j = \widetilde{F}_{p1} \otimes \tilde{x}_1 \oplus \widetilde{F}_{p2} \otimes \tilde{x}_2 \oplus \cdots \oplus \widetilde{F}_{pr} \otimes \tilde{x}_r
\end{aligned} \tag{2}
$$

subject to

$$
\widetilde{\Delta}(\widetilde{X}) = \{\tilde{A}\widetilde{X}(\preccurlyeq, \approx, \succcurlyeq)\tilde{b}\} = \Big\{ \sum_{j=1}^{r} \tilde{a}_{ij}\tilde{x}_j(\preccurlyeq, \approx, \succcurlyeq)\tilde{b}_i \Big\},
$$

where $\widetilde{F}_{kj}$, $\widetilde{X}$=$(\tilde{x}_j)$, $\widetilde{a}_{ij}$, $\widetilde{b}_i \in TFN(\mathbb{R}^+)$ for $k = 1, 2, \ldots, p,\ j = 1, 2, \ldots, r,\ i = 1, 2, \ldots, q$ and $\preccurlyeq, \succcurlyeq, \approx$, represent the fuzziness in the inequalities and equality, respectively. Moreover, $\widetilde{Z}, \widetilde{X}, \widetilde{\Delta}$ represent the fuzzy or interval valued characteristics whereas Z, X, Δ represent the real valued characteristics throughout this paper.

Definition 8. [22] $\widetilde{X}^* \in \widetilde{\Delta}(\widetilde{X})$ is a fuzzy Pareto optimal solution or fuzzy non-dominated solution of the FFMOLPP (2) if and only if there is no another feasible solution $\widetilde{X}^{**} \in \widetilde{\Delta}(\widetilde{X})$ such that $\widetilde{Z}_r(\widetilde{X}^{**}) \succcurlyeq \widetilde{Z}_r(\widetilde{X}^*)$ for all $(r = 1, 2, 3, \ldots, p)$ and $\widetilde{Z}_s(\widetilde{X}^{**}) \succ \widetilde{Z}_s(\widetilde{X}^*)$ for at least one $s \in \{1, 2, \ldots, p\}$.

3.2 Developed solution methodology:

On substituting the values of decision variables and coefficients/constants, which exist in form of TFNs, the objective functions and constraints of the model (2) can be expressed as follows:

$$
\begin{aligned}
&Max\ \widetilde{Z}_1(\widetilde{X}) = \sum_{j=1}^{n} (c_{1j}^l, c_{1j}^m, c_{1j}^n) \otimes (x_j^l, x_j^m, x_j^n) \\
&Max\ \widetilde{Z}_2(\widetilde{X}) = \sum_{j=1}^{n} (c_{2j}^l, c_{2j}^m, c_{2j}^n) \otimes (x_j^l, x_j^m, x_j^n) \\
&\quad \vdots
\end{aligned}
$$

$$Max\ \widetilde{Z}_p(\widetilde{X}) = \sum_{j=1}^{n}(c_{pj}^{l}, c_{pj}^{m}, c_{pj}^{n}) \otimes (x_j^l, x_j^m, x_j^n) \tag{3}$$

subject to

$$\widetilde{\Delta}(\widetilde{X}) = \begin{cases} \sum_{j=1}^{r}(a_{ij}^{l}, a_{ij}^{m}, a_{ij}^{n}) \otimes (x_j^l, x_j^m, x_j^n)(\preccurlyeq, \approx, \succcurlyeq)(b_i^l, b_i^m, b_i^n), \ i = 1, 2, \ldots, q, \\ \tilde{x}_j \succcurlyeq 0, \ j = 1, 2, \ldots, r. \end{cases}$$

In $\widetilde{\Delta}(\widetilde{X})$, $\tilde{x}_j = (x_j^l, x_j^m, x_j^n) \succcurlyeq 0$, that is , $\tilde{x}_j \in TFN(\mathbb{R}^+)$ implies $x_j^l \geq 0$, $x_j^m \geq x_j^l$, $x_j^n \geq x_j^m$ and $\tilde{a}_{ij} = (a_{ij}^l, a_{ij}^m, a_{ij}^n), \tilde{x}_j = (x_j^l, x_j^m, x_j^n) \in TFN(\mathbb{R}^+)$ implies $\tilde{a}_{ij} \otimes \tilde{x}_j = (a_{ij}^l x_j^l, a_{ij}^m x_j^m, a_{ij}^n x_j^n)$ based on the arithmetic operations in Definition 4. So, $\widetilde{\Delta}(\widetilde{X})$ is converted into $\widetilde{\Delta}_1(\widetilde{X})$ and the fuzzy model (3) is equivalently transformed into the following MOLPP using fuzzy α-cuts of TFNs in the objective functions:

$$\begin{aligned} Max\ \widetilde{Z}_1(\widetilde{X}) &= \sum_{j=1}^{n}[c_{1j}^l + \alpha(c_{1j}^m - c_{1j}^l), c_{1j}^n - \alpha(c_{1j}^n - c_{1j}^m)] \\ &\quad \times [x_j^l + \alpha(x_j^m - x_j^l), x_j^n - \alpha(x_j^n - x_j^m)] \\ Max\ \widetilde{Z}_2(\widetilde{X}) &= \sum_{j=1}^{n}[c_{2j}^l + \alpha(c_{2j}^m - c_{2j}^l), c_{2j}^n - \alpha(c_{2j}^n - c_{2j}^m)] \\ &\quad \times [x_j^l + \alpha(x_j^m - x_j^l), x_j^n - \alpha(x_j^n - x_j^m)] \\ &\vdots \\ Max\ \widetilde{Z}_p(\widetilde{X}) &= \sum_{j=1}^{n}[c_{pj}^l + \alpha(c_{pj}^m - c_{pj}^l), c_{pj}^n - \alpha(c_{pj}^n - c_{1j}^m)] \\ &\quad \times [x_j^l + \alpha(x_j^m - x_j^l), x_j^n - \alpha(x_j^n - x_j^m)] \end{aligned} \tag{4}$$

subject to

$$\widetilde{\Delta}_1(\widetilde{X}) = \begin{cases} \sum_{j=1}^{r}(a_{ij}^l x_j^l, a_{ij}^m x_j^m, a_{ij}^n x_j^n)(\preccurlyeq, \approx, \succcurlyeq)(b_i^l, b_i^m, b_i^n), & i = 1, 2, \ldots, q, \\ x_j^l, x_j^m - x_j^l, x_j^n - x_j^m \geq 0, & j = 1, 2, \ldots, r. \end{cases}$$

To convert the triangular fuzzy valued inequalities $\widetilde{\Delta}_1(\widetilde{X})$ to its equivalent deterministic form, the concept of centroid of TFNs discussed in Definition 7 is utilized in the set of constraints $\widetilde{\Delta}_1(\widetilde{X})$.

$\sum_{j=1}^{r}(a_{ij}^l x_j^l, a_{ij}^m x_j^m, a_{ij}^n x_j^n)(\preccurlyeq, \approx, \succcurlyeq)(b_i^l, b_i^m, b_i^n)$,

that is, $\Big(\sum_{j=1}^{r} a_{ij}^l x_j^l, \sum_{j=1}^{r} a_{ij}^m x_j^m, \sum_{j=1}^{r} a_{ij}^n x_j^n\Big)(\preccurlyeq, \approx, \succcurlyeq)\Big(b_i^l, b_i^m, b_i^n\Big)$, that is, $\left(\frac{\sum\limits_{j=1}^{r} a_{ij}^{l} x_j^{l} + \sum\limits_{j=1}^{r} a_{ij}^{m} x_j^{m} + \sum\limits_{j=1}^{r} a_{ij}^{n} x_j^{n}}{3}, \frac{1}{3}\right)(\preccurlyeq, \approx, \succcurlyeq)\left(\frac{b_i^l + b_i^m + b_i^n}{3}, \frac{1}{3}\right)$,

that is, $\Big(\sum\limits_{j=1}^{r} a_{ij}^{l} x_j^{l} + \sum\limits_{j=1}^{r} a_{ij}^{m} x_j^{m} + \sum\limits_{j=1}^{r} a_{ij}^{n} x_j^{n}\Big)(\preccurlyeq, \approx, \succcurlyeq)\Big(b_i^l + b_i^m + b_i^n\Big)$.

Consequently, the following deterministic form of inequalities $\Delta_2(X)$ is derived for the constraints:

$$\Delta_2(X) = \begin{cases} \Big(\sum\limits_{j=1}^{r} a_{ij}^{l} x_j^{l} + \sum\limits_{j=1}^{r} a_{ij}^{m} x_j^{m} + \sum\limits_{j=1}^{r} a_{ij}^{n} x_j^{n}\Big)(\leq, =, \geq)(b_i^l + b_i^m + b_i^n), \\ \qquad\qquad\qquad\qquad\qquad\qquad i = 1, 2, \dots, q \\ x_j^l, x_j^m - x_j^l, x_j^n - x_j^m \geq 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad j = 1, 2, \dots, r. \end{cases}$$

The objective functions of the model (4) can be transformed into interval valued forms. The mathematical formulation of $\widetilde{Z}_1(\widetilde{X})$ as interval valued function, is expressed below:

$$\begin{aligned} \widetilde{Z}_1(\widetilde{X}) =& \sum_{j=1}^{n} [(1-\alpha)c_{1j}^l + \alpha c_{1j}^m, (1-\alpha)c_{1j}^n + \alpha c_{1j}^m] \\ & \quad \times [(1-\alpha)x_{1j}^l + \alpha x_{1j}^m, (1-\alpha)x_j^n + \alpha x_j^m] \\ =& \Bigg[\sum_{j=1}^{n} \{(1-\alpha)c_{1j}^l + \alpha c_{1j}^m\}\{(1-\alpha)x_{1j}^l + \alpha x_{1j}^m\}, \\ & \quad \sum_{j=1}^{n} \{(1-\alpha)c_{1j}^n + \alpha c_{1j}^m\}\{(1-\alpha)x_j^n + \alpha x_j^m\}\Bigg] \\ =& [Z_1^L(X), Z_1^U(X)]. \end{aligned} \tag{5}$$

Similarly, the remaining objective functions can be expressed in interval valued forms,

$$\widetilde{Z}_2(\widetilde{X}) = [Z_2^L(X), Z_2^U(X)], \widetilde{Z}_3(\widetilde{X}) = [Z_3^L(X), Z_3^U(X)], \dots, \widetilde{Z}_p(\widetilde{X}) = [Z_p^L(X), Z_p^U(X)]. \tag{6}$$

The FFMOLPP (2) can be expressed now as the following optimization model with interval valued objective functions and deterministic constraints:

$$
\begin{aligned}
&Max\ \widetilde{Z}_1(\widetilde{X}) = [Z_1^L(X), Z_1^U(X)]\\
&Max\ \widetilde{Z}_2(\widetilde{X}) = [Z_2^L(X), Z_2^U(X)]\\
&\quad\vdots\\
&Max\ \widetilde{Z}_p(\widetilde{X}) = [Z_p^L(X), Z_p^U(X)]
\end{aligned} \tag{7}
$$

subject to

$$
\Delta_2(X) = \begin{cases} (\sum\limits_{j=1}^{r} a_{ij}^l x_j^l + \sum\limits_{j=1}^{r} a_{ij}^m x_j^m + \sum\limits_{j=1}^{r} a_{ij}^n x_j^n)(\leq, =, \geq)(b_i^l + b_i^m + b_i^n), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1, 2, \ldots, q, \\ x_j^l, x_j^m - x_j^l, x_j^n - x_j^m \geq 0, \qquad\qquad\qquad j = 1, 2, \ldots, r. \end{cases}
$$

The multi-objective interval valued optimization (7) can be scalarized into the following single objective interval valued optimization using weighting sum approach with the weight vectors $(w_1, w_2, \ldots, w_p)$:

$$
\begin{aligned}
Max\ \widetilde{Z}(\widetilde{X}) =& w_1[Z_1^L(X), Z_1^U(X)] + w_2[Z_2^L(X), Z_2^U(X)]\\
&+ \cdots + w_p[Z_p^L(X), Z_p^U(X)]
\end{aligned} \tag{8}
$$

subject to

$$
\Delta_2(X) = \begin{cases} (\sum\limits_{j=1}^{r} a_{ij}^l x_j^l + \sum\limits_{j=1}^{r} a_{ij}^m x_j^m + \sum\limits_{j=1}^{r} a_{ij}^n x_j^n)(\leq, =, \geq)(b_i^l + b_i^m + b_i^n), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1, 2, \ldots, q, \\ x_j^l, x_j^m - x_j^l, x_j^n - x_j^m \geq 0, w_k > 0, \sum\limits_{k=1}^{p} w_k = 1, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j = 1, 2, \ldots, r. \end{cases}
$$

The objective function of the optimization model (8) can be expressed in the following form.:

$$
\begin{aligned}
\widetilde{Z}(\widetilde{X}) =& \Big[w_1 Z_1^L(X) + w_2 Z_2^L(X) + \cdots + w_p Z_p^L(X),\ w_1 Z_1^U(X)\\
&+ w_2 Z_2^U(X) + \cdots + w_p Z_p^U(X)\Big].
\end{aligned}
$$

That is, $\tilde{Z}(\widetilde{X}) = \Big[\sum\limits_{i=1}^{p} w_i Z_i^L(X),\ \sum\limits_{i=1}^{p} w_i Z_i^U(X)\Big]$.

The optimization model (8) can be further simplified as,

$$Max_{X\in\Delta_2(X)} \ \ \widetilde{Z}(\widetilde{X}) = \Big[\sum_{i=1}^{p} w_i Z_i^L(X),\ \sum_{i=1}^{p} w_i Z_i^U(X)\Big]. \tag{9}$$

Based on Theorem 1, it can be clearly proved that the optimal solution of the following model (10) is a non-dominated solution of the optimization model (9):

$$Max_{X\in\Delta_2(X)} \ \ Z(X) = \sum_{i=1}^{p} w_i Z_i^L(X) + \sum_{i=1}^{p} w_i Z_i^U(X). \tag{10}$$

The FFMOLPP (2) is equivalently converted into the optimization model (8), which is next simplified as the model (9). Since an optimal solution of (10) is a non-dominated solution of the model (9), so it is also considered as non-dominated for the FFMOLPP (2).

Finally, the FFMOLPP (2) can be transformed into the following single objective deterministic LPP on applying Theorem 1 and simplifying the model (10):

$$\begin{aligned} Max\ Z(X) =& w_1[Z_1^L(X) + Z_1^U(X)] + w_2[Z_2^L(X) + Z_2^U(X)] \\ &+ \cdots + w_p[Z_p^L(X) + Z_p^U(X)] \end{aligned} \tag{11}$$

subject to

$$\Delta_2(X) = \begin{cases} (\sum\limits_{j=1}^{r} a_{ij}^l x_j^l + \sum\limits_{j=1}^{r} a_{ij}^m x_j^m + \sum\limits_{j=1}^{r} a_{ij}^n x_j^n)(\leq,=,\geq)(b_i^l + b_i^m + b_i^n), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1,2,\ldots,q, \\ x_j^l, x_j^m - x_j^l, x_j^n - x_j^m \geq 0, \\ w_k > 0, \sum\limits_{k=1}^{p} w_k = 1, \qquad\qquad\qquad\qquad j = 1,2,\ldots,r. \end{cases}$$

Solving the model (11), a set of non-dominated solutions can be generated for the FFMOLPP (2) by substituting different values of $\alpha \in [0,1]$ and weights $w_k > 0$ with $\sum\limits_{k=1}^{p} w_k = 1$.

3.3 Algorithm and flowchart presentation of the developed methodology:

An algorithm and a flowchart Figure 2 are presented below, describing the details of the proposed methodology.

Step 1. Use fuzzy α-cuts and arithmetic operations of TFNs in the objective functions and constraints of FFMOLPP (3), respectively.

Step 2. Simplify the objective functions in the form of the model (4) and the set of constraints as $\widetilde{\Delta}_1(\widetilde{X})$.

Step 3. Linearize the constraints of $\widetilde{\Delta}_1(\widetilde{X})$ as $\Delta_2(X)$ by using the concept of centroid of TFNs.

Step 4. Express the FFMOLPP as the optimization model (7) with interval valued objective functions and deterministic constraints.

Step 5. Use weighting sum method to scalarize the model (7) in form of single objective optimization (8).

Step 6. Apply Theorem 1 in order to transform the problem (8) to its equivalent model (11) comprising real valued objective function $Z(X)$ and deterministic constraints $\Delta_2(X)$.

Step 7. Solve the optimization model (11) substituting different values of $\alpha \in [0,1]$ and weights $w_k > 0$, $\sum\limits_{k=1}^{p} w_k = 1$. The optimal solutions obtained are considered as the non-dominated solutions of the FFMOLPP (2).

Step 8. If DM is unsatisfied with the obtained non-dominated solutions, then reformulate and solve model (11) by changing $\alpha \in [0,1]$ and weight vector (w_k).

4 Numerical examples

In order to illustrate and validate the feasibility of the proposed methodology, two existing numerical and one practical problems in form of FFMOLPP are solved and the comparative study on its result analysis is incorporated.

Example 1. Consider the following FFMOLPP, which is initially solved by Aggarwal and Sharma [1], Temelcan, Gonce Kocken, and Albayrak [27]:

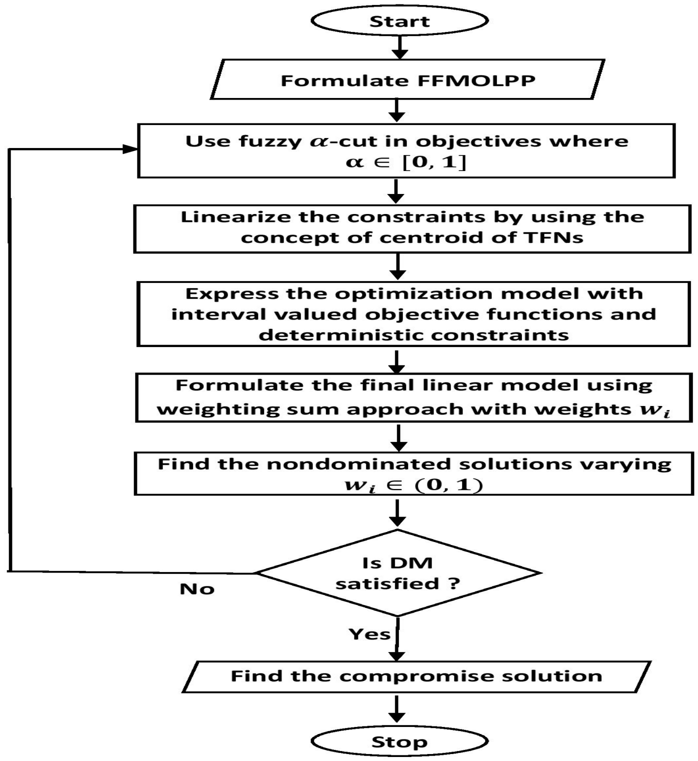


Figure 2: Flowchart of proposed solution approach

$$
\begin{aligned}
&Max\ \widetilde{Z}_1(\widetilde{X}) = (1,2,3)\otimes\tilde{x}_1 \oplus (2,4,5)\otimes\tilde{x}_2\\
&Max\ \widetilde{Z}_2(\widetilde{X}) = (2,3,4)\otimes\tilde{x}_1 \oplus (3,4,5)\otimes\tilde{x}_2\\
&\text{subject to}\\
&(0,1,2)\otimes\tilde{x}_1 \oplus (1,2,3)\otimes\tilde{x}_2 \preccurlyeq (1,10,27)\\
&(1,2,3)\otimes\tilde{x}_1 \oplus (0,1,2)\otimes\tilde{x}_2 \preccurlyeq (2,11,28)\\
&\tilde{x}_1, \tilde{x}_2 \succcurlyeq 0,
\end{aligned}
\tag{12}
$$

Solution:

According to the proposed methodology, the FFMOLPP (12) is equivalently transformed into the following optimization model with fuzzy constraints based on fuzzy α-cuts in objective functions and arithmetic operations of TFNs:

$$\begin{aligned}
Max\ \widetilde{Z}_1(\widetilde{X}) =& [1+\alpha, 3-\alpha][(1-\alpha)x_1^l + \alpha x_1^m, (1-\alpha)x_1^n + \alpha x_1^m] \\
&+ [2+2\alpha, 5-\alpha][(1-\alpha)x_2^l + \alpha x_2^m, (1-\alpha)x_2^n + \alpha x_2^m] \\
Max\ \widetilde{Z}_2(\widetilde{X}) =& [2+\alpha, 4-\alpha][(1-\alpha)x_1^l + \alpha x_1^m, (1-\alpha)x_1^n + \alpha x_1^m] \\
&+ [3+\alpha, 5-\alpha][(1-\alpha)x_2^l + \alpha x_2^m, (1-\alpha)x_2^n + \alpha x_2^m]
\end{aligned} \tag{13}$$

subject to

$$\begin{aligned}
&(x_2^l, x_1^m + 2x_2^m, 2x_1^n + 3x_2^n) \preccurlyeq (1, 10, 27) \\
&(x_1^l, 2x_1^m + x_2^m, 3x_1^n + 2x_2^n) \preccurlyeq (2, 11, 28) \\
&x_1^l, x_2^l, x_1^m - x_1^l, x_2^m - x_2^l, x_1^n - x_1^m, x_2^n - x_2^m \geq 0.
\end{aligned}$$

Using centroid concept of TFNs in constraints and arithmetic operations, the model (13) is expressed as the following interval valued optimization with deterministic linear constraints:

$$\begin{aligned}
Max\ \widetilde{Z}_1(\widetilde{X}) =& [(1-\alpha^2)x_1^l + \alpha(1+\alpha)x_1^m + (1-\alpha)(2+2\alpha)x_2^l + \alpha(2+2\alpha)x_2^m, \\
&(3-\alpha)(1-\alpha)x_1^n + \alpha(3-\alpha)x_1^m + (1-\alpha)(5-\alpha)x_2^n + \alpha(5-\alpha)x_2^m] \\
Max\ \widetilde{Z}_2(\widetilde{X}) =& [(1-\alpha)(2+\alpha)x_1^l + \alpha(2+\alpha)x_1^m + (1-\alpha)(3+\alpha)x_2^l + \alpha(3+\alpha)x_2^m, \\
&(1-\alpha)(4-\alpha)x_1^n + \alpha(4-\alpha)x_1^m + (1-\alpha)(5-\alpha)x_2^n + \alpha(5-\alpha)x_2^m]
\end{aligned} \tag{14}$$

subject to

$$\begin{aligned}
&x_1^m + 2x_1^n + x_2^l + 2x_2^m + 3x_2^n \leq 38 \\
&x_1^l + 2x_1^m + 3x_1^n + x_2^m + 2x_2^n \leq 41 \\
&x_1^l, x_2^l, x_1^m - x_1^l, x_2^m - x_2^l, x_1^n - x_1^m, x_2^n - x_2^m \geq 0.
\end{aligned}$$

The FFMOLPP (14) is finally formulated as the following single objective optimization using the weight vector (w_1, w_2):

$$\begin{aligned}
Max\ Z(X) =& w_1[(1-\alpha^2)x_1^l + 4\alpha x_1^m + (3-\alpha)(1-\alpha)x_1^n \\
&+ (1-\alpha)(2+2\alpha)x_2^l + \alpha(7+\alpha)x_2^m + (1-\alpha)(5-\alpha)x_2^n] \\
&+ w_2[(1-\alpha)(2+\alpha)x_1^l + 6\alpha x_1^m + (1-\alpha)(4-\alpha)x_1^n \\
&+ (1-\alpha)(3+\alpha)x_2^l + 8\alpha x_2^m + (1-\alpha)(5-\alpha)x_2^n]
\end{aligned}$$

subject to (15)

$$\begin{aligned}
&x_1^m + 2x_1^n + x_2^l + 2x_2^m + 3x_2^n \leq 38 \\
&x_1^l + 2x_1^m + 3x_1^n + x_2^m + 2x_2^n \leq 41 \\
&x_1^l, x_2^l, x_1^m - x_1^l, x_2^m - x_2^l, x_1^n - x_1^m, x_2^n - x_2^m \geq 0 \\
&w_1, w_2 > 0, w_1 + w_2 = 1.
\end{aligned}$$

On solving problem (15) with different $\alpha \in [0,1]$ and weights $w_1, w_2 > 0, w_1 + w_2 = 1$, a set of optimal solutions are derived, which are considered as the non-dominated solutions of the FFMOLPP (12).
The triangular fuzzy optimal objective values are computed along with their corresponding real valued expressions using ranking function and these values are shown in the following Table 1.

Example 2. Consider the following FFMOLPP, which is initially solved by Yang, Cao, and Lin [30]:

$$\begin{aligned}
&Min\ \widetilde{Z}_1(\widetilde{X}) = (7,10,11)\otimes\tilde{x}_1 \oplus (8,10,13)\otimes\tilde{x}_2\\
&Min\ \widetilde{Z}_2(\widetilde{X}) = (2,3,4)\otimes\tilde{x}_1 \oplus (4,7,12)\otimes\tilde{x}_2\\
&\text{subject to}\\
&(1,2,4)\otimes\tilde{x}_1 \oplus (2,8,10)\otimes\tilde{x}_2 \approx (3,22,36)\\
&(2,3,6)\otimes\tilde{x}_1 \oplus (4,10,15)\otimes\tilde{x}_2 \approx (6,29,54)\\
&\tilde{x}_1, \tilde{x}_2 \succcurlyeq 0.
\end{aligned} \tag{16}$$

Solution:
The FFMOLPP (16) is formulated into the following optimization model using fuzzy α-cuts in objective functions and arithmetic operations of TFNs in constraints:

$$\begin{aligned}
Min\ \widetilde{Z}_1(\widetilde{X}) =& [7+3\alpha, 11-\alpha][(1-\alpha)x_1^l + \alpha x_1^m, (1-\alpha)x_1^n + \alpha x_1^m]\\
&+ [8+2\alpha, 13-3\alpha][(1-\alpha)x_2^l + \alpha x_2^m, (1-\alpha)x_2^n + \alpha x_2^m]\\
Min\ \widetilde{Z}_2(\widetilde{X}) =& [2+\alpha, 4-\alpha][(1-\alpha)x_1^l + \alpha x_1^m, (1-\alpha)x_1^n + \alpha x_1^m]\\
&+ [4+3\alpha, 12-5\alpha][(1-\alpha)x_2^l + \alpha x_2^m, (1-\alpha)x_2^n + \alpha x_2^m]\\
\text{subject to}&\\
&(x_1^l + 2x_2^l, 2x_1^m + 8x_2^m, 4x_1^n + 10x_2^n) \approx (3,22,36)\\
&(2x_1^l + 4x_2^l, 3x_1^m + 10x_2^m, 6x_1^n + 15x_2^n) \approx (6,29,54)\\
&x_1^l, x_2^l, x_1^m - x_1^l, x_2^m - x_2^l, x_1^n - x_1^m, x_2^n - x_2^m \geq 0.
\end{aligned} \tag{17}$$

The above model (17) is equivalently expressed into the following model comprising interval valued objective functions and deterministic linear constraints based on the centroid concept:

Table 1: Non-dominated solutions, fuzzy valued optimal objective functions and their real valued expressions of Example 1

α	w	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{Z}_1$	$\tilde{Z}_2$	$R(\tilde{Z}_1)$	$R(\tilde{Z}_2)$
0.1	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(3.916667,3.916667, 3.916667)	(0,0,8.75)	(3.916667,7.8333, 55.5)	(7.8333,11.75, 59.4167)	18.7708	22.6875
	(0.8,0.2)	(0,0,0)	(0,0,12.667)	(0,0,63.335)	(0,0,63.335)	15.8338	15.8338
0.2	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(4.88889,4.88889, 4.88889)	(3.88889,3.88889, 3.88889)	(12.6667,25.3333, 34.1111)	(21.4445,30.2222, 39)	24.3611	30.2222
	(0.8,0.2)	(3.916667,3.916667, 3.916667)	(0,0,8.75)	(3.916667,7.8333, 55.5)	(7.8333,11.75, 59.4167)	18.7708	22.6875
0.3	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(4.88889,4.88889, 4.88889)	(3.88889,3.88889, 3.88889)	(12.6667,25.3333, 34.1111)	(21.4445,30.2222, 39)	24.3611	30.2222
0.4	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(4.88889,4.88889, 4.88889)	(3.88889,3.88889, 3.88889)	(12.6667,25.3333, 34.1111)	(21.4445,30.2222, 39)	24.3611	30.2222
0.5	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(4.88889,4.88889, 4.88889)	(3.88889,3.88889, 3.88889)	(12.6667,25.3333, 34.1111)	(21.4445,30.2222, 39)	24.3611	30.2222
0.6	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(4.88889,4.88889, 4.88889)	(3.88889,3.88889, 3.88889)	(12.6667,25.3333, 34.1111)	(21.4445,30.2222, 39)	24.3611	30.2222
0.7	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(4.333333,4.333333 ,4.333333)	(0,5,5)	(4.333333,28.6667, 38)	(8.6667,33, 42.3333)	24.9167	29.25
0.8	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(4.333333,4.333333 ,4.333333)	(0,5,5)	(4.333333,28.6667, 38)	(8.6667,33, 42.3333)	24.9167	29.25
0.9	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(4.333333,4.333333 ,4.333333)	(0,5,5)	(4.333333,28.6667, 38)	(8.6667,33, 42.3333)	24.9167	29.25
	(0.8,0.2)	(0,0,0)	(0,7.6,7.6)	(0,30.4,38)	(0,30.4,38)	24.7	24.7

$$\begin{aligned}
Min\ \widetilde{Z}_1(\widetilde{X}) =&[(7+3\alpha)(1-\alpha)x_1^l+\alpha(7+3\alpha)x_1^m+(8+2\alpha)(1-\alpha)x_2^l+\alpha(8+2\alpha)x_2^m,\\
&(11-\alpha)(1-\alpha)x_1^n+\alpha(11-\alpha)x_1^m+(13-3\alpha)(1-\alpha)x_2^n+\alpha(13-3\alpha)x_2^m]\\
Min\ \widetilde{Z}_2(\widetilde{X}) =&[(2+\alpha)(1-\alpha)x_1^l+\alpha(2+\alpha)x_1^m+(4+3\alpha)(1-\alpha)x_2^l+\alpha(4+3\alpha)x_2^m,\\
&(4-\alpha)(1-\alpha)x_1^n+\alpha(4-\alpha)x_1^m+(12-5\alpha)(1-\alpha)x_2^n+\alpha(12-5\alpha)x_2^m]
\end{aligned}$$

subject to

$$\begin{aligned}
&x_1^l+2x_1^m+4x_1^n+2x_2^l+8x_2^m+10x_2^n=61\\
&2x_1^l+3x_1^m+6x_1^n+4x_2^l+10x_2^m+15x_2^n=89\\
&x_1^l,x_2^l,x_1^m-x_1^l,x_2^m-x_2^l,x_1^n-x_1^m,x_2^n-x_2^m\geq 0.
\end{aligned} \tag{18}$$

The FFMOLPP (16) is finally formulated into the following single objective optimization according to the proposed methodology:

$$\begin{aligned}
Min\ Z(X) = w_1[&(7+3\alpha)(1-\alpha^2)x_1^l+\alpha(18+2\alpha)x_1^m+(11-\alpha)(1-\alpha)x_1^n\\
&+(8+2\alpha)(1-\alpha)x_2^l+\alpha(21-\alpha)x_2^m+(13-3\alpha)(1-\alpha)x_2^n]\\
&+w_2[(2+\alpha)(1-\alpha)x_1^l+6\alpha x_1^m+(4-\alpha)(1-\alpha)x_1^n\\
&+(4+3\alpha)(1-\alpha)x_2^l+\alpha(16-2\alpha)x_2^m+(12-5\alpha)(1-\alpha)x_2^n]
\end{aligned}$$

subject to

$$\begin{aligned}
&x_1^l+2x_1^m+4x_1^n+2x_2^l+8x_2^m+10x_2^n=61\\
&2x_1^l+3x_1^m+6x_1^n+4x_2^l+10x_2^m+15x_2^n=89\\
&x_1^l,x_2^l,x_1^m-x_1^l,x_2^m-x_2^l,x_1^n-x_1^m,x_2^n-x_2^m\geq 0\\
&w_1,w_2>0,w_1+w_2=1.
\end{aligned} \tag{19}$$

Substituting different $\alpha\in[0,1]$, weights $(w_1,w_2)>0$ with $w_1+w_2=1$ in (19) and solving each resulted optimization model, a set of optimal solutions are obtained, which are considered as the non-dominated solutions of the FFMOLPP (16). The triangular fuzzy optimal objective values and their corresponding real valued expressions are computed based on the concept of ranking function, which are shown in the following Table 2.

Example 3. Consider the following FFMOLPP, which is initially solved by Arana-Jiménez [5]:

Table 2: Non-dominated solutions, fuzzy valued optimal objective functions and their real valued expressions of Example 2

α	w	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{Z}_1$	$\tilde{Z}_2$	$R(\tilde{Z}_1)$	$R(\tilde{Z}_2)$
0.1	(0.2,0.8)	(0,6.41667,6.41667)	(0,1.25,1.25)	(0,76.6667,86.8333)	(0,28,40.6667)	60.0417	24.1667
	(0.4,0.6)	(0,0,0)	(2.5,2.5,3.6)	(20,25,46.8)	(10,17.5,43.2)	29.2	22.05
	(0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675
0.2	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675
0.3	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675
0.4	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675
0.5	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675
0.6	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675
0.7	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675
0.8	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675
0.9	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,1.25,5.1)	(0,12.5,66.3)	(0,8.75,61.2)	22.825	19.675

$$
\begin{aligned}
&Max\ \widetilde{Z}_1(\widetilde{X}) = (\frac{7}{5},4,\frac{43}{7}) \otimes \tilde{x}_1 \oplus (5,7,12) \otimes \tilde{x}_2 \oplus (\frac{39}{4},11,\frac{33}{2}) \otimes \tilde{x}_3\\
&Max\ \widetilde{Z}_2(\widetilde{X}) = (3,4,6) \otimes \tilde{x}_1 \oplus (\frac{10}{3},5,9) \otimes \tilde{x}_2 \oplus (4,5,10) \otimes \tilde{x}_3\\
&\text{subject to}\\
&(2,5,8) \otimes \tilde{x}_1 \oplus (3,\frac{41}{6},10) \otimes \tilde{x}_2 \oplus (5,\frac{31}{3},18) \otimes \tilde{x}_3 \preceq (6,\frac{50}{3},30) \qquad (20)\\
&(4,\frac{32}{3},12) \otimes \tilde{x}_1 \oplus (5,\frac{73}{6},20) \otimes \tilde{x}_2 \oplus (7,\frac{105}{6},30) \otimes \tilde{x}_3 \preceq (10,30,50)\\
&(3,5,7) \otimes \tilde{x}_1 \oplus (5,15,20) \otimes \tilde{x}_2 \oplus (5,10,15) \otimes \tilde{x}_3 \preceq (2,\frac{145}{6},30)\\
&\tilde{x}_1,\tilde{x}_2,\tilde{x}_3 \succcurlyeq 0.
\end{aligned}
$$

Solution:

The FFMOLPP (20) is formulated into the following model using fuzzy α-cuts in the objective functions and arithmetic operations of TFNs in the constraints:

$$
\begin{aligned}
Max\ \widetilde{Z}_1(\widetilde{X}) = {}&[\frac{7}{5}+2.6\alpha,\frac{43}{7}-2.1429\alpha][(1-\alpha)x_1^l+\alpha x_1^m,(1-\alpha)x_1^n+\alpha x_1^m]\\
&+[5+2\alpha,12-5\alpha][(1-\alpha)x_2^l+\alpha x_2^m,(1-\alpha)x_2^n+\alpha x_2^m]\\
&+[\frac{39}{4}+1.25\alpha,\frac{33}{2}-5.5\alpha][(1-\alpha)x_3^l+\alpha x_3^m,(1-\alpha)x_3^n+\alpha x_3^m]\\
Max\ \widetilde{Z}_2(\widetilde{X}) = {}&[3+\alpha,6-2\alpha][(1-\alpha)x_1^l+\alpha x_1^m,(1-\alpha)x_1^n+\alpha x_1^m]\\
&+[\frac{10}{3}+1.6667\alpha,9-4\alpha][(1-\alpha)x_2^l+\alpha x_2^m,(1-\alpha)x_2^n+\alpha x_2^m]\\
&+[4+\alpha,10-5\alpha][(1-\alpha)x_3^l+\alpha x_3^m,(1-\alpha)x_3^n+\alpha x_3^m]
\end{aligned}
$$

subject to

$$
\begin{aligned}
&(2x_1^l+3x_2^l+5x_3^l,5x_1^m+\frac{41}{6}x_2^m+\frac{31}{3}x_3^m,8x_1^n+10x_2^n+18x_3^n) \preceq (6,\frac{50}{3},30)\\
&(4x_1^l+5x_2^l+7x_3^l,\frac{32}{3}x_1^m+\frac{73}{6}x_2^m+\frac{105}{6}x_3^m,12x_1^n+20x_2^n+30x_3^n) \preceq (10,30,50)\\
&(3x_1^l+5x_2^l+5x_3^l,5x_1^m+15x_2^m+10x_3^m,7x_1^n+20x_2^n+15x_3^n) \preceq (2,\frac{145}{6},30)\\
&x_1^l,x_2^l,x_3^l,x_1^m-x_1^l,x_2^m-x_2^l,x_3^m-x_3^l,x_1^n-x_1^m,x_2^n-x_2^m,x_3^n-x_3^m \geq 0. \qquad (21)
\end{aligned}
$$

The model (21) is equivalently converted into the following interval valued optimization model, which contains the following deterministic linear constraints using the concept of centroid of TFNs:

$$
\begin{aligned}
Max\ \widetilde{Z}_1(\widetilde{X}) =&[(\frac{7}{5}+2.6\alpha)(1-\alpha)x_1^l+\alpha(\frac{7}{5}+2.6\alpha)x_1^m\\
&+(5+2\alpha)(1-\alpha)x_2^l+\alpha(5+2\alpha)x_2^m+(\frac{39}{4}+1.25\alpha)(1-\alpha)x_3^l\\
&+\alpha(\frac{39}{4}+1.25\alpha)x_3^m,\\
&(\frac{43}{7}-2.1429\alpha)(1-\alpha)x_1^n+\alpha(\frac{43}{7}-2.1429\alpha)x_1^m+(12-5\alpha)(1-\alpha)x_2^n\\
&+\alpha(12-5\alpha)x_2^m+(\frac{33}{2}-5.5\alpha)(1-\alpha)x_3^n+\alpha(\frac{33}{2}-5.5\alpha)x_3^m]\\
Max\ \widetilde{Z}_2(\widetilde{X}) =&[(3+\alpha)(1-\alpha)x_1^l+\alpha(3+\alpha)x_1^m+(\frac{10}{3}+1.6667\alpha)(1-\alpha)x_2^l\\
&+\alpha(\frac{10}{3}+1.6667\alpha)x_2^m+(4+\alpha)(1-\alpha)x_3^l+\alpha(4+\alpha)x_3^m,\\
&(6-2\alpha)(1-\alpha)x_1^n+\alpha(6-2\alpha)x_1^m+(9-4\alpha)(1-\alpha)x_2^n\\
&+\alpha(9-4\alpha)x_2^m+(10-5\alpha)(1-\alpha)x_3^n+\alpha(10-5\alpha)x_3^m]
\end{aligned} \tag{22}
$$

subject to

$$
\begin{aligned}
&2x_1^l+5x_1^m+8x_1^n+3x_2^l+\frac{41}{6}x_2^m+10x_2^n+5x_3^l+\frac{31}{3}x_3^m+18x_3^n\le 52.6667\\
&4x_1^l+\frac{32}{3}x_1^m+12x_1^n+5x_2^l+\frac{73}{6}x_2^m+20x_2^n+7x_3^l+\frac{105}{6}x_3^m+30x_3^n\le 90\\
&3x_1^l+5x_1^m+7x_1^n+5x_2^l+15x_2^m+20x_2^n+5x_3^l+10x_3^m+15x_3^n\le 56.1667\\
&x_1^l,x_2^l,x_3^l,x_1^m-x_1^l,x_2^m-x_2^l,x_3^m-x_3^l,x_1^n-x_1^m,x_2^n-x_2^m,x_3^n-x_3^m\ge 0.
\end{aligned}
$$

The FFMOLPP (22) is finally expressed as the following single objective deterministic linear optimization model based on the concepts of the proposed methodology:

$$
\begin{aligned}
&Max\ Z(X)\\
&=w_1[(\frac{7}{5}+2.6\alpha)(1-\alpha)x_1^l+\alpha(7.5429+0.4571\alpha)x_1^m+(\frac{43}{7}-2.1429\alpha)(1-\alpha)x_1^n\\
&\quad+(5+2\alpha)(1-\alpha)x_2^l+\alpha(17-3\alpha)x_2^m+(12-5\alpha)(1-\alpha)x_2^n\\
&\quad+(\frac{39}{4}+1.25\alpha)(1-\alpha)x_3^l+\alpha(26.25-4.25\alpha)x_3^m+(\frac{33}{2}-5.5\alpha)(1-\alpha)x_3^n]\\
&\quad+w_2[(3+\alpha)(1-\alpha)x_1^l+\alpha(9-\alpha)x_1^m+(6-2\alpha)(1-\alpha)x_1^n\\
&\quad+(\frac{10}{3}+1.6667\alpha)(1-\alpha)x_2^l+\alpha(12.3333-2.3333\alpha)x_2^m+(9-4\alpha)(1-\alpha)x_2^n\\
&\quad+(4+\alpha)(1-\alpha)x_3^l+\alpha(14-4\alpha)x_3^m+(10-5\alpha)(1-\alpha)x_3^n]
\end{aligned}
$$

subject to

$$2x_1^l + 5x_1^m + 8x_1^n + 3x_2^l + \frac{41}{6}x_2^m + 10x_2^n + 5x_3^l + \frac{31}{3}x_3^m + 18x_3^n \leq 52.6667$$
$$4x_1^l + \frac{32}{3}x_1^m + 12x_1^n + 5x_2^l + \frac{73}{6}x_2^m + 20x_2^n + 7x_3^l + \frac{105}{6}x_3^m + 30x_3^n \leq 90$$
$$3x_1^l + 5x_1^m + 7x_1^n + 5x_2^l + 15x_2^m + 20x_2^n + 5x_3^l + 10x_3^m + 15x_3^n \leq 56.1667$$
$$x_1^l, x_2^l, x_3^l, x_1^m - x_1^l, x_2^m - x_2^l, x_3^m - x_3^l, x_1^n - x_1^m, x_2^n - x_2^m, x_3^n - x_3^m \geq 0,$$
$$w_1, w_2 > 0, w_1 + w_2 = 1. \tag{23}$$

On solving the problem (23) with different values of $\alpha \in [0,1]$ and different weights $w_1,\ w_2 > 0$ satisfying $w_1 + w_2 = 1$, a set of solutions are generated, which are considered as the non-dominated solutions of the FFMOLPP (20). The triangular fuzzy valued optimal objective values are computed at the non-dominated solutions, which are shown in the Table 3.

Example 4 (Practical problem)**.** A company manufactures two types of products A and B with the profit amount of Rs. 5 and Rs. 3 per unit, respectively. The import of the products A and B are required as two units and eight units, respectively, in a day. The processing time required to manufacture the products A and B are 2hr. and 3hr. per unit, respectively, in a day not exceeding 5hr and the raw materials required are 4kg. and 5kg. per unit, respectively, not exceeding 10kg. The company aims to maximize the profit and maximize the import of products. The parameters and variables are considered as fuzzy numbers because the data varies due to insufficient labourers, supply of raw materials, defective machine, bad weather condition, etc.

Consider the formulation of the above problem as the following FFMOLPP:

$$\begin{aligned}
&Max\ \widetilde{Z}_1(\widetilde{X}) = (4,5,6)\otimes\tilde{x}_1 \oplus (2,3,4)\otimes\tilde{x}_2\\
&Max\ \widetilde{Z}_2(\widetilde{X}) = (1,2,3)\otimes\tilde{x}_1 \oplus (7,8,9)\otimes\tilde{x}_2\\
&\text{subject to}\\
&(3,4,5)\otimes\tilde{x}_1 \oplus (4,5,7)\otimes\tilde{x}_2 \preceq (9,10,11)\\
&(1,2,4)\otimes\tilde{x}_1 \oplus (1,3,5)\otimes\tilde{x}_2 \preceq (3,5,7)\\
&\tilde{x}_1, \tilde{x}_2 \succcurlyeq 0
\end{aligned} \tag{24}$$

Solution:

According to the proposed methodology, FFMOLPP (24) is expressed into

Table 3: Non-dominated solutions and fuzzy valued optimal objective values of Example 3

α	w	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\tilde{Z}_1$	$\tilde{Z}_2$	$R(\tilde{Z}_1)$	$R(\tilde{Z}_2)$
0.1	(0.2,0.8) (0.4,0.6) (0.5,0.5)	(0,0,5.462967)	(0,0,0.8962967)	(0,0,0)	(0,0,44.3138)	(0,0,40.8445)	11.0785	10.2111
	(0.6,0.4)	(0,0,0.90004)	(0,0,1.026666)	(0,0,1.95554)	(0,0,50.1152)	(0,0,34.1956)	12.5288	8.5489
	(0.8,0.2)	(0,0,0)	(0,0,1.11667)	(0,0,2.255553)	(0,0,50.6167)	(0,0,32.6056)	12.6542	8.1514
0.2	(0.2,0.8) (0.4,0.6)	(0,0,5.462967)	(0,0,0.8962967)	(0,0,0)	(0,0,44.3138)	(0,0,40.8445)	11.0785	10.2111
	(0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,0,0.79697)	(1.34091,1.34091,1.34091)	(13.0739,14.75,31.6887)	(5.3636,6.7046,20.5818)	18.5657	9.8387
0.3	(0.2,0.8)	(2.970483,2.970483,2.970483)	(0.2902363,0.2902363,0.2902363)	(0,0,0)	(5.6099,13.9136,21.7301)	(9.8789,13.3331,20.4350)	13.7918	14.2450
	(0.4,0.6)	(1.106286,1.106286,1.106286)	(0.3208687,0.3208687,0.3208687)	(0.8912555,0.8912555,0.8912555)	(11.8429,16.475,25.3519)	(7.9534,10.4858,18.4381)	17.5362	11.8408
	(0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0.3957865,0.3957865,0.3957865)	(1.344508,1.344508,1.344508)	(15.0879,17.5601,26.9338)	(6.6973,8.7015,17.0072)	19.2855	10.2769
0.4	(0.2,0.8)	(2.970483,2.970483,2.970483)	(0.2902363,0.2902363,0.2902363)	(0,0,0)	(5.6099,13.9136,21.7301)	(9.8789,13.3331,20.4350)	13.7918	14.2450
	(0.4,0.6)	(1.106286,1.106286,1.106286)	(0.3208687,0.3208687,0.3208687)	(0.8912555,0.8912555,0.8912555)	(11.8429,16.475,25.3519)	(7.9534,10.4858,18.4381)	17.5362	11.8408
	(0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0.3957865,0.3957865,0.3957865)	(1.344508,1.344508,1.344508)	(15.0879,17.5601,26.9338)	(6.6973,8.7015,17.0072)	19.2855	10.2769
0.5	(0.2,0.8)	(2.970483,2.970483,2.970483)	(0.2902363,0.2902363,0.2902363)	(0,0,0)	(5.6099,13.9136,21.7301)	(9.8789,13.3331,20.4350)	13.7918	14.2450
	(0.4,0.6)	(1.106286,1.106286,1.106286)	(0.3208687,0.3208687,0.3208687)	(0.8912555,0.8912555,0.8912555)	(11.8429,16.475,25.3519)	(7.9534,10.4858,18.4381)	17.5362	11.8408
	(0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0.3957865,0.3957865,0.3957865)	(1.344508,1.344508,1.344508)	(15.0879,17.5601,26.9338)	(6.6973,8.7015,17.0072)	19.2855	10.2769
0.6	(0.2,0.8) (0.4,0.6)	(2.970483,2.970483,2.970483)	(0.2902363,0.2902363,0.2902363)	(0,0,0)	(5.6099,13.9136,21.7301)	(9.8789,13.3331,20.4350)	13.7918	14.2450
	(0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0.3957865,0.3957865,0.3957865)	(1.344508,1.344508,1.344508)	(15.0879,17.5601,26.9338)	(6.6973,8.7015,17.0072)	19.2855	10.2769
0.7	(0.2,0.8)	(0,3.970588,3.970588)	(0,0,0)	(0,0,0)	(0,15.8824,24.3908)	(0,15.8824,23.8235)	14.0389	13.8971
	(0.4,0.6)	(0,3.297802,3.297802)	(0,0.4740881,0.4740881)	(0,0,0)	(0,16.5098,25.9470)	(0,15.5616,24.0536)	14.7417	13.7942
	(0.5,0.5) (0.6,0.4) (0.8,0.2)	(0,0,0)	(0,0.4774971,0.4774971)	(0.1290657,1.552359,1.552359)	(1.2584,20.4184,31.3439)	(0.5163,10.1493,19.8211)	18.3598	10.1590
0.8	(0.2,0.8)	(0,3.970588,3.970588)	(0,0,0)	(0,0,0)	(0,15.8824,24.3908)	(0,15.8824,23.8235)	14.0389	13.8971
	(0.4,0.6) (0.5,0.5)	(0,3.297802,3.297802)	(0,0.4740881,0.4740881)	(0,0,0)	(0,16.5098,25.9470)	(0,15.5616,24.0536)	14.7417	13.7942
	(0.6,0.4)	(0,0,0)	(0,0.486895,0.486895)	(0,1.565015,1.565015)	(0,20.6234,31.6655)	(0,10.2596,20.0322)	18.2281	10.1379
	(0.8,0.2)	(0,0,0)	(0,0.432284,0.432284)	(0,1.601997,1.601997)	(0,20.648,31.6204)	(0,10.1714,19.9105)	18.2291	10.0633
0.9	(0.2,0.8)	(0,3.970588,3.970588)	(0,0,0)	(0,0,0)	(0,15.8824,24.3908)	(0,15.8824,23.8235)	14.0389	13.8971
	(0.4,0.6) (0.5,0.5)	(0,3.297802,3.297802)	(0,0.4740881,0.4740881)	(0,0,0)	(0,16.5098,25.9470)	(0,15.5616,24.0536)	14.7417	13.7942
	(0.6,0.4)	(0,0,0)	(0,0.486895,0.486895)	(0,1.565015,1.565015)	(0,20.6234,31.6655)	(0,10.2596,20.0322)	18.2281	10.1379
	(0.8,0.2)	(0,0,0)	(0,0.432284,0.432284)	(0,1.601997,1.601997)	(0,20.648,31.6204)	(0,10.1714,19.9105)	18.2291	10.0633

the following model using fuzzy α-cuts in the objective functions and arithmetic operations of TFNs in the constraints:

$$\begin{aligned} Max\ \widetilde{Z}_1(\widetilde{X}) &= [4+\alpha, 6-\alpha][(1-\alpha)x_1^l + \alpha x_1^m, (1-\alpha)x_1^n + \alpha x_1^m] \\ &\quad + [2+\alpha, 4-\alpha][(1-\alpha)x_2^l + \alpha x_2^m, (1-\alpha)x_2^n + \alpha x_2^m] \\ Max\ \widetilde{Z}_2(\widetilde{X}) &= [1+\alpha, 3-\alpha][(1-\alpha)x_1^l + \alpha x_1^m, (1-\alpha)x_1^n + \alpha x_1^m] \\ &\quad + [7+\alpha, 9-\alpha][(1-\alpha)x_2^l + \alpha x_2^m, (1-\alpha)x_2^n + \alpha x_2^m] \\ \text{subject to} & \\ &(3x_1^l + 4x_2^l, 4x_1^m + 5x_2^m, 5x_1^n + 7x_2^n) \preceq (9, 10, 11) \\ &(x_1^l + x_2^l, 2x_1^m + 3x_2^m, 4x_1^n + 5x_2^n) \preceq (3, 5, 7) \\ &x_1^l, x_2^l, x_1^m - x_1^l, x_2^m - x_2^l, x_1^n - x_1^m, x_2^n - x_2^m \geq 0. \end{aligned} \tag{25}$$

The model (25) is equivalently converted into the following interval valued optimization model with deterministic linear constraints based on the centroid concept as discussed:

$$\begin{aligned} Max\ \widetilde{Z}_1(\widetilde{X}) &= [(4+\alpha)(1-\alpha)x_1^l + \alpha(4+\alpha)x_1^m + (2+\alpha)(1-\alpha)x_2^l + \alpha(2+\alpha)x_2^m, \\ &\quad (6-\alpha)(1-\alpha)x_1^n + \alpha(6-\alpha)x_1^m + (4-\alpha)(1-\alpha)x_2^n + \alpha(4-\alpha)x_2^m] \\ Max\ \widetilde{Z}_2(\widetilde{X}) &= [(1-\alpha^2)x_1^l + \alpha(1+\alpha)x_1^m + (7+\alpha)(1-\alpha)x_2^l + \alpha(7+\alpha)x_2^m, \\ &\quad (3-\alpha)(1-\alpha)x_1^n + \alpha(3-\alpha)x_1^m + (9-\alpha)(1-\alpha)x_2^n + \alpha(9-\alpha)x_2^m] \\ \text{subject to} & \\ &3x_1^l + 4x_1^m + 5x_1^n + 4x_2^l + 5x_2^m + 7x_2^n \leq 30 \\ &x_1^l + 2x_1^m + 4x_1^n + x_2^l + 3x_2^m + 55x_2^n \leq 15 \\ &x_1^l, x_2^l, x_1^m - x_1^l, x_2^m - x_2^l, x_1^n - x_1^m, x_2^n - x_2^m \geq 0. \end{aligned} \tag{26}$$

The FFMOLPP (24) is finally modeled as the following single objective deterministic linear optimization model based on the weighting sum approach and Theorem 1:

$$\begin{aligned} Max\ Z(X) &= w_1[(4+\alpha)(1-\alpha^2)x_1^l + 10\alpha x_1^m + (6-\alpha)(1-\alpha)x_1^n \\ &\quad + (2+\alpha)(1-\alpha)x_2^l + 6\alpha x_2^m + (4-\alpha)(1-\alpha)x_2^n] \\ &\quad + w_2[(1-\alpha^2)x_1^l + 4\alpha x_1^m + (3-\alpha)(1-\alpha)x_1^n \\ &\quad + (7+\alpha)(1-\alpha)x_2^l + 16\alpha x_2^m + (9-\alpha)(1-\alpha)x_2^n] \\ \text{subject to} & \end{aligned} \tag{27}$$

$$3x_1^l + 4x_1^m + 5x_1^n + 4x_2^l + 5x_2^m + 7x_2^n \leq 30$$
$$x_1^l + 2x_1^m + 4x_1^n + x_2^l + 3x_2^m + 55x_2^n \leq 15$$
$$x_1^l, x_2^l, x_1^m - x_1^l, x_2^m - x_2^l, x_1^n - x_1^m, x_2^n - x_2^m \geq 0$$
$$w_1, w_2 > 0, w_1 + w_2 = 1$$

On solving problem (27) by substituting different values of $\alpha \in [0,1]$ and different weights $w_1,\ w_2 > 0$ satisfying $w_1 + w_2 = 1$, a set of optimal solutions are generated, which are considered as the non-dominated solutions of the FFMOLPP (24). The triangular fuzzy valued optimal objective values are computed at the non-dominated solutions and also their real valued expressions are calculated using ranking function of TFNs, which are shown in Table 4.

4.1 Comparative result discussions

The FFMOLPP considered in the numerical examples are existing problems, which are solved using our proposed methodology. The computational results of Examples 1, 2, and 3 are incorporated in Tables 1, 2, and 3, respectively. The triangular fuzzy valued optimal objective values are computed at the fuzzy Pareto optimal solutions and their equivalent real valued expressions are determined using the concept of ranking function.

As it is observed, the ranking function values calculated for the optimal objective functions due to our proposed method and existing methods are closely approaching towards each other in certain cases. Besides, the methodology proposed in this paper derives a set of fuzzy Pareto optimal solutions from, which some of the solutions generate better optimal objective values based on the values of their ranking function.

In this context, some of the comparative results are incorporated in Tables 5, 6, and 7 for Examples 1, 2, and 3, respectively. The real-valued expressions of fuzzy optimal objective values are evaluated using ranking functions in Tables 5, 6, and 7 which are further comparatively discussed with existing methods through the following figures, Figures 3, 4, and 5, respectively. Since Exam-

Table 4: Non-dominated solutions, fuzzy valued optimal objective functions and their real valued expressions of Example 4

α	w	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{Z}_1$	$\tilde{Z}_2$	$R(\tilde{Z}_1)$	$R(\tilde{Z}_2)$
0.1	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(0,0,0)	(1.666667,1.666667, 1.666667)	(3.3333,5,6.6667)	(11.6667,13.3333,15)	5	13.3333
	(0.8,0.2)	(2.142857,2.142857, 2.142857)	(0,0,0)	(8.5714,10.7143, 12.8571)	(2.142857,4.2857, 6.4286)	10.7143	4.2857
0.2	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(0,0,0)	(1.666667,1.666667, 1.666667)	(3.3333,5,6.6667)	(11.6667,13.3333,15)	5	13.3333
	(0.8,0.2)	(2.142857,2.142857, 2.142857)	(0,0,0)	(8.5714,10.7143, 12.8571)	(2.142857,4.2857, 6.4286)	10.7143	4.2857
0.3	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(0,0,0)	(1.666667,1.666667, 1.666667)	(3.3333,5,6.6667)	(11.6667,13.3333,15)	5	13.3333
	(0.8,0.2)	(2.142857,2.142857, 2.142857)	(0,0,0)	(8.5714,10.7143, 12.8571)	(2.142857,4.2857, 6.4286)	10.7143	4.2857
0.4	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(0,0,0)	(1.666667,1.666667, 1.666667)	(3.3333,5,6.6667)	(11.6667,13.3333,15)	5	13.3333
	(0.8,0.2)	(2.142857,2.142857, 2.142857)	(0,0,0)	(8.5714,10.7143, 12.8571)	(2.142857,4.2857, 6.4286)	10.7143	4.2857
0.5	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(0,0,0)	(1.666667,1.666667, 1.666667)	(3.3333,5,6.6667)	(11.6667,13.3333,15)	5	13.3333
	(0.8,0.2)	(2.142857,2.142857, 2.142857)	(0,0,0)	(8.5714,10.7143, 12.8571)	(2.142857,4.2857, 6.4286)	10.7143	4.2857
0.6	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(0,0,0)	(1.666667,1.666667, 1.666667)	(3.3333,5,6.6667)	(11.6667,13.3333,15)	5	13.3333
	(0.8,0.2)	(2.142857,2.142857, 2.142857)	(0,0,0)	(8.5714,10.7143, 12.8571)	(2.142857,4.2857, 6.4286)	10.7143	4.2857
0.7	(0.2,0.8) (0.4,0.6) (0.5,0.5) (0.6,0.4)	(0,0,0)	(1.666667,1.666667, 1.666667)	(3.3333,5,6.6667)	(11.6667,13.3333,15)	5	13.3333
	(0.8,0.2)	(0,2.5,2.5)	(0,0,0)	(0,12.5,15)	(0,5,7.5)	10	4.375
0.8	(0.2,0.8) (0.4,0.6) (0.5,0.5)	(0,0,0)	(0,1.875,1.875)	(0,5.625,7.5)	(0,15,16.875)	4.6875	11.7188
	(0.6,0.4) (0.8,0.2)	(0,2.5,2.5)	(0,0,0)	(0,12.5,15)	(0,5,7.5)	10	4.375
0.9	(0.2,0.8) (0.4,0.6) (0.5,0.5)	(0,0,0)	(0,1.875,1.875)	(0,5.625,7.5)	(0,15,16.875)	4.6875	11.7188
	(0.6,0.4) (0.8,0.2)	(0,2.5,2.5)	(0,0,0)	(0,12.5,15)	(0,5,7.5)	10	4.375

ples 1, 2, and 3 are of maximization, minimization and maximization types, respectively, the figures validate and justify the feasibility and effectiveness of the proposed solution approach.

Table 5: Comparative results of Example 1

Methods	$\tilde{Z}_1$	$\tilde{Z}_2$	$R(\tilde{Z}_1)$	$R(\tilde{Z}_2)$
Agarwal and Sharma [1]	(4, 20, 43)	(7, 24, 49)	21.75	26
Temelcan, Gonce Kocken, and Albayrak [27]	(12, 24, 32.33)	(20.33, 28.67, 37)	23.08	28.67
Proposed method	(4.3333,28.6667,38)	(8.6667,33,42.3333)	24.9167	29.25
Proposed method	(3.8889,3.8889,3.8889)	(12.6667,25.3333,34.1111)	24.3611	30.2222

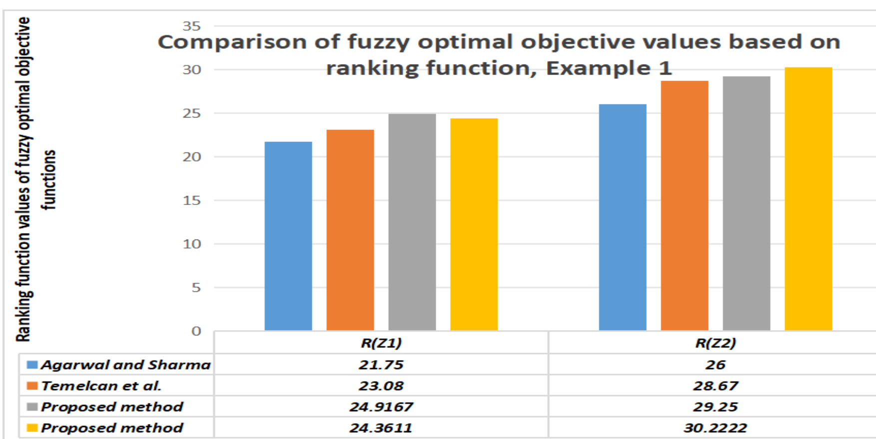


Figure 3: Comparative results on ranking function, Example 1 and Table 5

Table 6: Comparative results of Example 2

Methods	$\tilde{Z}_1$	$\tilde{Z}_2$	$R(\tilde{Z}_1)$	$R(\tilde{Z}_2)$
Yang, Cao, and Lin [30]	(12, 50, 66.3712)	(6, 23, 40.5008)	44.5928	23.1252
Proposed method	(20,25,46.8)	(10,17.5,43.2)	29.2	22.05

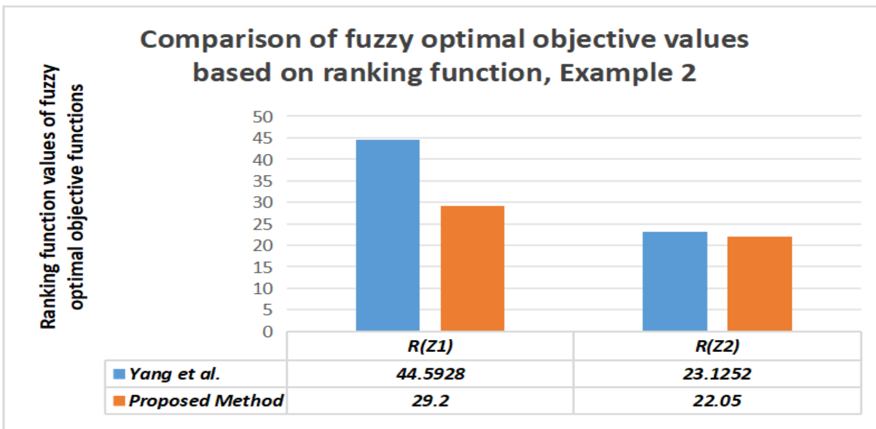


Figure 4: Comparative results on ranking function, Example 2 and Table 6

Table 7: Comparative results of Example 3

Methods	$\tilde{Z}_1$	$\tilde{Z}_2$	$R(\tilde{Z}_1)$	$R(\tilde{Z}_2)$
Arana-Jiménez [5]	(2.457143, 16.343755, 33.540027)	(2.4, 11.49366, 15.47613)	17.17117	10.2158625
Proposed method	(15.0879, 17.5601, 26.9338)	(6.6973, 8.7015, 17.0072)	19.2855	10.2769
Proposed method	(1.2584, 20.4184, 31.3439)	(0.5163, 10.1493, 19.8211)	18.3598	10.1590

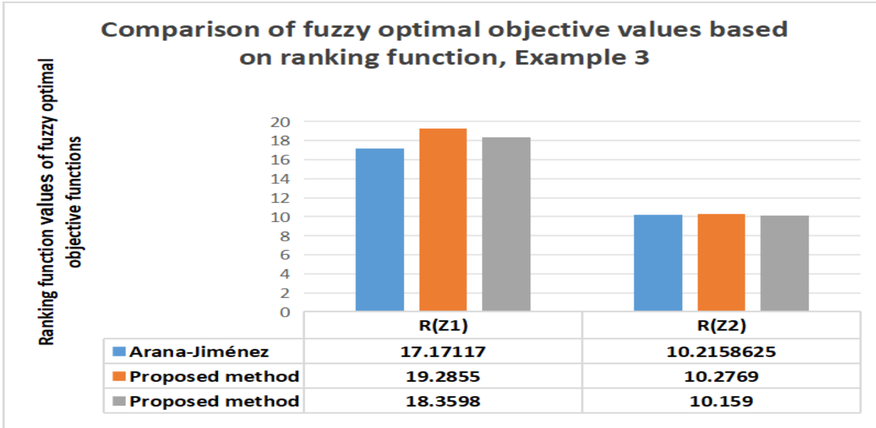


Figure 5: Comparative results on ranking function, Example 3 and Table 7

5 Conclusions

This paper developed a methodology that generates a set of fuzzy Pareto optimal solutions of an FFMOLPP, which is considered in a triangular fuzzy environment. The concepts of fuzzy cuts and centroid of TFNs are used to develop an equivalent interval valued MOLPP, which is further solved using weighting sum approach with different weight vectors and a theorem of interval analysis. Different fuzzy cuts and positive weight vectors derive the set of fuzzy Pareto optimal solutions. The solution approach developed is computationally simple and efficient, which provides a set of fuzzy valued solutions comprising the compromise solution of the FFMOLPP whereas most of the existing methods generate only one solution as the compromise solution. DM compares the resulted fuzzy optimal objective values using ranking function concept to decide the compromise solution. The incorporated algorithm and flowchart narrate the detailed steps of the proposed solution approach. In the numerical section, three existing problems along with one practical problem are solved and the results computed are discussed as compared to the existing methods, which shows the feasibility, effectiveness and acceptability of the proposed methodology. LINGO software is used for the computational works of the numerical section. The developed approach can also be applicable for FFMOLPP designed with trapezoidal, pentagonal and hexagonal fuzzy numbers etc. As the future scope of research, fully fuzzy nonlinear MOOP, fully fuzzy bi-level MOOP in various fuzzy environment can be studied to develop some new efficient solution methodologies.

Acknowledgements

Authors are grateful to the anonymous referees and the Editor for their valuable comments and suggestions to improve the quality of the paper. Authors are also grateful to KIIT University, Bhubaneswar to provide the platform and necessary support to carry out this research work.

References

[1] Aggarwal, S. and Sharma, U. *Implementing deviation degree of two closed intervals to decode fully fuzzy multiobjective linear programming problem*, J. Intell. Fuzzy Syst. 31(1) (2016), 443–455.

[2] Allahviranloo, T., Lotfi, F.H., Kiasary, M.K., Kiani, N.A. and Alizadeh, L. *Solving fully fuzzy linear programming problem by the ranking function*, Appl. Math. Sci. 2(1) (2008), 19–32.

[3] Arana-Jiménez, M. *Nondominated solutions in a fully fuzzy linear programming problem*, Math. Methods Appl. Sci. 41(17) (2018), 7421–7430.

[4] Arana-Jiménez, M. *Fuzzy Pareto solutions in fully fuzzy multiobjective linear programming. In Optimization of complex systems: theory, models, algorithms and applications*, Springer International Publishing. (2020) 509–517.

[5] Arana-Jiménez, M. *On generating fuzzy Pareto solutions in fully fuzzy multiobjective linear programming via a compromise method*, RAIRO Oper. Res. 56(6) (2022), 4035–4045.

[6] Bas, S. A. and Ozkok, B. A. *An iterative approach for the solution of fully fuzzy linear fractional programming problems via fuzzy multi-objective optimization*, AIMS Math., 9(6) (2024), 15361–15384.

[7] Collette, Y., and Patrick S. *Multiobjective optimization: principles and case studies*. Springer Science & Business Media, 2013.

[8] Daneshrad, R., and Jafari, D. *FFLP problem with symmetric trapezoidal fuzzy numbers*, Decision Science Letters, 4(2) (2015), 117–124.

[9] Das, S.K. *Modified method for solving fully fuzzy linear programming problem with triangular fuzzy numbers*, Int. J. Res. Ind. Eng. 6(4) (2017), 293–311.

[10] Das, S.K., Mandal, T. and Edalatpanah, S.A. *A mathematical model for solving fully fuzzy linear programming problem with trapezoidal fuzzy numbers*, Appl. Intell. 46 (2017), 509–519.

[11] Ebrahimnejad, A. *An effective computational attempt for solving fully fuzzy linear programming using MOLP problem*, J. Ind. Prod. Eng. 36(2) (2019), 59–69.

[12] Ezzati, R., Khorram, E. and Enayati, R. *A new algorithm to solve fully fuzzy linear programming problems using the MOLP problem*, Appl. Math. Model. 39(12) (2015), 3183–3193.

[13] Fathy, E. and Hassanien, A. E. *Fuzzy harmonic mean technique for solving fully fuzzy multilevel multiobjective linear programming problems*, Alex. Eng. J., 61(10) (2022), 8189–8205.

[14] Hamadameen, A.O. and Hassan, N. *A compromise solution for the fully fuzzy multiobjective linear programming problems*, IEEE Access. 6 (2018), 43696–43711.

[15] Hosseinzadeh, A. and Edalatpanah, S. *A new approach for solving fully fuzzy linear programming by using the lexicography method*, Adv. Fuzzy Syst. 2016(1) (2016), 1538496.

[16] Khalili, G.F., Nasseri, SH. and Taghi-Nezhad, N.A. *A new interactive approach for solving fully fuzzy mixed integer linear programming*, Yugosl J. Oper. Res. 30(1) (2020) 71–89.

[17] Khan, I.U., Ahmad, T. and Maan, N. *A simplified novel technique for solving fully fuzzy linear programming problems*, J. Optim. Theory Appl. 159 (2013), 536–546.

[18] Kumar, A., Kaur, J. and Singh, P. *A new method for solving fully fuzzy linear programming problems*, Appl. Math. Model. 35(2) (2011), 817–823.

[19] Lotfi, F.H., Allahviranloo, T., Jondabeh, M.A. and Alizadeh, L. *Solving a full fuzzy linear programming using lexicography method and fuzzy approximate solution*, Appl. Math. Model. 33(7) (2009), 3151–3156.

[20] Malik, M. and Gupta, S. K. *An application of fully intuitionistic fuzzy multi-objective linear fractional programming problem in e-education system*, Int. J. of Fuzzy Syst., 24(8) (2022), 3544–3563.

[21] Miettinen, K. *Nonlinear multiobjective optimization*. Vol. 12. Springer Science & Business Media, 1999.

[22] Nayak, S., and Maharana, S. *An efficient fuzzy mathematical approach to solve multi-objective fractional programming problem under fuzzy environment*, J. Appl. Math. Comp. (2023), 1–27.

[23] Niksirat, M. *A new approach to solve fully fuzzy multi-objective transportation problem*, Fuzzy Inform. Eng., 14(4) (2022), 456–467.

[24] Pérez-Cañedo, B., Rosete, A., Verdegay, J.L. and Concepción-Morales, E.R. *A fuzzy goal programming approach to fully fuzzy linear regression*, In International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems Cham: Springer International Publishing. (2020), 677–688.

[25] Pérez-Cañedo, B., Verdegay, J.L. and Miranda Perez, R. *An epsilon-constraint method for fully fuzzy multiobjective linear programming*, Int. J. Intell. Syst. 35(4) (2020), 600–624.

[26] Sharma, U. and Aggarwal, S. *Solving fully fuzzy multi-objective linear programming problem using nearest interval approximation of fuzzy number and interval programming* Int. J. Fuzzy Syst. 20 (2018), 488–499.

[27] Temelcan, G., Gonce Kocken, H. and Albayrak, I. *Finding compromise solutions for fully fuzzy multi-objective linear programming problems by using game theory approach*, J. Intell. Fuzzy Syst., 42(1) (2022), 283–293.

[28] Van Hop, N. *Solving fully fuzzy multi-objective linear programming problem with fuzzy dominant degrees*, J. Intell. Fuzzy Syst. 39(3) (2020), 3577–3595.

[29] Wu, H.C. *On interval-valued nonlinear programming problems*, J. Math. Anal. Appl. 338(1) (2008), 299–316.

[30] Yang, X.P., Cao, B.Y. and Lin, H.T. *Multi-objective fully fuzzy linear programming problems with triangular fuzzy numbers*, In 2014 11th International Conference on Fuzzy Systems and Knowledge Discovery FSKD (2014), 171–177.

[31] Zadeh, L.A. *Fuzzy sets*, Inf. Control, 8(3) (1965), 338–353.

[32] Zimmermann, H-J. *Fuzzy set theory and its applications*, Springer Science & Business Media, 2011.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1140–1167
https://doi.org/10.22067/ijnao.2024.87550.1421
https://ijnao.um.ac.ir/

Research Article

Adaptive mesh based Haar wavelet approximation for a singularly perturbed integral boundary problem

P. Shukla, S. Saini*, and V. Devi

Abstract

This research presents a nonuniform Haar wavelet approximation of a singularly perturbed convection-diffusion problem with an integral boundary. The problem is discretized by approximating the second derivative of the solution with the help of a nonuniform Haar wavelets basis on an arbitrary nonuniform mesh. To resolve the multiscale nature of the problem,

*Corresponding author

Received 9 April 2024; revised 22 July 2024; accepted 24 July 2024

Pragya Shukla
Department of Engineering Sciences, Indian Institute of Information Technology and Management Gwalior, Gwalior, Madhya Pradesh, 474015, India. e-mail: pragya@iiitm.ac.in

Sumit Saini
Department of Mathematics, Vellore Institute of Technology, Vellore, Tamilnadu, 632014, India. e-mail: sumit.iitv@gmail.com

Vinita Devi
Department of Mathematics, Bhakt Darshan Govt. P.G. College, Pauri Garhwal, Uttrakhand, 246193, India. e-mail: iitbhuvinita@gmail.com

How to cite this article
Shukla, P., Saini, S. and Devi, V., Adaptive mesh based Haar wavelet approximation for a singularly perturbed integral boundary problem. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1140-1167. https://doi.org/10.22067/ijnao.2024.87550.1421

adaptive mesh is generated using the equidistribution principle. This approach allows for the dynamical adjustment of the mesh based on the solution's behavior without requiring any information about the solution. The combination of nonuniform wavelet approximation and the use of adaptive mesh leads to improved accuracy, efficiency, and the ability to handle the multiscale behavior of the solution. On the adaptive mesh rigorous error analysis is performed showing that the proposed method is a second-order parameter uniformly convergent. Numerical stability and computational efficiency are validated in various tables and plots for numerical results obtained by the implementation of two test examples.

AMS subject classifications (2020): 34D15, 65L11, 65L20, 65T60, 65L50

Keywords: Singular perturbation problems, Adaptive mesh, Nonuniform Haar wavelet, Parameter uniform convergence

1 Introduction

Nowadays, singularly perturbed problems are among the most studied classes of differential equations because of their frequent appearance in various branches of science and engineering such as fluid mechanics, heat transfer, and problems in structural mechanics posed over thin domains [35, 31, 30]. These problems are characterized by a small positive parameter ε multiplied by the highest order derivative term, due to which in certain regions rapid change is observed in the solution, generally called as boundary layer. Actually, this multiscale behavior is the main attraction of this class of problems due to which the traditional numerical methods fail to give satisfactory information about the solution in this boundary layer region. We find several techniques in the literature considering various classes of singularly perturbed problems, for example, [31, 30, 37, 21, 41, 32, 1, 36, 39].

In this paper, one important subclass is addressed: Singularly perturbed convection-diffusion problems with integral boundary conditions. These problems appear in many areas such as fluid dynamics, plasma physics, thermo-elasticity, chemical engineering, underground water flow, oceanography, meteorology, water pollution problems, and so on [17]. In generalized form, this class is expressed as a differential equation with the perturbation parameter ε such that $0 < \varepsilon \ll 1$,

$$\begin{cases} \mathcal{L}_\varepsilon u := \varepsilon \frac{d^2u}{dx^2} + a(x)\frac{du}{dx} = f(x), \quad x \in G := (0,1), \\ \frac{du}{dx}(0) = \frac{\gamma_0}{\varepsilon}, \\ \int\limits_0^1 \xi(x)u(x)dx = \gamma_1. \end{cases} \tag{1}$$

Here, the function ξ and the constants γ_0 and γ_1 are known. The existence of unique solution can be obtained from the sufficient smoothness of the functions f and ξ and the condition $a(x) \geq \alpha > 0$ in G.

Furthermore, the following two remarks provide the stability of the model problem and derivative bounds of the solution, respectively [11].

Remark 1. Suppose that $y(x)$ is a sufficiently smooth function with both $y(0)$ and $y(1)$ being nonnegative. Then $\mathcal{L}_\varepsilon y(x) \leq 0, \quad x \in G$ implies that $y(x) > 0, \ x \in \bar{G}$.

Remark 2. For $0 \leq r \leq 2$, the solution $u(x)$ to the problem (1) satisfies the following bounds:

$$\left| \frac{d^r u}{dx^r} \right| \leq C \left(1 + \varepsilon^{-r} e^{-\frac{\alpha}{\varepsilon}x} \right). \tag{2}$$

From these solution bounds, the boundary layer behavior is quite visible near $x = 0$. So, in the numerical analysis of these problems, the main focus is to resolve this boundary layer region. The problem considered in this paper has been studied previously by some authors such as in [6, 11, 33]. In [6], the authors presented an exponentially fitted finite difference operator method with a linear order convergence rate. In [11], authors presented a simpler nonstandard finite difference operator method with the same linear order convergence. Recently, in [33], authors presented a second-order nonuniform Haar wavelet approach on an exponentially graded mesh. Indeed, the nonuniform graded mesh used is generated based on the knowledge of the width of the boundary layer, which is not available all the time in case of multiscale problems. Therefore, we need an adaptive mesh generation algorithm that works on a dynamic structure and does not require any information about the solution. This is the motivation for the present work.

For a long time, we have seen frequent applications of wavelets theory in various disciplines of science such as in time-frequency analysis, time series analysis, signal analysis, image and data compression, and so on. However, these days, we can see one very interesting application of wavelets theory in the construction of fast numerical algorithms due to their nice capacity for the representation of complex functions and operators. One such example is the Haar wavelets with some very nice properties such as theoretical simplicity, memory efficiency, orthogonality, compact support, and easy and fast implementation. Still, using the standard Haar wavelet method is not a good idea in case of problems exhibiting a multiscale nature in their solution. The uniform Haar wavelet methods working on the equidistant grid are not computationally efficient for such problems. Therefore applying the nonuniform Haar wavelet approach together with the adaptively generated nonuniform mesh is a nice idea for these problems. For the first time, nonuniform Haar wavelets were proposed in 2004 by Dubeau, Elmejdani, and Ksantini[12]. More details of Haar wavelets and their applications can be found in [24, 29, 13, 16].

Generally speaking, the idea of adaptive mesh is the movement of a fixed number of mesh points so that the regions of rapid variations attain more mesh points and those of smooth variation attain a comparatively smaller number of mesh points by some iterative process. In this work, the approximation of problem (1) is done using a nonuniform Haar wavelet approach on a non-equidistant adaptive mesh generated with the help of the equidistribution principle [18]. Starting with a uniform mesh, this technique aims to condense the maximum number of mesh points inside the layer regions working on a self-improving mechanism. The main advantage of the equidistribution approach is that it does not require any a priori information about the solution for the construction of adaptive mesh. A nonuniform mesh $\{x_i\}_{i=0}^{N}$ is defined as an equidistribution mesh for the monitor function $\mathcal{M}(x, u(x))$ if it satisfies the equidistribution principle, that is,

$$\int\limits_{x_{i-1}}^{x_i} \mathcal{M}(s, u(s))\, ds = \int\limits_{x_i}^{x_{i+1}} \mathcal{M}(s, u(s))\, ds, \quad 1 \leq i \leq N-1. \tag{3}$$

So far, this equidistribution mesh idea has been applied to many physical problems together with finite difference methods, for example, in [34, 28, 4, 19, 9, 10, 27, 8] and the references therein. However, to the best of the authors' knowledge, this idea has never been applied to any problem together with the wavelet methods. This is the first time the nonuniform Haar wavelet method is considered together with the solution adaptive equidistribution mesh. On this mesh, the highest order derivative is approximated by the Haar series and then obtained the other derivatives by integrating the Haar series. It is favorable to integrate the derivative approximation because the derivative of the Haar wavelet results in an impulse function. Finally, this approximation on collocation points gives a $2M \times 2M$ system of linear equations in wavelet coefficients as variables, which can be solved using any known method. Via a thorough analysis, the proposed method is proved to be a second-order parameter uniformly convergent and is validated by implementing it on two example problems.

The main highlights of this research can be listed as the following bullet points:

1. A new adaptive mesh-based numerical wavelet method is presented for singularly perturbed convection-diffusion problems with a nonlocal boundary.

2. The equidistribution mesh proposed, works on a self-improvement mechanism and does not require the a priori information about the solution, in comparison to other a priory adaptive meshes; such as Shishkin [37] and Bakhvalov meshes [2].

3. The advantage of the proposed method is that it uses the nice properties of nonuniform Haar wavelets and the adaptive mesh in the same framework.

4. Simple and effective uniform convergence analysis is presented for the proposed method.

5. The numerical algorithm is provided for constructing the adaptive mesh and solution.

The rest of the paper is organized as follows: Sections 2 and 3 briefly discuss the construction of the numerical method and the adaptive mesh, respectively. Error analysis in Section 4 demonstrates the second-order parameter uniform convergence of the proposed numerical method. Numerical stability and error results in Section 5 prove the efficient applicability of the proposed method and validate the theoretically proved results. Finally, Section 6 concludes the theoretical and numerical findings of the present research.

Notation: We use C as a generic positive constant, which is independent of ε and M and can take different values.

2 Wavelet-based numerical method

This section discusses the structure and theoretical aspects of nonuniform Haar wavelets and then provides a detailed numerical discretization of the problem with the help of nonuniform Haar wavelets [25].

2.1 Nonuniform Haar wavelet

Basically, Haar wavelets are characterized by a dilation parameter (j) and a translation parameter (l), where $j = 0, \dots, J$, and $l = 0, \dots, m-1$, with $m = 2^j$. Each wavelet is numbered by the index $i = m+l+1$. The minimum values for l and m are 0 and 1, respectively, which results in $i = 2$. Similarly, the maximum values of $i = 2M$, where $M = 2^J$, with J indicating the maximum level of resolution. Since the indexing here is starting from 2, we define the first Haar wavelet $h_1(x)$ as

$$h_1(x) = \begin{cases} 1, & x \in [0,1], \\ 0, & \text{otherwise}. \end{cases} \tag{4}$$

Let $\{x_k\}_{k=1}^{2M}$ be an arbitrary nonuniform mesh in $[0,1]$. Now taking $\nu = \frac{M}{m}$, choose three specific nodes on the discrete mesh, $\sigma_1(i) = x_{(2l\nu)}, \sigma_2(i) = x_{(2l\nu+\nu)}, \sigma_3(i) = x_{(2l\nu+2\nu)}$, and using the requisite $\int_0^1 h_i(x)dx = 0$, define the Haar wavelet, for $i = 2 \leq i \leq 2M$, by

$$h_i(x) = \begin{cases} 1, & x \in [\sigma_1(i), \sigma_2(i)], \\ -\psi_i, & x \in [\sigma_2(i), \sigma_3(i)], \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where

$$\psi_i = \frac{\sigma_2(i) - \sigma_1(i)}{\sigma_3(i) - \sigma_2(i)}. \tag{6}$$

Here, some particulars are provided on wavelets and their notations that will be used in upcoming sections:

1. In general one can write the integrals of Haar wavelets as

$$\mathcal{P}_{i,1}(x) = \int_0^x h_i(s)\, ds, \tag{7}$$

$$\mathcal{P}_{i,k+1}(x) = \int_0^x \mathcal{P}_{i,k}(s)\, ds, \quad 1 \leq i \leq 2M,\ k = 1, 2, \ldots. \tag{8}$$

The calculation of the first two integrals of the wavelet h_1 are $\mathcal{P}_{1,1}(x) = x$ and $\mathcal{P}_{1,2}(x) = \frac{x^2}{2}$.

In general, the first two integrals for each of the rest wavelets are

$$\mathcal{P}_{i,1}(x) = \begin{cases} x - \sigma_1(i), & x \in [\sigma_1(i), \sigma_2(i)], \\ -\psi_i(\sigma_3(i) - x), & x \in [\sigma_2(i), \sigma_3(i)], \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

$$\mathcal{P}_{i,2}(x) = \begin{cases} 0, & x < \sigma_1(i), \\ \frac{1}{2}(x - \sigma_1(i))^2, & x \in [\sigma_1(i), \sigma_2(i)], \\ \frac{1}{2}(\sigma_2(i) - \sigma_1(i))(\sigma_3(i) - \sigma_1(i)) - \frac{1}{2}\psi_i(\sigma_3(i) - x)^2, & x \in [\sigma_2(i), \sigma_3(i)], \\ \frac{1}{2}(\sigma_2(i) - \sigma_1(i))(\sigma_3(i) - \sigma_1(i)), & x \geq \sigma_3(i). \end{cases} \tag{10}$$

2. Piecewise orthogonality of the nonuniform Haar wavelets can be observed as

$$\int_0^1 h_i(x) h_{i'}(x)\, dx = \begin{cases} \psi_i(\sigma_3(i) - \sigma_1(i)), & i = i', \\ 0, & i \neq i'. \end{cases} \tag{11}$$

3. Also, we have the following bounds for $\mathcal{P}_{i,k}(x)$ [38]:

$$\mathcal{P}_{i,0}(x) \leq 1, \quad \mathcal{P}_{i,1}(x) \leq \frac{1}{2^j+1}, \quad i>1, \tag{12}$$

$$\mathcal{P}_{1,k}(x) \leq \frac{1}{k!}, \quad \mathcal{P}_{i,k}(x) < \frac{8}{3(\lfloor (k+1)/2 \rfloor !)^2}\left(\frac{1}{2^{j+1}}\right)^2, \quad k\geq 2, i>1. \tag{13}$$

2.2 Numerical method

Any square-integrable function $y(x)$ can be expressed as an infinite linear combination of Haar wavelets as [12]

$$y(x)=\sum_{i=1}^{\infty} c_i h_i(x), \quad x\in(0,1). \tag{14}$$

Here, if $y(x)$ is a piecewise constant function or if we can approximate it with a piecewise constant function, then this series may terminate in a finite sum.

If we denote the approximate solution by U, then following this discussion, we can consider the approximation of the second-order derivative as

$$\frac{d^2U}{dx^2}=\sum_{i=1}^{2M} p_i h_i(x), \tag{15}$$

where the coefficient p_i's are wavelet coefficients.

Integrating (15) from 0 to x, and using the boundary condition at $x=0$, we obtain the approximation for the first-order derivative as follows:

$$\frac{dU}{dx}=\frac{\gamma_0}{\varepsilon}+\sum_{i=1}^{2M} p_i \mathcal{P}_{i,1}(x). \tag{16}$$

Again integrating (16), from 0 to x, we obtain the following approximation of the exact solution u:

$$U(x)=U(0)+\frac{\gamma_0}{\varepsilon}x+\sum_{i=1}^{2M} p_i \mathcal{P}_{i,2}(x). \tag{17}$$

Next, to obtain the approximation of the integral boundary condition multiply $U(x)$ by $\xi(x)$ in (17) and integrate from 0 to 1 to get

$$\int_0^1 \xi(x)U(x)dx = U(0)\int_0^1 \xi(x)\,dx + \frac{\gamma_0}{\varepsilon}\int_0^1 x\xi(x)\,dx + \int_0^1 \xi(x)\left(\sum_{i=1}^{2M} p_i\mathcal{P}_{i,2}(x)\right)dx. \tag{18}$$

Rewriting (18), we get the approximate solution at $x=0$ as

$$U(0) = \frac{\gamma_1 - \frac{\gamma_0}{\varepsilon}\int_0^1 x\xi(x)\,dx - \int_0^1 \xi(x)\left(\sum_{i=1}^{2M} p_i\mathcal{P}_{i,2}(x)\right)dx}{\int_0^1 \xi(x)\,dx} \tag{19}$$

with $\int_0^1 \xi(x)\,dx \neq 0$.

Now, use the expressions for $\frac{d^2U}{dx^2}$, $\frac{dU}{dx}$, and $U(x)$ from (15), (16), and (17), respectively, in previous equation. Then substituting the collocation points considered at the midpoints of the mesh subintervals, that is,

$$x_k^c = \frac{x_{k-1}+x_k}{2}, \quad k = 1,\dots,2M, \tag{20}$$

we get the following system of linear equations

$$\sum_{i=1}^{2M} p_i[\varepsilon h_i(x_k^c) + a(x_k^c)\mathcal{P}_{i,1}(x_k^c)] = f(x_k^c) - a(x_k^c)\frac{\gamma_0}{\varepsilon}, \quad 1 \le k \le 2M. \tag{21}$$

In the standard form, we represent this system as $\mathcal{AX}=\mathcal{B}$, where $\mathcal{A} = [a_{k,i}]_{2M\times 2M}$ is the coefficients matrix in the associated system of linear equations, $\mathcal{X} = [p_k]_{2M\times 1}$ is the solution vector, and $\mathcal{B} = [b_k]_{2M\times 1}$ is the known vector, with $a_{k,i} = \varepsilon h_i(x_k^c) + a(x_k^c)\mathcal{P}_{i,1}(x_k^c)$ and $b_k = f(x_k^c) - a(x_k^c)\frac{\gamma_0}{\varepsilon}$. We can solve this system with the help of any known method and thus the solution is completely known using (17).

3 Construction and properties of adaptive mesh

This section explains the construction of a nonuniform adaptive mesh defined by using the equidistribution principle. Some insights into the adaptive mesh are also provided which will be used in the error analysis section.

Equivalently to (3), we define the nonuniform adaptive mesh $\bar{G}^{2M} = \{0 = x_0, \dots, x_{2N} = 1\}$ using the equidistribution principle as follows:

$$\int_{x_{i-1}}^{x_i} \mathcal{M}(s,u(s))\,ds = \frac{1}{2M}\int_0^1 \mathcal{M}(s,u(s))\,ds, \quad 1 \le i \le 2M. \tag{22}$$

In another form, the mesh equidistribution can be seen as a mapping $x = x(\xi)$, which relates the computational coordinate $\xi \in [0,1]$ to the physical coordinate $x \in [0,1]$, defined by

$$\int_0^{x(\xi)} \mathcal{M}(s,u(s))\,ds = \xi \int_0^1 \mathcal{M}(s,u(s))\,ds. \tag{23}$$

Motivated by [8, 3, 23], we choose the monitor function of the following form:

$$\mathcal{M}(x,u(x)) = \beta + \left|\frac{d^2u(x)}{dx^2}\right|^{1/2}, \tag{24}$$

where the positive constant β is introduced to maintain the reasonable distribution of the node points throughout the domain; if we do not include β, then the equidistribution results in node starvation outside the boundary layer part.

Now, consider the following approximation of $\frac{d^2u(x)}{dx^2}$,

$$\frac{d^2u(x)}{dx^2} \approx \frac{\eta}{\varepsilon^2} e^{-\frac{\alpha}{\varepsilon}x},$$

where the constant η is independent of x and ε.

Hence,

$$\int_0^1 \left|\frac{d^2u(x)}{dx^2}\right|^{1/2} dx \equiv \frac{2\,|\eta|^{1/2}}{\alpha}\left(1 - e^{-\frac{\alpha}{2\varepsilon}}\right) \approx \frac{2\,|\eta|^{1/2}}{\alpha} = \Lambda \quad \text{(say)}. \tag{25}$$

Now, using the definition (23), equidistribution of the monitor function (24) leads to the scaling

$$e^{\left(-\frac{\alpha}{2\varepsilon}x(\xi)\right)} - \frac{\beta}{\Lambda}x(\xi) = 1 - \left(\frac{\beta}{\Lambda} + 1 - e^{-\frac{\alpha}{2\varepsilon}}\right)\xi. \tag{26}$$

By this scaling, a nonuniform mesh in physical coordinates $\{x_k\}_{k=0}^{2M}$ corresponds from an equispaced mesh $\{\xi_k = k/2M\}_{k=0}^{2M}$ in computational coordinates. So, the above equation is written as

$$e^{\left(-\frac{\alpha}{2\varepsilon}x_k\right)} - \frac{\beta}{\Lambda}x_k = 1 - \left(\frac{\beta}{\Lambda} + 1 - e^{-\frac{\alpha}{2\varepsilon}}\right)\frac{k}{2M}. \tag{27}$$

Hence, the adaptively generated mesh points are given by the solution of the nonlinear algebraic equation (27).

Throughout the analysis, we take $\beta = \Lambda$. An important observation that we shall use in error analysis is obtained by using the bounds from (2) as follows:

$$\begin{aligned}
\int_0^1 \mathcal{M}(x, u(x))\,dx &= \int_0^1 \left(\beta + \left|\frac{d^2u(x)}{dx^2}\right|^{1/2}\right) dx \\
&\leq \int_0^1 \beta\,dx + C\int_0^1 \left(1 + \varepsilon^{-2} e^{-\frac{\alpha}{\varepsilon}x}\right)^{1/2} dx \\
&= \beta + C\int_0^1 \left(1 + \left(\varepsilon^{-1} e^{-\frac{\alpha}{2\varepsilon}x}\right)^2\right)^{1/2} dx \\
&\leq \beta + C\int_0^1 \left(1 + \varepsilon^{-1} e^{-\frac{\alpha}{2\varepsilon}x}\right) dx \\
&\leq \beta + C + \frac{2C}{\alpha}(1 - e^{\frac{\alpha}{2\varepsilon}}) \leq C.
\end{aligned} \tag{28}$$

4 Error analysis

In the proposed method, an approximation of the solution up to the Jth level of resolution is used, and the expression (17) represents the discrete version of the problem. So if the actual representation of the solution is taken in (17), then the expression for the exact solution is obtained as

$$u(x) = u(0) + \frac{\gamma_0}{\varepsilon}x + \sum_{i=1}^{\infty} p_i \mathcal{P}_{i,2}(x). \tag{29}$$

On comparing (17) and (29), we can write the error term as the truncated part of the infinite sum

$$\mathcal{E}^J = |u(x) - U(x)| = \left|\sum_{i=2M+1}^{\infty} p_i \mathcal{P}_{i,2}(x)\right|. \tag{30}$$

Now, to estimate the error bound completely, we calculate the bounds on both p_i's and $\mathcal{P}_{i,2}(x)$.

To calculate the bounds for the coefficients p_i's, multiply h_r in both sides of the expression

$$\frac{d^2u}{dx^2} = \sum_{i=1}^{\infty} p_i h_i(x),$$

for some particular natural number r, and integrate in 0 to 1. Then by using the orthogonality property of wavelets (11), we get

$$\int_0^1 \frac{d^2u}{dx^2} h_r(x)\, dx = p_r \psi_r(\sigma_3(r) - \sigma_1(r)).$$

This gives

$$\begin{aligned}
p_i &= \frac{1}{\psi_i\left(\sigma_3(i) - \sigma_1(i)\right)} \int_0^1 \frac{d^2u}{dx^2} h_i(x)\, dx \\
&= \frac{1}{\psi_i\left(\sigma_3(i) - \sigma_1(i)\right)} \left[\int_{\sigma_1(i)}^{\sigma_2(i)} \frac{d^2u}{dx^2}\, dx - \psi_i \int_{\sigma_2(i)}^{\sigma_3(i)} \frac{d^2u}{dx^2}\, dx \right] \\
&= \frac{\left(\sigma_3(i) - \sigma_2(i)\right)}{\left(\sigma_3(i) - \sigma_1(i)\right)\left(\sigma_2(i) - \sigma_1(i)\right)} \left[\int_{\sigma_1(i)}^{\sigma_2(i)} \frac{d^2u}{dx^2}\, dx \right. \\
&\quad \left. - \frac{\left(\sigma_2(i) - \sigma_1(i)\right)}{\left(\sigma_3(i) - \sigma_2(i)\right)} \int_{\sigma_2(i)}^{\sigma_3(i)} \frac{d^2u}{dx^2}\, dx \right] \\
&\le \frac{1}{\left(\sigma_2(i) - \sigma_1(i)\right)} \int_{\sigma_1(i)}^{\sigma_2(i)} \frac{d^2u}{dx^2}\, dx - \frac{1}{\left(\sigma_3(i) - \sigma_2(i)\right)} \int_{\sigma_2(i)}^{\sigma_3(i)} \frac{d^2u}{dx^2}\, dx \\
&= \frac{1}{\left(\sigma_2(i) - \sigma_1(i)\right)^2} \int_{\sigma_1(i)}^{\sigma_2(i)} \left(\sigma_2(i) - \sigma_1(i)\right) \frac{d^2u}{dx^2}\, dx \\
&\quad - \frac{1}{\left(\sigma_3(i) - \sigma_2(i)\right)^2} \int_{\sigma_2(i)}^{\sigma_3(i)} \left(\sigma_3(i) - \sigma_2(i)\right) \frac{d^2u}{dx^2}\, dx.
\end{aligned}$$

For a positive integer s and a positive decreasing function Φ on an interval $[t_1, t_2]$, we have [5]

$$\int_{t_1}^{t_2} (t - t_1)^{s-1} \Phi(t)\, dt \le \frac{1}{s} \left(\int_{t_1}^{t_2} \Phi(t)^{1/s}\, dt \right)^s. \tag{31}$$

So, using the above inequality (31), we get

$$\begin{aligned}
p_i &\le \frac{1}{2} \left(\frac{1}{\left(\sigma_2(i) - \sigma_1(i)\right)} \int_{\sigma_1(i)}^{\sigma_2(i)} \left(\frac{d^2u}{dx^2} \right)^{1/2} dx \right)^2 \\
&\quad + \frac{1}{2} \left(\frac{1}{\left(\sigma_3(i) - \sigma_2(i)\right)} \int_{\sigma_2(i)}^{\sigma_3(i)} \left(\frac{d^2u}{dx^2} \right)^{1/2} dx \right)^2
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{2}\left(\frac{1}{(\sigma_2(i)-\sigma_1(i))}\int_{\sigma_1(i)}^{\sigma_2(i)}\left[1+\left(\frac{d^2u}{dx^2}\right)^{1/2}\right]dx\right)^2\\
&\quad+\frac{1}{2}\left(\frac{1}{(\sigma_3(i)-\sigma_2(i))}\int_{\sigma_2(i)}^{\sigma_3(i)}\left[1+\left(\frac{d^2u}{dx^2}\right)^{1/2}\right]dx\right)^2.
\end{aligned}
$$

Next, we use the equidistribution principle (22) and the observation (28) to get

$$
\begin{aligned}
p_i &\leq \frac{1}{2}\left(\frac{1}{(\sigma_2(i)-\sigma_1(i))}\frac{(\sigma_2(i)-\sigma_1(i))}{2M}\int_0^1 \mathcal{M}\,dx\right)^2\\
&\quad+\frac{1}{2}\left(\frac{1}{(\sigma_3(i)-\sigma_2(i))}\frac{(\sigma_3(i)-\sigma_2(i))}{2M}\int_0^1 \mathcal{M}\,dx\right)^2\\
&\leq CM^{-2}=C2^{-2j}.
\end{aligned}
$$

Now, we can estimate the error bound from (30) by substituting these bounds of p_i and the bounds of $\mathcal{P}_{i,2}(x)$ from (13) as follows:

$$
\begin{aligned}
\mathcal{E}^J &= \left|\sum_{i=2M+1}^{\infty} p_i\mathcal{P}_{i,2}(x)\right|\\
&= \left|\sum_{j=J+1}^{\infty}\sum_{l=0}^{2^j-1} p_{2^j+l+1}\mathcal{P}_{2^j+l+1,2}(x)\right|\\
&\leq \left|\sum_{j=J+1}^{\infty} C\frac{8}{3\left(\lfloor(3)/2\rfloor!\right)^2}2^{-2j-2}\right|\\
&\leq C\left|\sum_{j=J+1}^{\infty}\sum_{l=0}^{2^j-1}2^{-2j-2}\right|\\
&\leq C2^{-2J}=\mathcal{O}(M^{-2}).
\end{aligned}
$$

This result ensures the second-order convergence of the proposed method.

5 Numerical experiments

In this section, two important elements of numerical experiments are discussed; the numerical algorithm followed for the computation process and the numerical stability.

5.1 Adaptive mesh generation

An important feature of this work is the construction of a robust adaptive nonuniform mesh based on the equidistribution principle. Here, we introduce a variant of De Boor's algorithm [5] for the construction of adaptive mesh. Based on finite difference methods, this algorithm has been applied to various classes of singularly perturbed problems (see [4, 14, 15, 23, 22] and the references therein). Also, for some classes of problems, the convergence of this algorithm is discussed in [40, 7]. In discrete form, we write the equidistribution condition (22) as

$$(x_k - x_{k-1})\mathcal{M}_k = \frac{1}{2M}\sum_{i=1}^{2M}(x_i - x_{i-1})\mathcal{M}_i, \quad k = 1, \ldots, 2M, \tag{32}$$

where $\mathcal{M}_i$ is the discrete version of the monitor function. However, in [20], it is remarked that the above discrete form is not necessary to be enforced strictly, rather it will work sufficiently well to stop the iterative algorithm to the following weakened form:

$$(x_k - x_{k-1})\mathcal{M}_k \leq \frac{\mathcal{C}_0}{2M}\sum_{i=1}^{2M}(x_i - x_{i-1})\mathcal{M}_i, \quad k = 1, \ldots, 2M, \tag{33}$$

with $\mathcal{C}_0 > 1$ taken as a relaxing constant. By the user-chosen value of $\mathcal{C}_0$, one can manage the number of iterations and the adaptiveness of the mesh during the solution process. The choice of $\mathcal{C}_0$ works such that a smaller value of $\mathcal{C}_0$, results in a more number of iterations, and a more accurate solution. For the example problems, we are taking $\mathcal{C}_0 = 1.05$. The numerical algorithm below provides a complete overview of the adaptive mesh generation mechanism.

Numerical algorithm for the adaptive mesh and solution

Input: $J, M \in \mathbb{N},\ 0 < \varepsilon \leq 1,$ and $\mathcal{C}_0 > 1.$

Output: Adaptive mesh $\{x_k\}$ and adaptive solution U_k.

1. Initialize the mesh taking $x_k^{(r)} = k/2M,\ k = 0, ..., 2M$, with $r = 0$.

2. Calculate $U_k^{(r)}$ from (21) on the collocation mesh points $x_k^{c,(r)}$ given by (20).

3. Compute the monitor function by

$$\mathcal{M}_k^{(r)} = \beta^{(r)} + \left| \frac{d^2 U_k^{(r)}}{dx^2} \right|^{1/2}, 1 \leq k \leq 2M-1, \tag{34}$$

by using approximation (15) for second-order derivative and define

$$\beta^{(r)} = \left(x_1^{c,(r)} - x_0^{c,(r)}\right) \left[\sum_{i=1}^{2M} p_i h_i(x_0^{c,(r)})\right]^{1/2} + \sum_{k=2}^{2M} \left[\left(x_k^{c,(r)} - x_{k-1}^{c,(r)}\right) \left(\frac{\left[\sum_{i=1}^{2M} p_i h_i(x_k^{c,(r)})\right]^{1/2} + \left[\sum_{i=1}^{2M} p_i h_i(x_{k-1}^{c,(r)})\right]^{1/2}}{2} \right) \right].$$

4. For $\{x_k^{(r)}\}$ and $U_k^{(r)}$, set

$$H_k^{(r)} = \left(\frac{\mathcal{M}_{k-1}^{(r)} + \mathcal{M}_k^{(r)}}{2} \right) (x_k^{(r)} - x_{k-1}^{(r)}), \quad 1 \leq k \leq 2M,$$

where $\mathcal{M}_k^{(r)}$ is calculated in Step 3 setting $\mathcal{M}_0^{(r)} = \mathcal{M}_1^{(r)}$.

5. Set $\mathcal{B}_0^{(r)} = 0$ and $\mathcal{B}_k^{(r)} = \sum_{i=1}^{k} H_i^{(r)}$, $1 \leq k \leq 2M$. Then for stopping criteria define

$$\mathcal{C}^{(r)} := \frac{2M}{\mathcal{B}_{2M}^{(r)}} \max_{1 \leq k \leq 2M} H_k^{(r)}.$$

6. If $\mathcal{C}^{(r)} \leq \mathcal{C}_0$, then go to Step 9.

7. Set $Y_k = k\mathcal{B}_{2M}^{(r)}/2M$, $0 \leq k \leq 2M$. Generate new mesh $\{x_k^{(r+1)}\}$ by interpolating the points $(\mathcal{B}_k^{(r)}, x_k)$ and evaluating the interpolant at Y_k, $0 \leq k \leq 2M$.

8. Set $r = r + 1$ and return to Step 2.

9. Take $\{x_k^{(r)}\}$ as the final adaptive mesh and $U_k^{(r)}$ as the final solution. Stop.

5.2 Numerical stability

To show the stability of the proposed method, we follow the following definition of stability [26].

Definition 1. A numerical method for solving differential equations is said to be stable if, for the associated system of linear equations $\mathcal{A}\mathcal{X} = \mathcal{B}$, $\mathcal{A}^{-1}$ exists and is bounded. That is,

$$\|\mathcal{A}^{-1}\| \leq C.$$

Now, we provide the implementation of the proposed method in the following two test problems to confirm the theoretical findings with the help of some tables and graphs.

Example 1. [33] Let us consider the first example as the following singularly perturbed differential equation:

$$\varepsilon u''(x) + u'(x) = 1, \quad \text{for} \;\; 0 < x < 1, \tag{35}$$

concerning the given conditions $u'(0) = \frac{1}{\varepsilon}$ and $\int_0^1 u(x)dx = \frac{1}{2}$.

The exact solution to the above problem is known

$$u(x) = -(\varepsilon^2 - \varepsilon)\Big(1 - e^{(-\frac{1}{\varepsilon})}\Big) + (\varepsilon - 1)e^{(\frac{-x}{\varepsilon})} + x.$$

Example 2. [33] Let us consider the second example as the following singularly perturbed differential equation:

$$\varepsilon u''(x) + u'(x) = (\varepsilon - 2)e^{-x}, \quad \text{for} \;\; 0 < x < 1, \tag{36}$$

subject to the given conditions $u'(0) = \frac{1}{\varepsilon}$ and

$$\int_0^1 u(x)dx = 1 - e^{-1} + c_1 + 0.5\varepsilon c_2\Big(1 - e^{(\frac{-2}{\varepsilon})}\Big).$$

The exact solution to the above problem is known

$$u(x) = c_1 + c_2 e^{\left(\frac{-2x}{\varepsilon}\right)} + e^{-x},$$

where

$$c_1 = -\frac{3}{7}\Big(e^{-1} + \frac{1}{3}e^{\frac{-1}{4}} + (1 + e^{\frac{-2}{\varepsilon}}) + \frac{1}{3}e^{\frac{-1}{2\varepsilon}}\Big)c_2,$$
$$c_2 = -\frac{1+\varepsilon}{2}.$$

To confirm the numerical stability of the proposed method for the example problems, we present the following tables of the norm of the inverse of the wavelet coefficient matrix $\mathcal{A}$ for both examples in Table 1. The finite condition numbers of matrix $\mathcal{A}$ depict that the matrix is well-conditioned even if we are increasing the resolution levels. Therefore, we can conclude that the method is reasonably stable.

Table 1: Condition number of coefficient matrix for various resolution levels for both examples.

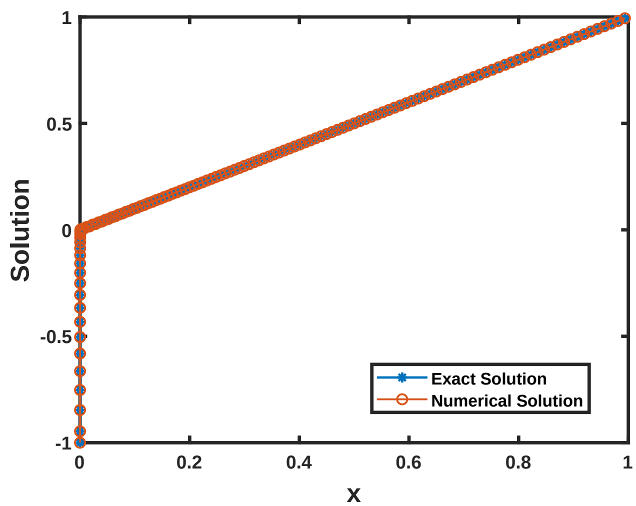
J	$\varepsilon = 10^{-2}$ Example 1	Example 2	$\varepsilon = 10^{-5}$ Example 1	Example 2
5	2.7119e+02	6.1539e+02	3.4215e+04	3.5816e+04
6	3.7465e+02	8.6552e+02	1.5364e+05	1.7826e+05
7	5.1816e+02	1.1982e+03	4.2509e+05	6.6213e+05
8	7.2116e+02	1.6866e+03	7.9665e+05	1.7912e+06
9	1.0101e+03	2.3338e+03	1.0983e+06	2.4694e+06

Figure 1 displays the comparison of exact and numerical solutions of both examples for $\varepsilon = 10^{-3}$ and $2M = 128$. These figures, clarify the effective resolution of the boundary layer phenomena by using the present adaptive mesh method. Furthermore, to demonstrate the relative applicability and effectiveness of the proposed method, we define the following formulas for the maximum absolute error $\mathcal{E}_\varepsilon^J$, and the order of convergence ρ_ε^J:

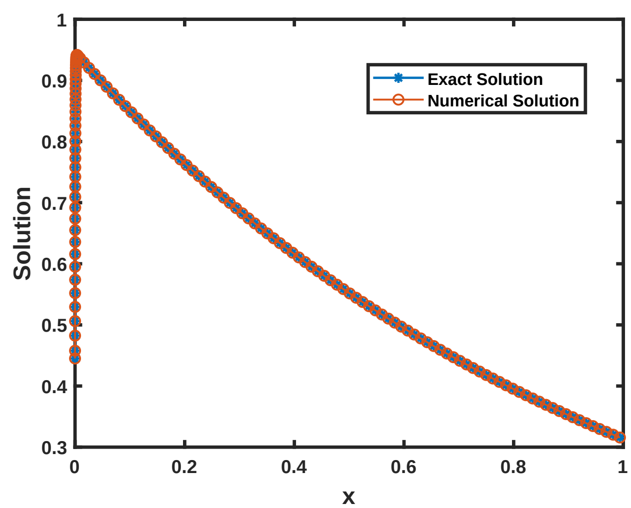
$$\mathcal{E}_\varepsilon^J = \max_{1\le i\le 2M} |U_k^J - u(x_k^c)|, \qquad \rho_\varepsilon^J = \log_2\left(\frac{\mathcal{E}_\varepsilon^J}{\mathcal{E}_\varepsilon^{J+1}}\right), \tag{37}$$

where U^J is the numerical solution at the resolution level J using the above algorithm, $u(x)$ is the exact solution. We also define the formulas for parameter-uniform errors and the parameter-uniform order of convergence

$$\mathcal{E}^J = \max_{\varepsilon} |\mathcal{E}_\varepsilon^J|, \qquad \rho^J = \log_2\left(\frac{\mathcal{E}^J}{\mathcal{E}^{J+1}}\right). \tag{38}$$



(a) Solution plot for Example 1



(b) Solution plot for Example 2

Figure 1: Comparison of the numerical solution obtained by the proposed method with the exact solution for $\varepsilon = 10^{-3}$ and $2M = 128$.

The maximum absolute errors and the order of convergence obtained by using these formulas for different resolution levels and different values of ε

are presented in Tables 2 and 3. From these tables, we see the second-order convergence rates, which are going to be stabilized as we move downwards along the columns in the tables. Thus, from the last few rows of these tables, we can observe the second-order parameter uniform nature of the proposed method.

Table 2: Maximum point-wise errors $\mathcal{E}_\varepsilon^J$, parameter-uniform errors $\mathcal{E}^J$, order of convergence ρ_ε^J, and parameter-uniform convergence order ρ^J using the proposed method for Example 1.

$\varepsilon = 10^{-r}$	$J=5$	$J=6$	$J=7$	$J=8$	$J=9$
$r=1$	1.3259e-04	3.2752e-05	8.1922e-06	2.0484e-06	5.1211e-07
	2.0173	1.9993	1.9998	2.0000	
$r=2$	2.8088e-04	6.9246e-05	1.7126e-05	4.2623e-06	1.0634e-06
	2.0202	2.0155	2.0065	2.0029	
$r=3$	3.6121e-04	8.0450e-05	1.8144e-05	4.3778e-06	1.0694e-06
	2.1667	2.1486	2.0512	2.0334	
$r=4$	4.9163e-04	8.3995e-05	1.8813e-05	4.4019e-06	1.0789e-06
	2.5492	2.1586	2.0955	2.0286	
$r=5$	5.0123e-04	8.5858e-05	1.9028e-05	4.4241e-06	1.0919e-06
	2.5454	2.1738	2.1047	2.0185	
$r=6$	1.9057e-03	8.8126e-05	1.9267e-05	4.4415e-06	1.0959e-06
	4.4346	2.1934	2.1170	2.0189	
$r=7$	4.9062e-03	1.6531e-04	1.9367e-05	4.4512e-06	1.0966e-06
	4.8914	3.0935	2.1213	2.0212	
$\mathcal{E}^J$	4.9062e-03	1.6531e-04	1.9367e-05	4.4512e-06	1.0966e-06
ρ^J	4.8914	3.0935	2.1213	2.0212	

In Tables 4 and 5, we compare the results obtained by the present method based upon a nonuniform equidistribution mesh with the second-order convergent numerical method in [33] based on a nonuniform graded mesh. We can observe that we can achieve better accuracy in the same efforts, that is, with the same discretization parameters from the same values of perturbation parameter. This also supports the suitability of the proposed method.

Table 6 shows the comparison of computational time (in seconds) in CPU for the wavelet approximation on the equidistribution mesh and the graded mesh. The CPU time is reported by averaging a few executions for a fixed value of $\varepsilon = 10^{-5}$ at each value of $J = 5, \ldots, 10$. It can be seen that the computational time is lesser for equidistribution mesh for smaller levels of

Table 3: Maximum point-wise errors $\mathcal{E}_\varepsilon^J$, parameter-uniform errors $\mathcal{E}^J$, order of convergence ρ_ε^J, and parameter-uniform convergence order ρ^J using the proposed method for Example 2.

$\varepsilon = 10^{-r}$	$J=5$	$J=6$	$J=7$	$J=8$	$J=9$
$r=1$	1.5675e-04	3.9077e-05	9.7696e-06	2.4427e-06	6.1071e-07
	2.0041	1.9999	1.9998	1.9999	
$r=2$	3.2592e-04	8.0649e-05	2.0186e-05	5.0492e-06	1.2643e-06
	2.0148	1.9983	1.9992	1.9977	
$r=3$	4.6697e-04	9.4866e-05	2.3479e-05	5.8828e-06	1.4740e-06
	2.2994	2.0145	1.9968	1.9968	
$r=4$	5.3535e-04	1.0452e-04	2.5524e-05	6.1883e-06	1.5182e-06
	2.3567	2.0339	2.0442	2.0272	
$r=5$	6.0230e-04	1.0858e-04	2.6228e-05	6.4441e-06	1.6019e-06
	2.4717	2.0496	2.0251	2.0082	
$r=6$	7.0127e-03	1.1826e-04	2.7867e-05	6.6985e-06	1.6719e-06
	5.8899	2.0853	2.0566	2.0023	
$r=7$	9.9962e-03	1.9131e-04	2.9267e-05	6.8812e-06	1.7166e-06
	5.7074	2.7086	2.0885	2.0031	
$\mathcal{E}^J$	9.9962e-03	1.9131e-04	2.9267e-05	6.8812e-06	1.7166e-06
ρ^J	5.7074	2.7086	2.0885	2.0031	

Table 4: Comparison of maximum point-wise errors for Example 1.

	$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-7}$	
J	Present method	Method in [33]	Present method	Method in [33]
5	3.6121e-04	1.7061e-03	4.9062e-03	1.7134e-03
6	8.0450e-05	4.2778e-04	1.6531e-04	4.2960e-04
7	1.8144e-05	1.0701e-04	1.9367e-05	1.0746e-04
8	4.3778e-05	2.6753e-05	4.4512e-06	2.6855e-05
9	1.0694e-06	6.6882e-06	1.0966e-06	6.7071e-06

resolutions. However as the resolution increases, because of the iterative algorithm of the adaptive mesh generation computational time increases fast. All the numerical experiments are performed using MATLAB R2019a (Mathworks Inc.) on a 64-bit Windows 10 machine, with Intel(R) Core(TM) I5 7th Gen 7200U processor running at 2.50 GHz and 8.00 GB RAM.

At last, we included two Figures 2 and 3 for both examples, respectively, showing the trajectory of the formation of adaptive mesh per iteration and the

Table 5: Comparison of maximum point-wise errors for Example 2.

J	$\varepsilon = 10^{-3}$ Present method	Method in [33]	$\varepsilon = 10^{-7}$ Present method	Method in [33]
5	4.6697e-04	1.6560e-03	9.9962e-03	1.6560e-03
6	9.4866e-05	4.1311e-04	1.9131e-04	4.1298e-04
7	2.3479e-05	1.0331e-04	2.9267e-05	1.0327e-04
8	5.8828e-06	2.5820e-05	6.8812e-06	2.5793e-05
9	1.4740e-06	6.4546e-06	1.7166e-06	6.4489e-06

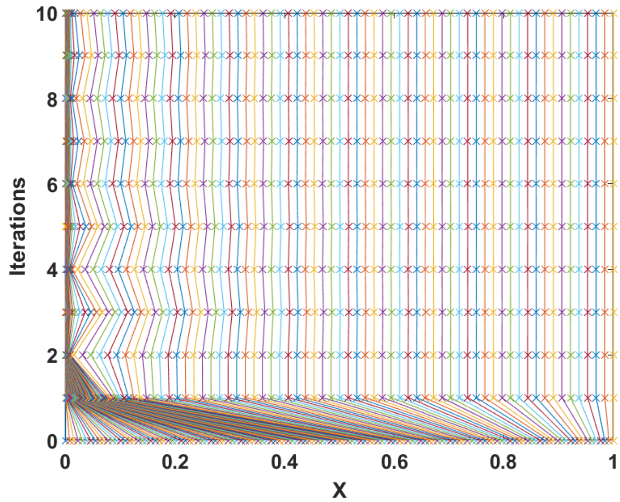
Table 6: Comparison of CPU computational times (in seconds) for the wavelet approximations on graded mesh [33] and equidistributed meshes with $\varepsilon = 10^{-5}$ for Example 1.

J	5	6	7	8	9	10
Equidistribution mesh	1.5844	1.6693	1.6898	1.7221	2.1088	17.2492
Graded mesh [33]	1.6718	1.6899	1.7241	1.7175	2.0794	14.9256

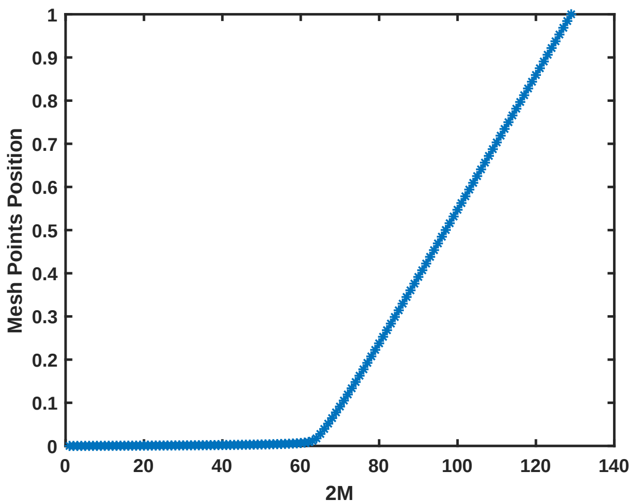
final position of adaptive mesh points. These figures show that the equidistribution mesh is highly adaptive towards the boundary layers present in singular perturbation problems.

6 Conclusions:

The purpose of this article was to introduce an efficient nonuniform wavelet-based numerical method on an adaptive mesh generated by equidistribution of a specially chosen monitor function for a class of singularly perturbed convection-diffusion problems with nonlocal boundary conditions. Rigorous error analysis for the proposed method was done and the parameter uniform convergence of the second-order is established. Furthermore, the computational stability of the proposed method was observed for two example problems, and the numerical results confirmed the convergence result. The application of a wavelet method together with a posteriorly generated equidistribution mesh is itself a novelty and the significant difference in the accuracy of results proves the best applicability of the proposed method over existing numerical methods. Applying the wavelet methods on a posteriorly gener-



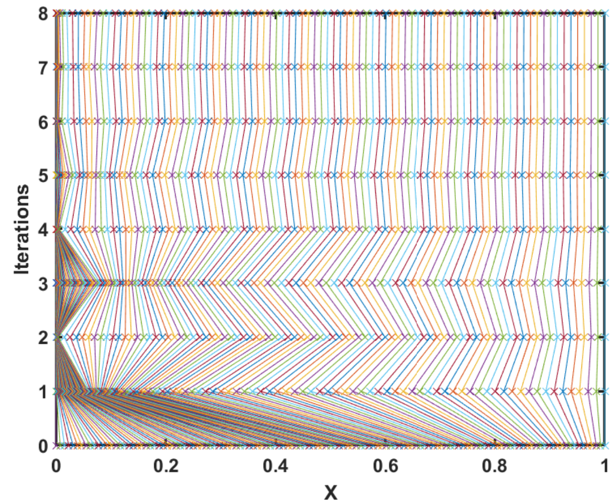
(a) Mesh trajectory per iteration



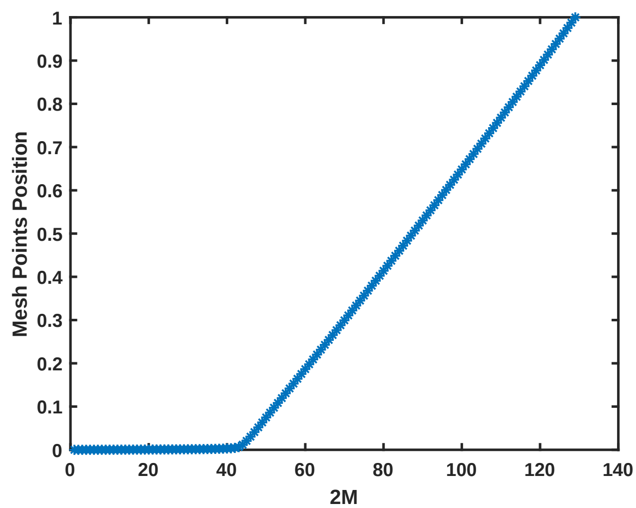
(b) Final position of mesh points

Figure 2: Moving mesh trajectory for each iteration and density of mesh points in the final adaptive mesh, respectively, for Example 1 with $\varepsilon = 10^{-3}$ and $2M = 128$.

ated equidistribution mesh for more complex problems can be the motivation for many future works.



(a) Mesh trajectory per iteration



(b) Final position of mesh points

Figure 3: Moving mesh trajectory for each iteration and density of mesh points in the final adaptive mesh, respectively, for Example 2 with $\varepsilon = 10^{-3}$ and $2M = 128$.

Acknowledgments

For the first author, this work is financially supported by FIG grant ABV-IIITM/DORC/FIG/012/2023-4008 from ABV-IIITM Gwalior, Madhya Pradesh, India.

Declarations

The authors declare that they have no conflicts of interest.

References

[1] Ahsan, M., Bohner, M., Ullah, A., Khan, A.A. and Ahmad, S. *A haar wavelet multiresolution collocation method for singularly perturbed differential equations with integral boundary conditions*, Math. Comput. Simul. 204 (2023), 166–180.

[2] Bakhvalov, N.S. *Towards optimization of methods for solving boundary value problems in the presence of boundary layers*, Zh. Vychisl. Mat. Mat. Fiz. 9, (1969), 841–859.

[3] Beckett, G. and Mackenzie, J.A. *Convergence analysis of finite difference approximations on equidistributed grids to a singularly perturbed boundary value problem*, Appl. Numer. Math. 35(2) (2000), 87–109.

[4] Beckett, G. and Mackenzie, J.A. *On a uniformly accurate finite difference approximation of a singularly perturbed reaction–diffusion problem using grid equidistribution*, J. Comput. Appl. Math. 131(1-2) (2001), 381–405.

[5] Boor, C. *Good approximation by splines with variable knots,* In: Spline functions and approximation theory: Proceedings of the Symposium Held at the University of Alberta, Edmonton May 29 to June 1, 1972, Springer, (1973) 57–72.

[6] Cakir, M. and Amiraliyev, G.M. *A finite difference method for the singularly perturbed problem with nonlocal boundary condition*, Appl. Math. Comput. 160(2) (2005), 539–549.

[7] Chadha, N.M. and Kopteva, N. *A robust grid equidistribution method for a one-dimensional singularly perturbed semilinear reaction–diffusion problem*, IMA J. Numer. Anal. 31(1) (2011), 188–211.

[8] Das, P. and Mehrmann, V. *Numerical solution of singularly perturbed convection-diffusion-reaction problems with two small parameters*, BIT Numer. Math. 56(1) (2016), 51–76.

[9] Das, P. and Natesan, S. *Higher-order parameter uniform convergent schemes for robin type reaction-diffusion problems using adaptively generated grid*, Int. J. Comput. Methods, 9(04) (2012), 1250052.

[10] Das, P., Rana, S. and Vigo-Aguiar, J. *Higher order accurate approximations on equidistributed meshes for boundary layer originated mixed type reaction diffusion systems with multiple scale nature*, Appl. Numer. Math. 148 (2020), 79–97.

[11] Debela, H.G. and Duressa, G.F. *Uniformly convergent numerical method for singularly perturbed convection-diffusion type problems with nonlocal boundary condition*, Int. J. Numer. Methods Fluids, 92(12) (2020), 1914–1926.

[12] Dubeau, F., Elmejdani, S. and Ksantini, R. *Non-uniform haar wavelets*, Appl. Math. Comput. 159(3) (2004), 675–693.

[13] Goswami, J.C. and Chan, A.K. *Fundamentals of wavelets: Theory, algorithms, and applications*, John Wiley & Sons, 2011.

[14] Gowrisankar, S. and Natesan, S. *The parameter uniform numerical method for singularly perturbed parabolic reaction–diffusion problems on equidistributed grids*, Appl. Math. 26(11) (2013), 1053–1060.

[15] Gowrisankar, S. and Natesan, S. *Robust numerical scheme for singularly perturbed convection–diffusion parabolic initial–boundary-value problems on equidistributed grids*, Comput. Phys. Commun. 185(7) (2014), 2008–2019.

[16] Haar, A. *Zur Theorie der Orthogonalen Funktionensysteme*. Georg-August-Universitat, Gottingen, 1909.

[17] Hirsch, C. *Numerical computation of internal and external flows: The fundamentals of computational fluid dynamics*, Elsevier, 2007.

[18] Huang, W. and Russell, R.D. *Adaptive moving mesh methods*, Springer, 2010.

[19] Kopteva, N., Madden, N. and Stynes, M. *Grid equidistribution for reaction–diffusion problems in one dimension*, Numer. Algorithms, 40(3) (2005), 305–322.

[20] Kopteva, N., and Stynes, M.*A robust adaptive method for a quasi-linear one-dimensional convection-diffusion problem*, SIAM J. Numer. Anal. 39(4) (2001), 1446–1467.

[21] Kumar, S. and Kumar, M. *Parameter-robust numerical method for a system of singularly perturbed initial value problems*, Numer. Algorithms, 59(2) (2012), 185–195.

[22] Kumar, S., Kumar, S. and Sumit, *A posteriori error estimation for quasilinear singularly perturbed problems with integral boundary condition*, Numer. Algorithms, 89(2) (2022), 791–809.

[23] Kumar, S., Sumit and Vigo-Aguiar, J. *A parameter-uniform grid equidistribution method for singularly perturbed degenerate parabolic convection–diffusion problems*, J. Comput. Appl. Math. 404 (2020), 113273.

[24] Lepik, U. *Numerical solution of differential equations using haar wavelets*. Math. Comput. Simul. 68(2), (2005) 127–143.

[25] Lepik, Ü. and Hein, H. *Haar wavelets with applications*, Springer, New York, 2014.

[26] LeVeque, R.J. *Finite difference methods for ordinary and partial differential equations: Steady-state and time-dependent problems*, SIAM, 2007.

[27] Liu, L.-B., Long, G. and Cen, Z. *A robust adaptive grid method for a nonlinear singularly perturbed differential equation with integral boundary condition*, Numer. Algorithms, 83(2) (2020), 719–739.

[28] Mackenzie, J.A. *Uniform convergence analysis of an upwind finite-difference approximation of a convection-diffusion boundary value problem on an adaptive grid*, IMA J. Numer. Anal. 19(2) (1999), 233–249.

[29] Mallat, S. *A wavelet tour of signal processing*, Elsevier 1999.

[30] Melenk, J.M. *Hp-finite element methods for singular perturbations*, Springer, 2002.

[31] Miller, J.J.H., O'Riordan, E. and Shishkin, G.I. *Fitted numerical methods for singular perturbation problems: Error estimates in the maximum norm for linear problems in one and two dimensions*, World Scientific, 2012.

[32] Pandit, S. and Kumar, M. *Haar wavelet approach for numerical solution of two parameters singularly perturbed boundary value problems*, Appl. Math. Inf. Sci. 8(6) (2014), 2965.

[33] Podila, P.C. and Sundrani, V. *A non-uniform haar wavelet method for a singularly perturbed convection–diffusion type problem with integral boundary condition on an exponentially graded mesh*. Comput. Appl. Math. 42(5) (2023), 216.

[34] Qiu, Y. and Sloan, D.M. *Analysis of difference approximations to a singularly perturbed two-point boundary value problem on an adaptively generated grid*, J. Comput. Appl. Math. 101(1-2) (1999), 1–25.

[35] Roos, H.G., Stynes, M. and Tobiska, L. *Numerical methods for singularly perturbed differential equations*, Springer, 1996.

[36] Sah, K.K. and Gowrisankar, S. *Richardson extrapolation technique on a modified graded mesh for singularly perturbed parabolic convection-diffusion problems*, Iran. J. Numer. Anal. Optim. 14(1) (2024), 219–264.

[37] Shishkin, G.I. *A difference scheme for a singularly perturbed equation of parabolic type with discontinuous boundary conditions*, USSR Comput. Math. Math. Phys. 28(6) (1988), 32–41.

[38] Wichailukkana, N., Novaprateep, B. and Boonyasiriwat, C. *A convergence analysis of the numerical solution of boundary-value problems by using two-dimensional haar wavelets*, Sci. Asia, 42 (2016), 346–355.

[39] Woldaregay, M.M. and Duressa, G.F. *Exponentially fitted tension spline method for singularly perturbed differential difference equations*, Iran. J. Numer. Anal. Optim. 11(2) (2021), 261–282.

[40] Xu, X., Huang, W., Russell, R.D. and Williams, J.F. *Convergence of de Boor's algorithm for the generation of equidistributing meshes*, IMA J. Numer. Anal. 31(2) (2011), 580–596.

[41] Yapman, O. and Amiraliyev, G.M. *A novel second-order fitted computational method for a singularly perturbed Volterra integro-differential equation*, Int. J. Comput. Math. 97(6) (2019), 1–10.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1168–1202
https://doi.org/10.22067/ijnao.2024.88052.1448
https://ijnao.um.ac.ir/

Research Article

Mathematical modeling and optimal control of customer's behavior toward e-commerce

H. Hassouni*, S. Belhdid and O. Balati

Abstract

The extensive influence of digital platforms has reshaped societal interactions and daily routines, integrating e-commerce into every aspect of modern life. This evolution not only redefines traditional business models

*Corresponding author

Received 13 May 2024; revised 29 July 2024; accepted 1 August 2024

Habib Hassouni
Laboratory of Fundamental Mathematics and their Applications, Department of Mathematics, Faculty of Sciences El Jadida, Chouaib Doukkali University, El Jadida, Morocco. e-mail: habibhassouni85@gmail.com

Salaheddine Belhdid
Laboratory of Fundamental Mathematics and their Applications, Department of Mathematics, Faculty of Sciences El Jadida, Chouaib Doukkali University, El Jadida, Morocco. e-mail: salaheddine.belhdid@gmail.com

Omar Balatif
Laboratory of Fundamental Mathematics and their Applications, Department of Mathematics, Faculty of Sciences El Jadida, Chouaib Doukkali University, El Jadida, Morocco. e-mail: balatif.maths@gmail.com

How to cite this article
Hassouni, H., Belhdid, S. and Balati, O., Mathematical modeling and optimal control of customer's behavior toward e-commerce. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1168-1202. https://doi.org/10.22067/ijnao.2024.88052.1448

but also fosters global connectivity and economic restructuring. However, despite its critical role in the global economy, e-commerce faces challenges, notably the hesitation of some consumers due to concerns about security and trust. To address this, we propose a novel mathematical model to examine customer behavior dynamics toward e-commerce, particularly the impact of the refusal behavior. Our study comprehensively examines the characteristics of our mathematical model, conducts a thorough stability analysis, and investigates the parameter sensitivity. Furthermore, control theory has been adopted to optimize the adoption of e-commerce using Pontryagin's maximum principle, with numerical simulations to evaluate the effectiveness of our proposed strategies.

AMS subject classifications (2020): Primary 93-10; Secondary 49J20, 34D20.

Keywords: Mathematical modeling, Stability analysis, E-commerce, Pontryagin Maximum principle, Sensitivity Analysis.

1 Introduction

The increasing dependence on digital platforms and the acceleration of technical progress have fundamentally transformed how people engage, converse, and go through their daily lives in modern society [7, 30]. Digital platforms have become an integral component in almost every aspect of contemporary life, ranging from interpersonal communication to career pursuits [16]. Global connectivity is fostered by social media, e-commerce, and online communication tools, which have become indispensable aspects of daily life. This significant change is redefining conventional corporate models and economic structures in addition to changing societal dynamics [36, 35]. The transformative potential of technology is clear as more people and organizations gravitate to digital solutions. This is because technology not only improves efficiency and convenience but also catalyzes paradigm shifts across varied industries, laying the groundwork for a future where everything is connected through the Internet. On the grounds of this, the major developments that e-commerce encountered made it an essential foundation for the growth of the nation's economy, as well as for the country's opening toward the worldwide market's accessibility of product presentation both domestically and abroad, and convenience of buying from any domestic or foreign market at any time or location [11, 39].

Nevertheless, in the contemporary landscape, e-commerce is one of the most important aspects of the global economy in the modern world. It has a significant impact on how businesses operate and how customers engage with goods and services [6]. In [29], the authors encountered the ever-expanding reach of digital platforms that have not only facilitated the seamless exchange of goods and services but has also blurred the traditional boundaries of market accessibility.

Online businesses or shops notice that their offered services can cease functioning due to some technical problems or more importantly, a decreasing number of customers, and sometimes it is difficult to control this behavior due to the lack of time or a misunderstanding of the customer's behavior toward online businesses or e-commerce in general. This lack of information can blur the future of the online business economy. Consequently, these situations serve as the inspiration behind the research question of this paper: How can we more significantly understand the customer's behavior toward e-commerce?

Regarding the enormous role e-commerce has in our lives and how it impacts people's behavior, scholars explored the interaction between people and e-commerce. In [8, 31], the authors analyzed the correlation between e-commerce and people's trust, forecasting the differences between a virtual and a conventional marketplace and how customers' trust affects their consumption and vision toward e-commerce. Moreover, some other researchers talked about and provided examples of the age, educational background, and income requirements of users on e-commerce platforms [26], and even the rumors dissemination impact on it [5]. Therefore, the significant attention that the public and governments accord e-commerce has created numerous opportunities for business owners. Thus, academics paid attention to this topic and studied the relationship between e-commerce and entrepreneurship [37, 15]. From another perspective, in [9], the authors investigated the impact of COVID-19 on e-commerce and its effects on consumer purchase behavior.

Following the pattern of modeling the customer's behavior, it is getting increasingly important to better understand the dynamic of customers toward these online businesses regarding the great evolution of digital platforms and the customer's needs and requirements. Consequently, various academics investigated the subject. In [41], the authors focused on the modeling of customer preferences using multi-network analysis, which gives a significant database to understand the customer's conduct. Furthermore, Park and Fader [32] developed a stochastic model of website visiting behavior to better understand what makes a customer change his conduct toward one web website to another. During the time customers spend on websites, they certainly leave behind a pattern trait that sums up their behavior. This process generates what we call big data information and web analytics that are significant in understanding customers' behavior when being on the website [13].

Perceiving the supreme role e-commerce took in people's lives and the correlation it has with our behavior and consumption aptitude, it is instinctive to better understand and analyze the process of how customers act toward e-commerce. Consequently, a proper and effective modeling approach must be used, such as the compartmental model approach, which was first introduced by Kermack and McKendrick [21] in 1927 in the Proceedings of the Royal Society of London, which is a seminal work that laid the groundwork for the mathematical modeling of infectious diseases, principally named the SIR models. Inspired by the epidemiological modeling approach, a population can be divided into three main compartments: susceptible, infected, and recovered, which make it convenient to use for customer's behavior modeling for a better and simpler understanding, interpreting, and controlling. In the literature, we can see that all the proposed mathematical modeling of customers' behavior toward e-commerce does not take into consideration the influence of customers with each other and the bad influence a refusal behavior can have on a group of customers toward these shopping websites. That is why a better modeling approach must take place. In our work, we formulate a new mathematical model to overview and understand the dynamic behavior of customers toward e-commerce and specifically to study the influence of customers' refusal to use e-commerce on others. To that end, we propose a compartmental model of four compartments (potential customers, refusers of e-commerce, users of e-commerce, and finally, those who temporally stop using e-commerce) that widely describe the interaction between customers regarding their behavior in the e-commerce environment. We study the main characteristics of the proposed model to discover the various aspects of the proposed modeling and to better understand the process of customers' behavior towards e-commerce. Nevertheless, the stability analysis of the proposed model is investigated in its local and global aspects for the equilibrium points subjected to this model. We also use control theory analysis and the Pontryagin maximum principle to seek our objective of maximization of users of e-commerce platforms and minimizing the refusers of e-commerce, which consequently will have a great impact on the global economy.

Control theory has been applied to various problems by a range of authors, such as the electoral behavior optimal control strategies regarding the political appurtenance [3]. In addition, in [18], the authors applied mathematical modeling and optimal control strategies to prevent the spread of the coronavirus. Further research on control theory can be found in the following works: [25, 22, 10, 2].

In our research, we formulate a new mathematical model to study the impact of the behavior of customers who refuse to use e-commerce on e-commerce. We study the basic characteristics of the proposed model. We will also apply control theory to this model to increase the number of customers using e-commerce.

The research is organized as follows: In Section 2, a mathematical model is presented that analyzes customers' behavior toward e-commerce. Then, in Section 3, a stability analysis is performed around the equilibrium points, in both the locally and globally aspects. Furthermore, a sensitivity analysis of the reproduction number parameters is also performed to find out the most effective parameters for the optimal control problem of the model. Following that, Section 4 presents the optimal control problem, where the existence and characterization of the proposed controls are proven. Section 5 showcases the numerical simulations of the model's different compartments to evaluate the effectiveness of the proposed strategies. Finally, the paper is concluded in Section 6.

2 The mathematical model formulation

In this research, the problem studied is that of customers who refuse to use electronic commerce because they are afraid of electronic crime or fraud, or because they are afraid that their personal data will be stolen. To study this problem mathematically, we have proposed a deterministic, non-linear model that describes the problem in all its aspects and its relationship with the rest of the parties that could be affected by the refusal to use electronic commerce. The proposed model has four compartments.

2.1 Description of the model

Compartment P: It represents potential customers for electronic commerce who have never made any purchases through electronic stores located on the Internet, knowing that these people are connected to the Internet. The behavior of this compartment is due to a group of reasons, including lack of knowledge of electronic commerce, fear of electronic crimes, fear of fraud regarding the quality of products, fear of leakage of personal data, or living in places far from means of transportation for transporting goods. The number of members of this compartment increases with the number Λ, which represents the rate of new people connected to the Internet. The number of individuals in this compartment is reduced by the number β, which represents the percentage of people affected by the refusal to use e-commerce, the number α, which represents the percentage of people who decided to use e-commerce as a result of positive thoughts about it, and the number μ, which represents the percentage of natural deaths of customers.

Compartment A: It represents customers who refuse to use e-commerce because of negative thoughts about it, knowing that the members of this compartment have never purchased through e-commerce stores. The number of members of this compartment increases with the number β, which represents the percentage of people affected by the refusal to use e-commerce, while the number of members of this compartment decreases with the numbers γ and μ, which, respectively, represent the percentage of people who refused to use e-commerce but decided to use it as a result of positive thoughts about it and the percentage of natural deaths of customers.

Compartment E**:** This category represents customers who use e-commerce, and they are people who purchase from electronic stores on the Internet. The number of members of this compartment increases by three numbers α, γ, and θ, which, respectively, represent the percentage of people who decided to use e-commerce as a result of positive thoughts about it, the percentage of people who refused to use e-commerce but decided to use it as a result of positive thoughts about it, and the rate of people who returned to using it. The number of people in this compartment is reduced by the number μ, which represents the percentage of natural deaths of customers, and the number δ, which represents the percentage of people who stopped using e-commerce as a result of a problem.

Compartment D**:** It represents customers who have temporarily stopped using e-commerce, and they are people who have temporarily stopped using e-commerce as a result of a group of factors, including exposure to fraud, traveling to places where the Internet is not available, or exposure to electronic crime. The number of members of this compartment increases with the number δ, which represents the percentage of people who stopped using e-commerce as a result of a problem. The number of people in this compartment is reduced by the number μ, which represents the percentage of natural deaths of customers, and by the number θ, which represents the percentage of people who returned to using e-commerce after solving their previous problem.

The schematic diagram in Figure 1 showcases the trends of customer flows between compartments. We will fill these trends with directional arrows to further clarify these flows. The parameters used in this model are summarized in Table 1.

The model outlines a dynamic flow of customers through four distinct compartments, each representing different stages of engagement with electronic commerce (e-commerce). Initially, individuals enter the system as potential customers (Compartment P), influenced by factors such as their familiarity with e-commerce, concerns about security, and geographical accessibility. From there, they may transition to become active customers

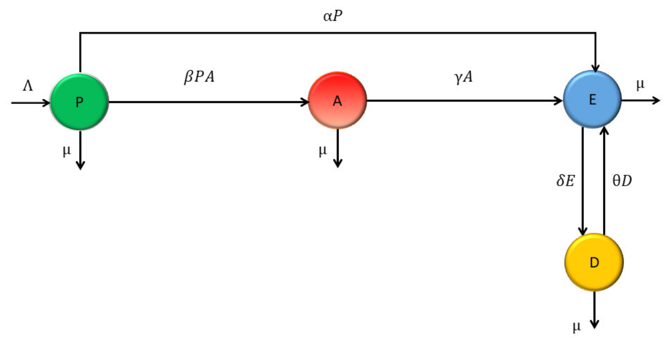


Figure 1: Compartments model diagram.

(Compartment E) through positive experiences or resolution of initial apprehensions. Alternatively, some individuals may initially refuse to engage with e-commerce (Compartment A) due to negative perceptions but may later be persuaded to join the active customer base through positive interventions. However, the model also accounts for fluctuations in customer behavior, as evidenced by the existence of temporarily inactive customers (Compartment D), who may pause their e-commerce activities due to encountered problems but could return to active engagement following issue resolution. Throughout this process, the flow of individuals between compartments is governed by rates of adoption, refusal, retention, and attrition, reflecting the complex interplay of factors shaping e-commerce participation. Following the hypothesis and description above, we present the mathematical model for customer's behavior, who refuses to use e-commerce when purchasing their daily needs. The model is governed by the following system of differential equations:

$$\begin{cases} \frac{dP(t)}{dt} = \Lambda - \beta P(t)A(t) - \alpha P(t) - \mu P(t), \\ \frac{dA(t)}{dt} = \beta P(t)A(t) - \gamma A(t) - \mu A(t), \\ \frac{dE(t)}{dt} = \alpha P(t) + \gamma A(t) + \theta D(t) - \delta E(t) - \mu E(t), \\ \frac{dD(t)}{dt} = \delta E(t) - \theta D(t) - \mu D(t), \end{cases} \tag{1}$$

Table 1: Model's parameter description

Parameter	Description
Λ	The rate of new people connected to the Internet.
β	The Rate of people affected by refusal to use e-commerce.
α	The rate of people who decided to use e-commerce as a result of positive thoughts about it.
γ	The rate of people who refused to use e-commerce but decided to use it as a result of positive thoughts about it.
δ	The rate of people who stopped using e-commerce as a result of a problem.
θ	The rate of people who returned to using e-commerce after solving their previous problem.
μ	Natural mortality rate of clients.

with the initial conditions $P(0) \geq 0$, $A(0) \geq 0$, $E(0) \geq 0$, and $D(0) \geq 0$. The total number of people who can buy from online stores at time t is $N(t)$, so we have $N(t) = P(t) + A(t) + E(t) + D(t)$.

2.2 Model's fundamental characteristics

Considering that the system model we proposed describes the dynamic of a human population, it is primordial to demonstrate that the system's solutions will remain positive for any initial positive condition and for all time $t > 0$. For the system model (1) and for $t > 0$, we have the following properties:

$$\begin{aligned}
\left.\frac{dP}{dt}\right|_{P=0} &= \Lambda \geq 0,\\
\left.\frac{dA(t)}{dt}\right|_{A=0} &= 0,\\
\left.\frac{dE(t)}{dt}\right|_{E=0} &= \alpha P(t) + \gamma A(t) + \theta D(t) \geq 0,\\
\left.\frac{dD(t)}{dt}\right|_{D=0} &= \delta E(t) \geq 0.
\end{aligned}$$

Thus, noting that all rates are nonnegative at the boundary of the nonnegative cone in $\mathbb{R}^4_+$ and considering the inward directions of the vector fields on these boundaries, starting the system from any interior point within this cone ensures that the solutions will remain inside the cone indefinitely.

2.2.1 Invariant region

In this section, we came to the conclusion that all possible solutions of the system described by model (1) are bounded.

Theorem 1. Let the region Ω be defined by
$\Omega = \left\{ (P(t),\ A(t),\ E(t),\ D(t)) \in \mathbb{R}_+^4,\ 0 \leq P(t) + A(t) + E(t) + D(t) \leq \frac{\Lambda}{\mu} \right\}.$
If the initial conditions $P(0)$, $A(0)$, $E(0)$, and $D(0)$ are positive, then the region Ω is positively invariant for the model (1).

Proof. To demonstrate the positivity invariance of Ω in the context of Theorem 1, let $t \geq 0$.

We have

$$\begin{aligned} \frac{dN(t)}{dt} &= \Lambda - \mu P(t) - \mu A(t) - \mu E(t) - \mu D(t), \\ &= \Lambda - \mu N(t). \end{aligned}$$

Hence

$$N(t) = N(0)\exp(-\mu t) + \left(\frac{\Lambda}{-\mu}\right)(\exp(-\mu t) - 1).$$

This implies

$$\lim_{t \to +\infty} \sup N(t) = \frac{\Lambda}{\mu}.$$

Therefore

$$N(t) \leq \frac{\Lambda}{\mu}, \tag{2}$$

where $N(t) = P(t) + A(t) + E(t) + D(t)$. This inequality ensures that trajectories starting within the region defined by

$$\Omega = \left\{ (P(t), A(t), E(t), D(t)) \in \mathbb{R}_+^4 \mid 0 \leq P(t) + A(t) + E(t) + D(t) \leq \frac{\Lambda}{\mu} \right\}$$

remain within Ω over time. Specifically, $N(t)$ cannot exceed $\frac{\Lambda}{\mu}$ due to the upper limit condition, thereby establishing the positivity invariance of Ω under the dynamics of model (1) [14]. This property allows us to focus our analysis exclusively on the behaviors within Ω. $\square$

2.2.2 Existence of solutions

In this subsection, we delve into the fundamental question of the existence of solutions within the context of our proposed compartment model. To this end, we present and prove the following theorem.

Theorem 2. System (1) that fulfills the initial conditions $P(0) \geq 0$, $A(0) \geq 0$, $E(0) \geq 0$, and $D(0) \geq 0$ has a unique solution.

Proof. Let

$$X = \begin{pmatrix} P(t) \\ A(t) \\ E(t) \\ D(t) \end{pmatrix} \quad \text{and} \quad \Psi(X) = \begin{pmatrix} \frac{dP(t)}{dt} \\ \frac{dA(t)}{dt} \\ \frac{dE(t)}{dt} \\ \frac{dD(t)}{dt} \end{pmatrix}.$$

Then

$$\Psi(X) = SX + Q(X),$$

in which

$$S = \begin{bmatrix} -\alpha-\mu & 0 & 0 & 0 \\ 0 & -\gamma-\mu & 0 & 0 \\ \alpha & \gamma & -\delta-\mu & \theta \\ 0 & 0 & \delta & -\theta-\mu \end{bmatrix} \quad \text{and} \quad Q(X) = \begin{bmatrix} \Lambda - \beta P(t)A(t) \\ \beta P(t)A(t) \\ 0 \\ 0 \end{bmatrix},$$

$$\begin{aligned} \|Q(X_1) - Q(X_2)\| &\leq 2\left|\, \beta P_1 A_1 - \beta P_2 A_2 \,\right|, \\ &\leq 2\beta \left|\, P_1 (A_1 - A_2) + A_2 (P_1 - P_2) \,\right|, \\ &\leq 2\beta \left|\, P_1 |A_1 - A_2| + A_2 |P_1 - P_2| \,\right|, \\ &\leq \tfrac{2\beta\Lambda}{\mu} \left|\, |A_1 - A_2| + |P_1 - P_2| \,\right|, \\ &\leq L \|X_1 - X_2\|, \end{aligned}$$

with

$$L = \frac{2\beta\Lambda}{\mu}.$$

Hence

$$\|\Psi(X_1) - \Psi(X_2)\| \leq K \|X_1 - X_2\|,$$

where $K = \max(L, \|S\|) < \infty$. Thus, it follows that the function Ψ is a Lipschitz continuous function, resulting at the system (1) admits a unique solution [23]. □

3 Stability analysis and model's parameters sensitivity

As we dig into the dynamics of our proposed compartment model, we now turn our analysis to two important aspects: stability analysis and sensitivity analysis of considered parameters in the model. These considerations are crucial in understanding the reliability of the model dynamics, putting light on its behavior under various conditions for model optimization.
This section will investigate the considered system (1) comportment at the rejection-free equilibrium point (RFEP) and rejection equilibrium point (REP).

1. The rejection-free equilibrium is given by $X_0 = (P_0, A_0, E_0, D_0)$, where
$$P_0 = \frac{\Lambda}{\alpha + \mu}, \quad A_0 = 0,$$
with
$$E_0 = \alpha \frac{\Lambda}{\mu} \frac{\theta + \mu}{(\alpha + \mu)(\theta + \delta + \mu)}$$
and
$$D_0 = \alpha \Lambda \frac{\delta}{\mu (\alpha + \mu)(\theta + \delta + \mu)}.$$
This equilibrium aligns with a scenario where customers exhibit a consistent acceptance and utilization of e-commerce platforms. In essence, it signifies a state where there is a harmonious balance between the demand for e-commerce services and the willingness of customers to engage with such platforms. This equilibrium suggests that customers are not inclined to reject or resist the adoption of e-commerce, indicating a stable and favorable environment for online transactions and interactions.

2. The rejection equilibrium corresponds to the situation in which the refusal to use electronic commerce is able to spread among the potential customers of electronic commerce. This equilibrium is given by $X_* = (P_*, A_*, E_*, D_*)$ with
$$P_* = \frac{1}{\beta} (\gamma + \mu), \quad A_* = \frac{1}{\beta} (R_0 - 1)(\alpha + \mu),$$

$$E_* = \frac{1}{\beta\mu}\frac{\theta+\mu}{\theta+\mu+\delta}\left(\alpha\mu+\alpha\gamma R_0+\gamma\mu\left(R_0-1\right)\right),$$

and

$$D_* = \frac{1}{\beta\mu}\frac{\theta+\mu}{\theta+\mu+\delta}\left(\alpha\mu+\alpha\gamma R_0+\gamma\mu\left(R_0-1\right)\right).$$

Here, the R_0 is the basic reproduction number corresponding to the number of secondary infections caused by an infected person during the course of their infection in a susceptible population. It is a measure widely used in epidemiology because an epidemic can only occur if $R_0 > 1$, and if $R_0 < 1$, the disease cannot invade and no outbreak is expected.

In our research, the basic reproduction number corresponds to the number of people who reject e-commerce, which was the result of a person who refuses to use e-commerce in a sensitive population.

We relied on the approach presented by P.Van Driessche and Watmough [33], in order to calculate R_0.

Letting $X = (A, E, P)$, then system (1) can be written as $X' = \Gamma(X) - \Upsilon(X)$.

We have

$$\Gamma(X) = \begin{pmatrix}\beta PA\\0\\0\end{pmatrix} \quad \text{and} \quad \Upsilon(X) = \begin{bmatrix}\gamma A+\mu A\\-\alpha P-\gamma A-\theta D+\delta E+\mu E\\-\Lambda+\beta PA+\alpha P+\mu P\end{bmatrix}.$$

Next, we can compute the Jacobian matrix of $\Gamma(X)$.

We have

$$J(\Gamma(X)) = \begin{bmatrix}\beta P & 0 & \beta A\\0&0&0\\0&0&0\end{bmatrix} \quad \text{so} \quad J(\Gamma(X_0)) = \begin{bmatrix}\beta P_0 & 0 & \beta A_0\\0&0&0\\0&0&0\end{bmatrix}. \tag{3}$$

We can then evaluate this Jacobian matrix at the equilibrium point X_0.

We have

$$J(\Gamma(X_0)) = \begin{bmatrix}\beta\frac{\Lambda}{\alpha+\mu} & 0 & 0\\0&0&0\\0&0&0\end{bmatrix}.$$

Next, we consider the function $\Upsilon(X)$.

We have

$$J(\Upsilon(X_0)) = \begin{bmatrix} \gamma+\mu & 0 & 0 \\ -\gamma & \delta+\mu & -\alpha \\ \beta P_0 & 0 & \beta A_0+\alpha+\mu \end{bmatrix}, \tag{4}$$

so that

$$J(\Upsilon(X_0)) = \begin{bmatrix} \gamma+\mu & 0 & 0 \\ -\gamma & \delta+\mu & -\alpha \\ \beta\frac{\Lambda}{\alpha+\mu} & 0 & \alpha+\mu \end{bmatrix}.$$

We can now define the matrices V and F as follows:

$$V = \begin{bmatrix} \gamma+\mu & 0 \\ -\gamma & \delta+\mu \end{bmatrix}$$

and

$$F = \begin{bmatrix} \beta\frac{\Lambda}{\alpha+\mu} & 0 \\ 0 & 0 \end{bmatrix}.$$

Finally, we compute the basic reproduction number R_0.

At the end, we get

$$R_0 = \rho\left(F \times V^{-1}\right) = \frac{\Lambda\beta}{(\alpha+\mu)(\gamma+\mu)}, \tag{5}$$

where $\rho\left(F \times V^{-1}\right)$ is the spectrum of the matrix $F \times V^{-1}(X_0)$.

3.1 The local stability study

In this subsection, we will delve into the local stability analysis, first of the RFEP. Consequently, we present the following theorem that characterizes local stability of the $RFEP$.

Theorem 3. The $RFEP$ is locally asymptotically stable if $R_0 < 1$.

Proof. We have the Jacobian matrix of J_0 in X_0:

$$J_0 = \begin{bmatrix} -\alpha-\mu & -\beta\frac{\Lambda}{\alpha+\mu} & 0 & 0 \\ 0 & \beta\frac{\Lambda}{\alpha+\mu}-\gamma-\mu & 0 & 0 \\ \alpha & \gamma & -\delta-\mu & \theta \\ 0 & 0 & \delta & -\theta-\mu \end{bmatrix}.$$

The Eigen values of the characteristic equation of J_0 are

$V_1 = -\alpha - \mu < 0,$
$V_2 = -\mu < 0,$
$V_3 = -\theta - \mu - \delta < 0,$
$V_4 = -\frac{1}{\alpha+\mu}\left(1 - R_0\right)\left(\alpha + \mu\right)\left(\gamma + \mu\right) < 0.$

Therefore, X_0 is locally asymptotically stable if $R_0 < 1$. □

Thereafter, let us now present the following theorem that concerns REP characterization.

Theorem 4. The REP is locally asymptotically stable if $R_0 > 1$.

Proof. We have the Jacobian matrix J_* evaluated at X_*:

$$J_* = \begin{bmatrix} -R_0(\alpha+\mu) & -\gamma-\mu & 0 & 0 \\ (R_0-1)(\alpha+\mu) & 0 & 0 & 0 \\ \alpha & \gamma & -\mu-\delta & \theta \\ 0 & 0 & \delta & -\theta-\mu \end{bmatrix}.$$

The characteristic polynomial of J_* is given by

$$P(X) = X^4 + a_3X^3 + a_2X^2 + a_1X + a_0,$$

where

$$\begin{aligned} a_3 &= \theta + 2\mu + \delta + R_0(\alpha+\mu), \\ a_2 &= \theta\mu + (R_0-1)(\alpha\gamma + \alpha\mu + \gamma\mu) + \mu\delta + R_0(3\mu^2 + \theta\alpha + \theta\mu + 2\alpha\mu \\ &\quad + \alpha\delta + \mu\delta), \\ a_1 &= (\alpha+\mu)\left[(R_0-1)(\gamma+\mu)(\theta+2\mu+\delta) + \mu\delta R_0 + \mu^2 R_0 + \theta\mu R_0\right], \\ a_0 &= \mu(R_0-1)(\alpha+\mu)(\gamma+\mu)(\theta+\mu+\delta). \end{aligned}$$

According to the Routh–Hurwitz criterion, for the system (1) to be locally asymptotically stable, all the coefficients a_i must be positive, and the following condition must hold:

$$a_1(a_2a_3 - a_1^2) > a_0a_3^2.$$

Let us verify these conditions step-by-step:

1. Positivity of Coefficients:

- $a_3 = \theta + 2\mu + \delta + R_0(\alpha + \mu) > 0$ since all parameters are positive and $R_0 > 1$.

- $a_2 = \theta\mu + (R_0 - 1)(\alpha\gamma + \alpha\mu + \gamma\mu) + \mu\delta + R_0(3\mu^2 + \theta\alpha + \theta\mu + 2\alpha\mu + \alpha\delta + \mu\delta) > 0$ given $R_0 > 1$ and all parameters are positive.

- $a_1 = (\alpha + \mu)[(R_0 - 1)(\gamma + \mu)(\theta + 2\mu + \delta) + \mu\delta R_0 + \mu^2 R_0 + \theta\mu R_0] > 0$ given $R_0 > 1$ and all parameters are positive.

- $a_0 = \mu(R_0 - 1)(\alpha + \mu)(\gamma + \mu)(\theta + \mu + \delta) > 0$ since $R_0 > 1$ and all parameters are positive.

2. Routh–Hurwitz Condition:

$$a_1(a_2 a_3 - a_1^2) > a_0 a_3^2.$$

By substituting the expressions for a_1, a_2, a_3, and a_0 and verifying algebraically, we confirm this inequality holds when $R_0 > 1$.

Thus, from the Routh–Hurwitz criterion, the system is locally asymptotically stable if $R_0 > 1$. □

3.2 The global stability study

Let us now investigate the global stability of the concerned states, the $RFEP$ and REP. In the following, we present the theorem showcasing the global stability characterization of $RFEP$.

Theorem 5. The RFEP is globally asymptotically stable if $R_0 \leq 1$.

Proof. Consider a Lyapunov function $V : \Omega \to \mathbb{R}$ given by

$$V(P, A) = \frac{1}{2}\left((P - P_0) + A\right)^2 + \frac{(\alpha + \gamma + 2\mu)}{\beta} A. \tag{6}$$

After deriving V, we get

$$V^{'}(P, A) = ((P - P_0) + A)\left(P^{'} + A^{'}\right) + \frac{(\alpha + \gamma + 2\mu)}{\beta} A^{'}.$$

Then

$$
\begin{aligned}
V^{'}(P,A) =&-(\alpha+\mu)\left(P-\frac{\Lambda}{\alpha+\mu}\right)^{2}-(\gamma+\mu)\,A^{2}\\
&-(\alpha+\gamma+2\mu)\times\frac{(\mu+\gamma)}{\beta}(1-\Lambda\frac{\beta}{(\alpha+\mu)\,(\gamma+\mu)})\times A\\
=&(\alpha+\mu)\left(P-\frac{\Lambda}{\alpha+\mu}\right)^{2}-(\gamma+\mu)\,A^{2}\\
&-(\alpha+\gamma+2\mu)\times\frac{(\mu+\gamma)}{\beta}(1-R_0)\times A\leq 0.
\end{aligned}
$$

Also, we obtain $V(P,A)=0 \Longleftrightarrow P=P_0$ and $A=0$.

Therefore, the previous theorem has been proved. Moreover, by La Salle's invariance principle [17], the RFEP is globally asymptotically stable on Ω. $\square$

In the following, we get to the global stability of the REP, which is given by the upcoming theorem

Theorem 6. The REP is globally asymptotically stable if $R_0 > 1$.

Proof. Consider Lyapunov function $V:\Omega\to\mathbb{R}$ given by

$$
V(P,A)=(P-P_*)-P_*\ln\left(\frac{P}{P_*}\right)+(A-A_*)-A_*\ln\left(\frac{A}{A_*}\right). \tag{7}
$$

After deriving V, we get

$$
\begin{aligned}
V^{'}(P,A) =&P^{'}+A^{'}-P_*\times\frac{P^{'}}{P}-A_*\times\frac{A^{'}}{A}.\\
V^{'}(P,A) =&-\frac{1}{P}\frac{\Lambda}{\beta\,(\gamma+\mu)}\,(\gamma+\mu-P\beta)^{2}\\
=&-\frac{1}{P}\frac{\Lambda}{P_*}\,(P-P_*)^{2}\leq 0.
\end{aligned}
$$

Also, we obtain

$V'(P,A)=0 \Longleftrightarrow P=P_*$.

Therefore, by La Salle's invariance principle [17], the REP is globally asymptotically stable on Ω.

$\square$

Remark 1. After thorough examination of local and global stability, particularly within our proposed model concerning customer rejection behavior, significant insights have emerged. Local stability analysis delves into the dynamics surrounding equilibrium points, revealing how the system behaves under small perturbations. Meanwhile, global stability analysis extends beyond local considerations, providing insights into the system's overall behavior across its entire phase space. This combined analysis offers a comprehensive understanding vision of the interaction between customer rejection behavior and e-commerce dynamics, aiding in strategic decision-making.

3.3 Model's parameters sensitivity analysis

The concept of sensitivity investigation and analysis is used to measure the prevailing parameters in the model that significantly influence the prevalence of people who refuse to use e-commerce. We know that the prevalence of people who refuse to use e-commerce in the rest of society is linked to the basic reproduction number R_0.

In this section, we study the effect of each parameter entered into the system (1) on the value of R_0. In order to identify the parameters that contribute most to the spread of people who refuse to use e-commerce within society, a sensitivity analysis is carried out in the sense of [28, 1]. We compute the partial derivatives of the value of R_0 with respect to the values of the system parameters (1) using the sensitivity index.

Definition 1. The sensitivity index $Z_q^{R_0}$ of the basic reproduction number R_0, varying with parameter q, is defined as

$$Z_q^{R_0} = \frac{\partial R_0}{\partial q} \cdot \frac{q}{R_0},$$

where $\frac{\partial R_0}{\partial q}$ denotes the partial derivative of R_0 with respect to q.

We have

$$\begin{cases} Z_\Lambda^{R_0} = \frac{\partial R_0}{\partial \Lambda} \cdot \frac{\Lambda}{R_0} = 1, \\ Z_\beta^{R_0} = \frac{\partial R_0}{\partial \beta} \cdot \frac{\beta}{R_0} = 1, \\ Z_\alpha^{R_0} = \frac{\partial R_0}{\partial \alpha} \cdot \frac{\alpha}{R_0} = \frac{-\alpha}{(\alpha+\mu)}, \\ Z_\gamma^{R_0} = \frac{\partial R_0}{\partial \gamma} \cdot \frac{\gamma}{R_0} = \frac{-\gamma}{(\gamma+\mu)}, \\ Z_\mu^{R_0} = \frac{\partial R_0}{\partial \mu} \cdot \frac{\mu}{R_0} = \frac{-\mu(\alpha+\gamma+2\mu)}{(\alpha+\mu)(\gamma+\mu)}. \end{cases} \tag{8}$$

The sign of the sensitivity index of the basic reproduction figure R_0 for the model parameters indicates that an increase (or decrease) in the value of each of the parameters, in this case, will lead to an increase (or decrease) in the value of the basic reproduction figure for people who refuse to use e-commerce. For example, $Z_{\beta}^{R_0} = 1$ means that increasing (or decreasing) the effective contact rate β by 10% will increase (or decrease) the basic reproduction number R_0 by 10%. On the other hand, the negative sign of the sensitivity index, for the basic reproduction number of the model parameters, means that an increase (or decrease) in the value of each of the parameters, in this case, leads to a corresponding decrease (or increase) in the value of the basic reproduction number for people who refuse to use e-commerce. For example, $Z_{\alpha}^{R_0}$ = -0.63 means that increasing (or decreasing) the parameter α by 10% decreases (or increases) R_0 by 6.3%. Figure 2 below shows the sensitivity indices of all the model parameters to R_0. The parameters are ordered from most to least sensitive. Through sensitivity analysis, it is therefore possible to obtain a comprehensive view of appropriate strategies to increase the number of people using e-commerce.

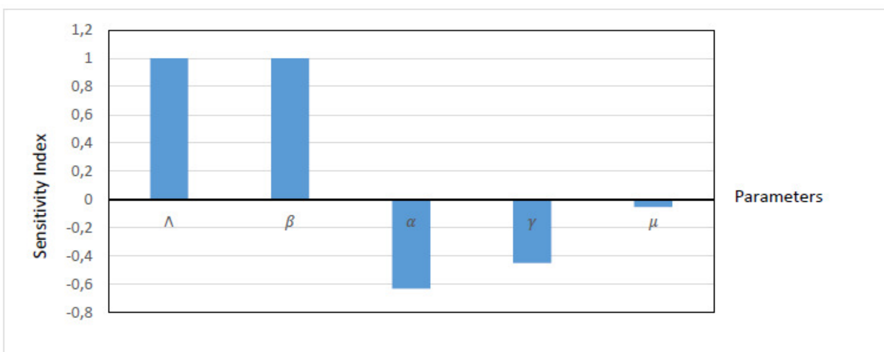


Figure 2: Sensitivity analysis of R_0.

3.4 Model's numerical simulation

This section presents some numerical solutions to the system (1) for different values of the parameters. We use different initial values for each state variable and use the following parameters:

Λ= 880750, $\mu = 1.3\times10^{-2}$, $\alpha = 1.7\times10^{-2}$, $\gamma = 8\times10^{-3}$, $\delta = 7\times10^{-4}$, $\theta = 3\times$

10^{-4}, and $\beta = 6 \times 10^{-11}$. We have the equilibrium point without e-commerce refusal $X_0 = (\ 2.9358{\times}10^7,\ 0,\ 3.6472{\times}10^7,\ 1.9196{\times}10^6)$ and $R_0 = 0.0839 < 1$. In this case, we observe that the number of people potentially refusing to use e-commerce approaches X_0 over time; that is, all state variables converge to the equilibrium point for four different initial values of each of the state variables studied. We also observe that the percentage of those who refuse to use e-commerce is close to zero. See Figure 3.

Also, for different initial values of each state variable, we have the following parameters: Λ= 880750, $\mu = 1.3 \times 10^{-2}$, $\alpha = 1.7 \times 10^{-2}$, $\gamma = 8 \times 10^{-3}$, $\delta = 7 \times 10^{-4}$, $\theta = 3 \times 10^{-4}$ and $\beta = 2 \times 10^{-9}$. Our equilibrium point, which corresponds to the situation where the refusal to use e-commerce is able to spread among potential e-commerce customers, is $X_* = (\ 1.0500{\times}10^7,\ 2.6940{\times}10^7,\ 2.8794{\times}10^7,\ 1.5155{\times}10^6)$ and $R_0 = 2.7960 > 1$, and the state variables converge to the equilibrium point X_*. In this case, we note that the number of people P and E converge over time to the same value for each of them. See Figure 4.

4 The optimal control problem

4.1 Problem synopsis

To increase the number of people using e-commerce for their daily needs and reduce the number of those who refuse to use it, we recommend two techniques.

Control v_1 involves encouraging customers to use e-commerce. The importance of e-commerce is introduced to customers, highlighting its advantages such as time-saving, reduced traffic accidents, and the convenience of shopping from anywhere at any time. Advertising is carried out through various means, such as audio and visual media, social media, and newspapers.

Control v_2 is to provide insurance and guarantees to individuals who refuse to use e-commerce for fear of cybercrime or online fraud. This will protect them if their money is stolen, they fall victim to fraud, or their personal information is stolen.

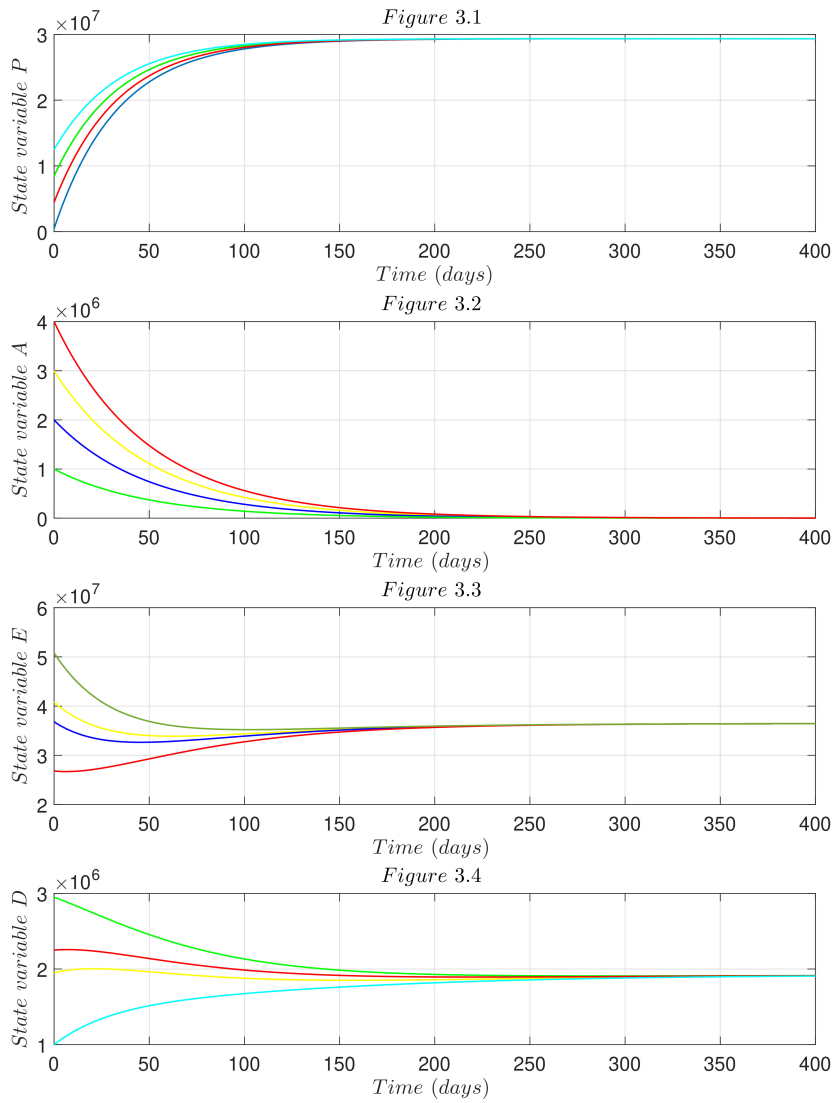


Figure 3: Rejection-free equilibrium model for e-commerce.

By adding the two controls $v_1(t)$ and $v_2(t)$ to the first system (1), we obtain the following control model:

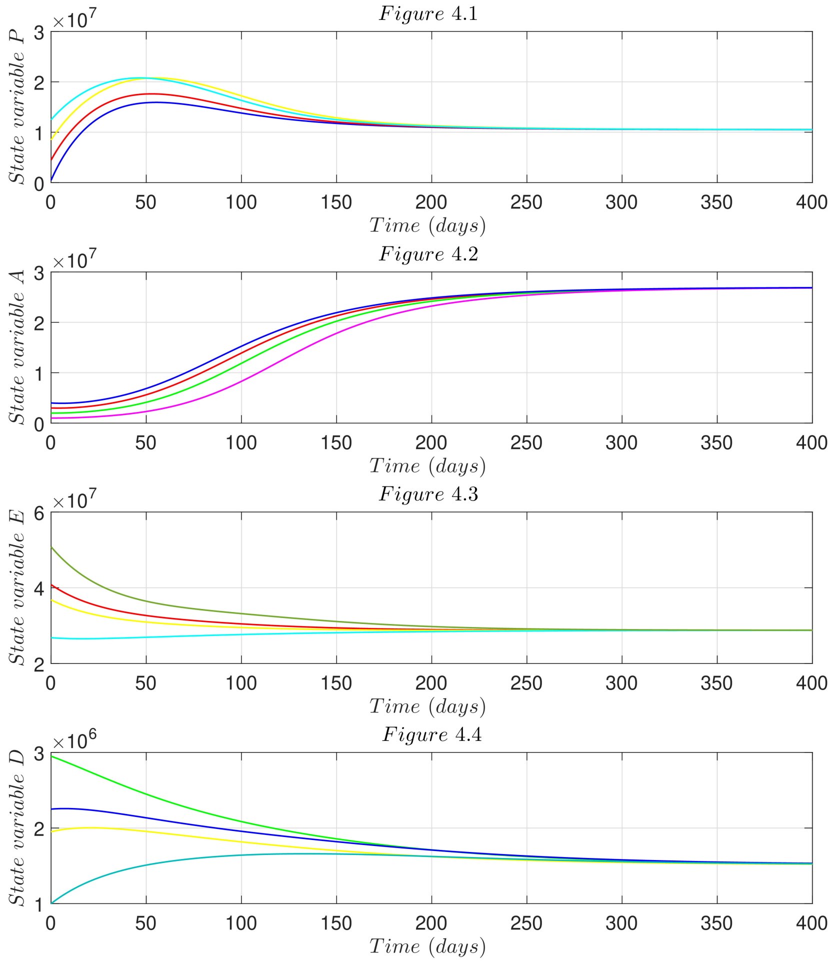


Figure 4: Equilibrium model with rejection of e-commerce.

$$\begin{cases} \frac{dP(t)}{dt} = \Lambda - \beta P(t)A(t) - \alpha P(t) - \mu P(t) - v_1(t)P(t), \\ \frac{dA(t)}{dt} = \beta P(t)A(t) - \gamma A(t) - \mu A(t) - v_2(t)A(t), \\ \frac{dE(t)}{dt} = \alpha P(t) + \gamma A(t) + \theta D(t) \\ \qquad -\delta E(t) - \mu E(t) - \mu E(t) + v_1(t)P(t) + v_2(t)A(t), \\ \frac{dD(t)}{dt} = \delta E(t) - \theta D(t) - \mu D(t), \end{cases} \tag{9}$$

with the initial conditions $P(0) \geq 0$, $A(0) \geq 0$, $E(0) \geq 0$, and $D(0) \geq 0$. Here, our goal is to reduce the total number of customers who refuse to use e-commerce using the control variables $v_1(t)$ and $v_2(t)$ in the model (1). We define the objective function as follows:

$$J(v_1, v_2) = A(t_f) - E(t_f) + \int_{t_0}^{t_f} \left[A(t) - E(t) + \frac{C_1}{2} (v_1(t))^2 + \frac{C_2}{2} (v_2(t))^2 \right] dt, \tag{10}$$

where $C_1 > 0$ and $C_2 > 0$ are the cost coefficients. They are selected to weigh the relative importance of $v_1(t)$ and $v_2(t)$ at time t, and t_f is the final time.

In other words, we seek the optimal controls v_1^* and v_2^* such that

$$J(v_1^*, v_2^*) = \min_{(v_1, v_2) \in E_d} J(v_1, v_2), \tag{11}$$

where E_d is the set of admissible controls defined by

$$E_d = \{(v_1(t), v_2(t)) : 0 \leq v_1(t) \leq 1 \text{ and } 0 \leq v_2(t) \leq 1 \ for \ t \in [t_0, t_f]\}. \tag{12}$$

4.2 The existence result of the optimal control

In this part, we present a theorem, which proves the existence of an optimal control (v_1^*,v_2^*) minimizing the cost function J.

Theorem 7. There exists an optimal control $(v_1^*, v_2^*) \in E_d$ such that

$$J(v_1^*, v_2^*) = \min_{(v_1, v_2) \in E_d} J(v_1, v_2).$$

Proof. To use the existence result in [12], we must check the following properties:
(Q_1): The set of controls and the corresponding state variables is nonempty.
(Q_2): The control set E_d is convex and closed.
(Q_3): The right-hand side of the state system is bounded by a linear function in the state and control variables.
(Q_4): The integrand $L(A, E, v_1, v_2)$ in the objective functional is convex on E_d, and there exist constants $\chi_1 > 0$, $\chi_2 > 0$, and $\sigma > 1$ such that

$$\begin{aligned} L(A, E, v_1, v_2) &= A(t) - E(t) + \frac{C_1}{2} (v_1(t))^2 + \frac{C_2}{2} (v_2(t))^2 \\ &\geq -\chi_1 + \chi_2 \left(|v_1|^2 + |v_2|^2 \right)^{\frac{\sigma}{2}}. \end{aligned}$$

The first condition (Q_1) is verified using the result in [27]. The set E_d is convex and closed by definition, thus the condition (Q_2) is verified. Our state system is linear in v_1 and v_2. Moreover the solutions of the system are bounded as proved in model (1) so the condition (Q_3). Also, we have the last condition needed,

$$L(A, E, v_1, v_2) \geq -\chi_1 + \chi_2 \left(|v_1|^2 + |v_2|^2\right)^{\frac{\sigma}{2}},$$

where

$$\chi_1 = 2 \sup_{t\in[t_0,t_f]} (A(t), E(t)), \ \chi_2 = \inf(\frac{C_1}{2}, \frac{C_2}{2}) \text{ and } \sigma = 2 \text{ since } C_1 > 0.$$

We conclude that there exists an optimal control $(v_1^*, v_2^*) \in E_d$ such that

$$J(v_1^*, v_2^*) = \min_{(v_1,v_2)\in E_d} J(v_1, v_2).$$

□

4.3 Characterization of the optimal controls

Pontryagin's maximum principle [24, 34, 4] transforms (9)–(11) and (12) into a problem of minimizing a Hamiltonian, H, point-wise with respect to v_1 and v_2:

$$H(t) = A(t) - E(t) + \frac{C_1}{2}\left(v_1(t)\right)^2 + \frac{C_2}{2}\left(v_2(t)\right)^2 + \sum_{i=1}^{4} \lambda_i g_i, \tag{13}$$

where g_i is the right side of the differential equations system (9).

In the following theorem, we give the necessary conditions for an optimal control to exist.

Theorem 8. Given the optimal controls (v_1^*, v_2^*) and solutions P^*, A^*, E^*, and D^* of the corresponding state system (9), there exists adjoint variables λ_1, λ_2, λ_3, and λ_4 that satisfy the following equations:

$$\begin{aligned}
\lambda_1' = -\frac{dH}{dP} = & -\lambda_1[-\beta A(t) - \alpha - \mu - v_1(t)] - \lambda_2[\beta A(t)] - \lambda_3[\alpha + v_1(t)],\\
\lambda_2' = -\frac{dH}{dA} = & -1 - \lambda_1\left(-\beta P(t)\right) - \lambda_2[\beta P(t) - \gamma - \mu - v_2(t)]\lambda_3[\gamma + v_2(t)],
\end{aligned}$$

$$\lambda_3' = -\frac{dH}{dE} = \ 1 - \lambda_3[-\delta - \mu] - \lambda_4\left(\delta\right),$$
$$\lambda_4' = -\frac{dH}{dD} = \ -\lambda_3\left(\theta\right) - \lambda_4\left(-\theta - \mu\right),$$

such that the transversality conditions at time t_f are: $\lambda_1(t_f) = 0$, $\lambda_2(t_f) = 1$, $\lambda_3(t_f) = -1$ and $\lambda_4(t_f) = 0$. Additionally we have, for $t \in [t_0, t_f]$, the optimal controls $v_1^*(t)$ and $v_2^*(t)$ are given by:

$$\begin{cases} v_1^*(t) = \min\left(1, \max\left(0, \frac{P(t)}{C_1}\left(\lambda_1 - \lambda_3\right)\right)\right), \\ \\ v_2^*(t) = \min\left(1, \max\left(0, \frac{A(t)}{C_2}\left(\lambda_2 - \lambda_3\right)\right)\right). \end{cases} \tag{14}$$

Proof. The Hamiltonian H is defined in the following way:

$$H(t) = A(t) - E(t) + \frac{C_1}{2}\left(v_1(t)\right)^2 + \frac{C_2}{2}\left(v_2(t)\right)^2 + \sum_{i=1}^{4}\lambda_i g_i,$$

where

$$\begin{cases} g_1 = \Lambda - \beta P(t)A(t) - \alpha P(t) - \mu P(t) - v_1(t)P(t), \\ g_2 = \beta P(t)A(t) - \gamma A(t) - \mu A(t) - v_2(t)A(t), \\ g_3 = \alpha P(t) + \gamma A(t) + \theta D(t) - \delta E(t) - \mu E(t) \\ \qquad -\mu E(t) + v_1(t)P(t) + v_2(t)A(t), \\ g_4 = \delta E(t) - \theta D(t) - \mu D(t). \end{cases}$$

For $t \in [t_0, t_f]$, the adjoint equations and transversality conditions can be obtained by using Pontryagin's maximum principle [24, 20] such that

$$\lambda_1' = -\frac{dH}{dP} = \ -\lambda_1[-\beta A(t) - \alpha - \mu - v_1(t)] - \lambda_2[\beta A(t)] - \lambda_3[\alpha + v_1(t)],$$
$$\lambda_2' = -\frac{dH}{dA} = \ -1 - \lambda_1\left(-\beta P(t)\right) - \lambda_2[\beta P(t) - \gamma - \mu - v_2\left(t\right)] - \lambda_3[\gamma + v_2(t)],$$
$$\lambda_3' = -\frac{dH}{dE} = \ 1 - \lambda_3[-\delta - \mu] - \lambda_4\left(\delta\right),$$
$$\lambda_4' = -\frac{dH}{dD} = \ -\lambda_3\left(\theta\right) - \lambda_4\left(-\theta - \mu\right).$$

For $t \in [t_0, t_f]$, the optimal controls v_1^* and v_2^* can be solved from the optimal conditions,

$$\frac{dH}{dv_1} = 0,$$
$$\frac{dH}{dv_2} = 0.$$

That is,

$$\begin{aligned}\frac{dH}{dv_1} &= C_1 v_1(t) - \lambda_1 P(t) + \lambda_3 P(t) = 0,\\ \frac{dH}{dv_2} &= C_2 v_2(t) + \lambda_2 \left(-A\left(t\right)\right) + \lambda_3 \left(A(t)\right) = 0.\end{aligned}$$

We have

$$\begin{aligned}v_1 &= \frac{P(t)}{C_1}\left(\lambda_1 - \lambda_3\right),\\ v_2 &= \frac{A(t)}{C_2}\left(\lambda_2 - \lambda_3\right).\end{aligned}$$

Thus, by the bounds in E_d, the set of admissible controls, we can smoothly obtain v_1^* and v_2^* in the form of (14). □

5 Numerical simulations and discussion

In our research, a set of preliminary data was collected for the country of France to serve as initial conditions for the system (1). Thus, in 2021, the population of France available on the Internet reached 86.1% [38] of the total population of this country. In the same year, the percentage of the French population using e-commerce was 69.1% [40] of the total French population. Below, we present a set of data for the application of numerical simulation in MATLAB:

Conditions for system (1) are as follows: $P(0) = 8467320$, $A(0) = 3000000$, $E(0) = 46815250$, $D(0) = 50000$, $\Lambda = 880750$, $\mu = 9 \times 10^{-3}$, $\alpha = 1.7 \times 10^{-2}$, $\gamma = 8 \times 10^{-3}$, $\delta = 7 \times 10^{-4}$, $\theta = 3 \times 10^{-4}$ and $\beta = 10^{-9}$.

5.1 Strategy v_1: Promotion of the importance of e-commerce through advertising

■ Figure 5.1: Before the implementation of the control element v_1, the number of people who refused to use e-commerce was 3 million. After applying the first strategy, from day 37, the number of people who refused to use e-commerce started to decrease, so that the percentage of this decrease reached about 6.92% on day 60. The analysis shows that after 100 days of implementing the first strategy, the number of merchants who refused to use e-commerce decreased by 35.20%.

■ Figure 5.2: Before applying the second strategy, the number of individuals using e-commerce was approximately 46815250 individuals. After applying the control element v_1, this number started to increase from day 17, so that the percentage of this increase reached about 6.33% on day 60. It is also clear that the number of people using electronic commerce increased by 18% after 100 days of application First strategy.

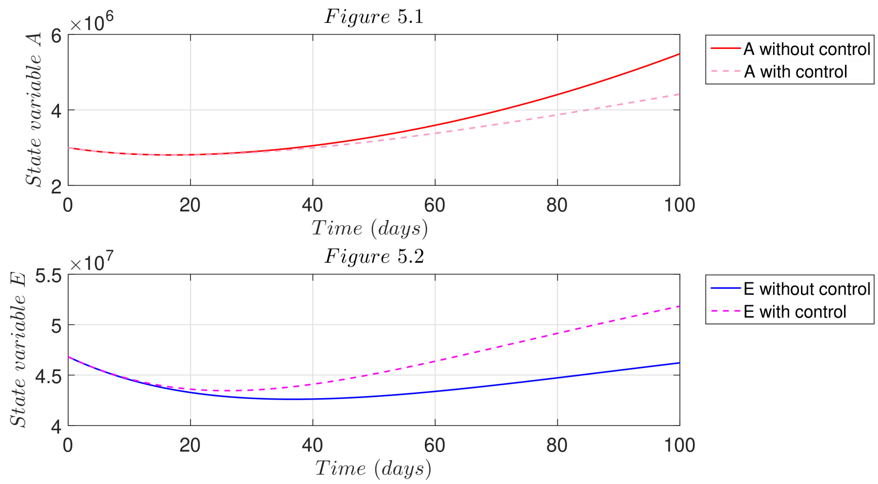


Figure 5: Dynamics with control v_1.

5.2 Strategy v_2 : Cybercrime insurance

■ Figure 6.1: Before implementing the v_2 control, the number of people who refused to use e-commerce was 3 million. After applying the second strategy, this number started to decrease from the seventh day. The percentage of this decrease reached 90.19% on the 60th day after the application of this strategy. From this analysis, we can conclude that the number of people who refused to use e-commerce decreased by 162.41% after 100 days of applying the second strategy.

■ Figure 6.2: Before applying the second strategy, the number of individuals using e-commerce was approximately 46815250 individuals. After applying the v_2 control, this number started to increase from day 7. The percentage of this increase reached 3.11% after 60 days of applying this strategy. According to this analysis, the number of people using e-commerce increased by 3.68% after 100 days of applying the second strategy.

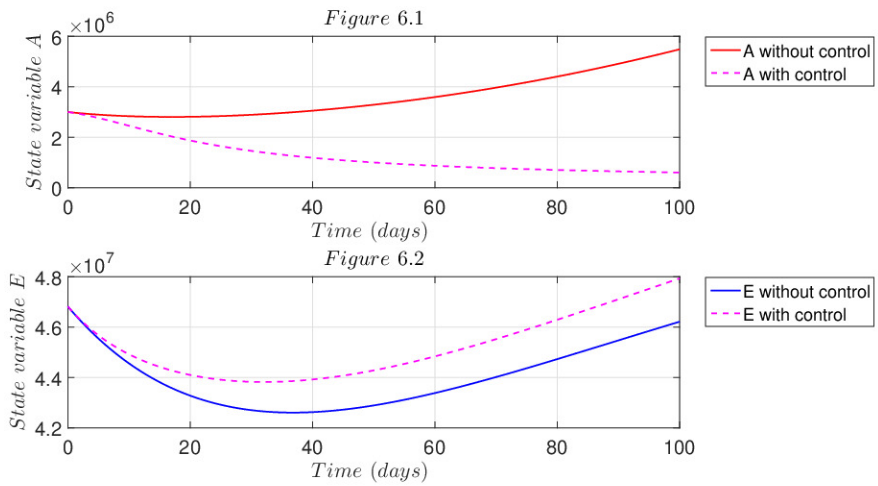


Figure 6: Dynamics with control v_2.

5.3 Strategy v_1 and v_2: Advertising and cybercrime insurance

■ Figure 7.1: Before applying controls v_1 and v_2, the number of people who refused to use e-commerce was 3 million. After applying the third strategy, this number started to decrease from the fourth day. The percentage of this decrease reached 91.56% on the 60th day after applying this strategy. From this analysis, we can conclude that the number of merchants who refused to use e-commerce decreased by 165.33% after 100 days of implementing the third strategy.

■ Figure 7.2: Before implementing the third strategy, the number of individuals using e-commerce was approximately 46815250 individuals. After applying the control elements v_1 and v_2, this number started to increase from the 4th day. The percentage of this increase reached 9.31% after 60 days of applying this strategy. From this analysis, we can conclude that the number of people using e-commerce increased by 14.61% after 100 days of implementing the second strategy.

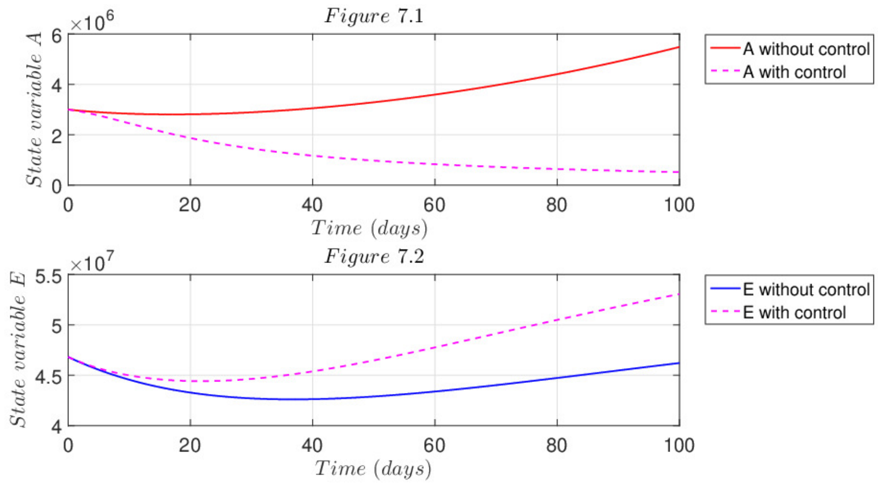


Figure 7: Dynamics with control v_1 and v_2.

In conclusion, the analysis of the three proposed strategies provides valuable insights into their effectiveness in addressing the issue of customer rejection behavior in e-commerce. Strategy v_1, focusing on promoting the importance of e-commerce through advertising, demonstrates promising results. It led to a significant decrease in the number of individuals rejecting e-commerce, with a reduction of 35.20% observed after 100 days of implementation. Moreover, there was a notable increase in the number of people utilizing e-commerce, indicating a positive impact on adoption rates. Strategy v_2, centered around cybercrime insurance, proved to be highly effective in mitigating customer rejection of e-commerce. The strategy resulted in a substantial decrease of 162.41% in the number of individuals rejecting e-commerce after 100 days. Additionally, there was a moderate increase in e-commerce adoption rates, further reinforcing the efficacy of this approach. Furthermore, combining both strategies, as in Strategy v_1 and v_2, yielded even more promising outcomes. The simultaneous implementation of advertising and cybercrime insurance led to a remarkable decrease of 165.33% in e-commerce rejection rates after 100 days. Furthermore, there was a notable increase in e-commerce adoption, highlighting the synergistic effects of employing multiple strategies.

Overall, these findings underscore the importance of tailored interventions in addressing customer rejection behavior in e-commerce. By strategically leveraging advertising and cybercrime insurance, policymakers and businesses can foster a more conducive environment for e-commerce adoption, ultimately driving growth and innovation in the digital marketplace.

6 Conclusion

In this research, we formulated a new mathematical model that describes the dynamics of people who refuse to use e-commerce and people who use e-commerce. We also computed the basic reproduction number R_0 and studied the parameters involved the sensitivity analysis to determine which parameters have the greatest influence on the R_0. We also analyzed the local and global stability of the behavior of those who refuse to use e-commerce. We also performed numerical simulations to verify the theoretical analysis using MATLAB. Finally, we proposed three different control strategies. This was done using Pontryagin's maximum principle to describe optimal controls. After analysis, it was found that the third strategy was the most effective in increasing the number of people using e-commerce and reducing the number of people rejecting it.

Declaration of interests

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.

References

[1] Aarabate, O., Belhdid, S. and Balatif, O., *Stability analysis and optimal strategies for controlling a boycotting behavior of a commercial product,*

Iranian Journal of Numerical Analysis and Optimization 14 (3) (2024), 708–735.

[2] Balatif, O., Khajji, B. and Rachik, M., *Mathematical modeling, analysis, and optimal control of abstinence behavior of registration on the electoral lists,* Discrete Dyn. Nature Soc. 2020 (2020), 9738934.

[3] Balatif, O., Labzai, A. and Rachik, M., *A discrete mathematical modeling and optimal control of the electoral behavior with regard to a political party,* Discrete Dyn. Nature Soc. 2018, (2018) 1–14.

[4] Benfatah, Y., Khaloufi, I., Boutayeb, H., Rachik, M. and Laarabi, H., *Optimal control for a discrete time epidemic model with zones evolution,* Commun. Math. Biol. Neurosci., 2022 (2022) 51.

[5] Belhdid, S., Balatif, O. and Khajji, B., *Optimal control and cost-effectiveness analysis of scam rumor propagation over social networks,* Results Control Optim. (2024) 100441.

[6] Behnam, M., Hollebeek, L.D., Clark, M.K. and Farabi, R., *Exploring customer engagement in the product vs. service context,* J. Retail. Consum. Serv. 60 (2021) 102456.

[7] Bonina, C., Koskinen, K., Eaton, B. and Gawer, A., *Digital platforms for development: Foundations and research agenda,* Inf. Syst. J. 31(6) (2021) 869–902.

[8] Corbitt, B.J., Thanasankit, T. and Yi, H., *Trust and e-commerce: A study of consumer perceptions,* Electron. Commer. Res. Appl.. 2(3) (2003) 203–215.

[9] Dionysiou, G., Fouskas, K. and Karamitros, D., *The impact of Covid-19 in E-commerce. effects on consumer purchase behavior,* In Strategic Innovative Marketing and Tourism in the COVID19 Era: 9th ICSIMAT Conference 2020 (pp. 199–210). Springer International Publishing, 2021.

[10] El Bhih, A., Benfatah, Y., Kouidere, A. and Rachik, M., *A discrete mathematical modeling of transmission of COVID-19 pandemic using*

optimal control, Commun. Math. Biol. Neurosci. 2020 (2020) Article-ID.

[11] Fedorko, M.R., *The current status and importance of e-commerce in a global perspective,* Online J. 4509 (2014) 1339–4509.

[12] Fleming, W.H. and Rishel, R.W., *Deterministic and stochastic optimal control* (Vol. 1). Springer Science and Business Media, 2012.

[13] Gour, A., Aggarwal, S. and Erdem, M., *Reading between the lines: analyzing online reviews by using a multi-method Web-analytics approach,* Int. J. Contemp. Hosp. Manag. 33(2) (2021) 490–512.

[14] Hethcote, H.W. and Thieme, H.R., *Stability of the endemic equilibrium in epidemic models with subpopulations,* Math. Biosci. 75(2), (1985) 205–227.

[15] Huang, L., Huang, Y., Huang, R., Xie, G. and Cai, W., *Factors influencing returning migrants' entrepreneurship intentions for rural e-commerce: An empirical investigation in China,* Sustainability, 14(6) (2022) 3682

[16] Jones, L.S. and Samples, T.R., *On the systemic importance of digital platforms.* U. Pa. J. Bus. L., 25 (2023) 141.

[17] LaSalle, J.P. *The stability of dynamical systems,* Regional Conference Series in Applied Mathematics, Vol. 25, SIAM, Philadelphia, PA, USA, 1976.

[18] Khajji, B., Boujallal, L., Balatif, O. and Rachik, M., *Mathematical modelling and optimal control strategies of a multistrain COVID-19 spread,* J. Appl. Math. 2022 (1) (2022): 9071890.

[19] Khajji, B., Boujallal, L., Elhia, M., Balatif, O. and Rachik, M., *A fractional-order model for drinking alcohol behaviour leading to road accidents and violence,* Math. Model. Comput. 9, (2022), 501–518.

[20] Khaloufi, I., Benfatah, Y., Laarabi, H. and Rachik, M., *A scenario to fight monkeypox using a mathematical model,* Commun. Math. Biol. Neurosci., 2022 (2022) 99.

[21] Kermack, W.O. and McKendrick, A.G., *A contribution to the mathematical theory of epidemics. Proceedings of the royal society of London,* Proc. R. soc. Lond. Ser. A-Contain. Pap. Math. Phys. Character. 115(772) (1927), 700–721.

[22] Kouidere, A., Labzai, A., Khajji, B., Ferjouchia, H., Balatif, O., Boutayeb, A. and Rachik, M., *Optimal control strategy with multi-delay in state and control variables of a discrete mathematical modeling for the dynamics of diabetic population,* Commun. Math. Biol. Neurosci. 2020 (2020) 14.

[23] Kouidere, A., Balatif, O. and Rachik, M., *Cost-effectiveness of a mathematical modeling with optimal control approach of spread of COVID-19 pandemic: A case study in Peru.* Chaos, Solitons Fractals. X, 10 (2023) 100090.

[24] Kouidere, A., Youssoufi, L.E., Ferjouchia, H., Balatif, O. and Rachik, M., *Optimal control of mathematical modeling of the spread of the COVID-19 pandemic with highlighting the negative impact of quarantine on diabetics people with cost-effectiveness,* Chaos, Solitons Fractals. 145 (2021) 110777.

[25] Labzai, A., Kouidere, A., Khajji, B., Balatif, O. and Rachik, M., *Mathematical modeling and optimal control strategy for a discrete time drug consumption model,* Discrete Dyn. Nature Soc. 2020 (2020) 1–10.

[26] Lightner, N.J., *What users want in e-commerce design: effects of age, education and income,* Ergonomics, 46(1-3), (2003) 153–168.

[27] Lukes, D.L., *Differential equations: classical to controlled,* Academic Press, New York, 1982.

[28] Makinde, O.D. and Okosun, K.O., *Impact of chemo-therapy on optimal control of malaria disease with infected immigrants,* Biosyst. 104(1), (2011) 32–41.

[29] Moorman, C., van Heerde, H.J., Moreau, C.P. and Palmatier, R.W., *Challenging the boundaries of marketing,* J. Mark. 83(5) (2019) 1–4.

[30] Nelson, J.L. and Lei, R.F., *The effect of digital platforms on news audience behavior,* Digit. Journal. 6(5), (2018) 619–633.

[31] Olson, J.S. and Olson, G.M., *I2i trust in e-commerce.* Commun. ACM, 43(12) (2000), 41–44.

[32] Park, Y.H. and Fader, P.S., *Modeling browsing behavior at multiple websites,* Mark. Sci. 23(3), (2004) 280–303.

[33] Driessche, P.V.and Watmough, J. *Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission,* Math.l Biosci. 180 (2002) 29–48.

[34] Pontryagin, L.S., *The mathematical theory of optimal processes and differential games,* Trudy Mat. Inst. Steklov, 169, (1985) 119–158.

[35] Rajala, T., Huhtamäki, J., Johanson, J.E. and Vakkuri, J., *Social impacts of digital platforms,* Public Innovation and Digital Transformation, (2022) 32–55.

[36] Salminen, J., Gach, N. and Kaartemo, V., *Platform as a social contract: An analytical framework for studying social dynamics in online platforms,* Collaborative value co-creation in the platform economy, (2018) 41–64.

[37] Shemi, A.P. and Procter, C., *E-commerce and entrepreneurship in SMEs: case of myBot,* Journal of Small Business and Enterprise Development, 25(3) (2018) 501–520.

[38] Statista.https://fr.statista.com/statistiques/472073/utilisateurs-internet-france/

[39] Svatosova, V., *The importance of online shopping behavior in the strategic management of e-commerce competitiveness* J. Compet. 12(4) (2020), 143.

[40] The National Institute of Statistics and Economic Studies collects, analyses and disseminates information on the French economy and society. https://www.insee.fr/en/accueil

[41] Wang, M., Chen, W., Huang, Y., Contractor, N.S. and Fu, Y., *Modeling customer preferences using multidimensional network analysis in engineering design,* Des. Sci. 2 (2016) e11.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1203–1223
https://doi.org/ 10.22067/ijnao.2024.87189.1405
https://ijnao.um.ac.ir/

Research Article

Modified hat functions: Application in space-time-fractional differential equations with Caputo derivative

E. Sedighi, O. Baghani*, and H. Azin

Abstract

The present article introduces an operational approach based on modified hat functions to solve the space-time-fractional differential equations in the Caputo sense. In this method, the derivative of the unknown function is considered as a linear combination of modified hat functions. We use the operational matrix of the Riemann–Liouville fractional integral of modified

*Corresponding author

Received 8 March 2024; revised 10 June 2024; accepted 7 August 2024

Elahe Sedighi
Department of Applied Mathematics, Faculty of Mathematics and Computer Science, Hakim Sabzevari University, Iran. e-mail: sedigi913@gmail.com

Omid Baghani
Department of Applied Mathematics, Faculty of Mathematics and Computer Science, Hakim Sabzevari University, Iran. e-mail: omid.baghani@gmail.com, o.baghani@hsu.ac.ir

Hadis Azin
Department of Applied Mathematics, Faculty of Mathematics and Computer Science, Hakim Sabzevari University, Iran. e-mail: h.azin1370@gmail.com

How to cite this article
Sedighi, E., Baghani, O. and Azin, H., Modified hat functions: Application in space-time-fractional differential equations with Caputo derivative. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1203-1223. https://doi.org/ 10.22067/ijnao.2024.87189.1405

hat functions to approximate the Caputo fractional derivative in order to reduce the problem to a system of Sylvester equations. The error of the mentioned method is of the order $\mathcal{O}(h^3)$. In addition, we examine several numerical examples to confirm the ability of the proposed approach.

AMS subject classifications (2020): Primary: 26A33; Secondary: 65M15, 65D25.

Keywords: Modified hat functions; Space-time-fractional differential equations; Operational matrix; Caputo fractional derivative.

1 Introduction

Fractional calculus including integral and derivative of arbitrary orders has effectively played an important role in engineering, physics, and various problems modeling during the past decades [26, 6]. The non-locality feature can be the main reason for using fractional calculus in applied problem modeling [21]. This characteristic refers to the property of a dynamical system where the future state of the system is influenced by all of its previous states. The growing usage of fractional differential equations in various applications has sparked significant interest in the development of numerical methods for their solutions. Several approaches have been devised to tackle these problems, as follows:

- Adomian decomposition method: This method involves decomposing fractional differential equations into a series of simpler equations and solving them iteratively. It has been applied in various fields such as physics, engineering, and biology [7, 11].

- Generalized differential transform method: This technique applies the differential transform method to fractional differential equations, enabling the solution through a series expansion. It has been used in solving a wide range of fractional problems [16, 20, 9].

- Variational iteration method: The variational iteration method constructs an iterative sequence of corrections to approximate the solution of a fractional differential equation. It has been successfully employed in various scientific and engineering problems [19].

- Finite difference method: The finite difference method, which is widely used in numerical analysis, discretizes the fractional differential equation on a grid and approximates the derivatives using finite difference formulas. It is a common and versatile approach for solving fractional differential equations [25, 15].

- Homotopy analysis method: The homotopy analysis method constructs a series solution by introducing an auxiliary parameter. By controlling this parameter, the solution can be refined iteratively, leading to accurate approximations for fractional differential equations [10].

- Wavelet method: The wavelet method applies wavelet theory to approximate the solution of fractional differential equations. It provides a robust framework for analyzing and solving fractional problems, particularly those with irregular or discontinuous solutions [4, 13, 5].

These methods offer various approaches to tackle the challenges posed by fractional differential equations, and their suitability depends on the specific characteristics of the problem. The cited references provide more detailed information on each method's theoretical foundations and practical applications. In recent decades, a limited number of algorithms have been proposed to solve numerical space-time-fractional differential problems [24, 12, 27].

In this work, we focus on a type of linear space-time-fractional model as follows:

$$D_x^{\alpha}\mathscr{U}(x,z) + D_z^{\beta}\mathscr{U}(x,z) = F(x,z), \qquad (x,z) \in [0,1]\times[0,1], \tag{1}$$

with the following initial and boundary conditions

$$\begin{aligned} \mathscr{U}(0,z) &= \theta_1(z), \qquad & \mathscr{U}(x,0) &= \theta_2(x),\\ D_z^1\mathscr{U}(0,z) &= \varphi_1(z), \qquad & D_x^1\mathscr{U}(x,0) &= \varphi_2(x), \end{aligned} \tag{2}$$

where $D_x^{\alpha}\mathscr{U}(x,z)$ and $D_z^{\beta}\mathscr{U}(x,z)$ are the Caputo fractional derivative with respect to x and z of order α and β with $0 < \alpha, \beta \leq 1$, respectively. Also,

$\mathscr{U}(x,z)$ is the unknown function and $F(x,z), \varphi_1(z), \varphi_2(x), \theta_1(z)$ and $\theta_2(x)$ are known continuous functions.

In the suggested approach of this investigation, the function $\mathscr{U}(x,z)$, which is not known, is expanded using the introduced modified hat functions. By employing integration and the fractional integral operational matrix, the approximations for $\mathscr{U}(x,z), D_x^{\alpha}\mathscr{U}(x,z)$ and $D_z^{\beta}\mathscr{U}(x,z)$ will be derived. This described method transforms the primary problem into an algebraic equation, enabling the computation of the unknown coefficients in the expansion.

It is important to note that the strengths and weaknesses of the operational matrix method depend on the specific problem and the implementation details. Consideration should be given to the specific requirements and characteristics of the problem at hand when choosing an appropriate numerical method. The operational matrix method provides efficient solutions for systems of ordinary and partial differential equations. It allows for the conversion of differential equations into algebraic equations, which can be solved using standard matrix techniques. By using higher-order approximation functions or increasing the number of terms in the expansion, the accuracy can be improved. The operational matrix methods are usually straightforward and well-suited for numerical computation. This makes the method relatively easy to implement with computationally efficient. Also, the convergence of the operational matrix method can be affected by the choice of approximation functions or the number of terms used in the expansion. In some cases, convergence may be slow, requiring additional computational resources or modifications to improve accuracy. On the other hand, incorporating boundary conditions into the operational matrix method can be challenging, especially for problems with complex or non-standard boundary conditions. Special techniques or modifications may be required to handle such cases effectively. This methodology has been widely utilized in numerous scholarly papers, showcasing its remarkable effectiveness and accuracy [3, 14, 1, 2, 22].

The structure of this paper is summarized as follows:

- Section 2 begins by presenting essential definitions and mathematical foundations for fractional calculus.

- In section 3, we conduct a thorough examination of modified hat functions.
- Section 4 is dedicated to the development of a numerical technique for addressing space-time-fractional partial differential equations.
- In section 5, an error estimate for the proposed method is derived.
- Moving on to section 6, we apply the devised method to solve a variety of fractional partial differential equations with numerical examples.
- Ultimately, the article concludes in section 7 with a summary of our findings and insights.

2 Fractional calculus

Herein, we express the information required for research work.

Definition 1 (see [23])**.** Let $0 < \alpha \leq 1$ and $0 < \beta \leq 1$ be real values and let $u : [0,1]^2 \to \mathbb{R}$ be a continuous function. The Caputo fractional derivatives of u concerning x and z are defined as

$$D_x^{\alpha} u(x,z) = \begin{cases} \dfrac{1}{\Gamma(1-\alpha)} \displaystyle\int_0^x (x-r)^{-\alpha} \frac{\partial u(r,z)}{\partial r} dr, & \alpha \in (0,1), \\ \dfrac{\partial u(x,z)}{\partial x}, & \alpha = 1, \end{cases} \tag{3}$$

and

$$D_z^{\beta} u(x,z) = \begin{cases} \dfrac{1}{\Gamma(1-\beta)} \displaystyle\int_0^z (z-s)^{-\beta} \frac{\partial u(x,s)}{\partial s} ds, & \beta \in (0,1), \\ \dfrac{\partial u(x,z)}{\partial z}, & \beta = 1. \end{cases} \tag{4}$$

Definition 2 (see [21])**.** For $0 < \alpha \leq 1$ and given function h, the Riemann–Liouville fractional integral with order of α is defined by

$$I_z^{\alpha}\ h(z) = \begin{cases} \dfrac{1}{\Gamma(\alpha)} \displaystyle\int_0^z (z-s)^{\alpha-1} h(s) ds, & \alpha \in (0,1), \\ h(z), & \alpha = 0. \end{cases} \tag{5}$$

Based on the next lemma, Caputo fractional derivatives and Riemann–Liouville fractional integrals can be considered as reciprocal operations.

Lemma 1 (see [18]). If $0 < z$ and $0 < \beta < \alpha \leq m \in \mathbb{N}$, then

$$I_z^{\alpha-\beta}\ D_z^{\alpha} f(z) = D_z^{\beta}\ f(z) - \sum_{l=\lceil\beta\rceil}^{m-1} f^{(l)}(0) \frac{z^{l-\beta}}{\Gamma(l-\beta+1)}, \tag{6}$$

where $\lceil\cdot\rceil$ is the ceiling function.

3 Modified hat functions

We now focus on modified hat functions and some of their properties. In this research study, the modified hat functions are used for approximating space and time variables.

Definition 3 (see [17, 18]). Let the interval $[0, t_f]$ be divided into n subintervals $[jh, (j+1)h]$ for $j = 0, 1, 2, \ldots, n-1$, where $2 \leq n$ and n is an even integer with $h = \frac{t_f}{n}$. The modified hat functions $\{\mathcal{H}_k(z)\}_{k=0}^{n}$ on $[0, t_f]$ are defined as follows:

$$\mathcal{H}_0(z) = \begin{cases} \dfrac{(z-h)(z-2h)}{2h^2}, & 0 \leq z \leq 2h, \\ 0, & \text{otherwise}, \end{cases} \tag{7}$$

for odd index k with $1 \leq k \leq n-1$,

$$\mathcal{H}_k(t) = \begin{cases} \dfrac{-(z-(k-1)h)(z-(k+1)h)}{h^2}, & (k-1)h \leq z \leq (k+1)h, \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

for even index k with $2 \leq k \leq n-2$,

$$\mathcal{H}_k(z) = \begin{cases} \dfrac{(z-(k-1)h)(z-(k-2)h)}{2h^2}, & (k-2)h \leq z \leq kh, \\ \dfrac{(z-(k+1)h)(z-(k+2)h)}{2h^2}, & kh \leq z \leq (k+2)h, \\ 0, & \text{otherwise}\ , \end{cases} \tag{9}$$

and

$$\mathcal{H}_n(z) = \begin{cases} \dfrac{(z-(t_f-h))(z-(t_f-2h))}{2h^2}, & t_f - 2h \leq z \leq t_f, \\ 0, & \text{otherwise}. \end{cases} \tag{10}$$

The modified hat functions exhibit noteworthy characteristics that render them highly advantageous for various applications. Several of these notable properties are as follows:

- These functions satisfy in the Kronecker delta condition as

$$\mathcal{H}_k(jh) = \delta_{kj}.$$

- The sum of the modified hat functions $\{\mathcal{H}_k(z)\}_{k=0}^{n}$ is one, that is,

$$\sum_{k=0}^{n} \mathcal{H}_k(z) = 1.$$

- The explicit formula for integral of these functions is obtained as follows:

$$\int_0^{t_f} \mathcal{H}_k(z)dz = \begin{cases} \frac{h}{3}, & k = 0, n, \\ \frac{4h}{3}, & k \text{ is odd and } 1 \leq k \leq n-1, \\ \frac{2h}{3}, & k \text{ is even and } \ 2 \leq k \leq n-2. \end{cases} \tag{11}$$

Any function $f(z) \in L^2[0, t_f]$ can be represented via the modified hat functions as

$$f(z) \simeq f_n(z) = \sum_{k=0}^{n} f_k \mathcal{H}_k(z) = F^T \mathscr{H}(z), \tag{12}$$

in which $f_i = f(ih)$ and

$$\mathscr{H}(z) = [\mathcal{H}_0(z),\ \mathcal{H}_1(z),\ \ldots,\ \mathcal{H}_n(z)]^T, \tag{13}$$

$$F = [f_0,\ f_1,\ \ldots,\ f_n]^T. \tag{14}$$

Moreover, every function $y(x, z)$ in $L^2([0, 1]^2)$ can be extended by the modified hat function as

$$y(x, z) \simeq y_n(x, z) = \sum_{k=0}^{n} \sum_{j=0}^{n} y_{kj} \mathcal{H}_i(x) \mathcal{H}_j(z) = \mathscr{H}^T(x) Y \mathscr{H}(z), \tag{15}$$

with $y_{kj} = y(x_k, z_j) = y(kh, jh)$.

Theorem 1 (see [17, 18]). For $\gamma > 0$, assume that $\mathscr{H}(z)$ is the vector introduced in the relation (13). Then, we get

$$I_z^\gamma \mathscr{H}(z) \simeq \mathcal{R}^{(\gamma)} \mathscr{H}(z), \tag{16}$$

in which $\mathcal{R}^{(\gamma)}$ is an $(n+1)\times(n+1)$ square matrix as

$$\mathcal{R}^{(\gamma)} = \frac{h^\gamma}{2\Gamma(\gamma+3)} \begin{pmatrix} 0 & \mathbf{p}_1^\gamma & \mathbf{p}_2^\gamma & \mathbf{p}_3^\gamma & \mathbf{p}_4^\gamma & \dots & \mathbf{p}_{n-1}^\gamma & \mathbf{p}_n^\gamma \\ 0 & \mathbf{q}_0^\gamma & \mathbf{q}_1^\gamma & \mathbf{q}_2^\gamma & \mathbf{q}_3^\gamma & \dots & \mathbf{q}_{n-2}^\gamma & \mathbf{q}_{n-1}^\gamma \\ 0 & \mathbf{w}_{-1}^\gamma & \mathbf{w}_0^\gamma & \mathbf{w}_1 & \mathbf{w}_2^\gamma & \dots & \mathbf{w}_{n-3}^\gamma & \mathbf{w}_{n-2}^\gamma \\ 0 & 0 & 0 & \mathbf{q}_0^\gamma & \mathbf{q}_1^\gamma & \dots & \mathbf{q}_{n-4}^\gamma & \mathbf{q}_{n-3}^\gamma \\ 0 & 0 & 0 & \mathbf{w}_{-1}^\gamma & \mathbf{w}_0^\gamma & \dots & \mathbf{w}_{n-5}^\gamma & \mathbf{w}_{n-4}^\gamma \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \mathbf{q}_0^\gamma & \mathbf{q}_1^\gamma \\ 0 & 0 & 0 & 0 & 0 & \dots & \mathbf{w}_{-1}^\gamma & \mathbf{w}_0^\gamma \end{pmatrix}, \tag{17}$$

where

$$\mathbf{p}_l^\gamma = \begin{cases} \gamma(3+2\gamma), & l=1, \\ l^{\gamma+1}(2l-6-3\gamma) + 2l^\gamma(1+\gamma)(2+\gamma) & \\ -(l-2)^{\gamma+1}(2l-2+\gamma), & 2 \le l \le n, \end{cases} \tag{18}$$

$$\mathbf{q}_l^\gamma = \begin{cases} 4(1+\gamma), & l=0, \\ 4\big[(l-1)^{\gamma+1}(l+1+\gamma) & \\ -(l+1)^{\gamma+1}(l-1-\gamma)\big], & 1 \le l \le n-1, \end{cases} \tag{19}$$

and

$$\mathbf{w}_l^\gamma = \begin{cases} -\gamma, & l=-1, \\ 2^{\gamma+1}(2-\gamma), & l=0, \\ 3^{\gamma+1}(4-\gamma) - 6(2+\gamma), & l=1, \\ (l+2)^{\gamma+1}(2l+2-\gamma) - 6l^{\gamma+1}(2+\gamma) & \\ -(l-2)^{\gamma+1}(2l-2+\gamma), & 2 \le l \le n-2. \end{cases} \tag{20}$$

4 Numerical method

This section is devoted to the implementation of a numerical approach for the problem (1)–(2) by the modified hat functions. For this purpose, we have

$$D_x^1(D_z^1\, \mathscr{U}(x,z)) \simeq \mathscr{H}^T(x) A \mathscr{H}(z), \tag{21}$$

where

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n1} & \dots & a_{nn} \end{pmatrix}. \tag{22}$$

From the relations (2), (16), and (21), we obtain

$$\begin{aligned} D_x^1 \mathscr{U}(x,z) &= \int_0^z \left(D_x^1 (D_z^1 \ \mathscr{U}(x,z)) \right) dz + D_x^1 u(x,0) \\ &\simeq \int_0^z \mathscr{H}^T(x) A \mathscr{H}(z) dz + \varphi_2(x) \\ &= \mathscr{H}^T(x) A \mathcal{R}^{(1)} \mathscr{H}(z) + \varphi_2(x). \end{aligned} \tag{23}$$

Similarly, it can be concluded

$$\begin{aligned} D_z^1 \mathscr{U}(x,z) &= \int_0^x \left(D_x^1 (D_z^1 \ \mathscr{U}(x,z)) \right) dx + D_z^1 \mathscr{U}(0,z) \\ &\simeq \int_0^z \mathscr{H}^T(x) A \mathscr{H}(z) dx + \varphi_1(z) \\ &= \mathscr{H}^T(x) \big(\mathcal{R}^{(1)}\big)^T A \mathscr{H}(z) + \varphi_1(z). \end{aligned} \tag{24}$$

So, we take the integral from $D_x^1 \mathscr{U}(x,z)$ with respect to x of order 1 to make an approximation of $\mathscr{U}(x,z)$ as follows:

$$\begin{aligned} \mathscr{U}(x,z) &\simeq \mathscr{H}^T(x) \big(\mathcal{R}^{(1)}\big)^T A \mathcal{R}^{(1)} \mathscr{H}(z) + \hat{u}(x,0) + u(0,z) \\ &= \mathscr{H}^T(x) \big(\mathcal{R}^{(1)}\big)^T A \mathcal{R}^{(1)} \mathscr{H}(z) + \hat{\theta}_2(x) + \theta_1(z), \end{aligned} \tag{25}$$

in which $\hat{\theta}_2(x)$ is $\theta_2(x)$ without its scalar. Replacing the relation (23) in (6), we have

$$\begin{aligned} D_x^\alpha \mathscr{U}(x,z) &= I_x^{1-\alpha} (D_x^1 \mathscr{U}(x,z)) - \vartheta_1 \\ &\simeq I_x^{1-\alpha} \left(\mathscr{H}^T(x) A \mathcal{R}^{(1)} \mathscr{H}(z) + \varphi_2(x) \right) - \vartheta_1 \\ &= \mathscr{H}^T(x) \big(\mathcal{R}^{(1-\alpha)}\big)^T A \mathcal{R}^{(1)} \mathscr{H}(z) + I_x^{1-\alpha}(\varphi_2(x)) - \vartheta_1, \end{aligned} \tag{26}$$

and

$$
\begin{aligned}
D_z^{\beta}\mathscr{U}(x,z) &= I_z^{1-\beta}(D_z^{1}\mathscr{U}(x,z)) - \vartheta_2 \\
&\simeq I_z^{1-\beta}\bigg(\mathscr{H}^T(x)\big(\mathcal{R}^{(1)}\big)^T A\mathscr{H}(z) + \varphi_1(z)\bigg) - \vartheta_2 \\
&= \mathscr{H}^T(x)\big(\mathcal{R}^{(1)}\big)^T A\mathcal{R}^{(1-\beta)}\mathscr{H}(z) + I_z^{1-\beta}(\varphi_1(z)) - \vartheta_2,
\end{aligned} \tag{27}
$$

where $\vartheta_1 = \mathscr{U}_x^{(1)}(0,0)$ and $\vartheta_2 = \mathscr{U}_z^{(1)}(0,0)$ are the known scalars. Therefore, by replacing the relation (26) and (27) in (1), we have

$$
\begin{aligned}
&(\mathscr{H}^T(x)\big(\mathcal{R}^{(1-\alpha)}\big)^T A\mathcal{R}^{(1)}\mathscr{H}(z) + I_x^{1-\alpha}(\varphi_2(x)) - \vartheta_1) \\
&+ (\mathscr{H}^T(x)\big(\mathcal{R}^{(1)}\big)^T A\mathcal{R}^{(1-\beta)}\mathscr{H}(z) + I_z^{1-\beta}(\varphi_1(t)) - \vartheta_2) = F(x,z).
\end{aligned} \tag{28}
$$

Consider the function $g(x,z)$ as follows:

$$
g(x,z) = F(x,z) - I_x^{1-\alpha}\big(\varphi_2(x)\big) - I_z^{1-\beta}\big(\varphi_1(z)\big) + \vartheta_1 + \vartheta_2. \tag{29}
$$

Now, we approximate the $g(x,z)$ via the modified hat functions as

$$
g(x,z) \simeq \mathscr{H}^T(x)G\mathscr{H}(z), \tag{30}
$$

in which G is a square matrix of rank $(n+1)$ with $g_{ij} = g(x_i, z_j)$. So, the relation (28) reduces to (31) as

$$
\begin{aligned}
\mathscr{H}^T(x)\big(\mathcal{R}^{(1-\alpha)}\big)^T A\mathcal{R}^{(1)}\mathscr{H}(z) + \mathscr{H}^T(x)\big(\mathcal{R}^{(1)}\big)^T A\mathcal{R}^{(1-\beta)}\mathscr{H}(z) \\
= \mathscr{H}^T(x)G\mathscr{H}(z).
\end{aligned} \tag{31}
$$

By removing the vectors $\mathscr{H}^T(x)$ and $\mathscr{H}(z)$ from both sides of the equation (31), we will have the following Sylvester-type equation [8]:

$$
\big(\mathcal{R}^{(1-\alpha)}\big)^T A\mathcal{R}^{(1)} + \big(\mathcal{R}^{(1)}\big)^T A\mathcal{R}^{(1-\beta)} = G. \tag{32}
$$

After solving (32) and replacing the values of a_{ij} in the relation (25), the approximate solution of $\mathscr{U}(x,z)$ is obtained.

5 Error estimation

In this section, we derive an error estimate for the proposed method. For this issue, we first recall the following theorem from the reference [17].

Theorem 2. If the function $f(\cdot) \in C^3[0,t_f]$ is approximated by the set of modified hat functions as $f_n(t) = \sum_{i=0}^{n} f(ih)\mathcal{H}_i(t)$, then $|f(t) - f_n(t)| = \mathcal{O}(h^3)$, where $2 \le n$ and n is an even integer with $h = \frac{t_f}{n}$.

Let sufficiently smooth function $u(x,z)$ be the exact solution of the original problem (1)–(2). Assume that $u(x,z)$ can be separated as a product of two functions $F(x)$ and $G(z)$, that is,

$$u(x,z) = F(x)G(z).$$

Now, by approximating $F(x)$ and $G(z)$ using modified hat functions, we get

$$F(x) = F^T \mathscr{H}(x) + \mathcal{O}(h^3), \quad G(z) = G^T \mathscr{H}(z) + \mathcal{O}(h^3),$$

where $F_i := F(x_i)$ and $G_j := G(z_j)$. So $u(x,z) = \mathscr{H}^T(x) F G^T \mathscr{H}^T(z) + \mathcal{O}(h^3)$. This shows that the error of a two-dimensional function estimated by modified hat functions is of the order $\mathcal{O}(h^3)$. With this explanation and by (21), we have thus derived the following relation:

$$|D_x^1(D_z^1 \mathscr{U}(x,z)) - \mathscr{H}^T(x) A \mathscr{H}(z)| = \mathcal{O}(h^3).$$

Now from (25), since the functions $\hat{\theta}_2(x)$ and $\theta_1(z)$ are known, by ignoring the operational matrix error, we have

$$\begin{aligned} &|\mathscr{U}(x,z) - \mathscr{U}_n(x,z)| \\ &= |I_z^1(I_x^1(D_x^1(D_z^1 \mathscr{U}(x,z)))) - \mathscr{H}^T(x)\big(\mathcal{R}^{(1)}\big)^T A \mathcal{R}^{(1)} \mathscr{H}(z)| \\ &= \mathcal{O}(h^3). \end{aligned} \tag{33}$$

This shows that the convergence rate of the suggested method is $\mathcal{O}(h^3)$.

6 Illustrative examples

In this section, we examine the presented numerical method to solve the problem (1)–(2) and compare the results with their exact solutions using Maple2020 software. In each example, the error E_n is defined by

$$E_n(\mathscr{U}) = \frac{1}{n^2}\left(\sum_{i=0}^{n}\sum_{j=0}^{n}\Big(\mathscr{U}(ih,jh) - \mathscr{U}_n(ih,jh)\Big)^2\right)^{\frac{1}{2}}, \tag{34}$$

where $\mathscr{U}$ and $\mathscr{U}_n$ are the exact and numerical solutions, respectively. Also, the error function is plotted for each example by $|\mathscr{U}(x,z) - \mathscr{U}_n(x,z)|$. All examples are numerically approximated with the help of *Maple* 2020 software on a laptop with CPU 3.1 GHz and Core i5.

Example 6.1. Consider the non-homogeneous space-time fractional problem with the following form:

$$D_x^{\frac{1}{4}} \mathscr{U}(x,z) + D_z^{\frac{1}{4}} \mathscr{U}(x,z) = F(x,z), \quad x, z \in [0,1], \tag{35}$$

in which

$$F(x,z) = \frac{4(x^{\frac{3}{4}} z + x z^{\frac{3}{4}})}{3\Gamma(\frac{3}{4})}, \tag{36}$$

with

$$D_x^1 \mathscr{U}(0,z) = D_z^1 \mathscr{U}(x,0) = \mathscr{U}(0,z) = \mathscr{U}(x,0) = 0. \tag{37}$$

The exact solution is $\mathscr{U}(x,z) = xz$. So, it can be written that

$$I_x^{\frac{3}{4}}(\varphi_2(x)) = I_z^{\frac{1}{4}}(\varphi_1(z)) = \vartheta_1 = \vartheta_2 = 0. \tag{38}$$

According to the relations (17) and (30) with $n = 2$, we get

$$\begin{gathered} G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \dfrac{2\ 2^{\frac{1}{4}}}{3\ \Gamma(\dfrac{3}{4})} & \dfrac{4(\dfrac{1}{2} + \dfrac{2^{\frac{1}{4}}}{2})}{3\ \Gamma(\dfrac{3}{4})} \\ 0 & \dfrac{4(\dfrac{1}{2} + \dfrac{2^{\frac{1}{4}}}{2})}{3\ \Gamma(\dfrac{3}{4})} & \dfrac{8}{3\ \Gamma(\dfrac{3}{4})} \end{pmatrix}, \\ \mathcal{R}^{(1)} = \begin{pmatrix} 0 & \dfrac{5}{24} & \dfrac{1}{6} \\ 0 & \dfrac{1}{3} & \dfrac{2}{3} \\ 0 & -\dfrac{1}{24} & \dfrac{1}{6} \end{pmatrix}, \quad \mathcal{R}^{(\dfrac{3}{4})} = \begin{pmatrix} 0 & \dfrac{27\ 2^{\frac{1}{4}}}{32\ \Gamma(\dfrac{15}{4})} & \dfrac{9}{16\ \Gamma(\dfrac{15}{4})} \\ 0 & \dfrac{7\ 2^{\frac{1}{4}}}{4\ \Gamma(\dfrac{15}{4})} & \dfrac{3}{\Gamma(\dfrac{15}{4})} \\ 0 & -\dfrac{3\ 2^{\frac{1}{4}}}{16\ \Gamma(\dfrac{15}{4})} & \dfrac{5}{4\ \Gamma(\dfrac{15}{4})} \end{pmatrix}. \end{gathered} \tag{39}$$

Then, we can rewritten (31) as

$$\mathscr{H}^T(x)\big(\mathcal{R}^{(\frac{3}{4})}\big)^T A\mathcal{R}^{(1)}\mathscr{H}(z)+\mathscr{H}^T(x)\big(\mathcal{R}^{(1)}\big)^T A\mathcal{R}^{(\frac{3}{4})}\mathscr{H}(z)=\mathscr{H}^T(x)G\mathscr{H}(z)), \tag{40}$$

which reduces to

$$\big(\mathcal{R}^{(\frac{3}{4})}\big)^T A\mathcal{R}^{(1)}+\big(\mathcal{R}^{(1)}\big)^T A\mathcal{R}^{(\frac{3}{4})}=G. \tag{41}$$

We solve (41) using Maple software and obtain the approximation $\mathscr{U}(x,z)$ with the relation (25). The exact and numerical solutions with $n=2$ are shown in Figure 1. The numerical solution with $n=8$ is displayed in Figure 2. Also, error functions with $n=2$ and $n=8$ are shown in Figure 3 in which $E_2(\mathscr{U})=0.00086$ and $E_8(\mathscr{U})=0$. We list the absolute error for different values of n in Example 6.1 in Table 1.

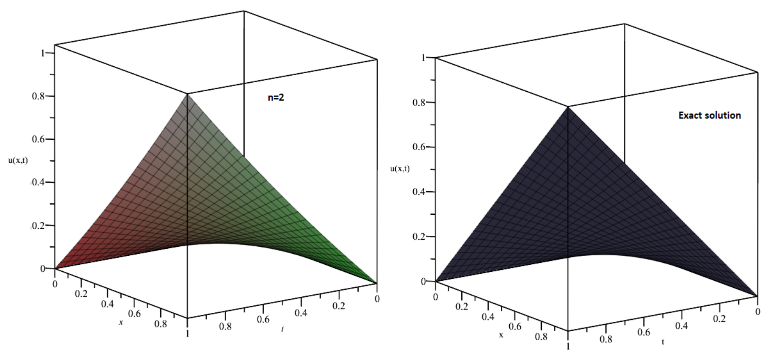


Figure 1: The exact and numerical solutions with $n=2$ for Example 6.1.

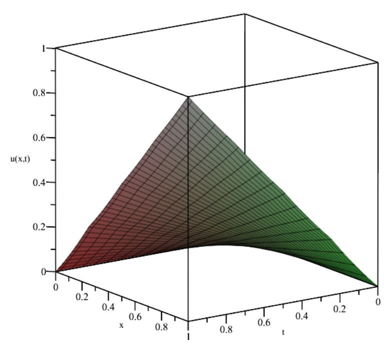


Figure 2: Numerical solution with $n=8$ for Example 6.1.

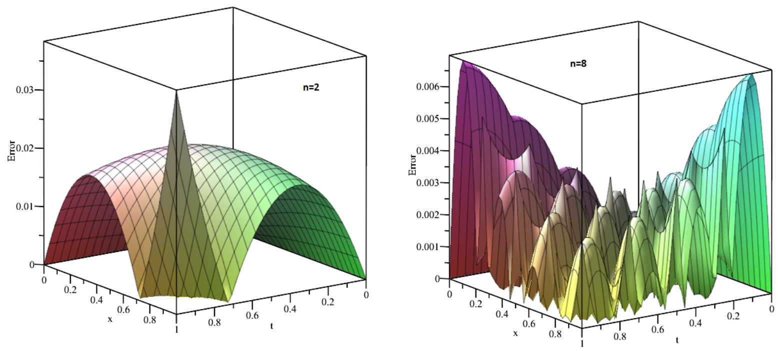


Figure 3: Error function $|\mathscr{U}(x,z) - \mathscr{U}_n(x,z)|$ with $n=2$ and $n=8$ for Example 6.1.

Table 1: The absolute error for different values of n in Example 6.1

(x,t)	$n=2$	$n=4$	$n=6$	$n=8$
$(0,0)$	0	0	0	0
$(0.2,0.2)$	0.0058	0.0042	0.0025	0.00060
$(0.4,0.4)$	0.017	0.0024	0.00035	0.0026
$(0.6,0.6)$	0.022	0.00079	0.0019	0.0023
$(0.8,0.8)$	0.0096	0.010	0.0041	0.000048
$(1,1)$	0.038	0.013	0.0059	0.0027

Example 6.2. Consider the following fractional space-time fractional model:

$$D_x^{\frac{1}{3}}\mathscr{U}(x,z) + D_z^{\frac{1}{2}}\mathscr{U}(x,z) = F(x,z), \quad 0 \le x, z \le 1, \tag{42}$$

where

$$F(x,z) = \frac{\Gamma(3)x^{\frac{5}{3}}}{\Gamma(\frac{8}{3})} + \frac{\Gamma(3)z^{\frac{3}{2}}}{\Gamma(\frac{5}{2})}, \tag{43}$$

with

$$D_z^1\mathscr{U}(0,z) = 2z,\ D_x^1\mathscr{U}(x,0) = 2x,\ \mathscr{U}(0,z) = z^2,\ \mathscr{U}(x,0) = x^2. \tag{44}$$

The exact solution is $\mathscr{U}(x,z) = x^2 + z^2$. Now, we calculate the following values:

$$
\begin{cases}
I_x^{\frac{2}{3}}(\varphi_2(x)) = I_x^{\frac{2}{3}}(2x) = \dfrac{2\Gamma(2)}{\Gamma(\frac{8}{3})} x^{\frac{5}{3}}, \\
I_z^{\frac{1}{2}}(\varphi_1(z)) = I_z^{\frac{1}{2}}(2z) = \dfrac{2\Gamma(2)}{\Gamma(\frac{5}{2})} z^{\frac{3}{2}}, \\
\vartheta_1 = \vartheta_2 = 0.
\end{cases} \tag{45}
$$

Therefore, we get

$$
g(x,z) = \left(\frac{\Gamma(3) x^{\frac{5}{3}}}{\Gamma(\dfrac{8}{3})} + \frac{\Gamma(3) z^{\frac{3}{2}}}{\Gamma(\dfrac{5}{2})} \right) - \frac{2\Gamma(2)}{\Gamma(\frac{8}{3})} x^{\frac{5}{3}} - \frac{2\Gamma(2)}{\Gamma(\frac{5}{2})} z^{\frac{3}{2}} = 0.
$$

Equation (32) is rewritten as follows:

$$
\left(\mathcal{R}^{(\frac{2}{3})}\right)^T A \mathcal{R}^{(1)} + \left(\mathcal{R}^{(1)}\right)^T A \mathcal{R}^{(\frac{1}{2})} = 0. \tag{46}
$$

After solving Sylvester equation (46), based on (25), we obtain $A = 0$ and $\mathscr{U}(x,z) = x^2 + z^2$, which is the exact solution. The exact and numerical solutions with $n = 2$ are plotted in Figure 4. Moreover, the error function is shown in Figure 5 in which $E_2(\mathscr{U}) = 0$.

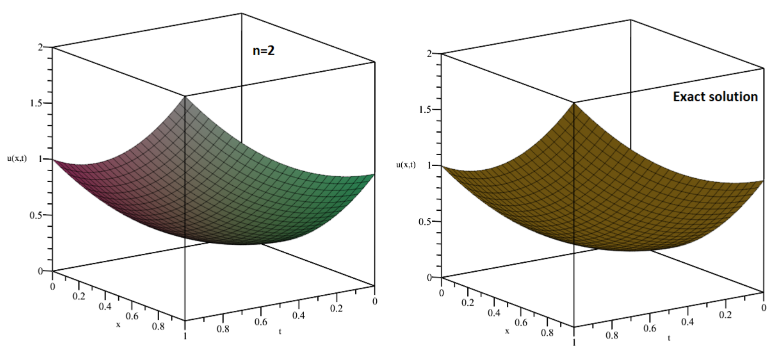


Figure 4: The exact and numerical solutions with $n = 2$ for Example 6.2.

Example 6.3. Consider the following non-homogeneous fractional space-time problem:

$$
D_x^{\frac{1}{2}} \mathscr{U}(x,z) + D_z^{\frac{1}{3}} \mathscr{U}(x,z) = F(x,z), \quad (x,z) \in [0,1]^2, \tag{47}
$$

with

$$
F(x,z) = \frac{(3t^5 \sqrt{\Pi} x)}{4} + \frac{729 z^{\frac{14}{3}}}{308 \Gamma(\frac{2}{3})}, \tag{48}
$$

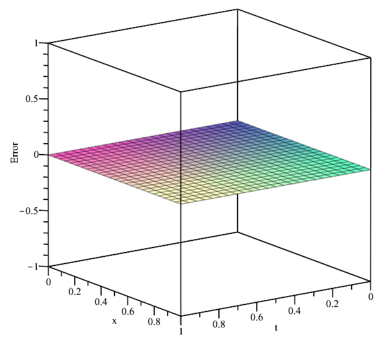


Figure 5: Error function with $n = 2$ for Example 6.2.

in which

$$D_z^1 \mathscr{U}(0,t) = D_x^1 \mathscr{U}(x,0) = \mathscr{U}(0,z) = \mathscr{U}(x,0) = 0. \tag{49}$$

The exact solution is $\mathscr{U}(x,z) = x^{\frac{3}{2}} t^5$. By calculating the operational matrices of the orders $\dfrac{1}{2}$ and $\dfrac{2}{3}$ and then substituting them into (32), it can be obtained a system of linear equations to calculate the unknown coefficients. Figure 6 shows the approximate and exact solutions of the fractional space-time equation for $n = 10$.

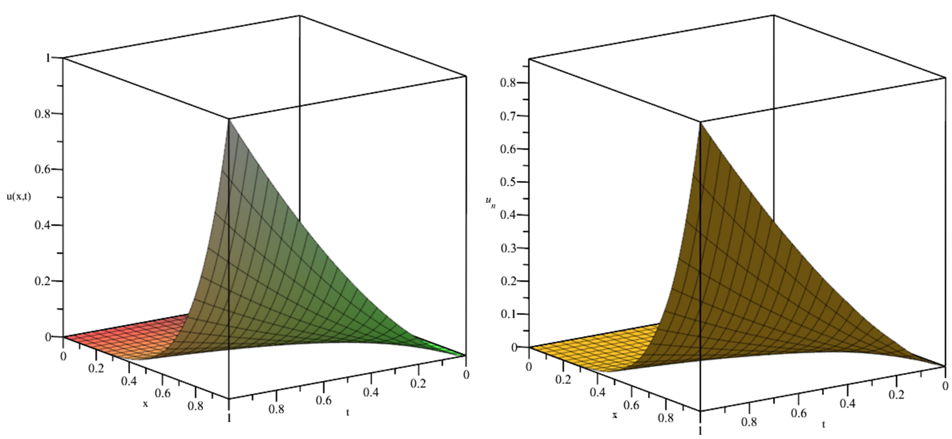


Figure 6: The exact and numerical solutions with n=10 for Example 6.3.

Example 6.4. At the end, consider the space-time-fractional differential equation as

$$D_x^{\alpha}\mathscr{U}(x,z) + D_z^{\beta}\mathscr{U}(x,z) = F(x,z), \quad (x,z) \in [0,1]^2, \tag{50}$$

with the term source

$$F(x,z) = \frac{\Gamma(3)x^{2-\alpha}(z^2+1)}{\Gamma(3-\alpha)} + \frac{\Gamma(3)z^{2-\beta}(x^2+1)}{\Gamma(3-\beta)}, \tag{51}$$

in which

$$D_z^1\mathscr{U}(0,t) = 2z,\ D_x^1\mathscr{U}(x,0) = 2x,\ \mathscr{U}(0,z) = z^2+1,\ \mathscr{U}(x,0) = x^2+1. \tag{52}$$

The analytical solution is $\mathscr{U}(x,z) = (x^2+1)(z^2+1)$ with $\alpha = \frac{1}{2}$ and $\beta = \frac{1}{3}$. So, we obtain

$$g(x,z) = \frac{\Gamma(3)x^{2-\alpha}(z^2)}{\Gamma(3-\alpha)} + \frac{\Gamma(3)z^{2-\beta}(x^2)}{\Gamma(3-\beta)}.$$

Now, for $n = 2$, we get

$$G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.2648 & 0.9036 \\ 0 & 0.8642 & 2.8337 \end{pmatrix}. \tag{53}$$

By replacing (53) in (32), this equation can be solved. The analytical and numerical solutions with $n = 2$ are shown in Figure 7 in which $E_2(\mathscr{U}) = 0.0159$.

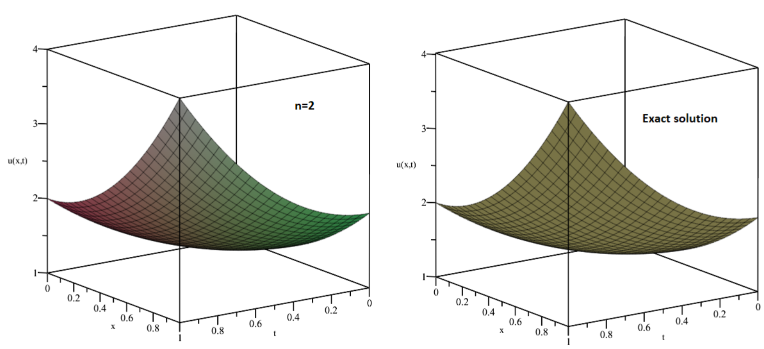


Figure 7: The exact and numerical solutions with $n = 2$ for Example 6.4.

7 Conclusion

In the present article, a numerical algorithm for solving space-time-fractional problems was given. Using the derivative in the Caputo sense and the Riemann–Liouville integral, we approximated the existing functions by modified hat functions. By replacing approximation with the main problem and simplifying it, the fractional equation was reduced to a system of algebraic equations. In the presented method, the linear space-time-fractional differential equation becomes a system of Sylvester equations, which is very easy to solve. The coefficients of the basic functions can be easily calculated, but it takes more time to find solutions closer to the real value. The solutions obtained from solving the mentioned system in (25) are an approximate solution of the main problem (1)–(2). The proposed approach has been used for several examples. In each example, the approximation function $\mathscr{U}_n(x, z)$ and the error from its exact solution were calculated. The results obtained demonstrated that the proposed technique has the capability to solve numerically the space-time-fractional partial differential problem.

Acknowledgements

The authors gratefully thank to the referees and editor for their constructive comments.

References

[1] Azin, H., Heydari, M.H. and Mohammadi, F., *Vieta–Fibonacci wavelets: Application in solving fractional pantograph equations*, Math. Methods Appl. Sci., 45 (2022) 411–422.

[2] Baghani, O., *SCW-iterative-computational method for solving a wide class of nonlinear fractional optimal control problems with Caputo derivatives*, Math. Comput. Simul., 202 (2022) 540–558.

[3] Jebreen, H.B. and Cattani, C., *Interpolating scaling functions Tau*

method for solving space–time fractional partial differential equations, Symmetry, 14 (2022) 2463.

[4] Chen, Y.M., Wu, Y.B., Cui, Y., Wang, Z. and Jin, D., *Wavelet method for a class of fractional convection-diffusion equation with variable coefficients*, J. Comput. Sci., 1 (2010) 146–149.

[5] Chen, Y.M., Yi, M.X. and Yu, C. X., *Error analysis for numerical solution of fractional differential equation by Haar wavelets method*, J. Comput. Sci., 5 (2012) 367–373.

[6] Diouf, M. and Sene N., *Analysis of the financial chaotic model with the fractional derivative operator*, Complexity, Article ID 9845031, 2020.

[7] El-Kalla, I.L., *Convergence of the Adomian method applied to a class of nonlinear integral equations*, Appl. Math. Comput., 21 (2008) 372–376.

[8] Gardiner, J.D., Laub, A.J., Amato, J.J. and Moler, C.B., *Solution of the Sylvester matrix equation* $AXB^T + CXD^T = E$, ACM Trans. Math. Softw., 18 (1992) 23–231.

[9] Guner, O. and Bekir, A., *Solving nonlinear space–time fractional differential equations via ansatz method*, Comput. Methods Differ. Equ., 6 (2018) 1–11.

[10] Hashim, I. and Abdulaziz, O., *Homotopy analysis method for fractional IVPs*, Commun. Nonlinear Sci. Numer. Simul., 14 (2009) 674-684.

[11] Hosseini, M.M., *Adomian decomposition method for solution of nonlinear differential algebraic equations*, Appl. Math. Comput., 181 (2006) 1737–1744.

[12] Hosseininia, M., Heydari, M.H., Maalek Ghaini, F.M. and Avazzadeh, Z., *A wavelet method to solve nonlinear variable-order time fractional 2D Klein-Gordon equation*, Comput. Math. Appl., 78 (2019) 3713–3730.

[13] Jafari, H. and Yousefi, S.A., *Application of Legendre wavelets for solving fractional differential equations*, Comput. Math. Appl., 62 (2011) 1038–1045.

[14] Lakestani, M., Dehghan, M., and Irandoust-pakchin, S., *The construction of operational matrix of fractional derivatives using B-spline functions*, Commun. Nonlinear Sci. Numer. Simul., 17 (2012) 1149–1162.

[15] Meerschaert, M.M., Scheffler, H.P. and Tadjeran, C., *Finite difference methods for two-dimensional fractional dispersion equation*, J. Comput. Phys., 211 (2006) 249–261.

[16] Momani, S. and Odibat, Z., *Generalized differential transform method for solving a space and time-fractional diffusion-wave equation*, Phys. Lett. A, 370 (2007) 379–387.

[17] Nemati, S. and Lima, P., *Numerical solution of nonlinear fractional integro-differential equations with weakly singular kernels via a modification of hat functions*, Appl. Math. Comput., 327 (2018) 79–92.

[18] Nemati, S., Lima, P. and Sedaghat, S., *An effective numerical method for solving fractional pantograph differential equations using modification of hat functions*, Appl. Numer. Math., 131 (2018) 174–89.

[19] Odibat, Z., *A study on the convergence of variational iteration method*, Math. Comput. Model., 51 (2010) 1181–1192.

[20] Odibat, Z. and Momani, S., *Generalized differential transform method: Application to differential equations of fractional order*, Appl. Math. Comput., 197 (2008) 467–477.

[21] Podlubny, I., *Fractional differential equations: Mathematics in science and engineering*, Vol. 198, San Diego, Academic Press, 1999.

[22] Pourbabaee, M. and Saadatmandi, A., *A new operational matrix based on Müntz–Legendre polynomials for solving distributed order fractional differential equations*, Math. Comput. Simul., 194 (2022) 210–235.

[23] Shen, S., Liu, F., Chen, J., Turner, I. and Anh, V., *Numerical techniques for the variable order time fractional diffusion equation*, Appl. Math. Comput., 218 (2012) 10861–10870.

[24] Soradi-Zeid, S., Jahanshahi, H., Yousefpour, A. and Bekiros, S., *King algorithm: A novel optimization approach based on variable-order fractional calculus with application in chaotic financial systems*, Chaos, Solit. Fract., 132 (2020) 109569.

[25] Sun, H., Chen W., Li, C. and Chen, Y., *Finite difference schemes for variable-order time fractional diffusion equation*, Int. J. Bifurc. Chaos Appl. Sci., 22 (2012) 1250085.

[26] Thiao, A. and Sene, A., *Fractional optimal economic control problem described by the generalized fractional order derivative*, 4th International Conference on Computational Mathematics and Engineering Sciences (CMES), 2019.

[27] Wang, L., Ma, Y. and Meng, Z., *Haar wavelet method for solving fractional partial differential equations numercally*, Appl. Math. Comput, 227 (2014) 66–76.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1224–1246
https://doi.org/10.22067/ijnao.2024.87501.1419
https://ijnao.um.ac.ir/

Research Article

An efficient collocation scheme for new type of variable-order fractional Lane–Emden equation

H. Azin*, A. Habibirad, E. Hesameddini, M.H. Heydari

Abstract

The fractional Lane–Emden model illustrates different phenomena in astrophysics and mathematical physics. This paper involves the Vieta–Lucas ($\mathcal{V}t$-$\mathcal{L}$) bases to solve types of variable-order ($\mathcal{V}$-$\mathcal{O}$) fractional Lane–Emden

*Corresponding author

Received 8 March 2024; revised 10 June 2024; accepted 7 August 2024

Hadis Azin
Department of Mathematics and Computer Sciences, Hakim Sabzevari University, Sabzevar, Iran. e-mail: h.azin1370@gmail.com

Ali Habibirad
Department of Mathematics and Computer Sciences, Hakim Sabzevari University, Sabzevar, Iran. e-mail: a.habibirad@sutech.ac.ir

Esmail Hesameddini
Department of Mathematics, Shiraz University of Technology, Shiraz, Iran. e-mail: hesameddini@sutech.ac.ir

Mohammad Hossein Heydari
Department of Mathematics, Shiraz University of Technology, Shiraz, Iran. e-mail: heydari@sutech.ac.ir

How to cite this article
Azin, H., Habibirad, A., Hesameddini, E. and Heydari, M.H., An efficient collocation scheme for new type of variable-order fractional Lane–Emden equation. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1224-1246. https://doi.org/10.22067/ijnao.2024.87501.1419

equation (linear and nonlinear). The operational matrix of the $\mathcal{V}$-$\mathcal{O}$ fractional derivative is obtained for the $\mathcal{V}t$-$\mathcal{L}$ polynomials. In the established approach, these polynomials are applied to transform the main problem into an algebraic equations system. To indicate the performance and capability of the scheme, a number of examples are presented for various types of $\mathcal{V}$-$\mathcal{O}$ fractional Lane–Emden equations. Also, for one example, a comparison is done between the calculated results by our technique and those obtained via the Bernoulli polynomials. Overall, this paper introduces a new methodology for solving $\mathcal{V}$-$\mathcal{O}$ fractional Lane–Emden equations using $\mathcal{V}t$-$\mathcal{L}$ bases. The derived operational matrix and the transformation to an algebraic equation system offer practical advantages in solving these equations efficiently. The presented examples and comparative analysis highlight the effectiveness and validity of the proposed technique, contributing to the understanding and advancement of fractional Lane–Emden models in astrophysics and mathematical physics.

AMS subject classifications (2020): 34A08; 65L60; 65N35; 35L10

Keywords: Lane–Emden equation; Vieta–Lucas polynomials; Variable-order fractional differential equation.

1 Introduction

Due to the helpful usage of fractional calculus in many scientific and engineering fields [7, 12, 25], from the first decade of this century until now, this issue has been challenging for researchers and will certainly continue in the coming decades. Finding the explicit solution for these equations is difficult and sometimes impossible. Therefore, many numerical techniques are extended for solving fractional differential equations. For example, interested readers can refer to [3, 17, 13, 22, 36, 35] and references therein.

Variable-order ($\mathcal{V}$-$\mathcal{O}$) operators are a natural generalization of operators with constant fractional order. In these types of operators, orders can be considered as a function of time, space, or both. In 1993, the first definition for this class of equations was provided by Samko and Ross [34]. Many dynamic procedures may vary by time or place. The fractional order role in these procedures shows that $\mathcal{V}$calculus is the normal prospect for supplying a useful mathematical plan to explain complicated dynamical models. For example, $\mathcal{V}$-$\mathcal{O}$ fractional derivatives have been used for the processing of geographical data in [9], diffusion in [38], signature verification in [40], and viscoelasticity in [8]. Several mapping effects on the fractional operators in types of cases in Hölder spaces expanded for the topic of $\mathcal{V}$-$\mathcal{O}$ in [32]. Researchers have done many studies on the numerical approaches for solving the $\mathcal{V}$-$\mathcal{O}$ fractional problems. For example in [24], the $\mathcal{V}$-$\mathcal{O}$ diffusion equation is discussed with the conditionally stable explicit finite difference method. Zhuang et al. [44] considered conditionally stable explicit and unconditionally stable implicit methods in the $\mathcal{V}$-$\mathcal{O}$ fractional advection-diffusion equation. The scholars of [13] introduced a meshless moving Kriging interpolation scheme to solve the two-dimensional $\mathcal{V}$-$\mathcal{O}$ fractional mobile/immobile advection-diffusion problem. The Adams–Bashforth–Moulton predictor-corrector technique was proposed in [39, 26] to simulate $\mathcal{V}$-$\mathcal{O}$ fractional differential equations with time delays. To see other properties and numerical schemes regarding $\mathcal{V}$-$\mathcal{O}$ fractional differential equations, see [29] and references therein.

The Lane–Emden equations play a significant role in the fields of engineering sciences and physics. These equations find wide application in addressing various phenomena across disciplines such as thermodynamics, fluid mechanics, mathematical physics, and astrophysics. Notable examples include modeling stellar structures, analyzing isothermal gas spheres, studying thermionic currents, and investigating the thermal behavior of spherical gas clouds [10, 5, 11, 6].

Various methods have been employed to tackle Lane–Emden problems effectively, including the Homotopy technique [41], the B-Spline approach [4], the modified variational method [37], Lie group analysis [21], the Adomian and modified Adomian decomposition techniques [42, 23], finite element techniques [18], transform differential procedures [43], Taylor's series method [14], the Legendre wavelets method [20], and the rational Bernoulli collocation approach [28]. For a more in-depth exploration of Lane–Emden equations, encompassing their variations, historical context, and practical applications, interested individuals are encouraged to consult the work [1] and its associated references.

In this study, consider the $\mathcal{V}$-$\mathcal{O}$ fractional Lane–Emden model as

$$
{}_{0}^{c}D_{\mathfrak{t}}^{\varrho(\mathfrak{t})}\nu(\mathfrak{t}) + \frac{\xi}{\mathfrak{t}^{\varrho(\mathfrak{t})-\varsigma(\mathfrak{t})}}\, {}_{0}^{c}D_{\mathfrak{t}}^{\varsigma(\mathfrak{t})}\nu(\mathfrak{t}) + F(\mathfrak{t}, \nu(\mathfrak{t})) = g(\mathfrak{t}), \qquad 0 \leq \mathfrak{t} \leq 1, \tag{1}
$$

subject to the initial conditions

$$
\begin{cases} \nu(0) = \nu_0, \\ \nu'(0) = \nu_1, \end{cases} \tag{2}
$$

where $1 < \varrho(\mathfrak{t}) < 2$, $0 < \varsigma(\mathfrak{t}) < 1$. Also, ξ is a positive constant, $F(\mathfrak{t}, \nu(\mathfrak{t}))$, and $g(\mathfrak{t})$ are given continuous functions. The approximate solution of the classical fractional order of this equation has been investigated by Bernoulli polynomials in [33]. In this work, we use the $\mathcal{V}t$-$\mathcal{L}$ polynomials to solve the above form of this equation. The $\mathcal{V}t$-$\mathcal{L}$ polynomials were introduced for the first time in the year 2020 and used to solve fractional advection-dispersion equations [2]. Also, the uniform convergence and error bound of the $\mathcal{V}t$-$\mathcal{L}$ polynomials has been checked in [2]. Despite the accuracy of these polynomials in approximating functions, they have not been used in constructing numerical methods for solving various problems. In [27], a numerical method was presented with the operational matrix of $\mathcal{V}t$-$\mathcal{L}$ polynomials to solve the fractional Bagley–Torvik equation, initial value problems, and nonhomogenous multi-order fractional problems. Researchers in [19] used these polynomials to calculate the approximate solution of the multi-Pantograph delay problems with singularity and compare the computed results with the exact one. The capability and performance of the numerical scheme established upon the $\mathcal{V}t$-$\mathcal{L}$ polynomials to solve the coupled nonlinear $\mathcal{V}$-$\mathcal{O}$ fractional Ginzburg-Landau equations [15] encouraged us to use these polynomials to solve the $\mathcal{V}$-$\mathcal{O}$ fractional Lane–Emden problem. For this purpose, we compute the $\mathcal{V}$-$\mathcal{O}$ fractional derivative and the classical derivative operational matrices of the $\mathcal{V}t$-$\mathcal{L}$ polynomials. The solution $\nu(\mathfrak{t})$ is expanded in terms of these basis polynomials, and the linear/nonlinear equation is converted to the linear/ nonlinear algebraic equation system. The capability of the proposed scheme is obtained through several examples.

Numerical methods based on operational matrices have proven to be highly effective in solving mathematical problems, particularly those involving differential and integral equations. Unlike traditional methods that directly calculate derivatives and integrals, the operational matrix method utilizes the operational matrices of derivatives and integrals. This approach offers several advantages, such as reducing CPU time due to the sparsity of the matrices and the presence of zero elements.

The operational matrix method is a powerful numerical technique for solving differential equations due to its simplicity, efficiency, accuracy, versatility, flexibility, and numerical stability. However, it is important to be aware of its limitations, including discretization errors, limited applicability to complex geometries, computational requirements, convergence issues, and limited support for discontinuous solutions. By considering these factors, researchers and practitioners can employ the operational matrix method effectively and make informed decisions regarding its application in various scientific and engineering fields.

The outline of the work is as follows: Several characteristics and concepts about $\mathcal{V}$-$\mathcal{O}$ fractional derivative in Caputo form are mentioned in section 2. The $\mathcal{V}t$-$\mathcal{L}$ polynomials are introduced in section 3. The formulation of the presented approach is explained in section 4. Several examples of linear and nonlinear types of the problem under study are examined in section 5. The conclusion of this work is briefly expressed in section 6.

2 $\mathcal{V}$-$\mathcal{O}$ fractional calculus

In this part, we introduce the indispensable relations and definitions of $\mathcal{V}$-$\mathcal{O}$ fractional calculus, which are required in our work.

Definition 1. [30] Let μ and $\zeta > 0$. The generalized Mittag-Leffler function is as follows:

$$\mathbf{E}_{\mu,\zeta}(\mathfrak{t}) = \sum_{i=0}^{\infty} \frac{\mathfrak{t}^i}{\Gamma(\mu i + \zeta)}, \tag{3}$$

where $\mathfrak{t} \in \mathbb{C}$.

Definition 2. [16] For a continuous function $\theta : \mathbb{R}^+ \cup \{0\} \longrightarrow (\hat{n}-1, \hat{n}]$ with $\hat{n} \in \mathbb{N}$, the $\mathcal{V}$-$\mathcal{O}$ fractional derivative of the function $\nu(\mathfrak{t})$ in the Caputo form is determined as

$$
{}_{0}^{c}D_{\mathfrak{t}}^{\theta(\mathfrak{t})}\nu(\mathfrak{t}) = \begin{cases} \dfrac{1}{\Gamma(\hat{n}-\theta(\mathfrak{t}))}\displaystyle\int_0^{\mathfrak{t}} (\mathfrak{t}-\rho)^{\hat{n}-\theta(\mathfrak{t})-1}\frac{d^{\hat{n}}\nu(\rho)}{d\rho^{\hat{n}}}d\rho, & \theta(\mathfrak{t}) \in (\hat{n}-1,\hat{n}), \\ \dfrac{d^{\hat{n}}\nu(\mathfrak{t})}{d\mathfrak{t}^{\hat{n}}}, & \theta(\mathfrak{t}) = \hat{n}. \end{cases} \tag{4}
$$

Lemma 1. Let the assumptions of the above definition be satisfied. Then, we have

$$
{}_{0}^{c}D_{\mathfrak{t}}^{\theta(\mathfrak{t})}\mathfrak{t}^{\mathbf{k}} = \begin{cases} 0, & \mathbf{k} = 0,1,\ldots,\hat{n}-1, \\ \dfrac{\Gamma(\mathbf{k}+1)}{\Gamma(\mathbf{k}+1-\theta(\mathfrak{t}))}\mathfrak{t}^{\mathbf{k}-\theta(\mathfrak{t})}, & \mathbf{k} \geq \hat{n}. \end{cases} \tag{5}
$$

3 The $\mathcal{V}t$-$\mathcal{L}$ polynomials

This section is dedicated to introduce the $\mathcal{V}t$-$\mathcal{L}$ polynomials and some of their properties.

Definition 3. [31, 15] The $\mathcal{V}t$-$\mathcal{L}$ polynomials are defined over $[0,1]$ as

$$
\mathcal{V}_j^*(\mathfrak{t}) = \begin{cases} 2, & j = 0, \\ \sum\limits_{l=0}^{j}(-1)^{j-l}\dfrac{2^{2l+1}j(j+l-1)!}{(2l)!(j-l)!}\mathfrak{t}^l, & j \geq 1. \end{cases} \tag{6}
$$

The orthogonal property of these polynomials is satisfied, which means that

$$
\int_0^1 \mathcal{V}_j^*(\mathfrak{t})\mathcal{V}_{\hat{j}}^*(\mathfrak{t})\varpi(\mathfrak{t})d\mathfrak{t} = \begin{cases} 4\pi, & j = \hat{j} = 0, \\ 2\pi, & j = \hat{j} \neq 0, \\ 0, & j \neq \hat{j}, \end{cases} \tag{7}
$$

where $\varpi(\mathfrak{t}) = \dfrac{1}{\sqrt{\mathfrak{t}-\mathfrak{t}^2}}$. Any function $\nu(\mathfrak{t}) \in L^2_{\varpi}[0,1]$ can be expanded with the $\mathcal{V}t$-$\mathcal{L}$ polynomials as

$$
\nu(\mathfrak{t}) \simeq \nu_N(\mathfrak{t}) = \sum_{j=0}^{N} c_j\mathcal{V}_j^*(\mathfrak{t}) \triangleq \mathbf{C}^T\Upsilon(\mathfrak{t}), \tag{8}
$$

where

$$
\begin{aligned} \mathbf{C} &= [\hat{c}_0 \quad \hat{c}_1 \quad \cdots \quad \hat{c}_N]^T, \\ \Upsilon(\mathfrak{t}) &= [\mathcal{V}_0^*(\mathfrak{t}) \quad \mathcal{V}_1^*(\mathfrak{t}) \quad \cdots \quad \mathcal{V}_N^*(\mathfrak{t})]^T, \end{aligned} \tag{9}
$$

and

$$\hat{c}_j = \frac{1}{\lambda_j}\int_0^1 \nu(\mathfrak{t})\mathcal{V}_j^*(\mathfrak{t})\varpi(\mathfrak{t})d\mathfrak{t}, \qquad j = 0, 1, \dots, N, \tag{10}$$

with

$$\lambda_j = \int_0^1 \left(\mathcal{V}_j^*(\mathfrak{t})\right)^2 \varpi(\mathfrak{t})d\mathfrak{t} = \begin{cases} 4\pi, & j = 0, \\ 2\pi, & j = 1, 2, \dots, N. \end{cases} \tag{11}$$

The following theorem demonstrates the uniform convergence of the $\mathcal{V}t$-$\mathcal{L}$ polynomials and its error bound for approximating the function $\nu(\mathfrak{t})$.

Theorem 1. [2] Assume that $\nu(\mathfrak{t}) \in L^2_{\varpi}$ with $\varpi(\mathfrak{t})$ and $|\nu''(\mathfrak{t})| < K$ such that K is a positive constant. Then, $\nu_N(\mathfrak{t}) \longrightarrow \nu(\mathfrak{t})$ as $N \longrightarrow \infty$. Moreover, the coefficients in relation (8) satisfy

$$|\hat{c}_j| \leq \frac{K}{4j(j^2-1)}, \qquad j > 2. \tag{12}$$

Also, the error bound will be obtained as

$$\|\nu(\mathfrak{t}) - \nu_N(\mathfrak{t})\|_{L^2_{\varpi}} < \frac{K}{12N^{\frac{3}{2}}}. \tag{13}$$

In the next theorem, the ordinary derivative matrix and the $\mathcal{V}$-$\mathcal{O}$ fractional of the $\mathcal{V}t$-$\mathcal{L}$ polynomials are derived.

Theorem 2. The differentiation of the vector $\Upsilon(\mathfrak{t})$ in (9) satisfies the following relation

$$\frac{d\Upsilon(\mathfrak{t})}{d\mathfrak{t}} \simeq \mathcal{D}^{(1)}\Upsilon(\mathfrak{t}), \tag{14}$$

in which $\mathcal{D}^{(1)}$ is an $(N+1)\times(N+1)$ matrix which its entries is computed by

$$[\mathcal{D}^{(1)}]_{ij} = \begin{cases} 0, & i = 1, \\ \varphi_{ij}, & i = 2, 3, \dots, N+1, \end{cases} \tag{15}$$

in which

$$\varphi_{ij} = \begin{cases} \dfrac{1}{4\pi}\displaystyle\sum_{l=1}^{i-1}(-1)^{i-l-1}\frac{2^{2(l+1)}(i-1)(i+l-2)!l}{(2l)!(i-l-1)!}\frac{\sqrt{\pi}\Gamma(l-\frac{1}{2})}{\Gamma(l)}, & j=1, \\ \dfrac{1}{2\pi}\displaystyle\sum_{l=1}^{i-1}\sum_{s=0}^{j-1}(-1)^{i+j-l-s-2}\left(\frac{2^{2(l+s+1)}(i-1)(j-1)(i+l-2)!(j+s-2)!l}{(2l)!(2s)!(i-l-1)!(j-s-1)!}\right. \\ \qquad\left.\dfrac{\sqrt{\pi}\Gamma(l+s-\frac{1}{2})}{\Gamma(l+s)}\right), & j=2,3,\dots,N+1. \end{cases} \tag{16}$$

Proof. For $\hat{i}=0$, the proof is obvious. For $\hat{i}=1,\dots,N$ from (6), we get

$$\frac{\mathcal{V}^*_{\hat{i}}(\mathfrak{t})}{d\mathfrak{t}} = \sum_{l=1}^{\hat{i}}(-1)^{\hat{i}-l}\frac{2^{2l+1}\hat{i}(\hat{i}+l-1)!l}{(2l)!(\hat{i}-l)!}\mathfrak{t}^{l-1}. \tag{17}$$

Approximating the above relation with the $\mathcal{V}t$-$\mathcal{L}$ polynomials results in

$$\frac{\mathcal{V}^*_{\hat{i}}(\mathfrak{t})}{d\mathfrak{t}} \simeq \sum_{\hat{j}=0}^{N}\mathcal{D}^{(1)}_{\hat{i}\hat{j}}\mathcal{V}^*_{\hat{j}}(\mathfrak{t}), \qquad \hat{i}=1,2,\dots,N, \tag{18}$$

in which, for $\hat{j}=0$, one obtains

$$\begin{aligned} \mathcal{D}^{(1)}_{10} &= \frac{1}{\lambda_0}\int_0^1 \frac{\mathcal{V}^*_{\hat{i}}(\mathfrak{t})}{d\mathfrak{t}}\mathcal{V}^*_0(\mathfrak{t})\varpi(\mathfrak{t})d\mathfrak{t} \\ &= \frac{1}{4\pi}\sum_{l=1}^{\hat{i}}(-1)^{i-l}\frac{2^{2l+1}\hat{i}(\hat{i}+l-1)!l}{(2l)!(\hat{i}-l)!}\int_0^1\frac{2\mathfrak{t}^{l-1}}{\sqrt{\mathfrak{t}-\mathfrak{t}^2}}d\mathfrak{t} \\ &= \frac{1}{4\pi}\sum_{l=1}^{\hat{i}}(-1)^{\hat{i}-l}\frac{2^{2(l+1)}\hat{i}(\hat{i}+l-1)!l}{(2l)!(\hat{i}-l)!}\frac{\sqrt{\pi}\Gamma(l-\frac{1}{2})}{\Gamma(l)}, \end{aligned} \tag{19}$$

and for $\hat{j}=1,2,\dots,N$,

$$\begin{aligned} \mathcal{D}^{(1)}_{\hat{i}\hat{j}} &= \frac{1}{\lambda_j}\int_0^1\frac{\mathcal{V}^*_{\hat{i}}(\mathfrak{t})}{d\mathfrak{t}}\mathcal{V}^*_{\hat{j}}(\mathfrak{t})\varpi(\mathfrak{t})d\mathfrak{t} \\ &= \frac{1}{2\pi}\sum_{l=1}^{\hat{i}}\sum_{s=0}^{\hat{j}}\left((-1)^{\hat{i}+\hat{j}-l-s}\frac{2^{2(l+s+1)}\hat{i}\hat{j}(\hat{i}+l-1)!(\hat{j}+s-1)!l}{(2l)!(2s)!(\hat{i}-l)!(\hat{j}-s)!}\right)\int_0^1\frac{\mathfrak{t}^{l+s-1}}{\sqrt{\mathfrak{t}-\mathfrak{t}^2}}d\mathfrak{t} \\ &= \frac{1}{2\pi}\sum_{l=1}^{\hat{i}}\sum_{s=0}^{\hat{j}}\left((-1)^{\hat{i}+\hat{j}-l-s}\frac{2^{2(l+s+1)}\hat{i}\hat{j}(\hat{i}+l-1)!(\hat{j}+s-1)!l}{(2l)!(2s)!(\hat{i}-l)!(\hat{j}-s)!}\right)\frac{\sqrt{\pi}\Gamma(l+s-\frac{1}{2})}{\Gamma(k+s)}. \end{aligned} \tag{20}$$

Therefore, by replacing $\hat{i}=i-1$ and $\hat{j}=j-1$, the proof is completed. □

For example, for $N = 5$, one can get

$$\mathcal{D}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 & 0 \\ 6 & 0 & 12 & 0 & 0 & 0 \\ 0 & 16 & 0 & 16 & 0 & 0 \\ 10 & 0 & 20 & 0 & 20 & 0 \end{bmatrix}.$$

Lemma 2. The vector $\Upsilon(\mathfrak{t})$ in (9) can be rewritten by

$$\Upsilon(\mathfrak{t}) = \mathfrak{R}\dot{\mathbf{X}}, \tag{21}$$

where

$$\dot{\mathbf{X}} = [1 \quad \mathfrak{t} \quad \mathfrak{t}^2 \quad \cdots \quad \mathfrak{t}^n]^T, \tag{22}$$

and $\mathfrak{R}$ is an $(n+1)$ square upper triangular matrix and

$$[\mathfrak{R}]_{ij} = \begin{cases} 2, & i = j = 1, \\ \mathfrak{u}_{ij}, & j \leq i, \end{cases} \tag{23}$$

in which

$$\mathfrak{u}_{ij} = \frac{(-1)^{i-j} 2^{2j+1} i(i+j-1)!}{(2j)!(i-j)!}. \tag{24}$$

Proof. By considering (6) and (9), one obtains

$$\begin{aligned} \Upsilon(\mathfrak{t}) &= [\mathcal{V}_0^*(\mathfrak{t}) \quad \mathcal{V}_1^*(\mathfrak{t}) \quad \mathcal{V}_2^*(\mathfrak{t}) \quad \cdots \quad \mathcal{V}_N^*(\mathfrak{t})]^T \\ &= [2 \quad \mathfrak{u}_{10} + \mathfrak{u}_{11}\mathfrak{t} \quad \mathfrak{u}_{20} + \mathfrak{u}_{21}\mathfrak{t} + \mathfrak{u}_{22}\mathfrak{t}^2 \quad \cdots \quad \mathfrak{u}_{n0} + \mathfrak{u}_{n1}\mathfrak{t} + \ldots + \mathfrak{u}_{nn}\mathfrak{t}^n]^T \\ &= \begin{bmatrix} 2 & 0 & 0 & 0 & \cdots & 0 \\ \mathfrak{u}_{10} & \mathfrak{u}_{11} & 0 & 0 & \cdots & 0 \\ \mathfrak{u}_{20} & \mathfrak{u}_{21} & \mathfrak{u}_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathfrak{u}_{n0} & \mathfrak{u}_{n1} & \mathfrak{u}_{n2} & \mathfrak{u}_{n3} & \cdots & \mathfrak{u}_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ \mathfrak{t} \\ \mathfrak{t}^2 \\ \vdots \\ \mathfrak{t}^n \end{bmatrix}, \end{aligned} \tag{25}$$

in which $\mathfrak{u}_{ij}$ is defined in (24). □

Lemma 3. Let $\dot{\mathbf{X}}$ be the vector defined in (22). Also, let $\mathbf{q} - 1 < \sigma(\mathfrak{t}) < \mathbf{q}$ be a continuous function given in $[0, 1]$. Then, the $\mathcal{V}$-$\mathcal{O}$ fractional derivative of $\dot{\mathbf{X}}$ is obtained as

$$ {}_0^c D_{\mathrm{t}}^{\sigma(\mathrm{t})} \dot{\mathbf{X}} = \mathbf{Q}^{\sigma(\mathrm{t})} \dot{\mathbf{X}}, \tag{26} $$

where $\mathbf{Q}^{\sigma(\mathrm{t})}$ is the square $(n+1)$-matrix whose first $\mathbf{q}$ columns are zero and

$$ \mathbf{Q}^{\sigma(\mathrm{t})} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \frac{\Gamma(\mathbf{q}+1)}{\Gamma(\mathbf{q}+1-\sigma(\mathrm{t}))}\mathrm{t}^{-\xi} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \frac{\Gamma(\mathbf{q}+2)}{\Gamma(\mathbf{q}+2-\sigma(\mathrm{t}))}\mathrm{t}^{-\xi} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \frac{\Gamma(n+1)}{\Gamma(n+1-\sigma(\mathrm{t}))}\mathrm{t}^{-\xi} \end{bmatrix}. \tag{27} $$

Proof. According to Lemma 1, the proof is obvious. □

Now, we compute the operational matrix of $\mathcal{V}$-$\mathcal{O}$ fractional derivative of the $\mathcal{V}t$-$\mathcal{L}$ polynomials.

Theorem 3. Let $\Upsilon(\mathrm{t})$ be the $\mathcal{V}t$-$\mathcal{L}$ polynomials vector defined in (9) and let $\sigma(\mathrm{t})$ be a function introduced in Lemma 3. Then, the $\mathcal{V}$-$\mathcal{O}$ fractional derivative of order $\sigma(\mathrm{t})$ is computed as follows

$$ {}_0^c D_{\mathrm{t}}^{\sigma(\mathrm{t})} \Upsilon(\mathrm{t}) = \mathfrak{D}^{\sigma(\mathrm{t})} \Upsilon(\mathrm{t}), \tag{28} $$

where $\mathfrak{D}^{\sigma(\mathrm{t})} = \mathfrak{R}\mathbf{Q}^{\sigma(\mathrm{t})}\mathfrak{R}^{-1}$ is the matrix of $\sigma(\mathrm{t})$ $\mathcal{V}$-$\mathcal{O}$ fractional derivative for the $\mathcal{V}t$-$\mathcal{L}$ polynomials.

Proof. Using Lemmas 2 and 3, one can get

$$ \begin{aligned} {}_0^c D_{\mathrm{t}}^{\sigma(\mathrm{t})} \Upsilon(\mathrm{t}) &= {}_0^c D_{\mathrm{t}}^{\sigma(\mathrm{t})} \big(\mathfrak{R}\dot{\mathbf{X}}\big) = \mathfrak{R}\, {}_0^c D_{\mathrm{t}}^{\sigma(\mathrm{t})} \dot{\mathbf{X}} \\ &= \mathfrak{R}\mathbf{Q}^{\sigma(\mathrm{t})}\dot{\mathbf{X}} = \big(\mathfrak{R}\mathbf{Q}^{\sigma(\mathrm{t})}\mathfrak{R}^{-1}\big)\Upsilon(\mathrm{t}) = \mathfrak{D}^{\sigma(\mathrm{t})}\Upsilon(\mathrm{t}), \end{aligned} \tag{29} $$

and the proof is performed. □

4 Description of the scheme

In this part, we state a spectral method via the $\mathcal{V}t$-$\mathcal{L}$ polynomials for the approximate solution of the $\mathcal{V}$-$\mathcal{O}$ fractional Lane–Emden equation expressed in (1). For this purpose, assume that

$$\nu(\mathfrak{t}) \simeq \nu_N(\mathfrak{t}) = \sum_{j=0}^{N} c_j \mathcal{V}_j^*(\mathfrak{t}) \triangleq \mathbf{C}^T \Upsilon(\mathfrak{t}), \tag{30}$$

where $\mathbf{C}$ is a vector with unknown elements and $\Upsilon(\mathfrak{t})$ is defined in relation (9). Using Theorem 3 implies that

$$ {}_0^c D_{\mathfrak{t}}^{\varrho(\mathfrak{t})} \nu(\mathfrak{t}) \simeq \mathbf{C}^T \mathfrak{D}^{\varrho(\mathfrak{t})} \Upsilon(\mathfrak{t}), \tag{31}$$

and

$$ {}_0^c D_{\mathfrak{t}}^{\varsigma(\mathfrak{t})} \nu(\mathfrak{t}) \simeq \mathbf{C}^T \mathfrak{D}^{\varsigma(\mathfrak{t})} \Upsilon(\mathfrak{t}). \tag{32}$$

Substituting (30)–(32) into (1) results in

$$\mathbf{C}^T \mathfrak{D}^{\varrho(\mathfrak{t})} \Upsilon(\mathfrak{t}) + \frac{\mu}{\mathfrak{t}^{\varrho(\mathfrak{t})-\varsigma(\mathfrak{t})}} \mathbf{C}^T \mathfrak{D}^{\varsigma(\mathfrak{t})} \Upsilon(\mathfrak{t}) + F(\mathfrak{t}, \mathbf{C}^T \Upsilon(\mathfrak{t})) - g(\mathfrak{t}) \triangleq \mathbf{R}(\mathfrak{t}). \tag{33}$$

Moreover, from (2), (30), and Theorem 2, we get

$$\begin{cases} \Lambda_0 \triangleq \mathbf{C}^T \Upsilon(0) - \nu_0 \simeq 0, \\ \Lambda_1 \triangleq \mathbf{C}^T \mathcal{D}^{(1)} \Upsilon(0) - \nu_1 \simeq 0. \end{cases} \tag{34}$$

Eventually, one obtains a numerical solution for (1) by solving relation (33) with initial conditions (34) in the collocation points $\mathfrak{t}_{\bar{i}} = \dfrac{2\bar{i}-1}{2(N+1)}$ with $\bar{i} = 1, 2, \ldots, N-1$, and through substituting in (30). The program of the proposed numerical method in pseudo-code format is designed in Algorithm 1. It is crucial to emphasize that the solution to the system of equations (33) and (34) is achieved by utilizing the "fsolve" command within the Maple 18 software.

Algorithm 1: The proposed method algorithm

Inputs: $N > 0$; $\varsigma(\mathfrak{t}) \in (0,1]$; $\varrho(\mathfrak{t}) \in (1,2]$; g; ν_0; ν_1.

Step 1: Define the functions $\mathcal{V}_j^*(\mathfrak{t})$ via (6).

Step 2: Construct the vector $\Upsilon(\mathfrak{t})$ and the matrix $\mathcal{D}^{(1)}$ by (9) and Theorem 2, respectively.

Step 3: Make of the matrices $\mathfrak{D}^{\varrho(\mathfrak{t})}$ and $\mathfrak{D}^{\varsigma(\mathfrak{t})}$ using Theorem (3).

Step 4: Construct the vector $\mathbf{C}$ in (30).

Step 5: Define $\mathbf{R}(\mathfrak{t})$, Λ_0, and Λ_1 in (33) and (34).

Step 6: Extract the algebraic system by the collocation points x_i.

Step 7: Solve the obtained system and compute the vector $\mathbf{C}$.

Outputs: The numerical solution $\nu_N(\mathfrak{t})$.

5 Numerical results

Here, four numerical examples are proposed to examine the accuracy of presented technique. To do this, we use the maximum absolute error as

$$L_\infty = \max_{0 \le x \le 1} |\nu(\mathfrak{t}) - \nu_N(\mathfrak{t})|, \tag{35}$$

where $\nu(\mathfrak{t})$ and $\nu_N(\mathfrak{t})$ are the analytical and numerical solution with $\mathcal{V}t$-$\mathcal{L}$ polynomials, respectively. Also, the calculations are carried out by Maple 17 with 25 digits.

Example 1. For the first example, consider the problem

$$ {}_0^cD_{\mathfrak{t}}^{\varrho(\mathfrak{t})}\nu(\mathfrak{t}) + \frac{1}{\mathfrak{t}^{\varrho(\mathfrak{t})-\varsigma(\mathfrak{t})}}\, {}_0^cD_{\mathfrak{t}}^{\varsigma(\mathfrak{t})}\nu(\mathfrak{t}) + e^{\nu(\mathfrak{t})} = g(\mathfrak{t}), \tag{36}$$

where

$$g(\mathfrak{t}) = 2\mathfrak{t}^{2-\varrho(\mathfrak{t})}\left(\frac{\Gamma(3-\varrho(\mathfrak{t})) + \Gamma(3-\varsigma(\mathfrak{t}))}{\Gamma(3-\varsigma(\mathfrak{t}))\Gamma(3-\varrho(\mathfrak{t}))}\right) + \frac{\mathfrak{t}^{1-\varrho(\mathfrak{t})}}{\Gamma(2-\varsigma(\mathfrak{t}))} + e^{\mathfrak{t}^2+\mathfrak{t}}, \tag{37}$$

subject to the initial conditions $\nu(0) = \nu'(0) = 0$. The analytic solution in the classical form of this problem is $\nu(\mathfrak{t}) = \mathfrak{t}^2 + \mathfrak{t}$. We implemented the presented approach with $N = 2$ for the approximate solution of this model for $\varrho(\mathfrak{t}) = \dfrac{3}{2}$ and $\varsigma(\mathfrak{t}) = \dfrac{3}{4}$. So, the approximate solution is derived as follows:

$$\begin{aligned}\nu_2(\mathfrak{t}) =& 1 \times 10^{-25} + 0.9999999999999999999999996\mathfrak{t} \\ &+ 0.9999999999999999999999994\mathfrak{t}^2. \end{aligned} \tag{38}$$

Therefore, the numerical solution is equal to the analytical. While, the maximum error reported in [33] by the Bernoulli polynomials with $N = 6$ is $1.02487E - 13$. This means that the presented scheme is more efficient and superior with respect to the method of [33] for solving this example.

Example 2. Suppose the Lane–Emden equation in the form

$$ {}_0^c D_{\mathfrak{t}}^{\varrho(\mathfrak{t})}\nu(\mathfrak{t}) + \frac{2}{\mathfrak{t}^{\varrho(\mathfrak{t})-\varsigma(\mathfrak{t})}}\, {}_0^c D_{\mathfrak{t}}^{\varsigma(\mathfrak{t})}\nu(\mathfrak{t}) - (\nu(\mathfrak{t}))^2 = g(\mathfrak{t}), \tag{39}$$

where

$$g(\mathfrak{t}) = \mathfrak{t}^{1-\varrho(\mathfrak{t})}\Big(\mathbf{E}_{2,2-\varrho(\mathfrak{t})}(-\mathfrak{t}^2) - \frac{1}{\Gamma(2-\varrho(\mathfrak{t}))}\Big) + 2\mathfrak{t}^{1-\varrho(\mathfrak{t})}\mathbf{E}_{2,2-\varsigma(\mathfrak{t})}(-\mathfrak{t}^2) - \sin(\mathfrak{t})^2, \tag{40}$$

in which, we apply the first 20 terms of the Mittag-Leffler series in our computational. The approximation also used in what follows in this paper. The exact solution is $\nu(\mathfrak{t}) = \sin(\mathfrak{t})$. So, $\nu(0) = 0$ and $\nu'(0) = 1$ are the initial conditions. We apply the proposed technique to compute the numerical solution for this example with different values of N for three selections $\varsigma_1(\mathfrak{t}) = 0.8 - 0.1\sin(\mathfrak{t})$, $\varsigma_2(\mathfrak{t}) = 0.8 - 0.3\sin(\mathfrak{t})$ and $\varsigma_3(\mathfrak{t}) = 0.8 - 0.5\sin(\mathfrak{t})$. Also, three selections $\varrho_1(\mathfrak{t}) = 1.8 - 0.35\exp(-\mathfrak{t}^2)$, $\varrho_2(\mathfrak{t}) = 1.8 - 0.55\exp(-\mathfrak{t}^2)$ and $\varrho_3(\mathfrak{t}) = 1.8 - 0.75\exp(-\mathfrak{t}^2)$ are used.

Table 1: The L_∞ errors in Example 2 with different values of N.

	$\varrho_1(\mathfrak{t})$			$\varrho_2(\mathfrak{t})$			$\varrho_3(\mathfrak{t})$		
N	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$
5	$4.7624E-05$	$4.7416E-05$	$4.7158E-05$	$5.0958E-05$	$5.0910E-05$	$5.0768E-05$	$5.3846E-05$	$5.4017E-05$	$5.4052E-05$
7	$1.7751E-07$	$1.7600E-07$	$1.7447E-07$	$1.9490E-07$	$1.9388E-07$	$1.9257E-07$	$2.1095E-07$	$2.1087E-07$	$2.1074E-07$
9	$3.7149E-10$	$3.6748E-10$	$3.6374E-10$	$4.1514E-10$	$4.1169E-10$	$4.0797E-10$	$4.5710E-10$	$4.5546E-10$	$4.5266E-10$
11	$4.9828E-13$	$4.9226E-13$	$4.8696E-13$	$5.6396E-13$	$5.5812E-13$	$5.5235E-13$	$6.2898E-13$	$6.2516E-13$	$6.2014E-13$
13	$4.6498E-16$	$4.5905E-16$	$4.5403E-16$	$5.3139E-16$	$5.2516E-16$	$5.1934E-16$	$5.9857E-16$	$5.9382E-16$	$5.8832E-16$

Table 1 reports the obtained results for this example. Column one of this table is the number of the $\mathcal{V}t$-$\mathcal{L}$ polynomials (N), which increases from top to bottom and other columns are L_∞ errors between the present method and the exact solution for different values of $\varsigma(\mathfrak{t})$ and $\varrho(\mathfrak{t})$. With increases in N, one can observe that the error decreases, which indicates the capability of the present technique. The results of this table show that our scheme is reliable for different values of $\mathcal{V}$-$\mathcal{O}$s.

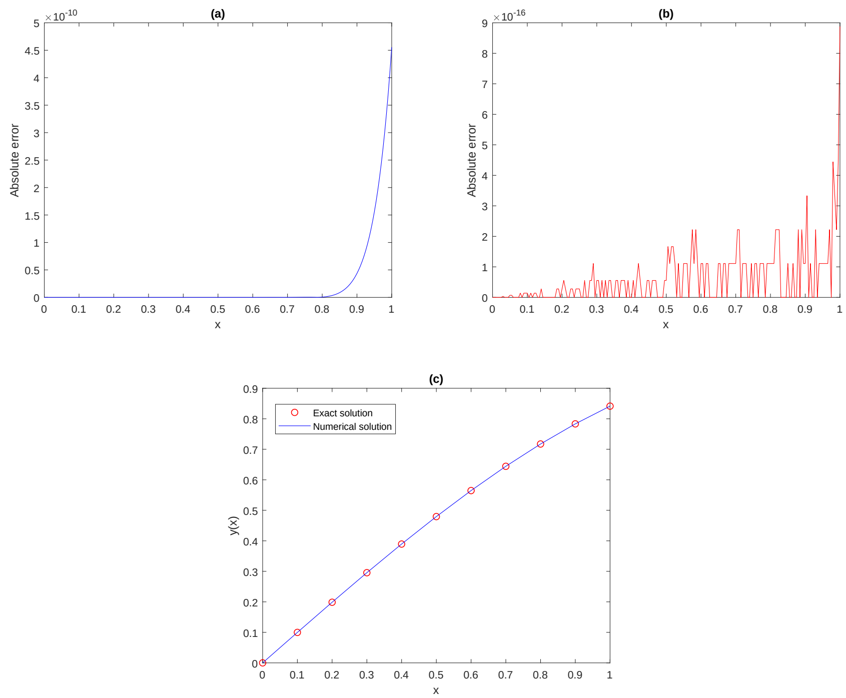


Figure 1: The absolute error of $\nu(\mathfrak{t})$ in Example 2 for $\varsigma_3(\mathfrak{t})$ and $\varrho_3(\mathfrak{t})$ with two values of N, (a) $N = 9$, (b) $N = 13$, and (c) the exact solution and numerical solution $\nu_{11}(\mathfrak{t})$.

Figure 1 (a) and (b) demonstrates the absolute errors for $\varsigma_3(\mathfrak{t})$ and $\varrho_3(\mathfrak{t})$ for different selections of N. Also, Figure 1(c) illustrates the numerical and exact solutions for this problem.

From the results of Table 1 and Figure 1, one can observe that the presented approach is very reliable and capable of calculating the approximate solution of this problem.

Example 3. In this example, we consider the Lane–Emden model as

$$ {}_0^cD_{\mathfrak{t}}^{\varrho(\mathfrak{t})}\nu(\mathfrak{t}) + \frac{4}{\mathfrak{t}^{\varrho(\mathfrak{t})-\varsigma(\mathfrak{t})}}\,{}_0^cD_{\mathfrak{t}}^{\varsigma(\mathfrak{t})}\nu(\mathfrak{t}) + \mathfrak{t}^2\nu(\mathfrak{t}) = g(\mathfrak{t}), \tag{41} $$

subject to the following initial conditions

$$\begin{cases} \nu(0) = 1, \\ \nu'(0) = -1, \end{cases}$$

with

$$\begin{aligned} g(\mathfrak{t}) =& \mathfrak{t}^{-\varrho(\mathfrak{t})}\Big(\mathbf{E}_{1,1-\varrho(\mathfrak{t})}(-\mathfrak{t}) - \frac{1}{\Gamma(1-\varrho(\mathfrak{t}))} + \frac{\mathfrak{t}}{\Gamma(2-\varrho(\mathfrak{t}))}\Big) \\ &+ 4\mathfrak{t}^{-\varrho(\mathfrak{t})}\Big(\mathbf{E}_{1,1-\varsigma(\mathfrak{t})}(-\mathfrak{t}) - \frac{1}{\Gamma(1-\varsigma(\mathfrak{t}))}\Big) + \mathfrak{t}^2\exp(-\mathfrak{t}). \end{aligned} \tag{42}$$

The analytical solution is $\nu(\mathfrak{t}) = \exp(-\mathfrak{t})$. Similar to the previous example, we use the following different selections for the $\mathcal{V}$-$\mathcal{O}$s $\varsigma(\mathfrak{t})$ and $\varrho(\mathfrak{t})$

$$\begin{aligned} &\varsigma_1(\mathfrak{t}) = 0.35 + 0.2\exp(-\mathfrak{t}), \qquad \varsigma_2(\mathfrak{t}) = 0.35 + 0.4\exp(-\mathfrak{t}), \\ &\varsigma_3(\mathfrak{t}) = 0.35 + 0.6\exp(-\mathfrak{t}), \qquad \varrho_1(\mathfrak{t}) = 1.8 - 0.25\cos(\mathfrak{t}), \\ &\varrho_2(\mathfrak{t}) = 1.8 - 0.45\cos(\mathfrak{t}), \qquad \varrho_3(\mathfrak{t}) = 1.8 - 0.65\cos(\mathfrak{t}), \end{aligned}$$

to test our method for approximating the solution of this example.

Table 2: The L_∞ errors in Example 3 with different values of N.

	$\varrho_1(\mathfrak{t})$			$\varrho_2(\mathfrak{t})$			$\varrho_3(\mathfrak{t})$		
N	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$
6	$5.6134E-06$	$5.6705E-06$	$5.6751E-06$	$6.4335E-06$	$6.4244E-06$	$6.3466E-06$	$7.2120E-06$	$7.1085E-06$	$6.9261E-06$
8	$1.4038E-08$	$1.4277E-08$	$1.4400E-08$	$1.6502E-08$	$1.6607E-08$	$1.6531E-08$	$1.9013E-08$	$1.8871E-08$	$1.8493E-08$
10	$2.1893E-11$	$2.2348E-11$	$2.2649E-11$	$2.6141E-11$	$2.6443E-11$	$2.6469E-11$	$3.0683E-11$	$3.0620E-11$	$3.0157E-11$
12	$2.3300E-14$	$2.3834E-14$	$2.4229E-14$	$2.8115E-14$	$2.8541E-14$	$2.8684E-14$	$3.3442E-14$	$3.3513E-14$	$3.3141E-14$
14	$1.7999E-17$	$1.8437E-17$	$1.8778E-17$	$2.1870E-17$	$2.2270E-17$	$2.2457E-17$	$2.6282E-17$	$2.6419E-17$	$2.6210E-17$

The L_∞ errors between the approximate and exact solution are provided in Table 2. The first column of this table is the various values of N from 6 to 14. The next columns are the errors for different selections of $\varsigma(\mathfrak{t})$ and $\varrho(\mathfrak{t})$. It is evident that augmenting the polynomial bases leads to a reduction in error for each $\varsigma(\mathfrak{t})$ and $\varrho(\mathfrak{t})$ value. The convergence of the outcomes presented in this table is readily discernible.

The diagram of the numerical solution and its absolute error is depicted in Figure 2 (a) and (b). Also, the numerical and exact solutions for $\varsigma_3(\mathfrak{t})$ and $\varrho_3(\mathfrak{t})$ are depicted in Figure 2 (c).

Table 2 and Figure 2 confirm that the efficiency of the obtained numerical results is improved by increasing the values of N. These results show the capability of the presented technique for solving this example.

Example 4. In the end, we examine the following Lane–Emden model:

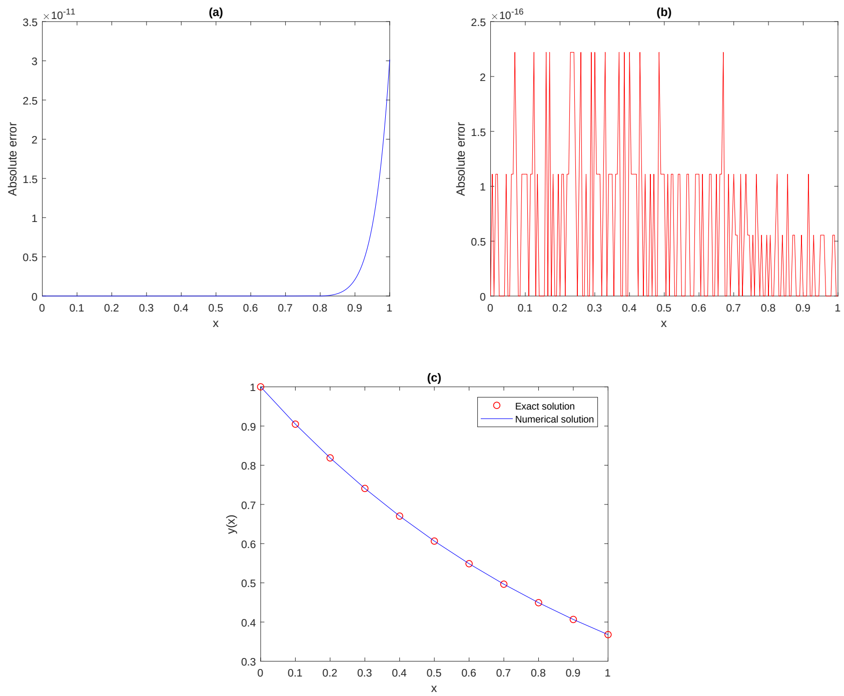


Figure 2: The absolute error of $\nu(\mathfrak{t})$ in Example 3 for $\varsigma_3(\mathfrak{t})$ and $\varrho_3(\mathfrak{t})$ with two values of N (a) $N = 10$, (b) $N = 14$, and (c) the analytical solution and numerical solutions $\nu_{12}(\mathfrak{t})$.

$$\begin{cases} {}_0^cD_{\mathfrak{t}}^{\varrho(\mathfrak{t})}\nu(\mathfrak{t}) + \dfrac{1}{\mathfrak{t}^{\varrho(\mathfrak{t})-\varsigma(\mathfrak{t})}}\,{}_0^cD_{\mathfrak{t}}^{\varsigma(\mathfrak{t})}\nu(\mathfrak{t}) - \dfrac{1}{\mathfrak{t}}\nu(\mathfrak{t}) = g(\mathfrak{t}), \\ \nu(0) = 0, \quad \nu'(0) = 1, \end{cases} \tag{43}$$

where

$$g(\mathfrak{t}) = \mathfrak{t}^{1-\varrho(\mathfrak{t})}\Big(\sum_{i=1}^{\infty}\frac{(2i+1)(-\mathfrak{t}^2)^i}{\Gamma(2i+2-\varrho(\mathfrak{t}))} + \sum_{i=0}^{\infty}\frac{(2i+1)(-\mathfrak{t}^2)^i}{\Gamma(2i+2-\varsigma(\mathfrak{t}))}\Big) + \cos(\mathfrak{t}). \tag{44}$$

The analytic solution is $\nu(\mathfrak{t}) = \mathfrak{t}\cos(\mathfrak{t})$ for any $0 < \varsigma(\mathfrak{t}) < 1$ and $1 < \varrho(\mathfrak{t}) < 2$. We apply the proposed approach for this problem with some different values of $\varsigma(\mathfrak{t})$ and $\varrho(\mathfrak{t})$ as follows:

$$\begin{aligned} &\varsigma_1(\mathfrak{t}) = 0.8 - 0.15\sin(\mathfrak{t})\cos(\mathfrak{t}), \qquad \varsigma_2(\mathfrak{t}) = 0.8 - 0.35\sin(\mathfrak{t})\cos(\mathfrak{t}), \\ &\varsigma_3(\mathfrak{t}) = 0.8 - 0.55\sin(\mathfrak{t})\cos(\mathfrak{t}), \end{aligned} \tag{45}$$

and

$$\varrho_1(\mathfrak{t}) = 0.55 + \frac{1}{1+\mathfrak{t}^2}, \qquad \varrho_2(\mathfrak{t}) = 0.75 + \frac{1}{1+\mathfrak{t}^2}, \qquad \varrho_3(\mathfrak{t}) = 0.95 + \frac{1}{1+\mathfrak{t}^2}. \tag{46}$$

We report the computed numerical results of the proposed scheme for some values of N in Table 3. The results obtained are compiled in Table 3, demonstrating that the outcomes improve as the values of N increase. These findings affirm that the numerical solutions converge toward the analytic solution. Figure 3 (a) and (b) depict the absolute errors for this problem with two values of N. Moreover, the analytical and approximate solutions are depicted in Figure 3 (c). The results of Figure 3 and Table 3 show that the presented technique accurately calculates the approximate solution for this example.

Table 3: The L_∞ errors in Example 4 with different values of N.

	$\varrho_1(\mathfrak{t})$			$\varrho_2(\mathfrak{t})$			$\varrho_3(\mathfrak{t})$		
N	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$	$\varsigma_1(\mathfrak{t})$	$\varsigma_2(\mathfrak{t})$	$\varsigma_3(\mathfrak{t})$
5	$3.3251E-04$	$3.3175E-04$	$3.3075E-04$	$3.0269E-04$	$3.0145E-04$	$3.0018E-04$	$2.7125E-04$	$2.6994E-04$	$2.6870E-04$
7	$1.6665E-06$	$1.6599E-06$	$1.6520E-06$	$1.4679E-06$	$1.4597E-06$	$1.4517E-06$	$1.2746E-06$	$1.2676E-06$	$1.2613E-06$
9	$4.4376E-09$	$4.4146E-09$	$4.3884E-09$	$3.8173E-09$	$3.7928E-09$	$3.7698E-09$	$3.2440E-09$	$3.2257E-09$	$3.2098E-09$
11	$7.2818E-12$	$7.2374E-12$	$7.1886E-12$	$6.1525E-12$	$6.1100E-12$	$6.0712E-12$	$5.1463E-12$	$5.1175E-12$	$5.0931E-12$
13	$8.0775E-15$	$8.0227E-15$	$7.9640E-15$	$6.7287E-15$	$6.6802E-15$	$6.6370E-15$	$5.5596E-15$	$5.5289E-15$	$5.5036E-15$

6 Conclusion

The $\mathcal{V}$-$\mathcal{O}$ fractional in the Caputo form was used to define the $\mathcal{V}$-$\mathcal{O}$ fractional Lane–Emden equation. So, the novelty of the article is the introduction of a novel equation type, which expands the existing knowledge base. The $\mathcal{V}t$-$\mathcal{L}$ polynomials were applied to solve this problem in linear and nonlinear cases. The presented method was based on these polynomials, under which the problem was transformed into an algebraic system of equations. Four examples were considered to show the capability of the obtained results, and they were compared to their analytical solution and the ones obtained by Bernoulli polynomials solutions in the first example. Through this comparison, the advantages and strengths of the constructed method were highlighted, showcasing its superior performance in terms of accuracy. The outcome confirmed the efficiency and performance of the proposed approach to solve the $\mathcal{V}$-$\mathcal{O}$ fractional Lane–Emden equation. Given the notable precision of $\mathcal{V}t$-$\mathcal{L}$ poly-

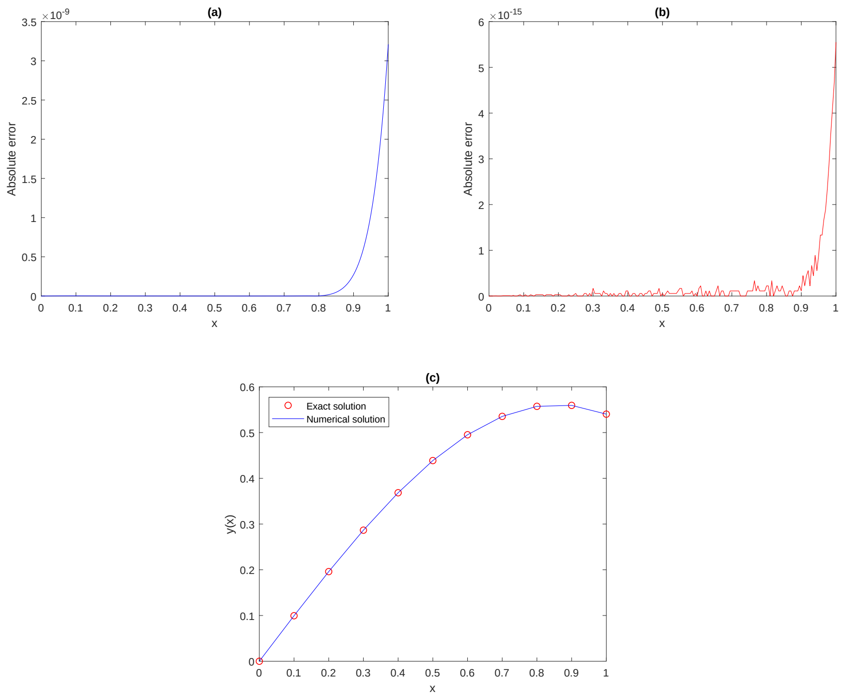


Figure 3: The absolute error of $\nu(\mathrm{t})$ in Example 4 for $\varsigma_3(\mathrm{t})$ and $\varrho_3(\mathrm{t})$ with two values of N (a) $N = 9$, (b) $N = 13$, and (c) the exact solution and numerical solution $\nu_{11}(\mathrm{t})$.

nomials in numerically solving the $\mathcal{V}$-$\mathcal{O}$ fractional Lane–Emden equation, future research should focus on several areas: extending the application of this derivative to various other equations and integrating the operator matrix method with other numerical methodologies.

References

[1] Abu Arqub, O., El-Ajou, A., Bataineh, A.S., and Hashim, I. *A representation of the exact solution of generalized Lane–Emden equations using a new analytical method*, In Abstr. Appl. Anal. (Vol. 2013, No. 1, p. 378593), Hindawi, 2013.

[2] Agarwal, P. and El-Sayed, A.A. *Vieta–Lucas polynomials for solving a*

fractional-order mathematical physics model, Adv. Diff. Equ. 2020(1) (2020) 1–18.

[3] Bu, W., Tang, Y., Wu, Y. and Yang, J. *Finite difference/finite element method for two-dimensional space and time fractional Bloch–Torrey equations*, J. Comput. Phys. 293 (2015) 264–279.

[4] Caglar, N. and Caglar, H. *B-spline solution of singular boundary value problems*, Appl. Math. Comput. 182(2) (2006) 1509–1513.

[5] Chambré, P.L. *On the solution of the Poisson-Boltzmann equation with application to the theory of thermal explosions*, J. Chem. Phys. 20(11) (1952) 1795–1797.

[6] Chandrasekhar, S. *Eddington: The most distinguished astrophysicist of his time*, Cambridge University Press, 1983.

[7] Choudhury, M.D., Chandra, S., Nag, S., Das, S. and Tarafdar, S. *Forced spreading and rheology of starch gel: Viscoelastic modeling with fractional calculus*, Colloids Surf. A: Physicochem. Eng. Asp. 407 (2012) 64–70.

[8] Coimbra, C.F.M. *Mechanics with variable-order differential operators*, Annal. Phys. 12(11-12) (2003) 692–703.

[9] Cooper, G.R.J. and Cowan, D.R. *Filtering using variable order vertical derivatives*, Comput. Geosci. 30(5) (2004) 455–459.

[10] Davis, H.T. *Introduction to nonlinear differential and integral equations*, US Government Printing Office, 1961.

[11] Dehghan, M. and Shakeri, F. *Approximate solution of a differential equation arising in astrophysics using the variational iteration method*, New Astron. 13(1) (2008) 53–59.

[12] Flandoli, F. and Tudor, C.A. *Brownian and fractional Brownian stochastic currents via Malliavin calculus*, J. Funct. Anal. 258(1) (2010) 279–306.

[13] Habibirad, A., Hesameddini, E., Heydari, M.H. and Roohi, R. *An efficient meshless method based on the moving Kriging interpolation for two-dimensional variable-order time fractional mobile/immobile advection-diffusion model*, Math. Method. Appl. Sci. 44(4) (2021) 3182–3194.

[14] He, J. H. and Ji, F. Y. *Taylor series solution for Lane–Emden equation,* J. Math. Chem. 57 (2019) 1932–1934.

[15] Heydari, M.H., Avazzadeh, Z. and Razzaghi, M. *Vieta–Lucas polynomials for the coupled nonlinear variable-order fractional Ginzburg–Landau equations*, Appl. Numer. Math. 165 (2021) 442–458.

[16] Heydari, M.H., Avazzadeh, Z. and Yang, Y. *A computational method for solving variable-order fractional nonlinear diffusion-wave equation*, Appl. Math. Comput. 352 (2019) 235–248.

[17] Heydari, M.H., Hooshmandasl, M.R. and Ghaini, F.M.M. *An efficient computational method for solving fractional biharmonic equation*, Comput. Math. Appl. 68(3) (2014) 269–287.

[18] Izadi, M. *A discontinuous finite element approximation to singular Lane–Emden type equations*, Appl. Math. Comput. 401 (2021) 126115.

[19] Izadi, M., Yüzbaşı, Ş. and Ansari, K.J. *Application of Vieta–Lucas series to solve a class of multi-pantograph delay differential equations with singularity*, Symmetry, 13(12) (2021) 2370.

[20] Karimi Dizicheh, A., Salahshour, S., Ahmadian, A. and Baleanu, D. *A novel algorithm based on the Legendre wavelets spectral technique for solving the Lane–Emden equations*, Appl. Numer. Math. 153 (2020) 443–456.

[21] Khalique, C. M. and Muatjetjeja, B. *Lie group classification of the generalized Lane–Emden equation*, Appl. Math. Comput. 210(2) (2009) 405–410.

[22] Kilicman, A., Shokhanda, R. and Goswami, P. *On the solution of (n+1)-dimensional fractional m-burgers equation*, Alex. Eng. J. 60(1) (2021) 1165–1172.

[23] Kumar, M. and Singh, N. *Modified Adomian decomposition method and computer implementation for solving singular boundary value problems arising in various physical problems*, Comput. Chem. Eng. 34(11) (2010) 1750–1760.

[24] Lin, R., Liu, F., Anh, V. and Turner, I. *Stability and convergence of a new explicit finite-difference approximation for the variable-order nonlinear fractional diffusion equation*, Appl. Math. Comput. 212(2) (2009) 435–445.

[25] Magin, R.L., Ingo, C., Colon-Perez, L., Triplett, W. and Mareci, T.H. *Characterization of anomalous diffusion in porous biological tissues using fractional order derivatives and entropy*, Micropor. Mesopor. Mat. 178 (2013) 39–43.

[26] Moghaddam, B.P., Yaghoobi, S. and Machado, J.A.T. *An extended predictor-corrector algorithm for variable-order fractional delay differential equations*, J. Comput. Nonlinear Dyn. 11(6) (2016).

[27] Noor, Z.A., Talib, I., Abdeljawad, T. and Alqudah, M.A. *Numerical study of Caputo fractional-order differential equations by developing new operational matrices of Vieta–Lucas polynomials*, Fract. Fract. 6(2) (2022) 79.

[28] Parand, K., Yousefi, H. and Delkhosh, M. *A numerical approach to solve Lane–Emden type equations by the fractional order of rational Bernoulli functions*, Rom. J. Phys. 62(104) (2017) 1–24.

[29] Patnaik, S., Hollkamp, J.P. and Semperlotti, F. *Applications of variable-order fractional operators: A review,* Proc. R. Soc. A, 476(2234) (2020) 20190498.

[30] Podlubny, I. *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*, Elsevier, 1998.

[31] Robbins, N. *Vieta's triangular array and a related family of polynomials*, Int. J. Math. Math. Sci. 14(2) (1991) 239–244.

[32] Ross, B. and Samko, S. *Fractional integration operator of variable order in the holder spaces hλ (x),* Int. J. Math. Math. Sci. 18(4) (1995) 777–788.

[33] Sahu, P.K. and Mallick, B. *Approximate solution of fractional order Lane–Emden type differential equation by orthonormal Bernoulli's polynomials*, Int. J. Appl. Comput. Math. 5(3) (2019) 1–9.

[34] Samko, S.G. and Ross, B. *Integration and differentiation to a variable fractional order*, Integral Transforms Spec. Funct. 1(4) (1993) 277–300 1993.

[35] Shokhanda, R. and Goswami, P. *Solution of generalized fractional burgers equation with a nonlinear term,* Int. J. Appl. Comput. Math. 8(5) (2022) 1–14.

[36] Shokhanda, R., Goswami, P., He, J.H. and Althobaiti, A. *An approximate solution of the time-fractional two-mode coupled burgers equation*, Fract. Fract. 5(4) (2021) 196.

[37] Singh, R., Das, N. and Kumar, J. *The optimal modified variational iteration method for the Lane–Emden equations with Neumann and Robin boundary conditions*, Eur. Phys. J. Plus, 132 (2017) 1–11.

[38] Sun, H., Chen, W. and Chen, Y.Q. *Variable-order fractional differential operators in anomalous diffusion modeling*, Physica A Stat. Mech. Appl. 388(21) (2009) 4586–4592.

[39] Sweilam, N.H., Nagy, A.M., Assiri, T.A. and Ali, N.Y. *Numerical simulations for variable-order fractional nonlinear delay differential equations*, J. Frac. Calc. Appl. 6(1) (2015) 71–82.

[40] Tseng, C.C. *Design of variable and adaptive fractional order fir differentiators*, Signal Process. 86(10) (2006) 2554–2566.

[41] VanGorder, R.A. *An elegant perturbation solution for the Lane–Emden equation of the second kind*, New Astron. 16(2) (2011) 65–67.

[42] Wazwaz, A.M. and Rach, R. *Comparison of the Adomian decomposition method and the variational iteration method for solving the Lane–Emden equations of the first and second kinds*, Kybernetes, 40(9-10) (2011) 1305–1318.

[43] Xie, L.J., Zhou, C.L. and Song Xu, S. *Solving the systems of equations of Lane–Emden type by differential transform method coupled with Adomian polynomials*, Math. 7(4) (2019) 377.

[44] Zhuang, P., Liu, F., Anh, V. and Turner, I. *Numerical methods for the variable-order fractional advection-diffusion equation with a nonlinear source term*, SIAM J. Numer. Anal. 47(3) (2009) 1760–1781.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1247–1279
https://doi.org/10.22067/ijnao.2024.88781.1472
https://ijnao.um.ac.ir/

Research Article

A novel mid-point upwind scheme for fractional-order singularly perturbed convection-diffusion delay differential equation

N.A. Endrie*, and G.F. Duressa

Abstract

This study presents a numerical approach for solving temporal fractional-order singularly perturbed parabolic convection-diffusion differential equations with a delay using a uniformly convergent scheme. We use the asymptotic analysis of the problem to offer a priori bounds on the exact solution and its derivatives. To discretize the problem, we use the implicit Euler technique on a uniform mesh in time and the midpoint upwind finite difference approach on a piece-wise uniform mesh in space. The proposed

*Corresponding author

Received 2 July 2024; revised 5 August 2024; accepted 19 August 2024

Nuru Ahmed Endrie
Department of Mathematics, College of Natural and Computational Science, Arba Minch University, Arba Minch, Ethiopia. e-mail: nuruahmed222@gmail.com

Gemechis File Duressa
Department of Mathematics,College of Natural and Computational Science, Arba Minch University, Arba Minch, Ethiopia and
Department of Mathematics, College of Natural and Computational Science, Jimma University, Jimma, Ethiopia. e-mail: gammeef@gmail.com

How to cite this article
Endrie, N.A. and Duressa, G.F., A novel mid-point upwind scheme for fractional-order singularly perturbed convection-diffusion delay differential equation. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1247-1279. https://doi.org/10.22067/ijnao.2024.88781.1472

technique has a nearly first-order uniform convergence order in both spatial and temporal dimensions. To validate the theoretical analysis of the scheme, two numerical test situations for various values of ε are explored.

AMS subject classifications (2020): Primary 65L11; Secondary 65N12.

Keywords: Caputo–Fabrizio derivative operator, midpoint upwind scheme, fractional-order differential equation, delay differential equation, singularly perturbed problem.

1 Introduction

This work examines the singularly perturbed delay parabolic differential equation involving fractional order in time

$$\begin{aligned}\mathcal{L}w(s,t) \equiv^{Cf} D_t^{\gamma} w(s,t) - \varepsilon \frac{\partial^2 w(s,t)}{\partial s^2} + a(s)\frac{\partial w(s,t)}{\partial s} + b(s,t)w(s,t)\\ = -c(s,t)w(s,t-\delta) + f(s,t), \qquad (s,t) \in \Omega = (0,1)\times(0,\mathfrak{T}]\end{aligned} \tag{1}$$

with

$$\begin{cases} w(s,t) = \varphi_b(s,t), & \text{for}(s,t) \in [0,1]\times[-\delta,0],\\ w(0,t) = \varphi_l(t), w(1,t) = \varphi_r(t), & \text{for} \;\; t \in (0,\mathfrak{T}),\end{cases} \tag{2}$$

where $\varepsilon (0 < \varepsilon \ll 1)$ is a singular perturbation parameter, δ is a delay parameter, and the Caputo–Fabrizio fractional derivative of order $0 < \gamma < 1$ is denoted by ${}^{CF}D_t^{\gamma}$. The solution of the problem (1) possesses a boundary layer of regular type at $s = 1$ of width $\mathcal{O}(\varepsilon)$ if $a(s) \geq \beta > 0, b(s,t) \geq 0$, $c(s,t) \geq \alpha > 0$ and they are smooth and bounded on the domain $\bar{\Omega} = [0,1] \times [0,\mathfrak{T}]$, and to avoid oscillations in the solution, we take $b(s,t) + c(s,t) \geq \theta$.

In contrast to its classical equivalents, a number of scientists and academics are showing the effectiveness of fractional derivative and integral operators in analyzing and modeling real-world occurrences. Throughout recent years, fractional calculus has been studied because it illustrates memory quality. Time-fractional advection-diffusion equations, time-fractional Burgers equations, and other models have been developed by applying fractional partial differential equations to model problems in a wide range of fields, such as control theory, chemical physics, engineering, medicine, stochastic processes, and biology [20, 38, 32].

Caputo fractional derivatives are employed to explain most mathematical models for such systems. The main argument in favor of the Caputo fractional derivative is that the beginning conditions have the same form as the differential equations of integer order at time $t = 0$. On the other hand, the Riemann–Liouville technique necessitates initial conditions that include the limit values of the Riemann–Liouville fractional derivative at the initial time $t = 0$. These criteria are not entirely apparent in terms of their physical interpretation. On the contrary, the Riemann–Liouville fractional derivative may give the Caputo fractional derivative with certain assumptions about the regularity of the function [27]. To overcome the singularity of the kernel, Caputo and Fabricio devised a new fractional derivative in response to these difficulties and inconveniences. One of the intriguing characteristics of this novel fractional derivative is the presence of an exponential kernel, which can construct and characterize structures at various scales. Further benefits are gained in certainly published journals using this new fractional derivative operator for the analysis of the groundwater contamination problem, the nonlinear Fisher's reaction-diffusion problem, and the magneto hydrodynamics free convection flow [4, 19, 29].

Singularly perturbed delay differential equations (SPDDEs) are employed to model physical problems that evolve based on both their current condition and history. To make a model more realistic, it may be important to represent former system states in addition to the current state. Delay differential equations (DDEs) are useful for describing time-dependent phenomena that rely on a past state [11].

SPDDEs are used to model physical problems whose growth is controlled not only by the system's present state but also by its history. It is occasionally required to comprise past states of the system rather than only the current state in order to make the model more accurate. When the rate of change of a time-dependent phenomenon is dependent on a preceding state, DDEs play an important role in mathematical modeling [10]. The study of DDEs has grown in popularity over the last several decades due to their numerous applications in domains such as bio-sciences, control theory, economics, material science, medicine, robotics, and others; see Cooke [5], Diekmann et al. [6], Driver [7], Norkin [26], Kolmanovskii, and Myshkis [15], Hale and Lunel [13], Kolmanovskii and Nosov [16], Kuang [17]. There are various real-world examples of DDEs in the works by Nelson and Perelson [25], Villasana and Radunskaya [37], and Zhao [39].

Because the boundary layer exists in the singularly perturbed problem solution, standard finite difference approaches produce unsuitable numerical results. Various numerical algorithms have been devised to solve these problems, which require a technique known as a uniformly convergent method independent of the perturbation parameter's value. Fitted mesh and fitted operator are the two most used numerical methods to construct a uniformly convergent scheme. Fitted mesh approaches use a layer-adapted nonuniform mesh that is dense in the layer(s), whereas the fitted operator employs an exponentially fitted scheme [31, 14]. However, this highlights the numerical method's computational inefficiency. When the number of mesh points grows, the resulting algebraic system of equations may become ill-conditioned. The shortcoming encourages the creation of a suitable numerical approach whose accuracy is independent of the perturbation parameter, highlighting the key advantage of the proposed method [12].

The number of papers addressing numerical solutions to time-fractional singularly perturbed ordinary differential equations (ODEs) and partial differential equations (PDEs) is rather modest. Bijura [3] offered higher-order asymptotic solutions for fractionally-order nonlinear singular perturbation problems. Roop [30] solved fractional ODEs numerically using the finite element method. Al-Mdallal and Syam [1] applied the Pade approximation to solve nonlinear singularly perturbed two-point boundary-value problems of fractional order. Atangana and Goufo [2] extended the matched asymptotic technique to address fractional-order boundary layer problems. Sayevand and Pichaghchi [35] proposed a method to solve singularly perturbed ODEs in their fractional-order boundary value problem. They defined the local fractional derivative and expanded the matching asymptotic expansion technique using its properties. Sayevand and Pichaghchi [34] developed a linear B-spline operational matrix with fractional derivatives for singularly perturbed ODEs and PDEs. Sahoo and Gupta [33] proposed a finite difference approach to tackle time-fractional, singularly perturbed convection-diffusion problems. Kumar and Vigo-Aguiar [18] studied delay parabolic and time-fractional SPDEs and created discretized time domains with uniform step size and piece-wise-uniform Shishkin meshes in space domains.

To the best of our knowledge, the development and analysis of a numerical scheme for the class of singularly perturbed fractional-order delay differential equations under discussion is covered in only one publication in the literature [18]. This paper aims to present and analyze a new class of midpoint upwind finite difference methods using piece-wise uniform Shishkin mesh. By preserving significant aspects of the corresponding continuous problems, these techniques produce reliable numerical results.

The succeeding sections of the paper are organized as follows: Section 2 defines the preliminary concepts. Section 3 examines the continuous problem description, including its analytic solution and derivative behavior within stated constraints. Section 4 presents a midpoint upwind numerical scheme based on the Shishkin mesh developed. The method's uniform convergence analysis is discussed in Section 5. Section 6 details the numerical experimentation conducted to validate theoretical results and demonstrate the accuracy of the method. The final section provides a summary of the main conclusions drawn in the paper.

2 Preliminaries

We now offer the definitions and approaches that are required for this study.

Definition 1 (Singularly-perturbed problem)**.** [21] When the highest-order derivative of a differential equation is multiplied by a small parameter ε (where ε is the perturbation parameter, such that $0 < \varepsilon << 1$), the equation is considered singularly perturbed.

Definition 2 (Gamma function)**.** [28] For a complex number z with real part nonnegative ($\mathbb{R}(z) > 0$), the gamma function is defined as

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx. \tag{3}$$

Definition 3 (Caputo fractional derivative operator)**.** [28] For $n \in \mathbb{N}$ and $\gamma \in (n-1, n)$, the Caputo fractional derivative of a function $g(t)$ with lower limit zero can be defined as

$$D_0^{\gamma} g(t) = \frac{1}{\Gamma(n-\gamma)} \int_0^t \frac{g^{(n)}(\xi)}{(t-\xi)^{\gamma-n+1}} d\xi. \tag{4}$$

Definition 4 (Caputo–Fabrizio Operator)**.** [4] Considering fractional derivatives with order $\gamma > 0$, a novel operator known as the Caputo–Fabrizio operator is defined as follows:

$$^{CF}D_{0,t}^{\gamma}(g(t)) = \frac{L(\gamma)}{1-\gamma} \int_0^t g'(\xi) \exp\left[-\gamma \frac{t-\xi}{1-\gamma}\right] d\xi, \tag{5}$$

where $g \in H^1(a,b), b > a$, and $L(\gamma)$ is the normalization functions (any smooth positive function) such that $L(0) = L(1) = 1$. Moreover, we observe that the aforementioned definition does not contain a singular kernel.

3 Properties of continuous problem

Assuming sufficiently smoothness of $\varphi_l(t), \varphi_r(t),$ and $\varphi_b(x,t)$ and satisfying the following compatibility conditions at the corner points $(0,0), (1,0)$ and $(0,-\delta)$ as well as $(1,-\delta)$, the existence and uniqueness of the solution of (1)–(2) can be established.

$$\begin{cases} \varphi_b(0,0) = \varphi_l(0), \\ \varphi_b(1,0) = \varphi_r(0), \end{cases} \tag{6}$$

and

$$\begin{cases} \left.\dfrac{d^{\gamma}\varphi_l}{d^{\gamma}t}\right|_{t=0} - \varepsilon \left.\dfrac{\partial^2 \varphi_b}{\partial x^2}\right|_{(0,0)} + a(0) \left.\dfrac{\partial \varphi_b}{\partial x}\right|_{(0,0)} + b(0,0)\varphi_b(0,0) \\ \quad = -c(0,0)\varphi_b(0,-\delta) + f(0,0), \\ \left.\dfrac{d^{\gamma}\varphi_l}{d^{\gamma}t}\right|_{t=0} - \varepsilon \left.\dfrac{\partial^2 \varphi_b}{\partial x^2}\right|_{(1,0)} + a(1) \left.\dfrac{\partial \varphi_b}{\partial x}\right|_{(1,0)} + b(1,0)\varphi_b(1,0) \\ \quad = -c(1,0)\varphi_b(1,-\delta) + f(1,0). \end{cases} \tag{7}$$

Now, we will show that the operator $\mathfrak{L}$ satisfies the maximum principle.

Lemma 1 (Continuous maximum principle)**.** Consider the function $\phi(s,t) \in C^2(\Omega) \cap C^0(\bar{\Omega})$, with $\mathfrak{L}\phi(s,t) \geq 0$ in Ω and $\phi(s,t) \geq 0,$ for all $(s,t) \in \Lambda = \{0\} \times (0,\mathfrak{T}] \cup \{1\} \times (0,\mathfrak{T}] \cup [0,1] \times [-\delta, 0]$. Then $\phi(s,t) \geq 0,$ for all $(s,t) \in \bar{\Omega}$.

Proof. Let us assume that there exists $(\varsigma, \iota) \in \bar{\Omega}$ with

$$\phi(\varsigma,\iota) = \min_{(s,t)\in\bar{\Omega}} \phi(s,t), \quad \text{and} \quad \phi(\varsigma,\iota) < 0$$

With respect to these assumptions, we can deduce that $(\varsigma,\iota) \notin \Lambda$, implying that $(\varsigma,\iota) \in \Omega$. Using the operator $\mathfrak{L}$ on $\phi(s,t)$, we get

$$\mathfrak{L}\phi(s,t) = D_t^\gamma \phi(x,t) - \varepsilon\phi_{ss}(s,t) + a(s)\phi_s(s,t) + b(s,t)\phi(s,t),$$

At the point of minimum (ς,ι), we obtain

$$\mathfrak{L}\phi((\varsigma,\iota)) = D_t^\gamma \phi(\varsigma,\iota) - \varepsilon\phi_{ss}(\varsigma,\iota) + a(\varsigma)\phi_s(\varsigma,\iota) + b(\varsigma,\iota)\phi(\varsigma,\iota).$$

The function ϕ has minima at the point (ς,ι), so $D_t^\gamma\phi \geq 0$, $\phi_s = 0$, $\phi_{ss} \geq 0$ at point (ς,ι), and $b(\varsigma,\iota) \geq 0$ for $(\varsigma,\iota) \in \Omega$. Therefore, we have

$$\mathfrak{L}\phi(\varsigma,\iota) < 0.$$

This contradicts our assumption $\mathfrak{L}\phi(s,t)$ in Ω.

Consequently, we conclude that $\phi(s,t) \geq 0,$ for all $(s,t) \in \bar{\Omega}$. □

Lemma 2. The differential equation (1)–(2) has a solution $y(s,t)$ that satisfies this estimate:

$$|w(x,t) - \varphi_b(s,0)| \leq Ct, \qquad (x,t) \in \bar{\Omega},$$

in which C is a constant that does not depend on ε.

Proof. See reference [23]. □

Lemma 3. With its initial and boundary conditions in (2), the solution to problem (1) is bounded as

$$|w(s,t)| \leq C, \qquad \text{for all } (s,t) \in \bar{\Omega}. \tag{8}$$

Proof. From Lemma 2, we have

$$
\begin{aligned}
|w(s,t)| &= |w(s,t) - \varphi_b(s,0) + \varphi_b(s,0)| \\
&\leq |w(s,t) - \varphi_b(s,0)| + |\varphi_b(s,0)| \\
&\leq Ct + |\varphi_b(s,0)| \\
&\leq Ct + C \\
&\leq C \qquad \text{since } t \in (0, \mathfrak{T}], \text{ t is bounded.}
\end{aligned}
$$

□

Lemma 4. [9] For any nonnegative integers n, m with $0 \leq n+m \leq 5$ and appropriate compatible conditions at the edges, both the regular $v(x,t)$ and singular component $w(x,t)$ satisfy the subsequent bounds:

$$
\left| \frac{\partial^{n+m} v(s,t)}{\partial s^n \partial t^m} \right| \leq C \left(1 + \varepsilon^{4-n}\right), \tag{9}
$$

$$
\left| \frac{\partial^{n+m} w(s,t)}{\partial s^n \partial t^m} \right| \leq C\varepsilon^{-n} \exp\left(-\beta(1-s)/\varepsilon\right), \quad (s,t) \in \Omega. \tag{10}
$$

4 Numerical schemes

We are going to develop the numerical scheme in this section as well. After discretizing the temporal derivative with implicit Euler's method, the spatial derivative is discretized using the midpoint upwind approach by employing Shishkin mesh and obtaining the system of linear equations. This equation is then solved using any classical methods.

4.1 Temporal semi-discretization

We first partition the temporal domain $[0, \mathfrak{T}]$ to M_τ subintervals with uniform step size $\tau = \mathfrak{T}/M_\tau$. We chose M_τ such that for any positive integer $k \in (0, M_\tau)$, $\delta = k\tau$ must be a mesh point. The set Ω^{M_τ} represents all mesh points in the temporal direction. We then have $\Omega^{M_\tau} = \{t_0 = 0 < t_1 < t_2 < \cdots < t_k = \delta < t_{M_\tau - 1} < t_{M_\tau} = \mathfrak{T}\}$. We employ that $\Omega_\delta^{M_\tau}$ is the collection of all mesh points between zero and $-\delta$; that is, $\Omega_\delta^{M_\tau} = \{t_{-k} = -\delta < t_{-k+1} <$

$\cdots < t_{-1} < t_0 = 0\}$.

Consider the exact solution $w(s, t_{j+1})$ at $(j+1)^{th}$ mesh point and denote the approximation solution by $W^{j+1}(s)$. Let $z(s, t_{j+1}) = \frac{\partial^\gamma W(s,t_{j+1})}{\partial t^\gamma}$. According to Definition 4, then

$$
\begin{aligned}
z(s, t_{j+1}) &= \frac{\partial^\gamma W(s, t_{j+1})}{\partial t^\gamma} \\
&= \frac{L(\gamma)}{1-\gamma} \int_0^t \frac{W(s,\xi)}{\partial \xi} \exp\left[-\gamma \frac{t_{j+1}-\xi}{1-\gamma}\right] d\xi \\
&= \frac{L(\gamma)}{1-\gamma} \sum_{i=0}^{j} \int_{i\tau}^{(i+1)\tau} \frac{W(s,\xi)}{\partial \xi} \exp\left(-\gamma \frac{t_{j+1}-\xi}{1-\gamma}\right) d\xi \\
&= \frac{L(\gamma)}{1-\gamma} \sum_{i=0}^{j} \frac{(W(s,t_{i+1}) - W(s,t_i))}{\tau} \int_{(j-i)\tau}^{(j-i+1)\tau} \exp\left(-\gamma \frac{t_{j+1}-\xi}{1-\gamma}\right) d\xi \\
&\quad + \mathcal{R}_\tau \\
&= \frac{L(\gamma)}{1-\gamma} \sum_{i=0}^{j} \frac{(W(s,t_{j-i+1}) - W(s,t_{j-i}))}{\tau} \\
&\quad \int_{(i)\tau}^{(i+1)\tau} \exp\left(-\gamma \frac{t_{j+1}-\xi}{1-\gamma}\right) d\xi + \mathcal{R}_\tau \\
&= \eta \Bigg[(W(s,t_{j+1}) - W(s,t_j)) \\
&\quad + \sum_{i=1}^{j} (W(s,t_{j-i+1}) - W(s,t_{j-i})) \exp\left(-\frac{\gamma\tau}{1-\gamma} i\right) \Bigg] + \mathcal{R}_\tau.
\end{aligned}
$$

Let us consider $L(\gamma) = 1$ the normalization functions such that $L(0) = L(1) = 1$. Then

$$
\begin{aligned}
\eta &= \frac{\exp\left(\frac{-\gamma\tau}{1-\tau}\right) - 1}{\gamma\tau}, \\
\mathcal{R}_\tau &= \frac{O(\tau)}{1-\gamma} \int_0^{t_{j+1}} \exp\left(-\gamma \frac{t_{j+1}-\xi}{1-\gamma}\right) d\xi \quad \text{is the truncation error.}
\end{aligned}
$$

Hence we obtain

$$
\begin{aligned}
z(s, t_{j+1}) =& \eta \left(W(s,t_{j+1}) - W(s,t_j)\right) \\
&+ \eta \sum_{i=1}^{j} (W(s,t_{j-i+1}) - W(s,t_{j-i})) \exp\left(-\frac{\gamma\tau}{1-\gamma} i\right) + \mathcal{R}_\tau.
\end{aligned} \tag{11}
$$

Substituting (11) into (1) on Ω^{M_τ}, we get

$$\begin{aligned}&z(s,t_{j+1}) - \varepsilon\frac{\partial^2 W^{j+1}(s)}{\partial s^2} + a(s)\frac{\partial W^{j+1}(s)}{\partial s} + b^{j+1}(s)W^{j+1}(s)\\&= -c^{j+1}(s)W^{j-k+1}(s) + f^{j+1}(s).\end{aligned}$$

Once the expressions are rearranged and the operator form has been put in, we get

$$\tilde{\mathfrak{L}}W^{j+1}(s) = -\varepsilon\frac{\partial^2 W^{j+1}(s)}{\partial s^2} + a(s)\frac{\partial W^{j+1}(s)}{\partial s} + \nu^{j+1}(s)W^{j+1}(s) = F^j(s), \tag{12}$$

for $j = 1, 2, \ldots, M_\tau$ with

$$\begin{cases} w(s,t) = \varphi_b(s,t_j), & \text{for}(s,t) \in [0,1]\times[-\delta,0],\\ w(0,t_j) = \varphi_l(t_j), w(1,t_j) = \varphi_r(t_j), & \text{for } \ t \in (0,\mathfrak{T}),\end{cases} \tag{13}$$

where

$$\nu(s,t_{j+1}) = b^{j+1}(s) + \eta,$$

$$F^{j+1}(s) = \begin{cases} -c^{j+1}(s)\varphi_b(s,t_{j-k+1}) + f(s,t_{j+1}) \\ +\eta\sum_{i=1}^{j}\left(W(s,t_{j-i+1}) - W(s,t_{j-i})\right)\exp\left(-\frac{\gamma\tau}{1-\gamma}i\right) \\ +\eta W(s,t_j), & \text{for } \ j = 1,2,\ldots,k,\\ -c^{j+1}(s)W^{j-k+1}(s) + f^{j+1}(s) \\ +\eta\sum_{i=1}^{j}\left(W(s,t_{j-i+1}) - W(s,t_{j-i})\right)\exp\left(-\frac{\gamma\tau}{1-\gamma}i\right) \\ +\eta W(s,t_j), & \text{for } \ j = k+1, k+2, \ldots, M_\tau.\end{cases}$$

After some rearrangement of (12) we obtain

$$\left(1 + \alpha_0\mathfrak{L}^*_{\varepsilon,\delta}\right)W^{j+1}(x) = F^{j+1}(s), \tag{14}$$

where

$$\alpha_0 = \frac{\gamma\tau}{\exp\left(\frac{-\gamma\tau}{1-\gamma}\right) - 1},$$

$$\mathfrak{L}^*_{\varepsilon,\delta} = -\varepsilon\frac{\partial^2}{\partial s^2} + a(s)\frac{\partial}{\partial s} + b^{j+1}(s)$$

$$F^{j+1}(s)=\begin{cases}-\alpha_0 b^{j+1}(s)\varphi_b(s,t_{j-k+1})+\alpha_0 fj+1(s)\\ +\varphi_b(s,t_j)+\sum_{i=1}^{j}\left(W(s,t_{j-i+1})-W(s,t_{j-i})\right)\exp\left(-\frac{\gamma\tau}{1-\gamma}i\right)\\ +W(s,t_j), & \text{for} \quad j=1,2,\dots,k,\\ -\alpha_0 s^{j+1}(s)W^{j-k+1}(s)+\alpha_0 fj+1(s)\\ +\varphi_b(x,t_j)+\sum_{i=1}^{j}\left(W(s,t_{j-i+1})-W(s,t_{j-i})\right)\exp\left(-\frac{\gamma\tau}{1-\gamma}i\right)\\ +W(s,t_j), & \text{for} \quad j=k+1,k+2,\dots,M_\tau.\end{cases}$$

Lemma 5 (Semi-discrete maximum principle)**.** Let $\psi(s,t_{j+1})$ be a smooth function such that $\psi(0,t_{j+1})\geq 0$ and, $\psi(1,t_{j+1})\geq 0$, for all $(s,t_{j+1})\in \Lambda=\{0\}\times(0,\mathbb{T}]\cup\{1\}\times(0,\mathbb{T}]\cup[0,1]\times[-\delta,0]$. Then $(1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})\psi(s,t_{j+1})\geq 0$ in Ω implies that $\psi(s,t_{j+1})\geq 0,$ for all $(s,t_{j+1})\in\Omega$.

Proof. Suppose there exists $(\iota,t_{j+1})\in\bar{\Omega}$ with

$$\psi(\iota,t_{j+1})=\min_{(s,t_{j+1})\in\bar{\Omega}}\psi(s,t_{j+1}), \quad \text{and} \quad \psi(\iota,t_{j+1})<0.$$

Then we have

$$(\iota,t_{j+1})\notin\{(0,t_{j+1}),(1,t_{j+1})\} \quad \text{and} \quad \psi_s(\iota,t_{j+1})=0,\psi_{ss}(\iota,t_{j+1})>0.$$

Applying the operator $\mathfrak{L}^*_{\varepsilon,\delta}$ on $\psi(s,t_{j+1})$, we get

$$\begin{aligned}(1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})\psi(\iota,t_{j+1})=&\psi(\iota,t_{j+1})+\alpha_0(-\varepsilon\psi_{xx}(\iota,t_{j+1})+a(\iota)\varphi_s(\iota,t_{j+1})\\ &+b(\iota,t_{j+1})\psi(\iota,t_{j+1})).\end{aligned}$$

The function ψ has a minimum at (ι,t_{j+1}), resulting in $\psi_s=0$, $\psi_{ss}\geq 0$, and $r(\iota,t_{j+1})\geq 0$ for $(\iota,t_{j+1})\in\Omega$. Therefore, we have

$$(1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})\psi(\iota,t_{j+1})<0,$$

which contradicts our assumption of $(1+\mathfrak{L}^*_{\varepsilon,\delta})\psi(s,t_{j+1})$ in Ω.
Therefore, we conclude that $\psi(s,t_{j+1})\geq 0,$ for all $(s,t_{j+1})\in\bar{\Omega}$.
Hence from the above, we prove that the operator $(1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})$ satisfies the maximum principle and consequently

$$\left\|(1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})^{-1}\right\|\leq\frac{1}{1+\theta\tau}. \tag{15}$$

□

Lemma 6 (Truncation error)**.** The local truncation error corresponding to the semi-discretized problem (13) satisfies

$$\left|\mathcal{R}_\tau^{j+1}\right| \leq C\tau. \tag{16}$$

Proof. From semi-discretized problem, we have

$$\begin{aligned}
\mathcal{R}_\tau^{j+1} &= \frac{O(\tau)}{1-\gamma}\int_0^{t_{j+1}} \exp\left(-\gamma\frac{t_{j+1}-\xi}{1-\gamma}\right)d\xi\\
&= \frac{O(\tau)}{1-\gamma}\int_0^{(j+1)\tau} \exp\left(-\gamma\frac{(j+1)\tau-\xi}{1-\gamma}\right)d\xi\\
&= \frac{O(\tau)\gamma}{(1-\gamma)^2}\left(\exp\left(-\gamma\frac{(j+1)\tau}{1-\gamma}\right)-\exp\left(\frac{-\gamma}{1-\gamma}\right)\right)\\
&\leq \frac{\gamma}{(1-\gamma)^2}\left(\exp\left(-\gamma\frac{j+1}{1-\gamma}\right)-\exp\left(\frac{-\gamma}{1-\gamma}\right)\right)\tau\\
&\leq C\tau,
\end{aligned}$$

since $\exp\left(-\gamma\frac{(j+1)\tau}{1-\gamma}\right) \leq \exp\left(-\gamma\frac{j+1}{1-\gamma}\right)$ and

$$C = \frac{\gamma}{(1-\gamma)^2}\left(\exp\left(-\gamma\frac{j+1}{1-\gamma}\right)-\exp\left(\frac{-\gamma}{1-\gamma}\right)\right).$$

Therefore, we obtain

$$\left|\mathcal{R}_\tau^{j+1}\right| \leq C\tau.$$

□

Lemma 7 (Error bound)**.** The global error estimation at t_{j+1} satisfies

$$\|E_{j+1}\| \leq C\tau.$$

Proof. Since the function $y(x,t_{j+1})$ satisfies

$$(1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})y(x,t_{j+1}) = F^{j+1}(x), \tag{17}$$

also the solution of the continuous problem (1)–(2) is smooth enough, then we have

$$\begin{aligned}
F^{j+1}(s) &= (1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})W(s,t_{j+1}) + \mathcal{R}_\tau^{j+1}\\
&= (1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})W(s,t_{j+1}) + C\tau
\end{aligned}$$

$$\Longrightarrow F^{j+1}(s) = (1+\alpha_0 \mathfrak{L}^*_{\varepsilon,\delta}) W(s,t_{j+1}) + C\tau. \tag{18}$$

From (17)–(18), the error corresponding to (14) satisfies the following problem

$$\begin{aligned}(1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})E_{j+1} &= C\tau\\ \Longrightarrow E_{j+1} &= (1+\alpha_0\mathfrak{L}^*_{\varepsilon,\delta})^{-1}\tau\end{aligned}$$

$$\|E_{j+1}\| \leq \frac{1}{1+\theta\tau} C\tau.$$

Hence, we obtain the result

$$\|E_{j+1}\| \leq C\tau.$$

□

Theorem 1. The semi-discretize solution $W(s,t_{j+1})$ and its derivatives meet the following bounds:

$$\left|\frac{d^i W(s,t_{j+1})}{ds^i}\right| \leq C(1+\varepsilon^{-i}\exp(-\beta(1-s)/\varepsilon)), \quad \text{for} \quad i = 0,1,2,3,4.$$

Proof. For the proof, refer [8]. □

We can rewrite (12) as operator form

$$\tilde{\mathfrak{L}}^\tau_\varepsilon W^{j+1}(s) = F^{j+1}(s), \tag{19}$$

where $\tilde{\mathfrak{L}}^\tau_\varepsilon W^{j+1}(s) = -\varepsilon\frac{\partial^2 W^{j+1}(s)}{\partial s^2} + a(s)\frac{\partial W^{j+1}(s)}{\partial s} + \nu(s)W^{j+1}(s)$ and

$$F^j(x) = \begin{cases} -s^{j+1}(s)\varphi_b(s,t_{j-k+1}) + f(s,t_{j+1}) \\ +\sigma B_j\varphi_b(s,t_j) + \sigma\sum_{i=1}^{j} B_i\left(W(s,t_{j-i+1})\right. \\ \left.-W(s,t_{j-i})\right), \qquad \text{for } j = 1,2,\ldots,k, \\ -s^{j+1}(s)W^{j-k+1}(s) + f^{j+1}(s) \\ +\sigma\psi_b(s,t_j) + \sigma\sum_{i=1}^{j} B_i\left(W(s,t_{j-i+1})\right. \\ \left.-W(s,t_{j-i})\right), \qquad \text{for } j = k+1,k+2,\ldots,M_\tau. \end{cases}$$

Lemma 8. [9] Consider the regular V^{j+1} and singular U^{j+1} components of W^{j+1} and their derivatives, which holds the following bounds:

$$\left|\frac{\partial^n V^{j+1}(x)}{\partial x^n}\right| \leq C\left(1+\varepsilon^{4-n}\right), \tag{20}$$

$$\left|\frac{\partial^n U^{j+1}(s)}{\partial s^n}\right| \leq C\varepsilon^{-n}\exp\left(-\alpha(1-s)/\varepsilon\right), \qquad n=0,1,\ldots,4. \tag{21}$$

4.2 Spatial discretization

4.2.1 Mesh generation

To address the semi-discretized problem (12), we used a piecewise-uniform Shishkin mesh to partition the space domain. This ensures that the boundary layer region had more mesh points than the outer region. The domain $[0,1]$ is separated into two subdomains, $[0,1-\sigma]$ and $(1-\sigma,1]$, where σ is given as

$$\sigma = \min\{0.5, \sigma_0\varepsilon\ln N\},$$

with boundary point $1-\sigma$, whereas $\sigma_0 \leq 1/\beta$ and N is the number of subintervals. The mesh sizes in the inner and outer boundary layer are

$$h_i = \begin{cases} \frac{2(1-\sigma)}{N}, & i=0,1,\ldots,\frac{N}{2}, \\ \frac{2\sigma}{N}, & i=\frac{N}{2}+1,\ldots,N. \end{cases}$$

The mesh points have been defined by

$$x_i = \begin{cases} i\frac{2(1-\sigma)}{N}, & i=0,1,\ldots,\frac{N}{2}, \\ 1-\sigma+(i-\frac{N}{2})\frac{2\sigma}{N}, & i=\frac{N}{2}+1,\ldots,N. \end{cases}$$

The set Ω^N represents all mesh points along the spatial direction. Hence, we have

$$\Omega^N = \{x_0 = 0 < x_1 < \cdots < x_{N-1} < x_N = 1\}.$$

The set Ω^N is the set that includes all mesh points along the spatial direction with equal step size $h = 1/N$ over $\sigma = 1/2$.

In the next section, we derive stable finite difference method.

4.2.2 Upwind numerical scheme

We utilize a midpoint upwind finite difference approach on the semi-discrete scheme (12) along the spatial direction. The difference operators with $\hat{h}_i = h_i + h_{i+1}$ are given as

$$\begin{aligned} D_s^- W_i^{j+1} &= \frac{W_i^{j+1} - W_{i-1}^{j+1}}{h_i}, \quad D_s^+ W_i^{j+1} = \frac{W_{i+1}^{j+1} - W_i^{j+1}}{h_{i+1}}, \\ D_s^0 W_i^{j+1} &= \frac{W_{i+1}^{j+1} - W_{i-1}^{j+1}}{\hat{h}_i}, \quad \delta_s^2 W_i^{j+1} = \frac{2}{\hat{h}_i}\left(D^+ W_i^{j+1} - D^- W_i^{j+1}\right). \end{aligned} \tag{22}$$

Additionally, we define $W_{i-1/2}^{j+1} = \frac{W_i^{j+1} + W_{i-1}^{j+1}}{2}$. We similarly define other terms.

Therefore using these difference operators, we derive the complete finite difference scheme:

$$\begin{cases} W_0^{j+1} = \varphi_l(0, t_{j+1}), \\ \tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} W_{i-1/2}^{j+1} = F_{i-1/2}^{j+1}, \quad i = 1, 2, \ldots, N, \\ W_N^{j+1} = \varphi_r(1, t_{j+1}), \end{cases} \tag{23}$$

where

$$\begin{aligned} \tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} &= -\varepsilon\delta_s^2 + a_{i-1/2} D_s^- + \nu_{i-1/2}, \\ F_{i-1/2}^{j+1} &= -c_{i-1/2}^{j+1} W_{i-1/2}^{j-k+1} + f_{i-1/2}^{j+1} \\ &\quad + \eta \sum_{m=1}^{j} \left(W_{i-1/2}^{j-m+1} - W_{i-1/2}^{j-m}\right) \exp\left(-\frac{\gamma\tau}{1-\gamma} m\right) + \eta W_{i-1/2}^{j}. \end{aligned}$$

Through the rearrangement of equation (23), we deduce the subsequent system of equations:

$$\Upsilon_i^- W_{i-1}^{j+1} + \Upsilon_i^0 W_i^{j+1} + \Upsilon_{i+1}^+ W_{i-1}^{j+1} = F_{i-1/2}^{j+1}, \qquad i = 1, 2, \ldots, N, \tag{24}$$

$$\begin{cases} \Upsilon_i^- = -\frac{2\varepsilon}{h_i \hat{h}_i} - \frac{q_{i-1/2}}{h_i} + \frac{\nu_{i-1/2}^{j+1}}{2}, \\ \Upsilon_i^0 = \frac{2\varepsilon}{h_i h_{i+1}} + \frac{q_{i-1/2}}{h_i} + \frac{\nu_{i-1/2}^{j+1}}{2}, \\ \Upsilon_i^+ = -\frac{2\varepsilon}{h_{i+1} \hat{h}_i}. \end{cases}$$

5 Convergence analysis

This section discusses the stability of the discretized scheme (24). To determine the ε-uniform error estimate, the truncation error is used. The stability of the developed scheme can be studied using the following lemma.

Lemma 9 (Discrete maximum principle)**.** Let φ_i^j be any function satisfying $\varphi_0 \geq 0$, $\varphi_{N+1} \geq 0$. Then $\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\Phi_i^{j+1} \geq 0$, for all $i = 1, 2, 3, \ldots, N$, implies that $\varphi_i^j \geq 0$ for all $i = 1, 2, 3, \ldots, N$.

Proof. This follows the same arguments as in Lemma 5. □

Lemma 10. Consider the difference operator of (24) $\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}W_i^{j+1}, i = 1, 2, \ldots . N-1$, with boundaries W_0^{j+1} and W_N^{j+1} has a solution. If $\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}W_i^{j+1} \leq \tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}Z_i^{j+1}$, $W_0^{j+1} \leq Z_0^{j+1}$ and $W_N^{j+1} \leq Z_N^{j+1}$, then $W_i^{j+1} \leq Z_i^{j+1}, i = 1, 2, \ldots, N-1$.

Proof. The coefficient matrix for the operator $\tilde{\mathfrak{L}}_{\varepsilon}^{h,\tau}$ has a size of $(N_h+1)(N_h+1)$ with its constituents, for $i = 1, 2, \ldots, N_h - 1$, are

$$\Upsilon_i^- = -\frac{2\varepsilon}{h_i\hat{h}_i} - \frac{q_{i-1/2}}{h_i} + \frac{\nu_{i-1/2}^{j+1}}{2} < 0, \quad \text{since all terms are positive,}$$

$$\Upsilon_i^0 = \frac{2\varepsilon}{h_i h_{i+1}} + \frac{q_{i-1/2}}{h_i} + \frac{\nu_{i-1/2}^{j+1}}{2} > 0, \quad \text{since all terms are positive,}$$

$$\Upsilon_i^+ = -\frac{2\varepsilon}{h_{i+1}\hat{h}_i} < 0, \quad \text{since all terms are positive .}$$

The proposed method's coefficient matrix (24) for the differential equation (1)–(2) meets the properties of an M-matrix. This suggests that the inverse matrix exists and is not negative. This ensures the existence and uniqueness of the discrete solution. A similar procedure was used to demonstrate the stability of a discrete scheme in [36, 24]. □

A direct consequence of Lemma 10 is the discrete stability result as follows.

Lemma 11 (The discrete stability result)**.** If W_i^{j+1} is the solution of the fully discretized equation corresponding to the original problem in (1), such that $W_0^{j+1} = W_N^{j+1} = 0$, then the following inequality is satisfied:

$$\left|W_i^{j+1}\right| \leq \left\|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} W_i^{j+1}\right\| + \max_{0\leq i,j\leq N,M}\{|\varphi l|\,,|\varphi r|\}.$$

Proof. Consider two barrier functions

$$\prod^{\pm} = \zeta \pm W_i^{j+1}, \quad \text{where} \quad \zeta = \left\|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} W_i^{j+1}\right\| + \max_{1\leq i,j\leq N,M}\{|\varphi_l|\,,|\varphi_r|\}$$

At the boundaries
$\prod_0^{\pm} = \zeta \pm w(0,t_j) = \left\|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} W_i^{j+1}\right\| + \max_{1\leq j\leq M}\{|\varphi_l|\,,|\varphi_r|\} \pm \varphi_l \geq 0$ and
$\prod_N^{\pm} = \zeta \pm w(1,t_j) = \left\|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} W_i^{j+1}\right\| + \max_{1\leq j\leq M}\{|\varphi_l|\,,|\varphi_r|\} \pm \varphi_r \geq 0.$
Using the operator $\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}$ on the barrier function, we have

$$\begin{aligned}
\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} \prod_i^{\pm} &= \left(-\varepsilon\delta_s^2 + a_{i-1/2}D_s^- + \nu_{i-1/2}\right)\left(\zeta \pm W_i^{j+1}\right)\\
&= -\varepsilon\delta_s^2\left(\zeta \pm W_i^{j+1}\right) + a_{i-1/2}D_s^-\left(\zeta \pm W_i^{j+1}\right) + \nu_{i-1/2}\left(\zeta \pm W_i^{j+1}\right)\\
&\geq \left\|\tilde{\mathfrak{L}}_{\varepsilon}^{h,\tau} W_i^{j+1}\right\| + \max_{1\leq i,j\leq N,M_\tau}\{|\varphi_l|\,,|\varphi_r|\} \pm W_i^{j+1}\\
&\geq 0.
\end{aligned}$$

By using the maximum principle, Lemma 9, we get

$$\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} \prod^{\pm} \geq 0 \quad \Longrightarrow \left|W_i^{j+1}\right| \leq \left\|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} W_i^{j+1}\right\| + \max_{0\leq i,j\leq NM_\tau}\{|\varphi l| + |\varphi r|\}.$$

□

Lemma 12. [31] Consider the mesh function Λ_i, for $i = 0,1,\ldots,N$ defined by

$$\Lambda_i^{j+1} = \prod_{l=1}^{i}\left(1 + \frac{\alpha h_l}{2\varepsilon}\right),$$

with the familiar convection that if $i = 0$, then $\Lambda_0^{j+1} = 1$. Then, for $i = 1,2,\ldots,N-1$, a positive constant C exists such that

$$\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\Lambda_i^{j+1} \geq \frac{C}{\max(\varepsilon,h_i)}\Lambda_i^{j+1}.$$

Theorem 2. Let $w(s_i,t_j)$ represent the exact solution for the problem (1) and let W_i represent the full discretization solution for problem (19). The error estimate is then provided by

$$|W_i - w(s_i, t_j)| \leq CN^{-1}(\ln N)^2.$$

Proof. The discrete problem's solution, W_i, can be divided into regular V_i and singular U_i components.
Thus

$$W_i^{j+1} = V_i^{j+1} + U_i^{j+1}.$$

Moreover, V_i and U_i represent the solutions to the following problems, respectively:

$$\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} V_i = F_{i-1/2}, V_i(0) = v(0), V_i(1) = v(1) (\text{nonhomogeneous problem})$$

$$\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} U_i = 0, U_i(0) = u(0), U_i(1) = u(1) \quad (\text{homogeneous problem})$$

Consider $W_i^{j+1} = W_i, V_i^{j+1} = V_i$, and $U_i^{j+1} = U_i$.

The error can be expressed as

$$W_i - w(x_i) = (V_i + U_i) - (v(x_i) + u(x_i)),$$

$$W_i - w(x_i) = (V_i - v(x_i)) + (U_i - u(x_i)). \tag{25}$$

Using the triangle inequality, the error can be estimated as

$$\|W_i - w(s_i)\| \leq \|V_i - v(s_i)\| + \|U_i - u(s_i)\|.$$

Next, we need to estimate the errors in the regular and singular components individually.

The following classical reasoning can be used to calculate the error estimate for a regular component. From the differential and difference equations, we get

$$\begin{aligned}
\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} (V_i - v(s_i)) &= \tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} V_i - \tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} v(s_i) \\
&= F_i - \tilde{\mathfrak{L}}_{mpu}^{\tau,h_i} v(s_i) \\
&= \left(\tilde{\mathfrak{L}} - \tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\right) v(s_i) \\
&= -\varepsilon \left(\frac{d^2}{dx^2} - \delta^2\right) v(s_i) + a \left(\frac{d}{ds} - D^-\right) v(s_i).
\end{aligned}$$

By using the estimate in [22], we get

$$
\begin{aligned}
\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\left(V_i - v(x_i)\right)\right| &\leq \frac{1}{4}\left(x_{i+1} - x_{i-1}\right)|v|_3 + \frac{1}{2}\left(x_i - x_{i-1}\right)|v|_2 \\
&\leq \frac{1}{4}\left(x_{i+1} - x_{i-1}\right)|v|_3 + \frac{1}{2}\left(x_{i+1} - x_{i-1}\right)|v|_2 \\
&\qquad\qquad\qquad \text{since } \; x_i - x_{i-1} \leq x_{i+1} - x_{i-1}, \\
&\leq C\left(x_{i+1} - x_{i-1}\right)\left(|v|_3 + |v|_2\right).
\end{aligned}
$$

Therefore, we obtain

$$
\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\left(V_i - v(s_i)\right)\right| \leq C\left(s_{i+1} - s_{i-1}\right)\left(|v|_3 + |v|_2\right). \tag{26}
$$

Indeed $s_{i+1} - s_{i-1} \leq 2N^{-1}$, for all $i, 0 \leq i \leq N-1$ and applying the discrete stability lemma for the regular component, v satisfies the following bound of the derivatives:

$$
\left|v^{(k)}(s)\right| \leq C\left(1 + \varepsilon^{-\frac{(k-2)}{2}}\right), \qquad 0 \leq k \leq 4. \tag{27}
$$

Then from (27) we obtain

$$
|v|_2 = |v''| \leq C, \quad \text{and} \quad |v|_3 = |v'''| \leq C(1 + \varepsilon^{-\frac{1}{2}}). \tag{28}
$$

By substituting (28) into (26), for the estimates of the derivatives of v, we have

$$
\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\left(V_i - v(s_i)\right)\right| \leq CN^{-1}.
$$

Next, we apply maximum principle result for the operator $\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}$ to the function $V_i - v(x)$. Then we obtain

$$
|(V_i - v(s))| \leq CN^{-1}. \tag{29}
$$

The error estimate regarding the singular component $U_i - u(x_i)$ is determined by the transition parameter value, which is either $\sigma = 1/2$ or $\sigma = \sigma_0\varepsilon \ln N$.

Case-I: For $\sigma = \frac{1}{2}$, the grids are equal and $\varepsilon\sigma_0 \ln N \geqslant \frac{1}{2}$, resulting in

$$
s_{i+1} - s_{i-1} = 2N^{-1} \quad \text{and} \quad \varepsilon^{-1} \leq C \ln N. \tag{30}
$$

The classical argument, which was used earlier to estimate the regular argument, concludes

$$\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}(U_i - u(s_i))\right| \leq C\left(s_{i+1} - s_{i-1}\right)\left(|u|_3 + |u|_2\right).$$

We use a similar fashion with the regular component. Also, $s_{i+1} - s_{i-1} = 2N^{-1}$ and $\varepsilon^{-1} \leq C \ln N$. Then we have

$$\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}(U_i - u(s_i))\right| \leq C\varepsilon^{-2}N^{-1}.$$

Therefore, we obtain

$$\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}(U_i - u(s_i))\right| leq C\varepsilon^{-2}N^{-1}. \tag{31}$$

Substituting (30) into (31), then we get

$$\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}(U_i - u(s_i))\right| \leq CN^{-1}(\ln N)^2.$$

An application to ε-uniform stability result for the operator $\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}$ to the function $U_i - u(s_i)$ then gives

$$|U_i - u(s_i)| \leq CN^{-1}(\ln N)^2. \tag{32}$$

Case-II: If $\sigma = \varepsilon\sigma_0 \ln N < \frac{1}{2}$, the mesh is piece-wise uniform, with the mesh spacing $\frac{2(1-\sigma)}{N}$ in the subinterval $[0, 1-\sigma]$ and $\frac{2\sigma}{N}$ in the subinterval $[1-\sigma, 1]$, then we estimate the error in $[0, 1-\sigma]$ and $[1-\sigma, 1]$ individually. In the subinterval $[0, 1-\sigma]$ without boundary layer both U_i and $u(s_i)$ are small and since by triangle inequality $|U_i - u(s_i)| \leq |U_i| + |u(x_i)|$, it helps to bound U_i and u individually.
We assumed that $\sigma = \sigma_0\varepsilon \ln N (> 2\varepsilon \ln N/\alpha)$, which gives $-\alpha\sigma/\varepsilon < \ln N^{-2}$. From lemma 8, we have

$$\begin{aligned}
|U_i| &\leq Ce^{-\alpha\sigma/\varepsilon} \\
&\leq Ce^{-\alpha(\frac{\varepsilon \ln N}{\alpha})/\varepsilon} \\
&\leq Ce^{-\ln N} \\
&\leq CN^{-1},
\end{aligned}$$

$$|U_i| \leq CN^{-1}. \tag{33}$$

Next, we need to show $|U_i| \leq CN^{-1}$ for a suitable constant C.
Let $\Lambda_i^{j+1} = \Lambda_i$. From the Taylor series expansion $e^s \geq 1 + s$, then for the mesh function Λ_i,

$$\begin{aligned}
\Lambda_N &= \prod_{l=1}^{N}\left(1 + \frac{\alpha h_l}{2\varepsilon}\right)\\
&= \prod_{l=1}^{i}\left(1 + \frac{\alpha h_l}{2\varepsilon}\right)\prod_{l=i+1}^{N}\left(1 + \frac{\alpha h_l}{2\varepsilon}\right)\\
&= \Lambda_i \prod_{l=i+1}^{N}\left(1 + \frac{\alpha h_l}{2\varepsilon}\right),\\
\Lambda_i/\Lambda_N &= \prod_{l=i+1}^{N}\left(1 + \frac{\alpha h_l}{2\varepsilon}\right)^{-1} \geq \prod_{l=i+1}^{N}\exp\left(-\frac{\alpha h_l}{2\varepsilon}\right) = \exp\left(-\frac{\alpha(1-x_i)}{2\varepsilon}\right).
\end{aligned}$$

Let $\hat{\Lambda}_i = C\Lambda_i/\Lambda_N$ for $i = 0, 1, \ldots, N$. Hence by using the definition of U_i and Lemma 12,

$$\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\hat{\Lambda}_0 = \left(\frac{C}{\Lambda_N}\right)\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\Lambda_i \geq \left(\frac{C}{\Lambda_N}\right)\frac{\Lambda_i}{\max\{\varepsilon, h_i\}} \geq 0 = \left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}U_N\right|,$$

$$\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\hat{\Lambda}_0 \geq 0. \tag{34}$$

For a suitably large value of C, we get that

$$\hat{\Lambda}_0 = \frac{C\Lambda_0}{\Lambda_N} \geq C\exp\left(\frac{-\alpha}{2\varepsilon}\right) \geq C\exp\left(\frac{-\alpha}{\varepsilon}\right) \geq |U_0|, \tag{35}$$

$$\hat{\Lambda}_0 = C \geq |U_N|. \tag{36}$$

Using (34)–(36), by the discrete comparison principle, Lemma 10, we obtain

$$|U_i| \leq \hat{\Lambda}_i = C\Lambda_i/\Lambda_N.$$

When $i = 0, 1, \ldots, N/2$,

$$\Lambda_i/\Lambda_N \leq \prod_{l=N/2+1}^{N}\left(1 + \frac{\alpha(x_{i+1} - x_{i-1})}{2\varepsilon}\right)^{-1} = \left(1 + 2N^{-1}\ln N\right)^{-N/2}. \tag{37}$$

After some simplification and the Taylor series expansion, $\ln(1+s) > s - \frac{s^2}{2}, s > 0$, to the right end expression of (37), we get

$$|U_i| \leq CN^{-1}, \quad \text{for} \quad i = 0, 1, \dots, N/2. \tag{38}$$

Combining (33) and (38), we obtain our result

$$|U_i| \leq CN^{-1}. \tag{39}$$

Triangle inequality results

$$|U_i - u(s_i)| \leq CN^{-1}, \quad \text{for} \quad 0 \leq i \leq N/2. \tag{40}$$

Next, we consider the subinterval $[1-\sigma, 1]$.

Using the classical argument, we get

$$\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\left(U_i - u(s_i)\right)\right| \leq C\varepsilon^{-2}(s_{i+1} - s_{i-1}), \quad \text{for} \quad N/2+1, N/2+2, \dots, N-1.$$

For the second interval, the mesh spacing is $2\sigma/N$. Since $s_{i+1} - s_{i-1} = 4\sigma/N$, we obtain that

$$\left|\tilde{\mathfrak{L}}_{mpu}^{\tau,h_i}\left(U_i - u(s_i)\right)\right| \leq C\varepsilon^{-2}\sigma N^{-1}, \quad \text{for} \quad N/2+1, N/2+2, \dots, N-1, \tag{41}$$

and

$$|U_N - u(1)| = 0.$$

Also, using the inequality (40), we obtain

$$\left|U_{N/2} - u(s_{N/2})\right| \leq CN^{-1}. \tag{42}$$

Now, we introduce the barrier function, with the suitable choices of C_1 and C_2 as

$$\xi_i = (s_i - (1-\sigma))C_1\varepsilon^{-2}\sigma N^{-1} + C_2 N^{-1},$$

and the mesh function

$$\chi_i^{\pm} = \xi_i \pm (u(s_i) - U_i), \quad N/2 \leq i \leq N.$$

Thus, we have

$$\begin{aligned}
\chi_{N/2}^{\pm} &= \xi_{N/2} \pm (u(s_{N/2}) - U_{N/2}) \\
&\geq (s_{N/2} - (1-\sigma))C_1\varepsilon^{-2}\sigma N^{-1} + C_2 N^{-1} \pm (\mp CN^{-1}) \\
&\geq C_2 N^{-1} \pm (\mp CN^{-1}) \quad \text{since} \quad s_{N/2} = 1 - \sigma
\end{aligned}$$

$$\geq (C_2 \mp C)N^{-1}.$$

Choose C_2 as $C_2 \mp C \geq 0$ and then

$$\chi^{\pm}_{N/2} \geq 0.$$

Next, we want to show that $\chi_N \pm \geq 0$. Therefore,

$$\begin{aligned}
\chi^{\pm}_N &= \xi_N \pm (u(s_N) - U_N) \\
&\geq (s_N - (1-\sigma))C_1\varepsilon^{-2}\sigma N^{-1} + C_2 N^{-1} \\
&\geq C_1\sigma\sigma_0(\ln N)^2 N^{-1} + C_2 N^{-1} \\
&\geq 0.
\end{aligned}$$

Lastly, we need to show that $\tilde{\mathfrak{L}}^{\tau,h_i}_{mpu}\chi^{\pm}_i \geq 0$. Hence

$$\begin{aligned}
\tilde{\mathfrak{L}}^{\tau,h_i}_{mpu}\chi^{\pm}_i &= \tilde{\mathfrak{L}}^{\tau,h_i}_{mpu}\left[\xi_i \pm (u(s_i) - U_i)\right] \\
&= \tilde{\mathfrak{L}}^{\tau,h_i}_{mpu}\xi_i \pm \tilde{\mathfrak{L}}^{\tau,h_i}_{mpu}(u(s_i) - U_i) \\
&\geq \left(\beta + \frac{2}{\tau}\right)\left((s_i - (1-\sigma))C_1\varepsilon^{-2}\sigma N^{-1} + C_2 N^{-1}\right) \\
&\quad \pm \tilde{\mathfrak{L}}^{\tau,h_i}_{mpu}(u(s_i) - U_i) \\
&\geq (s_i - (1-\sigma))C_1\varepsilon^{-2}\sigma N^{-1} + C_2 N^{-1} + C_2 N^{-1} \\
&\quad \pm (\mp C\varepsilon^{-2}\sigma N^{-1}), \quad \text{by (41)} \\
&\geq (s_i - \sigma)C_1\varepsilon^{-2}\sigma N^{-1} + C_2 N^{-1} + C_2 N^{-1} \pm (\mp C\varepsilon^{-2}\sigma N^{-1}) \\
&\geq ((s_i - \sigma)C_1 \mp C)\varepsilon^{-2}\sigma N^{-1} + C_2 N^{-1} + C_2 N^{-1} \\
&\geq 0, \quad \text{since} \quad (s_i - \sigma) \geq 0 \quad \text{then} \quad (s_i - \sigma)C_1 \mp C \geq 0.
\end{aligned}$$

Therefore by using the discrete maximum principle, we have

$$\chi^{\pm}_i \geq 0, \quad N/2 \leq i \leq N.$$

Thus we obtain the result

$$\left| u^{j+1}(s_i) - U^{j+1}_i \right| \leq CN^{-1}(\ln N)^2. \tag{43}$$

Then combining (32) and (43), we get

$$\left| u^{j+1}(s_i) - U^{j+1}_i \right| \leq CN^{-1}(\ln N)^2, \quad \text{for} \quad 0 \leq i \leq N. \tag{44}$$

Therefore, using (25) the inequalities in (29) and (44), gives

$$\left|(W_i^{j+1} - w(s_i, t_{j+1})\right| \leq CN^{-1}(\ln N)^2.$$

$\square$

Theorem 3. If w and W are exact and approximation solution of the problem (1), respectively, then the following error bound holds:

$$\left|w(s_i, t_j) - W_i^{j+1}\right| \leq C\left(N^{-1}(\ln N)^2 + \tau\right). \tag{45}$$

Proof. By combining Lemma 7 and Theorem 2 we get our result. $\square$

6 Numerical result

In this section, we present two numerical examples that illustrate the method's accuracy as well as the error analysis results. Separate tables show the error and convergence rates for each of these two instances of the test. This article will employ double mesh to determine the numerical solution's accuracy, as the exact answer is uncertain. The maximum point-wise absolute error is calculated as

$$E_\varepsilon^{N,M} = \max_{0\leq i,j\leq N,M}\left|W_{i,j}^{N,M} - W_{2i,2j}^{2N,2M}\right|,$$

where N and M represent the number of mesh points in spatial and temporal directions, respectively. The parameter for uniform error estimation is

$$e^{N,M} = \max_{\varepsilon}\{E_\varepsilon^{N,M}\}$$

. The method's rate of convergence is determined using the formula

$$RoC_\varepsilon^{N,M} = \log_2\left(\frac{E_\varepsilon^{N,M}}{E_\varepsilon^{2N,2M}}\right).$$

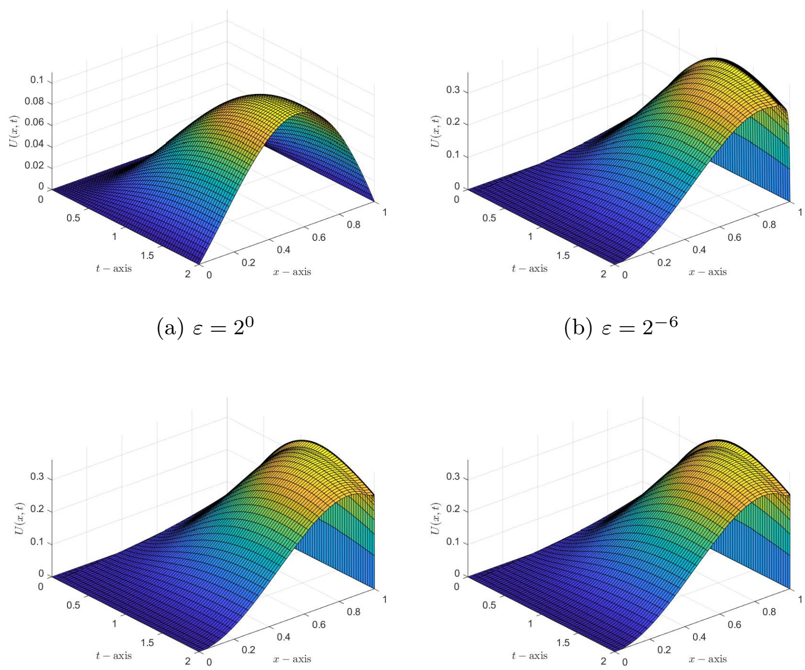
. The parameter uniform rate of convergence can be stated as

$$R^{N,M} = \max_{\varepsilon}\{RoC_\varepsilon^{N,M}\}$$

Example 1. Consider the time-fractional partial differential equation

$$D_t^{\gamma} y(s,t) - \varepsilon \frac{\partial^2 y(s,t)}{\partial s^2} + \left(2 - s^2\right) \frac{\partial y(s,t)}{\partial s} + ((s+1)(t+1))y(s,t)$$
$$= y(s,t-1) + 10t^2 \exp(-t) s(1-s),$$

on $(s,t) \in \Omega = (0,1) \times (0,\mathfrak{T}]$, with initial and boundary conditions $\varphi_b(s,t) = 0$, $\varphi_l(t) = 0$, and $\varphi_r(t) = 0$.



(a) $\varepsilon = 2^0$ (b) $\varepsilon = 2^{-6}$

(c) $\varepsilon = 2^{-10}$ (d) $\varepsilon = 2^{-20}$

Figure 1: A three-dimensional depiction of the numerical solution for Example 1 with various values of ε with $\gamma = 0.5$, $N_h = 32$, and $M_\tau = 40$.

Example 2. Consider the time-fractional partial differential equation

$$D_t^{\gamma} y(s,t) - \varepsilon \frac{\partial^2 y(s,t)}{\partial s^2} + \left(2 - s^2\right) \frac{\partial y(s,t)}{\partial s} + s y(s,t)$$
$$= y(s,t-1) + 10t^2 \exp(-t) s(1-s),$$

on $(s,t) \in \Omega = (0,1) \times (0,\mathfrak{T}]$, with initial and boundary conditions $\varphi_b(s,t) = 0$, $\varphi_l(t) = 0$, and $\varphi_r(t) = 0$.

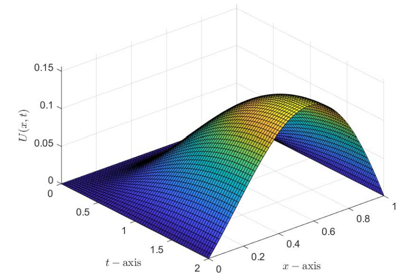
Table 1: The absolute maximum error and rate of convergence for Example 1 for various values of ε, with fix $\gamma = 0.5$

$N_h = M_\tau \Rightarrow$	32	64	128	256	512
$\varepsilon = 1$	7.2009e-03	3.9138e-03	2.0434e-03	1.0436e-03	5.3354e-04
	0.8796	0.9376	0.9693	0.9679	-
$\varepsilon = 2^{-4}$	1.3920e-02	8.8730e-03	4.8151e-03	2.4093e-03	1.1793e-03
	0.6497	0.8819	0.9990	1.0307	-
$\varepsilon = 2^{-8}$	8.6296e-02	5.0505e-02	2.7245e-02	1.3603e-02	6.3990e-03
	0.7729	0.8904	1.0020	1.0881	-
$\varepsilon = 2^{-12}$	4.3419e-03	1.9191e-03	9.1791e-04	4.6674e-04	2.3530e-04
	1.1779	1.0640	0.9757	0.9881	-
$\varepsilon = 2^{-16}$	1.0750e-01	6.3792e-02	3.5339e-02	1.8579e-02	9.5447e-03
	0.7529	0.8521	0.9276	0.9609	-
$\varepsilon = 2^{-20}$	1.0772e-01	6.3935e-02	3.5424e-02	1.8631e-02	9.5791e-03
	0.7526	0.8519	0.9270	0.9597	-
$\varepsilon = 2^{-24}$	1.0773e-01	6.3940e-02	3.5427e-02	1.8633e-02	9.5802e-03
	0.7526	0.8519	0.9270	0.9597	-
$\varepsilon = 2^{-30}$	1.0773e-01	6.3940e-02	3.5427e-02	1.8633e-02	9.5803e-03
	0.7526	0.8519	0.9270	0.9597	-
e^{N_h, M_τ}	1.0773e-01	6.3940e-02	3.5427e-02	1.8633e-02	9.5803e-03
R^{N_h, M_τ}	0.7526	0.8519	0.9270	0.9597	-

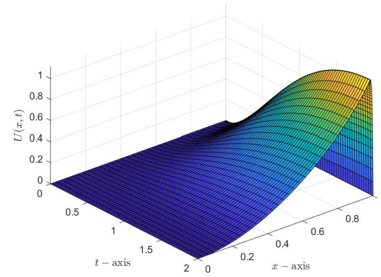
A boundary layer, as shown in Figures 1 and 2, is located at the right side of the space domain in the numerical solution of Examples 1 and 2 above. Figures 1 and 2 also display the computed solutions $W_{i,j}$ for various perturbation parameter values, along with the influence of fractional order. Figure 3 displays the log-log plots of the maximum absolute errors against the number of meshes for both cases, demonstrating the developed numerical scheme's convergent nature regardless of the perturbation value. The suggested scheme is ε-uniformly convergent, as illustrated by the numerical results shown in Tables 1 and 2, by combining the midpoint upwind numerical method in the spatial direction with the implicit Euler's method in the temporal direction. We can see that, for each value of ε, the maximum pointwise error decreases as N and M grow, from the results in Tables 1 and 2. It is evident that, for every N, M, the maximum point-wise error is $\varepsilon \longrightarrow 0$ stable. By utilizing these two examples, we verify that the suggested numerical technique is more accurate, stable, and ε-uniformly convergent, with a convergence rate that is almost one. The maximum point-wise error is sta-

Table 2: Maximum error and rate of convergence for various ε values with a fixed $\gamma = 0.5$ for Example 2.

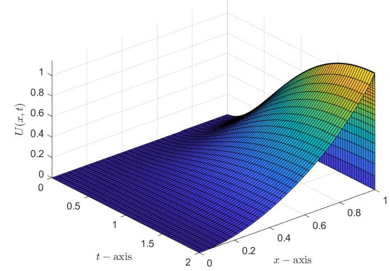
$N_h = M_\tau \Rightarrow$	32	64	128	256	512
$\varepsilon = 1$	4.6381e-03	2.0344e-03	9.4019e-04	4.5021e-04	2.2007e-04
	1.1889	1.1136	1.0623	1.0327	-
$\varepsilon = 2^{-4}$	4.7943e-02	2.3827e-02	1.2645e-02	6.8350e-03	3.6952e-03
	1.0088	0.9140	0.8875	0.8873	-
$\varepsilon = 2^{-8}$	6.3770e-02	2.8254e-02	1.2921e-02	6.1518e-03	3.0787e-03
	1.1744	1.1288	1.0706	0.9987	-
$\varepsilon = 2^{-10}$	6.6336e-02	3.0540e-02	1.3973e-02	6.7869e-03	3.3299e-03
	1.1191	1.1280	1.0418	1.0272	-
$\varepsilon = 2^{-12}$	6.7001e-02	3.1194e-02	1.4619e-02	6.9617e-03	3.4446e-03
	1.1029	1.0934	1.0703	1.0151	-
$\varepsilon = 2^{-14}$	6.7211e-02	3.1406e-02	1.4841e-02	7.1616e-03	3.5058e-03
	1.0976	1.0815	1.0512	1.0306	-
$\varepsilon = 2^{-20}$	6.7224e-02	3.1420e-02	1.4855e-02	7.1765e-03	3.5212e-03
	1.0973	1.0808	1.0496	1.0272	
$\varepsilon = 2^{-24}$	6.7225e-02	3.1421e-02	1.4856e-02	7.1775e-03	3.5222e-03
	1.0973	1.0807	1.0495	1.0270	-
$\varepsilon = 2^{-26}$	6.7225e-02	3.1421e-02	1.4856e-02	7.1775e-03	3.5223e-03
	1.0973	1.0807	1.0495	1.0270	-
$\varepsilon = 2^{-30}$	6.7225e-02	3.1421e-02	1.4856e-02	7.1775e-03	3.5223e-03
	1.0973	1.0807	1.0495	1.0270	-
e^{N_h,M_τ}	6.7225e-02	3.1421e-02	1.4856e-02	7.1775e-03	3.5223e-03
R^{N_h,M_τ}	1.0973	1.0807	1.0495	1.0270	-



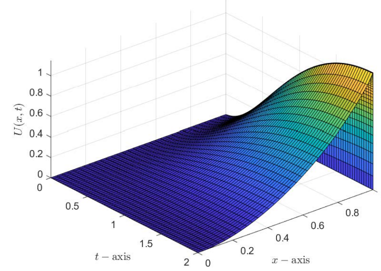
(a) $\varepsilon = 2^0$



(b) $\varepsilon = 2^{-6}$

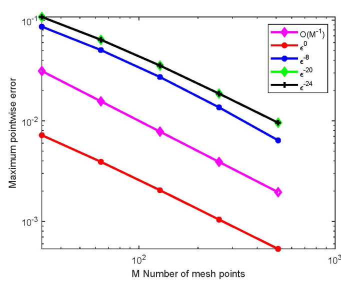


(c) $\varepsilon = 2^{-10}$

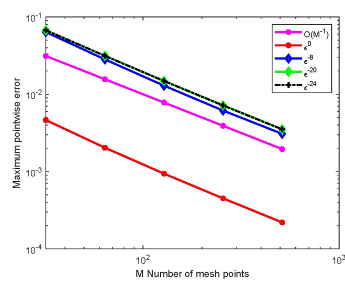


(d) $\varepsilon = 2^{-20}$

Figure 2: A three-dimensional representation of the numerical solution for Example 2 for various ε values with $\gamma = 0.5$, $M_x = 32$, and $M_t = 40$.



(a) Log-log plot of the max error for Example 1



(b) Log-log plot of the max error for Example 2

Figure 3: Log-log plot of maximum absolute errors for Examples 1 and 2 for different values of ε.

ble for all N, M. Using these two cases, we demonstrate that the proposed numerical technique is accurate, stable, and ε-uniformly convergent, with a convergence rate of almost one.

7 Conclusion

The time delay singularly perturbed parabolic convection-diffusion problem with time-fractional derivative order was solved using the midpoint upwind numerical approach. The solution to the problem depicted a boundary layer on the right side of the spatial domain. The solution's layer region has a steep gradient due to the presence of ε. Because of the frequently changing solution behavior in the layer region, it is computationally difficult to calculate the solution analytically or using traditional numerical methods. To mitigate this effect, we devised a strategy that employs the midway upwind scheme in the spatial direction and an implicit Euler's scheme in the temporal direction. The established numerical technique has been proven to be stable and uniformly convergent. To confirm the method's compatibility, two model problems were considered for numerical testing at varied perturbation parameters and fractional-order derivative values. Tables 1 and 2 provide a summary of the numerical results, including maximum absolute errors and numerical rate of convergence. Furthermore, the log-log plot (Figure 3) and the numerical solution for Examples 1 and 2 (refer to Figures 1 and 2) indicate the ε-uniform convergence of the scheme. The scheme's order of convergence is $O(N^{-1} + \tau)$, and it is uniformly convergent at ε. The devised scheme provides more stable, accurate, and uniformly convergent numerical results.

Acknowledgments

The authors would like to thank the anonymous reviewers for their helpful feedback and recommendations that helped to make the article more successful.

References

[1] Al-Mdallal, Q. and Syam, M. *An efficient method for solving non-linear singularly perturbed two points boundary-value problems of fractional order,* Commun. Nonlinear Sci. Numer. Simul. 17(6) (2012) 2299–2308.

[2] Atangana, A. and Goufo, E. *Extension of matched asymptotic method to fractional boundary layers problems,* Math. Probl. Eng. (2014) 2014.

[3] Bijura, A.M. *Nonlinear singular perturbation problems of arbitrary real orders,* 2003.

[4] Caputo, M. and Fabrizio, M. *A new definition of fractional derivative without singular kernel,* Progr. Fract. Differ. Appl. 1(2) (2015) 73–85.

[5] Cooke, K. *Differential—difference equations,* In International symposium on nonlinear differential equations and nonlinear mechanics, pages 155–171. Elsevier, 1963.

[6] Diekmann, O., Gils, S., Lunel, S. and Walther, H. Delay equations: functional-, complex-, and nonlinear analysis, volume 110. Springer Science & Business Media, 2012.

[7] Driver, R. *Ordinary and delay differential equations*, volume 20. Springer Science & Business Media, 2012.

[8] Gelu, F. and Duressa, G. *Hybrid method for singularly perturbed robin type parabolic convection–diffusion problems on shishkin mesh,* Partial Differ. Equ. Appl. Math. 8:100586, 2023.

[9] Govindarao, L. and Mohapatra, J. *A second order numerical method for singularly perturbed delay parabolic partial differential equation,* Eng. Comput. 36(2) (2018) 420–444.

[10] Hailu, W. and Duressa, G. Accelerated parameter-uniform numerical method for singularly perturbed parabolic convection-diffusion problems with a large negative shift and integral boundary condition. Result. Appl. Math. 18 (2023) 100364.

[11] Hailu, W. and Duressa, G. *Uniformly convergent numerical scheme for solving singularly perturbed parabolic convection-diffusion equations with integral boundary condition,* Differ. Equ. Dyn. Syst. (2023) 1–27 .

[12] Hailu, W. and Duressa, G. *A robust collocation method for singularly perturbed discontinuous coefficients parabolic differential difference equations,* Research in Mathematics, 11(1) (2024) 2301827.

[13] Hale, J. and Lunel, S. *Introduction to functional differential equations*, volume 99. Springer Science & Business Media, 2013.

[14] Hassen, Z. and Duressa, G. *Parameter uniform hybrid numerical method for time-dependent singularly perturbed parabolic differential equations with large delay,* Appl. Math. Sci. Eng. 32(1) (2024) 2328254.

[15] Kolmanovskii, V. and Myshkis, A. *Applied theory of functional differential equations*, volume 85. Springer Science & Business Media, 2012.

[16] Kolmanovskii, V. and Nosov, V. *Stability of functional differential equations*, volume 180. Elsevier, 1986.

[17] Kuang, Y. *Delay differential equations: with applications in population dynamics.* Academic press, 1993.

[18] Kumar, K. and Vigo-Aguiar, J. *Numerical solution of time-fractional singularly perturbed convection–diffusion problems with a delay in time,* Math. Method. Appl. Sci. 44(4) (2021) 3080–3097.

[19] Losada, J. and Nieto, J. *Properties of a new fractional derivative without singular kernel,* Progr. Fract. Differ. Appl. 1(2) (2015) 87–92.

[20] Meerschaert, M. and Tadjeran, C. *Finite difference approximations for two-sided space-fractional partial differential equations,* Applied numerical mathematics, 56(1) (2006) 80–90.

[21] Miller, J., O'riordan, E, and Shishkin, G. *Fitted numerical methods for singular perturbation problems: error estimates in the maximum norm for linear problems in one and two dimensions.* World scientific, 2012.

[22] Miller, P. *Applied asymptotic analysis*, volume 75. American Mathematical Soc., 2006.

[23] Negero, N. and Duressa, G. *A method of line with improved accuracy for singularly perturbed parabolic convection–diffusion problems with large temporal lag,* Results Appl. Math. 11 (2021) 100174.

[24] Negero, N. and Duressa, G. *An exponentially fitted spline method for singularly perturbed parabolic convection-diffusion problems with large time delay,* Tamkang J. Math., 54(4) (2023) 313–338.

[25] Nelson, P. and Perelson, A. *Mathematical analysis of delay differential equation models of HIV-1 infection,* Math.Biosci., 179(1) (2002) 73–94.

[26] Norkin, S. *Introduction to the theory and application of differential equations with deviating arguments.* Academic Press, 1973.

[27] Podlubny, I. *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications.* elsevier, 1998.

[28] Podlubny, I. *An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications,* Math. Sci. Eng. 198 (340) (1999) 0924–34008.

[29] Rangaig, N. and Pido, A. *Finite difference approximation method for two-dimensional space-time fractional diffusion equation using nonsingular fractional derivative.* Prog. Fract. Differ. Appl, 5(4) (2019) 1–11.

[30] Roop, J. *Numerical approximation of a one-dimensional space fractional advection–dispersion equation with boundary layer,* Comput. Math. Appl. 56(7) (2008) 1808–1819.

[31] Roos, H., Stynes, M. and Tobiska, L. *Robust numerical methods for singularly perturbed differential equations: convection-diffusion-reaction and flow problems*, volume 24. Springer Science & Business Media, 2008.

[32] Sadri, K. and Aminikhah, H. *An efficient numerical method for solving a class of variable-order fractional mobile-immobile advection-dispersion*

equations and its convergence analysis, Chaos, Solitons Fractals, 146 (2021) 110896.

[33] Sahoo, S. and Gupta, V. *A robust uniformly convergent finite difference scheme for the time-fractional singularly perturbed convection-diffusion problem,* Comput. Math. Appl. 137 (2023) 126–146.

[34] Sayevand, K. and Pichaghchi, K. *Efficient algorithms for analyzing the singularly perturbed boundary value problems of fractional order,* Commun. Nonlinear Sci. Numer. Simul., 57 (2018) 136–168.

[35] Sayevand, K. and Pichaghchi, K. *A novel operational matrix method for solving singularly perturbed boundary value problems of fractional multi-order,* Int. J. Comput. Math. 95(4) (2018) 767–796.

[36] Shakti, D., Mohapatra, J., Das, P. and Vigo-Aguiar, J. *A moving mesh refinement based optimal accurate uniformly convergent computational method for a parabolic system of boundary layer originated reaction–diffusion problems with arbitrary small diffusion terms,* J. Comput. Appl. Math. 404 (2022) 113167.

[37] Villasana, M. and Radunskaya, A. *A delay differential equation model for tumor growth,* J. Math. Bio. 47 (2003) 270–294.

[38] Yuste, S. and Acedo, L. *An explicit finite difference method and a new von neumann-type stability analysis for fractional diffusion equations,* SIAM J. Numer. Anal. 42(5) (2005) 1862–1874.

[39] Zhao, T. *Global periodic-solutions for a differential delay system modeling a microbial population in the chemostat,* J. Math. Anal. Appl. 193(1) (1995) 329–352.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1280–1308
https://doi.org/10.22067/ijnao.2024.89077.1484
https://ijnao.um.ac.ir/

Research Article

Advanced mathematical modeling and prognosication of regulated spatio-temporal dynamics of Monkeypox

M. Baroudi*,[1], I. Smouni[1], H. Gourram[1], A. Labzai[2] and M. Belam[1]

Abstract

This study explores a continuous spatio-temporal mathematical model to illustrate the dynamics of Monkeypox virus spread across various regions, considering both human and animal hosts. We propose a comprehensive

*Corresponding author

Received 25 July 2024; revised 30 August 2024; accepted 10 September 2024

Mohamed Baroudi, (m.mohamed.baroudi@gmail.com)
Imane Smouni, (imanesmouni23@gmail.com)
Hicham Gourram, (gourramhicham03@gmail.com)
Abderrahim Labzai, (labzaiabdo1977@gmail.com)
Mohamed Belam, (m.belam@gmail.com)

[1]Laboratory LMACS, Sultan Moulay Slimane University, MATIC Research Team: Applied Mathematics and Information and Communication Technologie, Department of Mathematics and Computer Science, Khouribga Polydisciplinary Faculty, Morocco.
[2]Laboratory of Analysis Modeling and Simulation, Department of Mathematics and Computer Science, Faculty of Sciences Ben M'sik, Hassan II University of Casablanca, Morocco.

How to cite this article
Baroudi, M., Smouni,I., Gourram,H., Labzai, A. and Belam, M., Advanced mathematical modeling and prognosication of regulated spatio-temporal dynamics of Monkeypox. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1280-1308. https://doi.org/10.22067/ijnao.2024.89077.1484

strategy that includes awareness campaigns, security measures, and health interventions in areas where the virus is prevalent. The goal is to reduce transmission between humans and animals, thereby decreasing human infections and eradicating the virus in animal populations. Our model, which integrates spatial variables, accurately reflects the geographical spread of the virus and the impact of interventions, followed by the implementation and analysis of an applicable optimal control problem. Optimal control theory methods are applied in this work to demonstrate the existence of optimal control and the necessary conditions for optimality. We conduct numerical simulations using **MATLAB** with the forward-backward sweep method, revealing the efficiency of strategies focused on protecting vulnerable populations, preventing contact with infected individuals and animals, and promoting the use of quarantine facilities as the most effective means to control the spread of the Monkeypox virus. Additionally, the study examines the socio-economic impacts of the virus and the benefits of timely intervention. This approach provides valuable insights for policymakers and public health officials in managing and controlling the spread of Monkeypox.

AMS subject classifications (2020): 49J15, 93C10, 92B05, 93A30.

Keywords: Monkeypox; Optimal control; Spatio-temporal model; Mathematical model; optimization .

1 Introduction

Monkeypox (MPX) is a rare and severe disease caused by a virus closely related to the smallpox virus (MPOX). It primarily affects animals and humans (anthropozoonosis) living in regions near the dense forests of Central and West Africa. Among the affected countries, the Democratic Republic of Congo (DRC) is particularly impacted, reporting nearly 85% of known human cases [3, 16]. The DRC has experienced several epidemics in recent years[9]. Currently, most researchers agree that after 30 years of cessation of smallpox vaccination campaigns, there has been a significant resurgence of MPX cases in several tropical regions, including the DRC. This resurgence is becoming a major public health concern.

Most researchers concur that the re-emergence of MPX is closely linked to the discontinuation of smallpox vaccination campaigns, leading to a significant increase in MPX cases over the last three decades in several tropical areas, particularly in the DRC. Despite the growing concern, MPX remains an understudied disease, with its transmission dynamics and spatial and temporal distribution not fully understood[4, 24, 25].

This research aims to contribute to understanding MPX by developing an estimator for MPX prevalence in the DRC and analyzing the determinants of its spatial and temporal distribution. Initially, a score was developed to assess the level of agreement between morbidity data reported by the integrated disease surveillance and response (IDSR) in the DRC and actual morbidity. Subsequently, spatial and temporal clusters of MPX were identified at the Health Zone (HZ) level using retrospective scan statistics. Finally, an investigation into the environmental factors associated with MPX occurrence in the DRC was conducted [4, 12, 21].

A straightforward and practical score was designed to measure the reliability of data produced by the IDSR in the DRC. The analysis of spatial clusters of MPX revealed higher reported case rates in the traditional hotspots of the Sankuru and Tshuapa districts, with the disease spreading to neighboring areas over the years. This pattern indicates the central Congo basin as a probable original epicenter for MPX and highlights its spread dynamics over two decades[4, 21].

Annual temporal analysis shows a seasonal trend with an increase in MPX cases during the dry season. The developed model indicates that several environmental factors are positively correlated with MPX incidence, although these factors alone do not fully explain the emergence and persistence of MPX epidemics in the DRC. These findings are critical as they emphasize the need to incorporate various environmental variables into analyses to achieve a more comprehensive understanding of disease dynamics [10, 12, 14].

Future models must include socio-economic and anthropological factors to better capture the complex interactions between humans and their environment, thus affecting their exposure and risk levels. These factors are essential for developing a nuanced model that can accurately predict MPX outbreak patterns and offer insights into effective intervention strategies[12, 20].

The spatial analysis in this study enabled the targeting of high-risk areas and periods for MPX in the DRC, facilitating the generation of vital spatio-temporal information required to prioritize prevention and control interventions against the disease[10, 12]. Identifying these high-risk zones and periods allows public health officials to allocate resources more effectively and implement targeted measures to curb the disease's spread.

The motivation for this study stems from the significant increase in MPX cases, particularly in the DRC, following the cessation of smallpox vaccination campaigns. Despite growing public health concerns, MPX remains an understudied disease, with its transmission dynamics and spatial-temporal distribution not fully understood. This study aims to fill this gap by developing an estimator for MPX prevalence and analyzing its spatial and temporal determinants in the DRC, as it represents an unexplored aspect in current studies.

Additionally, understanding the spatial distribution and the environmental and social factors contributing to the spread of MPX is crucial for developing comprehensive strategies to control and prevent future outbreaks. Incorporating spatial variables into models allows for a more detailed and accurate depiction of how the disease spreads across different regions, considering both natural and human-induced factors[10, 12, 14, 20].

These findings highlight the necessity of a nuanced approach to modeling the spatial and temporal dynamics of MPX. By including spatial data, we can better understand the localized patterns of disease spread and the impact of various environmental and socio-economic factors[1, 11]. This approach can inform more effective public health interventions and policies tailored to the specific conditions of different regions[6].

In addition to identifying high-risk areas, this study emphasizes the importance of continuous surveillance and data collection. High-resolution spatial data can reveal subtle changes in disease patterns and assist in the early detection of potential outbreaks. This proactive approach is essential for preventing the spread of MPX and other emerging infectious diseases[1, 6].

Furthermore, these results open avenues for investigation in other affected African countries. By leveraging spatio-temporal data, predictive models can be enhanced, and more effective public health policies can be developed,

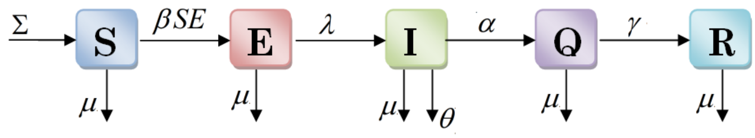


Figure 1: Model description.

tailored to the specific needs and conditions of each region. This approach not only improves our understanding of MPX but also provides a framework for studying other zoonotic diseases that threaten global health.

In conclusion, integrating spatial analysis in the study of MPX provides valuable insights into the disease's dynamics and guides the development of targeted interventions. Understanding the spatial and temporal aspects of disease spread allows for better preparation and response to future public health challenges[1, 6, 11].

2 Model formulation

We utilize the $SEIQR$ mathematical model developed by Imane Smouni et al. [18], which describes the spread of MPX.

In the remainder of this paper, we set $S = Z_1$, $E = Z_2$, $I = Z_3$, $Q = Z_4$, and $R = Z_5$.

Therefore, we introduce the mathematical model for MPX, which is described by the following system of differential equations:

$$\begin{cases} \frac{dZ_1(t)}{dt} = \Sigma - \beta Z_1(t)Z_2(t) - \mu Z_1(t), \\ \frac{dZ_2(t)}{dt} = \beta Z_1(t)Z_2(t) - (\lambda + \mu)Z_2(t), \\ \frac{dZ_3(t)}{dt} = \lambda Z_2(t) - (\alpha + \mu + \theta)Z_3(t), \\ \frac{dZ_4(t)}{dt} = \alpha Z_3(t) - (\gamma + \mu)Z_4(t), \\ \frac{dZ_5(t)}{dt} = \gamma Z_4(t) - \mu Z_5(t), \end{cases} \tag{1}$$

where the initial conditions $Z_i(0) \geq 0$ are nonnegative for each $i \in \{1, 2, 3, 4, 5\}$.

The various components of our spatio-temporal model are detailed in Table 1 below, which includes their descriptions (refer to [18]).

Table 1: The delineation of the distinct compartments within our model.

Compartment	Description
S	The potential number of individuals who may contract the virus.
E	The count of individuals infected with the virus who do not show symptoms.
I	The count of individuals who are infected and displaying symptoms.
Q	The count of individuals admitted to the hospital.
R	The count of individuals who have recovered.

The temporal dynamic system (1) clearly lacks the capability to comprehensively describe the spatial spread of the MPX virus. To address this shortcoming, we propose the integration of the Laplacian operator, which transforms it into a more complex model (2) that can account for spatial heterogeneity and diffusion. This transformation is typically achieved by adding a diffusion term, represented by the Laplacian, to the original differential equations of the model. This term allows the model to capture the rate at which diseases spread geographically, reflecting local interactions and movements within the population. Specifically, we augment the deterministic epidemic model for MPX with the following modification to include spatial diffusion effects. This adjustment will allow us to better capture the complexities of virus transmission across different regions, providing a more robust framework for understanding and predicting the spread of the disease. Hence

$$\begin{cases} \frac{dZ_1(t,y)}{dt} = d_{Z_1}\Delta Z_1(t,y) + \Sigma - \beta Z_1(t,y)Z_2(t,y) - \mu Z_1(t,y), \\ \frac{dZ_2(t,y)}{dt} = d_{Z_2}\Delta Z_2(t,y) + \beta Z_1(t)Z_2(t,y) - (\lambda+\mu)Z_2(t,y), \\ \frac{dZ_3(t,y)}{dt} = d_{Z_3}\Delta Z_3(t,y) + \lambda Z_2(t,y) - (\alpha+\mu+\theta)Z_3(t,y), \\ \frac{dZ_4(t,y)}{dt} = d_{Z_4}\Delta Z_4(t,y) + \alpha Z_3(t,y) - (\gamma+\mu)Z_4(t,y), \\ \frac{dZ_5(t,y)}{dt} = d_{Z_5}\Delta Z_5(t,y) + \gamma Z_4(t,y) - \mu Z_5(t,y). \end{cases} \tag{2}$$

The initial conditions and the no-flux boundary conditions are delineated as follows: Initially, the state of the system is defined by the values of the variables at time $t = 0$. Additionally, to ensure that there is no net flux of the variables across the boundaries of the domain, the no-flux boundary conditions are applied. These conditions are critical for accurately modeling the system's behavior and ensuring that the variables remain within the defined spatial boundaries throughout the simulation period. The precise mathematical expressions for these initial and boundary conditions are detailed below:

$$
\begin{aligned}
&Z_1(0,y) = Z_{1,0}, \qquad Z_2(0,y) = Z_{2,0}, \qquad Z_3(0,y) = Z_{3,0}, \\
&Z_4(0,y) = Z_{4,0}, \qquad Z_5(0,y) = Z_{5,0}, \qquad z \in \Omega.
\end{aligned} \tag{3}
$$

Here, Ω represents a bounded domain within $\mathbb{R}^2$ characterized by a smooth boundary denoted as $\partial\Omega$. For a comprehensive analysis, it is essential to incorporate the following equation into our model framework. This equation will account for the dynamic interactions within the defined spatial domain and ensure that our model accurately reflects the physical processes under consideration. The inclusion of this equation is crucial for enhancing the model's capability to simulate the behavior of the system accurately. Therefore,

$$
\frac{\partial Z_1}{\partial \zeta} = \frac{\partial Z_2}{\partial \zeta} = \frac{\partial Z_3}{\partial \zeta} = \frac{\partial Z_4}{\partial \zeta} = \frac{\partial Z_5}{\partial \zeta} = 0, \tag{4}
$$

where ζ denotes the outward unit normal vector on the boundary $\partial\Omega$, and $\dfrac{\partial}{\partial \zeta}$ signifies the outward normal derivative on $\partial\Omega$. Additionally, we define a control function $w : [0;1] \times \Omega \longrightarrow [0;t_f]$, aimed at reducing the number of infected individuals while increasing the number of susceptible individuals. This objective is achieved through strategies such as minimizing social interactions, launching safety campaigns to curb migration, and implementing health precautions to mitigate virus transmission.

Our goal is to balance the number of individuals affected by the disease with the cost of the treatment program. Therefore, we propose the following formulation: Minimizing of the cost functional represented by

$$
N(Z_1, Z_2, Z_3, Z_4, Z_5, w) = \int_\Omega \int_0^{t_f} a_1 Z_3(t,y)\,dt\,dy + \frac{b_1}{2} \left\| w(t,y) \right\|^2_{L^2([0,t_f])}, \tag{5}
$$

subject to the constraints of the controlled system (2).

Furthermore, we define W_{ad} as the set of permissible controls, given by

$$W_{ad} = \{w \in L^{\infty}(Q),\ 0 \leq w \leq 1 \quad a.e.\ \ on\, Q\}, \tag{6}$$

where $Q = [0, t_f] \times \Omega$, with Ω represents a bounded region within $\mathbb{R}^2$ that has a smooth boundary $\partial\Omega$. To ensure the model's comprehensiveness, we include detailed parameter descriptions, which can be found in Table 2.

In our approach, it is essential to consider not only the direct health impact but also the socio-economic factors influencing the disease dynamics. By incorporating these elements, we aim to develop a holistic and effective control strategy that can be practically implemented. Moreover, the integration of feedback mechanisms and continuous monitoring will enhance the adaptability and responsiveness of the control measures, ensuring optimal outcomes over time.

Table 2: The description of the parameters of our spatio-temporal model (2).

Symbol	Description
Σ	Recruitment rate
μ	Natural mortality rate
α	The rate at which a virus is transmitted from an asymptomatic infected individual to a susceptible person.
β	The rate of infection for individuals who are asymptomatic.
λ	The mortality rate of individuals infected with the virus.
γ	The mortality rate of a hospitalized individual infected with the virus.
θ	Rate of recovery
d_S	Spread of susceptible individuals
d_E	Spread of asymptomatic infected individuals
d_I	Spread of symptomatic infected individuals
d_Q	Dissemination of hospitalized patients
d_R	Spread of individuals who have recovered

3 Presence and singularity of a global solution

In this section, we demonstrate that our model (2) possesses a robust global solution. For more information, you can refer to the following references: [2, 15, 17, 22, 23].

We examine $M(\Omega) = (L_2(\Omega))^5$, $H^1(\Omega) = \left\{ w \in L_2(\Omega) : \dfrac{\partial w}{\partial x} \in L_2(\Omega), \dfrac{\partial w}{\partial y} \in L_2(\Omega) \right\}$, and $H^2(\Omega) = \left\{ w \in H^1(\Omega) : \dfrac{\partial^2 w}{\partial y^2}, \dfrac{\partial^2 w}{\partial y^2}, \dfrac{\partial^2 w}{\partial x \partial y}, \dfrac{\partial^2 w}{\partial y \partial x} \in L_2(\Omega) \right\}$ the Hilbert spaces.

Let $L^2\left(0, t_f; H^2(\Omega)\right)$ be the space of all strongly measurable functions $\zeta : [0, t_f] \longmapsto H^2(\Omega)$ such that

$$\int_0^{t_f} \|\zeta(t,y)\|_{H^2(\Omega)} \mathrm{d}t < \infty. \tag{7}$$

Furthermore, we introduce the space $L^\infty(0, t_f; H^1(\Omega))$, which consists of all functions υ mapping from the interval $[0, t_f]$ to the Sobolev space $H^1(\Omega)$ [5, 8]. These functions must satisfy the following conditions:

$$\sup_{t \in [0, t_f]} \left(\|\zeta(t,y)\|_{H^1(\Omega)} \right) \prec \infty. \tag{8}$$

The norm within the space $L^\infty\left(0, t_f; H^1(\Omega)\right)$ is defined as follows:

$$\|\zeta\|_{L^\infty(0,t_f;H^1(\Omega))} = \inf\left\{ m \in \mathbb{R}_+ : \|\zeta(t,z)\|_{H^1(\Omega)} < m \right\}. \tag{9}$$

Our spatio-temporal model given by (2) can be represented in the following form:

$$\frac{\partial z(t,y)}{\partial t} = Bz(t,y) + g(t, z(t,y)). \tag{10}$$

In the context, $z = (z_1, z_2, z_3, z_4, z_5) = (S, E, I, Q, R)$ and $g = (g_1, g_2, g_3, g_4, g_5)$ is specified as follows:

$$\begin{cases} g_1 = \Sigma - (1-w)\beta z_1 z_2 - \mu z_1, \\ g_2 = (1-w)\beta z_1 z_2 - (\lambda + \mu) z_2, \\ g_3 = \lambda z_2 - (\alpha + \mu + \theta) z_3, \\ g_4 = \alpha z_3 - (\gamma + \mu) z_4, \\ g_5 = \gamma z_4 - \mu z_5. \end{cases} \tag{11}$$

For every i in the set $\{1,2,3,4,5\}$., we have

$$\frac{\partial z_i}{\partial t} = d_i \Delta z_i + g_i(z(t,y)). \tag{12}$$

Define B as the linear operator with its domain $D(B)$, which is a subset of $M(\Omega)$. This operator is characterized by the following formula:

$$Bz = (d_S \Delta z_1, d_E \Delta z_2, d_I \Delta z_3, d_Q \Delta z_4, d_R \Delta z_5) \tag{13}$$

with

$$z \in D(B) = \Big\{ z = (z_1, z_2, z_3, z_4, z_5) \in \left(H^2(\Omega)\right)^5 : \\ \frac{\partial z_1}{\partial \zeta} = \frac{\partial z_2}{\partial \zeta} = \frac{\partial z_3}{\partial \zeta} = \frac{\partial z_4}{\partial \zeta} = \frac{\partial z_5}{\partial \zeta} = 0 \quad \text{on} \quad \partial\Omega \Big\}. \tag{14}$$

Theorem 1. Consider a bounded domain Ω within $\mathbb{R}^2$ that has a smoothly defined boundary. Assume that the initial values z_i^0 are nonnegative across Ω for each $i \in \{1,2,3,4,5\}$, and that the parameters $\Sigma, \mu, \alpha, \beta, \lambda, \gamma$, and θ are also nonnegative. Given an admissible control $w \in W_{ad}$ and an initial condition $z^0 \in D(A)$, the system described in (2) is guaranteed to have a unique strong nonnegative solution z belonging to $z \in W^{1,2}([0,t_f]; M(\Omega))$. This solution adheres to the following conditions:

$$z_1, z_2, z_3, z_4, z_5 \in L^2\left(0, t_f; H^2(\Omega)\right) \cap L^\infty\left(0, t_f; H^1(\Omega)\right) \cap L^\infty(Q). \tag{15}$$

Additionally, there is a constant $M_0 > 0$, which does not depend on the control v, such that for every $t \in [0, t_f]$ and for each $i \in \{1,2,3,4,5\}$,

$$\left\| \frac{\partial z_i}{\partial t} \right\|_{L^2(Q)} + \|z_i\|_{L^2(0,t_f,H^2(\Omega))} + \|z_i\|_{H^1(\Omega)} + \|z_i\|_{L^\infty(Q)} \le M_0. \tag{16}$$

The symbols L^2, L^∞, and $\|\cdot\|$ are used without definitions. These symbols are standard notations in functional analysis as follows:

- L^2: The space of square-integrable functions, that is, functions f such that $\int |f(x)|^2 \, dx < \infty$.
- L^∞: The space of essentially bounded functions, that is, functions f such that there exists a bound M where $|f(x)| \le M$ almost everywhere.
- $\|\cdot\|$: Typically denotes a norm, with $\|f\|_2$ indicating the L^2 norm and $\|f\|_\infty$ indicating the L^∞ norm.

Proof. Since the operator Δ exhibits dissipative characteristics, is self-adjoint, and can generate a C_0 semigroup of contractions on $M(\Omega)$ [13], it follows that the function g represented by $(g_1, g_2, g_3, g_4, g_5)$ is Lipschitz continuous with respect to the variable z denoted by $(z_1, z_2, z_3, z_4, z_5)$, and this continuity is uniform in t over the interval $[0, t_f]$. Consequently, the system admits a unique strong solution $z \in W^{1,2}([0, t_f]; M(\Omega))$ with

$$z_i \in L^2\left(0, t_f; H^2(\Omega)\right) \quad \text{for all} \quad i \in \{1, 2, 3, 4, 5\}. \tag{17}$$

We now demonstrate that for all $i \in \{1, 2, 3, 4, 5\}, z_i \in L^\infty(Q)$.

Set $c = \max\left\{\|g_i\|_{L^\infty(Q)}, \left\|z_i^0\right\|_{L^\infty(\Omega)} : i \in \{1, 2, 3, 4, 5\}\right\}$ and define

$$V_i(t, y) = z_i(t, y) - ct - \left\|z_i^0\right\|_{L^\infty(\Omega)}.$$

It is evident that V_i satisfies the system

$$\begin{cases} \dfrac{\partial V_i(t, y)}{\partial t} = d_i \Delta V_i(t, y) + g_i(t, z(t, y)) - c, \quad t \in [0, t_f], \\ V_i(0, y) = z_i^0 - \left\|z_i^0\right\|_{L^\infty(\Omega)}. \end{cases} \tag{18}$$

This system admits a unique strong solution given by

$$V_i(t, y) = \Gamma(t)\left(z_i^0 - \left\|z_i^0\right\|_{L^\infty(\Omega)}\right) + \int_0^t \Gamma(t - x)\left(g_i(z(x)) - c\right) dx, \tag{19}$$

where $\Gamma(t)$ is an infinitesimal semigroup associated with the operator $d_i \Delta$.

It follows that $V_i(t, y) \leq 0$, so $z_i \leq ct + \left\|z_i^0\right\|_{L^\infty(\Omega)}$.

Similarly, we can show that $V_i(t, y) = z_i(t, y) + ct + \left\|z_i^0\right\|_{L^\infty(\Omega)}$ is nonnegative. Thus, $z_i \geq -ct - \left\|z_i^0\right\|_{L^\infty(\Omega)}$, which implies

$$|z_i(t, y)| \leq ct + \left\|z_i^0\right\|_{L^\infty(\Omega)}, \tag{20}$$

and hence

$$z_i \in L^\infty(Q) \quad \text{for all } (t, y) \in [0, t_f] \times \Omega, \quad \text{for } i \in \{1, 2, 3, 4, 5\}. \tag{21}$$

Next, we establish that $z_i \in L^\infty\left(0, t_f; H^1(\Omega)\right)$ for all $i \in \{1, 2, 3, 4, 5\}$. Considering $i \in \{1, 2, 3, 4, 5\}$, we start from the equation

$$\frac{\partial z_i(t, y)}{\partial t} - d_i \Delta z_i(t, y) = g_i(t, z(t, y)), \quad (t, y) \in [0, t_f] \times \Omega. \tag{22}$$

We have

$$\int_0^t \int_\Omega \left(\frac{\partial z_i(t, y)}{\partial t} - d_i \Delta z_i(t, y)\right)^2 dyds = \int_0^t \int_\Omega \left(g_i(t, z(t, y))\right)^2 dyds. \tag{23}$$

Using Green's formula, we obtain

$$\begin{aligned}
&\int_0^t \int_\Omega \left(\frac{\partial z_i}{\partial t}\right)^2 dyds + d_i^2 \int_0^t \int_\Omega (\Delta z_i)^2 \, dyds \\
&\quad = 2d_i \int_0^t \int_\Omega \frac{\partial z_i}{\partial t} \times \Delta z_i dyds + \int_0^t \int_\Omega (g_i(t, z_i))^2 \, dyds \\
&\quad = d_i \int_\Omega |\nabla z_i^0|^2 dy - d_i \int_\Omega |\nabla z_i|^2 dy + \int_0^t \int_\Omega (g_i(t, z_i))^2 \, dyds.
\end{aligned}$$

Since $z_i^0 \in H^2(\Omega)$ and $\|z_i\|_{L^\infty(Q)}$ are bounded independently of v, it follows that

$$z_i \in L^\infty \left(0; t_f; H^1(\Omega)\right), \quad \text{for } i \in \{1, 2, 3, 4, 5\}. \tag{24}$$

Combining (17), (21), and (24), we conclude that the inequality presented in (16) holds. Furthermore, using arguments similar to those employed for the Field–Noyes equations in [19], we deduce that the solution $(z_1, z_2, z_3, z_4, z_5)$ remains nonnegative. Consider the set

$$\Theta = \{(z_1, z_2, z_3, z_4, z_5) : 0 \le z_i \le D \text{ for } i \in \{1, 2, 3, 4, 5\}\},$$

and define the convex functions G_i over Π as $G_i(z_1, z_2, z_3, z_4, z_5) = -z_i$. This leads to the following relationships:

$$\begin{aligned}
\nabla(G_1) \cdot g|_{z_1=0} &= \nabla(-z_1) \cdot g|_{z_1=0} = -\Sigma \le 0, \\
\nabla(G_2) \cdot g|_{z_2=0} &= \nabla(-z_2) \cdot g|_{z_2=0} = 0 \le 0, \\
\nabla(G_3) \cdot g|_{z_3=0} &= \nabla(-z_3) \cdot g|_{z_3=0} = -\lambda z_2 \le 0, \\
\nabla(G_4) \cdot g|_{z_4=0} &= \nabla(-z_4) \cdot g|_{z_4=0} = -\alpha z_3 \le 0, \\
\nabla(G_5) \cdot g|_{z_5=0} &= \nabla(-z_5) \cdot g|_{y_5=0} = -\gamma z_4 \le 0.
\end{aligned} \tag{25}$$

As demonstrated in [6], the set Θ is positively invariant, ensuring the nonnegativity of the solutions over time. □

4 Existence of an optimal control

In this section, our primary goal is to establish the existence of an optimal control for the problem described in (5). This problem is constrained by the reaction-diffusion system detailed in (2)–(4), with the control variable belonging to the admissible set $w \in W_{ad}$. Our objective is to show that under these constraints, there exists a control that optimizes the given performance criterion. The main result we intend to demonstrate in this section is the following key theorem, which provides a rigorous foundation for the existence of such an optimal control. By proving this theorem, we will significantly advance our understanding of the underlying reaction-diffusion system and its controllability properties.

Theorem 2. Based on the conditions specified in Theorem 1, the optimal control problem described in (2)-(4) admits a solution, which we denote by (z^*, w^*). This pair represents the optimal state and control variables that satisfy the given system and minimize the cost function.

Proof. Since both the control function w and the state variables z_i (for i in the set $\{1, 2, 3, 4, 5\}$) are uniformly bounded in the space $L^\infty(Q)$, it is ensured that the infimum of the cost function exists. Let us denote this infimum by C^*, which is defined as $C^* = \inf_{w \in W_{ad}} C(z, w)$.

Now, consider a sequence $\{w_m\} \subset W_{ad}$ that minimizes the cost function. Specifically, this means that as m approaches infinity, the cost function evaluated at the sequence converges to the infimum: $\lim\limits_{n \to +\infty} C(z^m, w_m) = C^*$.

Here, $(z_1^m, z_2^m, z_3^m, z_4^m, z_5^m)$ represents the solution of the given system of equations (denoted by (2)–(4)) when the control input is u_n. This scenario leads to the formulation of the following system of equations:

$$\begin{cases} \frac{\partial z_1^m(t,y)}{\partial t} = d_S \Delta z_1^m(t, y) + \Sigma - (1 - v_m)\beta z_1^m(t, y) z_2^m(t, y) - \mu z_1^m(t, y), \\ \frac{\partial z_2^m(t,y)}{\partial t} = d_E \Delta z_2^m(t, y) + (1 - v_m)\beta z_1^m(t) z_2^m(t, y) - (\lambda + \mu) z_2^m(t, y), \\ \frac{\partial z_3^m(t,y)}{\partial t} = d_I \Delta z_3^m(t, y) + \lambda z_2^m(t, y) - (\alpha + \mu + \theta) z_3^m(t, y), \\ \frac{\partial z_4^m(t,y)}{\partial t} = d_Q \Delta z_4^m(t, y) + \alpha z_3^m(t, y) - (\gamma + \mu) z_4^m(t, y), \\ \frac{\partial z_5^m(t,y)}{\partial t} = d_R \Delta z_5^m(t, y) + \gamma z_4^m(t, y) - \mu z_5^m(t, y). \end{cases} \tag{26}$$

In the region Q, the partial derivatives with respect to ζ are zero for all state variables, specifically:

$$\frac{\partial z_1^m}{\partial \zeta} = \frac{\partial z_2^m}{\partial \zeta} = \frac{\partial z_3^m}{\partial \zeta} = \frac{\partial z_4^m}{\partial \zeta} = \frac{\partial z_5^m}{\partial \zeta} = 0. \tag{27}$$

This condition indicates that the state variables do not change with respect to the variable v within the domain Q.

Given that the Sobolev space $H^1(\Omega)$ is compactly embedded in the Lebesgue space $L^2(\Omega)$, we can deduce that the sequence $z_i^m(t,y)$ possesses compactness properties in $L^2(\Omega)$ for each i in the set $\{1,2,3,4,5\}$. This compact embedding ensures that sequences that are bounded in $H^1(\Omega)$ have subsequences that converge in $L^2(\Omega)$.

Additionally, it is necessary to establish the equicontinuity of the sequence $\{z_i^m(t,y)\}_{m\geq 1}$ in the space $C([0,t_f],L^2(\Omega))$. Equicontinuity implies that the variations in the sequence y_k^n over time are uniformly controlled, ensuring that the sequence behaves regularly over the interval $[0,t_f]$. The boundedness of the time derivatives $\frac{\partial z_i^m}{\partial t}$ in $L^2(Q)$ implies that there exists a positive constant h such that the difference in the integrals of the squared state variables over the domain Ω at two different times t and x is bounded by h times the absolute difference between t and x:

$$\left|\int_\Omega (z_i^m)^2(t,y)\,dz - \int_\Omega (z_i^m)^2(x,y)\,dz\right| \leq h|t-x|. \tag{28}$$

This bound indicates that the changes in the state variables over time are controlled, which is a key aspect of equicontinuity.

Therefore, by applying the Ascoli–Arzelà Theorem, which provides criteria for the compactness of a set of continuous functions, we can confirm that the sequence y_i^m is compact in $C([0,t_f],L^2(\Omega))$. As a result, z_i^m converges uniformly to a limit function z_i^* in $L^2(\Omega)$ with respect to time t. This uniform convergence implies that the difference between z_i^m and y_k^* becomes arbitrarily small as n increases, uniformly for all t in $[0,t_f]$.
Given the boundedness of Δz_i^m in $L^2(Q)$, there exists a subsequence, still denoted by Δz_i^m, that converges weakly in $L^2(Q)$. This weak convergence implies that for any test function φ in the space of distributions, the following integral equality holds:

$$\int_Q \varphi\Delta z_i^m = \int_Q z_i^m \Delta\varphi \to \int_Q z_i^* \Delta\varphi = \int_Q \varphi\Delta z_i^*. \tag{29}$$

This result indicates that Δz_i^m converges to Δz_i^* weakly in $L^2(Q)$. In addition to this, we can also deduce the weak convergence of the time derivatives and the state variables themselves: $\frac{\partial z_i^m}{\partial t}$ converges weakly to $\frac{\partial z_i^*}{\partial t}$ in $L^2(Q)$, and z_i^m converges weakly to z_i^* in both $L^2(0,t_f;H^2(\Omega))$ and $L^\infty(0,t_f;H^1(\Omega))$.

Moreover, considering the product of state variables, we observe that

$$z_1^m z_2^m - z_1^* z_2^* = (z_1^m - z_1^*) z_2^m + z_1^* (z_2^m - z_2^*). \tag{30}$$

This decomposition allows us to deduce that $z_1^m z_2^m$ converges to $z_1^* z_2^*$ in $L^2(Q)$. Consequently, the control sequence w_m converges to w^* in $L^2(Q)$. Since the admissible control set W_{ad} is closed by definition, it follows that w^* also belongs to W_{ad}.

By taking the limit as $m \to \infty$ in (26), we can demonstrate that z^* is indeed a solution to (12) associated with the optimal control w^*. Thus, the optimal cost function value can be expressed as

$$\begin{aligned} C(z^*, w^*) &= \int_0^{t_f} a z_3^*(t,y)\, dt\, dy + \frac{b}{2} \left\| w^*(t,y) \right\|_{L^2(Q)}^2 \\ &\le \liminf \int_0^{t_f} a z_3^m(t,y)\, dt\, dy + \frac{b}{2} \left\| w^m(t,y) \right\|_{L^2(Q)}^2 \\ &\le \lim \int_0^{t_f} a z_3^m(t,y)\, dt\, dy + \frac{b}{2} \left\| w^m(t,y) \right\|_{L^2(Q)}^2 = C^*. \end{aligned} \tag{31}$$

This chain of inequalities confirms that the cost function C attains its minimum value at the optimal pair (z^*, w^*). Thus, (z^*, w^*) is indeed the optimal solution to the given control problem. $\square$

5 Necessary optimality conditions

In this segment, we will explore the necessary conditions for optimality for the problem (2)–(5) with $w \in W_{ad}$. Our goal is to provide comprehensive insights into the nature and characterization of the optimal control strategy. Let us consider an optimal pair represented by (z^*, w^*). To investigate the behavior of the system under slight perturbations, we introduce a perturbed control w^ϵ defined as $w^* + \epsilon w$, where $\epsilon > 0$. This perturbed control must satisfy the conditions $w \in L^2(Q)$ and $w \in W_{ad}$.

The corresponding solutions to (2) associated with the controls w^ϵ and w^* are denoted by $z^\epsilon = (z_1^\epsilon, z_2^\epsilon, z_3^\epsilon, z_4^\epsilon, z_5^\epsilon)$ and $z^* = (z_1^*, z_2^*, z_3^*, z_4^*, z_5^*)$, respectively. These solutions represent the state of the system under the perturbed and optimal controls.

To proceed, we define the matrices F and B, which will help us in formulating the conditions for optimality:

$$F = \begin{pmatrix} -\mu-(1-w^*)\beta z_2 & -(1-w^*)\beta z_1 & 0 & 0 & 0 \\ (1-w^*)\beta z_2 & (1-w^*)\beta z_1-(\lambda+\mu) & 0 & 0 & 0 \\ 0 & \lambda & -(\alpha+\mu+\theta) & 0 & 0 \\ 0 & 0 & \alpha & -(\gamma+\mu) & 0 \\ 0 & 0 & 0 & \gamma & -\mu \end{pmatrix}$$

and

$$B = \begin{pmatrix} \beta z_1^* z_2^* \\ -\beta z_1^* z_2^* \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The matrix F encapsulates the dynamics of the system influenced by the optimal control w^*, while the vector B captures specific interactions within the system states z_i^*.

By investigating the differences between the perturbed state z^ϵ and the optimal state z^*, we derive necessary conditions for optimality. This involves analyzing how the state z^ϵ changes as ϵ varies and identifying the impact of the perturbation on the system's performance.

We can calculate the derivative of the performance index with respect to ϵ and set it to zero to find the conditions for optimality. This process ensures that we identify the optimal control strategy that either minimizes or maximizes the desired performance index.

Moreover, understanding the system's behavior under slight variations in control allows us to refine our control strategies, ensuring they are both effective and efficient. Through this rigorous analysis, we aim to provide a deeper understanding of the optimal control mechanisms and their implications for the system's behavior.

In conclusion, our detailed examination provides critical insights into the optimal control strategy, enhancing the development of more effective control methodologies for complex systems.

Theorem 3. The mapping $y : W_{ad} \to W^{1,2}([0,t_f], M(\Omega))$ for $i \in \{1,2,3,4,5\}$ is Gateaux differentiable with respect to w^*. For any $w \in W_{ad}$, the derivative of z_i at w^*, represented by $z_i'(w^*)w$, is denoted by Z_i. Thus, $Z = (Z_1, Z_2, Z_3, Z_4, Z_5)$ is the unique solution to the differential equation given by

$$\frac{\partial Z}{\partial t} = AZ + FZ + wB \quad \text{with initial condition} \quad Z(0,y) = 0. \tag{32}$$

Proof. We introduce $Z_i^\varepsilon = \frac{z_i^\varepsilon - z_i^*}{\varepsilon}$ for $i \in \{1,2,3,4,5\}$. By subtracting the differential equations corresponding to z_i^ε and z_i^*, we obtain the following equation:

$$\frac{\partial Z^\varepsilon}{\partial t} = AZ^\varepsilon + F^\varepsilon Z^\varepsilon + wB \quad \text{subject to} \quad Z^\varepsilon(0,y) = 0 \quad \text{for all} \quad y \in \Omega, \tag{33}$$

where

$$F^\varepsilon = \begin{pmatrix} -\mu - (1-w^\varepsilon)\beta z_2^* & -(1-w^\varepsilon)\beta z_1^* & 0 & 0 & 0 \\ (1-w^\varepsilon)\beta z_2^* & (1-w^*)\beta z_1^* - (\lambda+\mu) & 0 & 0 & 0 \\ 0 & \lambda & -(\alpha+\mu+\theta) & 0 & 0 \\ 0 & 0 & \alpha & -(\gamma+\mu) & 0 \\ 0 & 0 & 0 & \gamma & -\mu \end{pmatrix}.$$

Let $(\Gamma(t), t \geq 0)$ be the semigroup generated by A. The solution to this system can be expressed as

$$Z^\varepsilon(t,y) = \int_0^t \Gamma(t-x) F^\varepsilon Z^\varepsilon(x,y)\, dx + \int_0^t \Gamma(t-x) wP\, dx. \tag{34}$$

Given that the coefficients of the matrix F^ε are uniformly bounded with respect to ε, we can use Grönwall's inequality to show that Z_i^ε is bounded in $L^2(Q)$. Consequently, z_i^ε converges to z_i^* in $L^2(Q)$. Taking the limit as $\varepsilon \to 0$, we get the following result:

To ensure the solution's stability and uniqueness, we impose that the matrix A is dissipative. This property allows the use of semigroup theory to handle the evolution of Z^ε over time. The boundedness of F^ε coefficients further guarantees the application of Grönwall's inequality, leading to the necessary boundedness in $L^2(Q)$.

Moreover, the convergence of z_i^ε to z_i^* is uniform, meaning that the error between z_i^ε and z_i^* diminishes uniformly as ε approaches zero. This uniform convergence is critical for the robustness of our solution, ensuring that our model accurately predicts the behavior of the system under various perturbations.

Thus, by analyzing the behavior of Z^ε through the semigroup approach and leveraging the boundedness properties of the involved matrices, we achieve a comprehensive understanding of the system's dynamics and their stability over the defined domain. Hence

$$\frac{\partial Z}{\partial t} = AZ + FZ + wB \quad \text{subject to} \quad Z(0, y) = 0, \quad \text{for all} \quad y \in \Omega \tag{35}$$

By using a similar method, we conclude that Z_i^ε converges to Z_i^* as ε approaches 0. □

Let us introduce the adjoint variable $q = (q_1, q_2, q_3, q_4, q_5)$ and denote F^* as the adjoint of the Jacobian matrix F. The dual system corresponding to our problem can be expressed as:

$$-\frac{\partial q}{\partial t} - Aq - F^* q = D^* D\psi \quad \text{with the condition} \quad q(t_f, y) = 0, \tag{36}$$

where

$$D = \begin{pmatrix} 0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0 \\ 0\,0\,1\,0\,0 \\ 0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0 \end{pmatrix} \quad \text{and} \quad \psi = \begin{pmatrix} 0 \\ 0 \\ a \\ 0 \\ 0 \end{pmatrix}.$$

Lemma 1. Under the assumptions of Theorem 1, let (z^*, w^*) be an optimal pair. There is a unique strong solution for the dual system (36), represented by $q \in W^{1,2}([0, t_f], M(\Omega))$. Specifically, the components q_i are elements of $L^2\left(0, t_f; H^2(\Omega)\right)$ and $L^\infty\left(0, t_f; H^1(\Omega)\right)$, for $i \in \{1, 2, 3, 4, 5\}$. Therefore,

$$\text{for all } i \in \{1, 2, 3, 4, 5\}, \quad q_i \in L^2(0, t_f; H^2(\Omega)) \cap L^\infty(0, t_f; H^1(\Omega)). \tag{37}$$

This uniqueness guarantees the stability of the solution within the prescribed function spaces. It also implies that the solution q maintains the required regularity, ensuring accurate representation and predictability of the system's behavior. The regularity of q_i provides significant insights into the system's dynamics and is crucial for the development of effective control strategies.

Proof. This lemma can be derived by applying a variable transformation $t' = t_f - t$ and following a comparable approach as utilized in the proof of Theorem 3. □

By utilizing the connections between the state and adjoint equations, the objective functional, and conventional optimality methods, we can comprehensively define the optimal control. This process allows us to derive the necessary conditions required to solve the optimal control problem effectively.

Theorem 4. If an optimal control w^* exists along with its corresponding solution $z^* \in W^{1,2}([0,t_f]; M(\Omega))$, then the optimal control w^* can be represented as:

$$w^* = \min\left(w_{\max}, \max\left(0, \frac{(q_2 - q_1)\,\beta z_1^* z_2^*}{b}\right)\right). \tag{38}$$

Proof. Let (z^*, w^*) be an optimal pair. Suppose $w^\varepsilon = w^* + \varepsilon h \in W_{ad}$ where $h \in L^2(\Omega)$, and let z^ε be the associated state solution. Then, we have

$$\begin{aligned}
C'(w^*)(h) &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left(C(w^\varepsilon) - C(w^*)\right)\\
&= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left(a\int_0^{t_f}\int_\Omega \left(z_3^\varepsilon - z_3^*\right) dydt\right.\\
&\quad \left.+\frac{b_1}{2}\int_0^1\int_\Omega \left((w^\varepsilon)^2 - (w^*)^2\right) dydt\right)\\
&= \lim_{\varepsilon\to 0}\left(a\int_0^{t_f}\int_\Omega \left(\frac{z_3^\varepsilon - z_3^*}{\varepsilon}\right) dydt\right.\\
&\quad \left.+\frac{b_1}{2}\int_0^1\int_\Omega \left(2hw^* + \varepsilon h^2\right) dydt\right).
\end{aligned}$$

Given that $\lim_{\varepsilon\to 0} \dfrac{z_3^\varepsilon - z_3^*}{\varepsilon} = \lim_{\varepsilon\to 0} \dfrac{z_3(u^* + \varepsilon h) - z_3^*}{\varepsilon} = Z_3$, $\lim_{\varepsilon\to 0} z_3^\varepsilon = z_3^*$, and $z_3^\varepsilon, z_3^* \in L^\infty(Q)$, it follows that C is Gateaux differentiable with respect to w^*, and we obtain

$$\begin{aligned}
C'(w^*)(h) &= \int_0^{t_f}\int_\Omega aZ_3 dydt + b_1\int_0^{t_f}\int_\Omega hw^* dydt\\
&= \int_0^{t_f} \langle D\psi, DZ\rangle dt + \int_0^1 \langle b_1 w^*, h\rangle_{L^2(\Omega)} dt.
\end{aligned}$$

If we set $h = v - w^*$, then we have

$$C'(w^*)(v - w^*) = \int_0^{t_f} \langle D\psi, DY\rangle dt + \int_0^1 \langle b_1 w^*, v - w^*\rangle_{L^2(\Omega)} dt.$$

Considering that

$$\begin{aligned}
\int_0^{t_f} \langle D\psi, DZ\rangle dt &= \int_0^{t_f} \langle D^* D\psi, Z\rangle \, dt \\
&= \int_0^{t_f} \left\langle -\frac{\partial q}{\partial t} - Aq - F^* q, Z \right\rangle dt \\
&= \int_0^{t_f} \left\langle q, \frac{\partial Z}{\partial t} - AZ - FZ \right\rangle dt \\
&= \int_0^{t_f} \langle q, B(v - w^*)\rangle \, dt \\
&= \int_0^{t_f} \langle B^* q, v - w^*\rangle_{L^2(\Omega)} \, dt,
\end{aligned}$$

and since W_{ad} is convex, we have $C'(w^*)(v - w^*) \geq 0$ for every $w \in W_{ad}$, which is equivalent to

$$\int_0^{t_f} \langle B^* q + b_1 w^*, v - w^*\rangle_{L^2(\Omega)} \, dt \geq 0 \text{ for every } w \in W_{ad}.$$

Thus, $bw^* = -B^* q$, which implies $w^* = \frac{(q_2 - q_1)\beta z_1^* z_2^*}{b_1}$. Given that $w^* \in W_{ad}$, it follows that (38) holds.

This result is important as it provides the explicit form of the optimal control u^* in terms of the adjoint variables q_1 and q_2, as well as the state variables z_1^*, z_2^*, and z_3^*. This formulation allows for a deeper understanding of the interaction between the state and adjoint systems in the context of the optimal control problem. The conditions under which this optimal control is derived ensure that the solution is both feasible and optimal within the given constraints. □

6 Computational modeling and outcome display

To accomplish the objectives of optimal control and evaluate the precision of our spatio-temporal model, both with and without control interventions, we executed a numerical simulation of our optimal control problem within the realm of disease spread modeling. This study adopts a stringent methodological framework. Our optimality system is delineated by state equations with initial and boundary conditions (2)-(4), adjoint equations with transversality conditions (36), and an optimal control formulation (38). To resolve this system, we employ the forward-backward scanning method in an iterative manner. Temporal discretization of the state equations is performed using the explicit Euler method, while the adjoint equations are solved retrospectively. The spatial domain considered, symbolizing a city $\Omega = 75\,\mathrm{Km}^2$, is initially populated randomly at $t = 1$ day, under the assumption that the epidemic commenced spreading randomly across various sections of the city. The outcomes will be delineated as follows.

To deepen our understanding, we performed additional simulations to evaluate the impact of various control strategies on disease dynamics. These simulations offer further insights into the effectiveness of different intervention measures under various scenarios, enhancing the robustness of our model. The data and results obtained from these simulations will be meticulously analyzed and discussed to provide comprehensive insights into the model's performance under different conditions.

The parameter values associated with our spatio-temporal model, with and without control measures, are presented in Table 3 below.

Table 3: Model parameters in scenarios with and without control measures.

Parameters	Values
$(\Sigma, \mu, \alpha, \beta, \lambda, \gamma, \theta)$	(0.07,0.12,0.75,0.5,0.0002,0.001,0.03)
$(d_S, d_E, d_I, d_Q, d_R)$	(0.25,0.25,0.25,0.25,0.25)

Figure 2 reveals that before the imposition of control measures, asymptomatic infections were prevalent throughout numerous locations in the region. This observation implies a broad dissemination of the epidemic within the city. Furthermore, Figure 4 depicts a consistently high proportion of asymptomatic cases, sustaining a steady rate of 13% of the total population over time. This continuous presence of asymptomatic individuals facilitated the emergence of a substantial cluster of symptomatic cases across the city, representing 34% of the total population. The visual representation underscores the extensive transmission of the infection in various areas within the city.

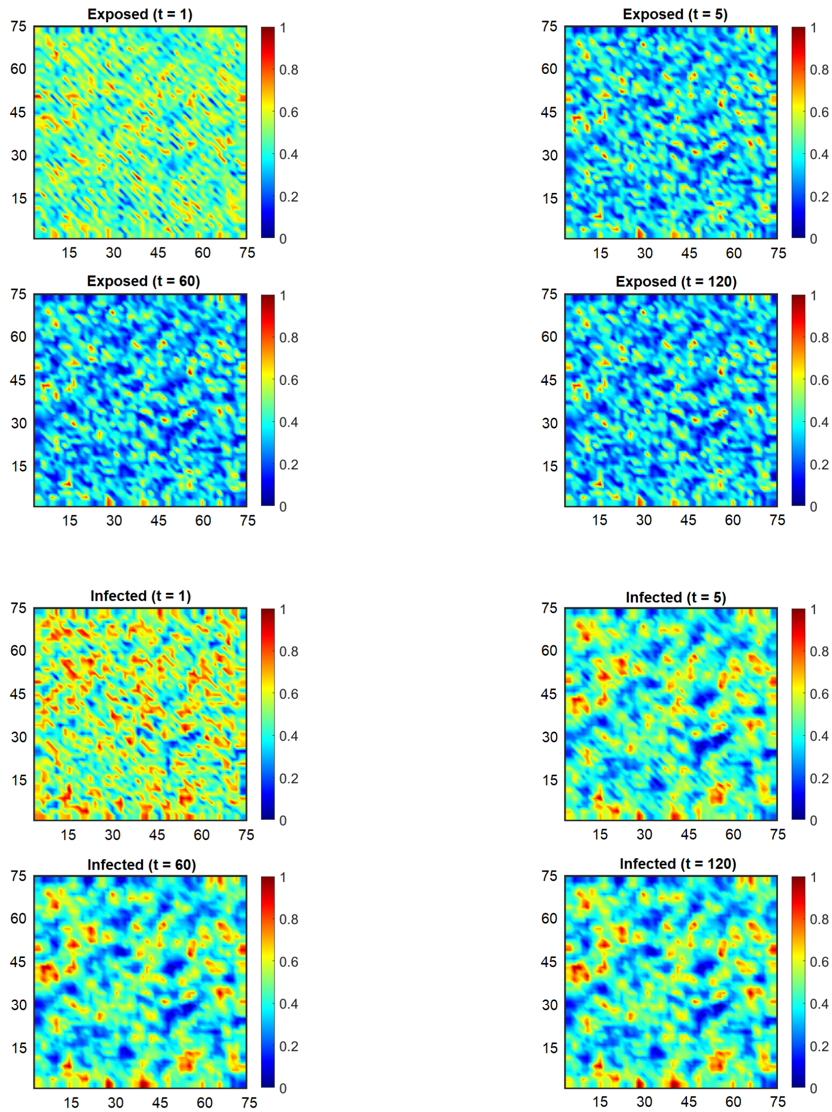
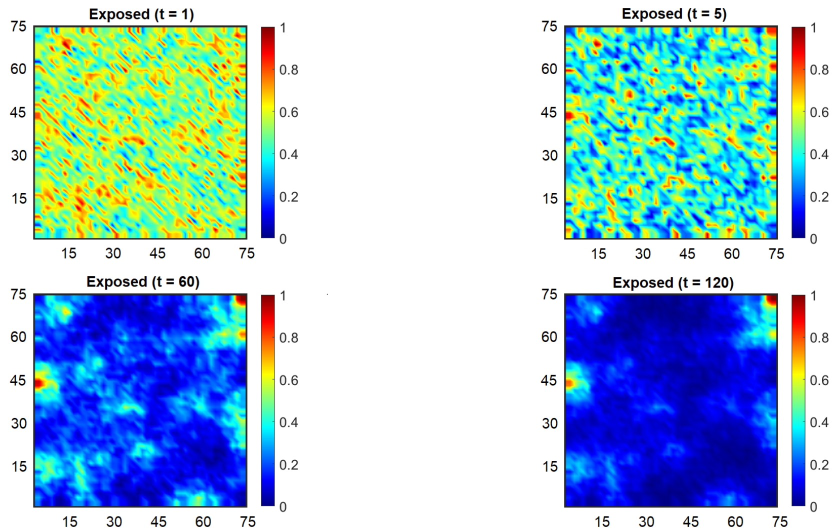


Figure 2: The dynamics of compartments E and I in the absence of control measures

Furthermore, the data reveals that the uncontrolled spread of asymptomatic infections is a key factor in the rapid escalation of the epidemic. This situation underscores the importance of early intervention and the implementation of control measures to curb the virus's spread. The figures collectively provide a clear illustration of how the infection permeated var-



ious sections of the city, resulting in a substantial number of symptomatic cases and emphasizing the necessity for timely and effective public health responses.

To elaborate further, the persistence of asymptomatic cases not only increased the number of symptomatic individuals but also placed a strain on the healthcare system. The widespread nature of the infection required urgent and comprehensive measures to control its transmission. The figures thus provide valuable insights into the dynamics of disease spread and the critical need for strategic interventions.

Figure 3 illustrates a significant shift in the epidemic landscape following the implementation of stringent control measures. These carefully planned interventions aimed to drastically reduce the widespread incidence of asymptomatic infections across the region. Key elements of these measures included limiting social interactions, launching safety campaigns to discourage migration, and promoting health precautions to reduce virus transmission. As depicted in Figure 4, these proactive efforts led to a dramatic decrease in the overall percentage of asymptomatic cases, dropping from an initial 13% to just 0.25%. Consequently, there was a marked decline in the number of symptomatic cases, with the affected population now representing only 0.3% of the total. This positive outcome highlights the effectiveness of the control

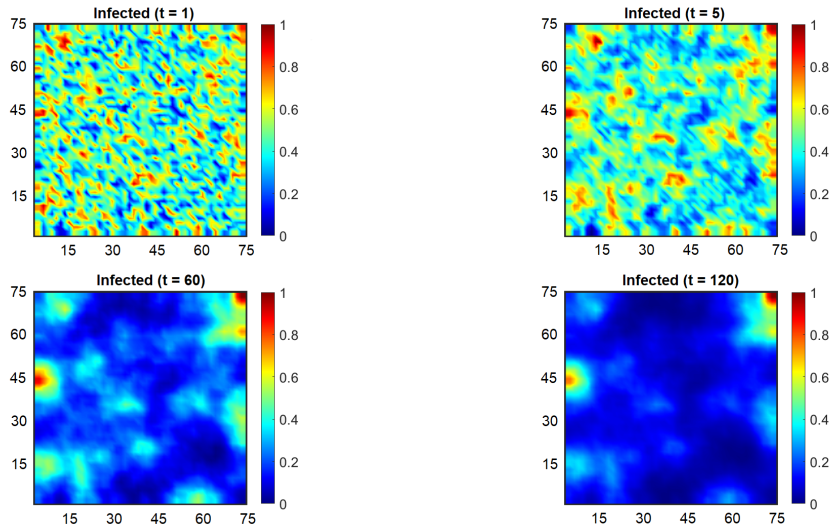


Figure 3: Dynamics of compartments E and I under the influence of control measures

measures in curbing the spread of the disease and reducing its impact on the population. The resulting scenario reflects a more controlled and manageable situation, underscoring the success of the interventions in limiting the transmission of the infection within the city.

Furthermore, the findings underscore the importance of timely interventions in managing the epidemic. The substantial reduction in both asymptomatic and symptomatic cases demonstrates the significant impact of early and decisive action. These control measures not only slowed the spread of the virus but also eased the burden on the healthcare system, allowing it to operate more efficiently and effectively.

In addition, the data suggests that ongoing monitoring and adjustment of control strategies are essential to maintaining the low infection rates that have been achieved. The evolving nature of the epidemic requires adaptive response measures to address new challenges as they emerge. This approach ensures sustained control over the virus, helping to prevent future outbreaks and ultimately protecting public health and well-being.

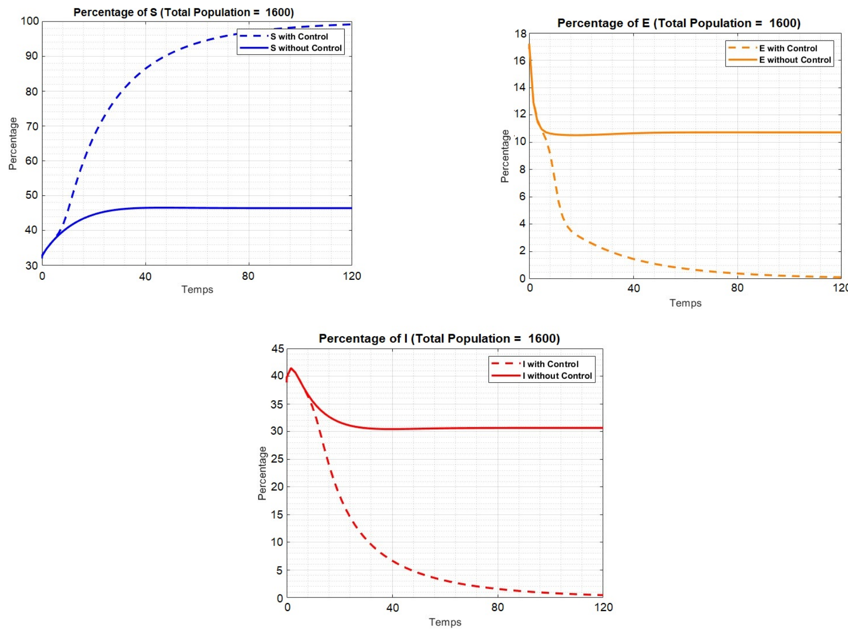


Figure 4: The proportional distribution of compartments S, E, and I in relation to the total population

7 Conclusion

The spatio-temporal model employed in this investigation adeptly encapsulates essential elements such as population density, movement trajectories, and social interactions, which are paramount for comprehending the transmission dynamics of the MPX virus. Our research underscores the critical importance of control measures via diverse intervention strategies, including advocacy campaigns, the creation of social barriers, and targeted testing, to thwart the virus's proliferation while mitigating associated costs and societal upheavals.

Moreover, our examination of this spatio-temporal epidemiological model makes a substantial contribution to the decision-making process and elucidates fundamental mechanisms in epidemic progression. This methodology has the potential to amalgamate rigorous mathematical analysis with empirical data to discern effective targeted therapies, thereby bolstering global endeavors to combat the virus and safeguard public health.

To further refine the accuracy and effectiveness of these models, future iterations should incorporate additional variables such as age, socio-economic status, and behavioral patterns. Enhancing data collection frameworks and developing comprehensive reports will yield current, reliable data for precise sampling, which is indispensable for informing efficacious epidemic mitigation strategies.

Moreover, integrating advanced computational techniques and machine learning algorithms could further refine the model's predictive capabilities. This integration would enable more nuanced simulations and better preparedness for potential future outbreaks. Collaborative efforts across various disciplines, including epidemiology, data science, and public health, will be essential to develop and implement these sophisticated models.

In conclusion, the continuous evolution and refinement of spatio-temporal models are essential for addressing the complex dynamics of epidemic spread. By leveraging these advanced tools, we can enhance our understanding and management of infectious diseases, ultimately safeguarding human health on a global scale.

Acknowledgements

Authors are grateful to there anonymous referees and editor for their constructive comments.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Amaral, A.V.R., González, J.A. and Moraga, P. *Spatio-temporal modeling of infectious diseases by integrating compartment and point process*

models, Stoch. Environ. Res. Risk Assess. 37 (4) (2023) 1519–1533.

[2] Barbu, V. *Mathematical methods in optimization of differential systems*, Vol. 310: Springer Science and Business Media, 2012.

[3] Baroudi, M., Gourram, H., Labzai, A. and Belam, M. *Mathematical modeling and Monkeypox's optimal control strategy*, Commun. Math. Biol. Neurosci. 2023 (2023) Article ID 110.

[4] Bleichrodt, A., Dahal, S., Maloney, K., Casanova, L., Luo, R. and Chowell, G. *Real-time forecasting the trajectory of Monkeypox outbreaks at the national and global levels, July–October 2022*, BMC Med. 21 (1) (2023) 19.

[5] Bourdaud, G. *Le calcul fonctionnel dans les espaces de Sobolev*, Séminaire Équations aux dérivées partielles (Polytechnique) dit aussi Séminaire Goulaouic-Schwartz (1991) 1–4.

[6] Chen, D., Moulin, B. and Wu, J. *Analyzing and modeling spatial and temporal dynamics of infectious diseases*: Wiley Online Library, 2015.

[7] Falb, P.L. and Athans, M. *Optimal control: An introduction to the theory and its applications*, Dover Publications, 2006

[8] Gérard, P. *Description du défaut de compacité de l'injection de Sobolev*, ESAIM: Contr. Optim. Calculus Var. 3 (1998) 213–233.

[9] Hoff, N.A., Doshi, R.H., Colwell, B., Kebela-Illunga, B., Mukadi, P., Mossoko, M., Muyembe-Tamfum, J.J., Okitolonda-Wemakoy, E., Lloyd-Smith, J. and Rimoin, A.W. *Evolution of a disease surveillance system: an increase in reporting of human Monkeypox disease in the Democratic Republic of the Congo, 2001–2013*, Int. J. Trop. Dis. Health, 25 (2) (2017).

[10] Islam, M.A., Sangkham, S., Tiwari, A., Vadiati, M., Hasan, M.N., Noor, S.T.A., Mumin, J., Bhattacharya, P. and Sherchan, S.P. *Association between global Monkeypox cases and meteorological factors*, Int. J. Environ. Res. Public Health, 19 (23) (2022) 15638.

[11] Liu, B., Farid, S., Ullah, S., Altanji, M., Nawaz, R. and Wondimagegnhu Teklu, S. *Mathematical assessment of Monkeypox disease with the impact of vaccination using a fractional epidemiological modeling approach*, Sci. Rep. 13 (1) (2023) 13550.

[12] Mandja, B.A.M., Brembilla, A., Handschumacher, P., Bompangue, D., Gonzalez, J.P., Muyembe, J.J. and Mauny, F. *Temporal and spatial dynamics of Monkeypox in Democratic Republic of Congo, 2000–2015*, EcoHealth 16 (2019) 476–487.

[13] Moumine, E.M., Balatif, O. and Rachik, M. *Mathematical analysis and forecasting of controlled Spatio-temporal dynamics of the EG. 5 virus*, Iranian Journal of Numerical Analysis and Optimization 14 (2) (2024) 500–521.

[14] Munir, T., Khan, M., Cheema, S.A., Khan, F., Usmani, A. and Nazir, M. *ime series analysis and short-term forecasting of Monkeypox outbreak trends in the 10 major affected countries*, BMC Infect. Dis. 24 (1) (2024) 16.

[15] Ozoliņš, V., Lai, R., Caflisch, R. and Osher, S. *Compressed plane waves yield a compactly supported multiresolution basis for the Laplace operator*, Proc. Natl. Acad. Sci. 111 (5) (2014) 1691–1696.

[16] Pal, M., Mengstie, F. and Kandi, V. *Epidemiology, diagnosis, and control of Monkeypox disease: A comprehensive review*, Am. J. Infect. Dis. Microbiol. 5 (2) (2017) 94–99.

[17] Rodriguez, N. and Bertozzi, A. *Local existence and uniqueness of solutions to a PDE model for criminal behavior*, Math. Models Methods Appl. Sci. 20 (supp01) (2010) 1425–1457.

[18] Smouni, I., El Mansouri, A., Khajji, B., Labzai, A., Belam, M. and Tidli, Y. *Mathematical modeling and analysis of a Monkeypox model*, Commun. Math. Biol. Neurosci. 2023 (2023) Article ID 85.

[19] Stengel, R.F. *Optimal control and estimation:* Courier Corporation, 1994.

[20] Takhar, S.S. *Epidemiologic and clinical characteristics of Monkeypox cases–United States, May 17-July 22, 2022*, Ann. Emerg. Med. 81, (1) (2023) 31–34.

[21] Talisuna, A.O., Okiro, E.A., Yahaya, A.A., Stephen, M., Bonkoungou, B., Musa, E.O., Minkoulou, E.M., Okeibunor, J., Impouma, B., Djingarey, H.M. and YAO, N.D.K.M. *Spatial and temporal distribution of infectious disease epidemics, disasters and other potential public health emergencies in the World Health Organisation Africa region, 2016–2018*, Glob. Health. 16 (2020) 1–12.

[22] Wittenberg, R.W. *Local dynamics and spatiotemporal chaos. The Kuramoto-Sivashinsky equation: A case study*, Princeton University, 1998.

[23] Yang, S. *Global solutions of nonlinear wave equations with large data*, Sel. Math. 21 (2015) 1405–1427.

[24] Yinka-Ogunleye, A., Aruna, O., Dalhat, M., Ogoina, D., McCollum, A., Disu, Y., Mamadu, I., Akinpelu, A., Ahmad, A., Burga, J. and Ndoreraho, A., *Outbreak of human Monkeypox in Nigeria in 2017–2018: a clinical and epidemiological report*, Lancet Infect. Dis. 19 (8) (2019) 872–879.

[25] Zebardast, A., Latifi, T., Shafiei-Jandaghi, N.Z., Gholami Barzoki, M. and Shatizadeh Malekshahi, S. *Plausible reasons for the resurgence of Mpox (formerly Monkeypox): An overview*, Trop. Dis. Travel Med. Vaccines. 9 (1) (2023) 23.

Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1309–1333
https://doi.org/10.22067/ijnao.2024.87548.1422
https://ijnao.um.ac.ir/

Research Article

A Petrov–Galerkin approach for the numerical analysis of soliton and multi-soliton solutions of the Kudryashov–Sinelshchikov equation

H. Samy[1], W. Adel[*,2], I. Hanafy[1] and M. Ramadan[1]

Abstract

This study delves into the potential polynomial and rational wave solutions of the Kudryashov–Sinelshchikov equation. This equation has multiple applications including the modeling of propagation for nonlinear waves in

*Corresponding author

Received 9 April 2024; revised 29 July 2024; accepted 30 July 2024

Hayah Samy, (h.samy.aly.z@gmail.com)
Waleed Adel, (waleed.ouf@ufe.edu.eg)
Ibrahim Hanafy, (ihanafy@hotmail.com)
Moutaz Ramadan, (motaz_ramadan@sci.psu.edu.eg)

[1] Department of Mathematics and Computer Sciences, Faculty of Science, Port-Said University, Egypt.

[2] Laboratoire Interdisciplinaire de l'Universite Francaise d'Egypte (UFEID Lab), Universite Francaise d'Egypte, Cairo 11837, Egypt. Department of Mathematics and Engineering Physics, Faculty of Engineering, Mansoura University, Mansoura, 35516, Egypt.

How to cite this article

Samy, H., Adel, W., Hanafy, I. and Ramadan, M., A Petrov–Galerkin approach for the numerical analysis of soliton and multi-soliton solutions of the Kudryashov–Sinelshchikov equation. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1309-1333. https://doi.org/10.22067/ijnao.2024.87548.1422

various physical systems. Through detailed numerical simulations using the finite element approach, we present a set of accurate solitary and soliton solutions for this equation. To validate the effectiveness of our proposed method, we utilize a collocation finite element approach based on quintic B-spline functions. Error norms, including L_2 and L_∞, are employed to assess the precision of our numerical solutions, ensuring their reliability and accuracy. Visual representations, such as graphs derived from tabulated data, offer valuable insights into the dynamic changes of the equation over time or in response to varying parameters. Furthermore, we compute conservation quantities of motion and investigate the stability of our numerical scheme using Von Neumann theory, providing a comprehensive analysis of the Kudryashov–Sinelshchikov equation and the robustness of our computational approach. The strong alignment between our analytical and numerical results underscores the efficacy of our methodology, which can be extended to tackle more complex nonlinear models with direct relevance to various fields of science and engineering.

AMS subject classifications (2020): 65N12; 65N30.

Keywords: Quintic B-spline; Finite element method; Error analysis.

1 Introduction

The concept of solitons has long been regarded as a captivating nonlinear phenomenon that has attracted significant attention from researchers across various scientific disciplines and technological fields. It can be defined as a self-reinforcing solitary wave possessing unique properties, such as maintaining its shape and speed during propagation, even when encountering obstacles or interacting with other waves. These remarkable features stem from the delicate balance between nonlinear and dispersive effects within the underlying physical system. With these pivotal characteristics, solitons find numerous applications across diverse domains. For instance, in the realm of optical fibers, solitons demonstrate the ability to propagate over long distances without distortion, rendering them ideal for high-capacity data transmission [19, 6, 27]. Moreover, solitons have found utility in simulating nonlinear systems, such as the Maccari system [24], and in modeling Rossby waves in geophysical fluid mechanics within fluid dynamics [39]. Further exploration of potential applications of solitons can be found in [11] and references therein.

Nonlinear evolution equations (NLEEs) represent a category of mathematical equations capturing the temporal evolution of nonlinear phenomena spanning diverse scientific domains. These equations exhibit a strong association with solitons, as certain types of NLEEs accurately describe their behavior. Characterized by nonlinear terms capturing interactions and feedback among different variables or quantities, NLEEs have found wide-ranging applications in physics, biology, chemistry, and engineering. Prominent among these equations are the Korteweg–de Vries (KdV) equation [18], as discussed by Johnson [16]; the nonlinear Schrödinger equation, examined by Rabinowitz [29]; and the sine-Gordon equation, explored by Rubinstein [32], among others. Governing diverse physical phenomena, such as water waves, optical fibers, Bose–Einstein condensates, and magnetic materials, NLEEs hold significant importance. Given their wide-ranging applications, researchers have actively pursued analytical and numerical solutions for NLEEs. For instance, Hyder and Barakat [12] employed the improved Kudryashov method to obtain exact solutions for NLEEs. Additionally, Bildik and Deniz [4] adapted perturbation iteration techniques to simulate solutions for nonlinear Klein–Gordon equations. Gepreel [8] employed a range of algebraic methods to tackle the nonlinear (1 + 1) Ito integral differential equation as well as the (1 + 1) nonlinear Schrödinger equation. Ghanbari et al. [9] employed the Lie symmetry method to simulate solutions for Kawahara–KdV type equations, obtaining exact Jacobi elliptic solutions. In addition, Pal, Chatterjee, and Saha [28] employed the Darboux transformation method for simulating the multiple soliton solutions for the general Lax equation. These are only a few of the models that can be found to simulate the NLEEs, and others can be found in [5] and references therein.

On the other hand, numerical techniques have been employed to approximate solutions to such problems. For instance, the finite difference method has been applied to solve Caputo–Hadamard fractional differential equations [10]. The literature abounds with extensive efforts in simulating soliton solutions for these problems, and further research in this area continues to flourish. One of the most important forms of these equations is the

Kudryashov–Sinelshchikov equation, first introduced in 2003 by Kudryashov and Sinelshchikov; see [1]. This equation, belonging to the class of integrable equations, finds numerous applications in studying various physical phenomena [20]. It is considered a generalization of the well-known KdV equation, involving the interaction of dispersion and nonlinearity [15]. The study of the Kudryashov–Sinelshchikov equation has led to the development of various analytical and numerical techniques, including the Painlevé test, Hirota bilinear method [14], and Darboux transformation [35]. This equation has garnered significant attention from researchers due to its rich mathematical structure and its ability to model a wide range of phenomena, including soliton dynamics, wave interactions, and nonlinear wave phenomena. In 2010, Kudryashov and Sinelshchikov developed a nonlinear partial differential equation to describe pressure waves in liquid and gas bubble mixtures, considering liquid viscosity and heat transfer. Researchers continue to explore the properties and solutions of this equation to deepen our understanding of nonlinear dynamics and its implications in physics and related fields. For example, Ryabov [33] discovered a new exact solution for this type of equation using a modification of the truncated method. Additionally, Ak, Osman, and Kara [1] established polynomial and rational wave solutions for the Kudryashov–Sinelshchikov equation. They validated these findings through the application of an appropriate numerical technique. Other methods for solving this equation may include the G/G' polynomial expansion method [23] and the fractional novel analytic method [36]. It should be noted that relatively little work has been done on solving this type of equation in the literature, which motivated our investigation into an accurate solution for such a model equation.

One of the recently adopted approaches for addressing analogous issues is the finite element method. This technique is recognized for its effectiveness and draws from classical methodologies like the Ritz method [21], Petrov–Galerkin method [18, 26], and the least squares method to approximate solutions for boundary value problems encountered in the theory of elliptic partial differential equations. These methods have been extensively applied in solving complex engineering problems [40] and physics phenomena [25, 30]. The finite element method involves partitioning a complex geometry into smaller, simpler regions known as elements, which are interconnected by nodes. The governing equations are then solved for each element, and the solutions are combined to obtain an approximation of the solution for the entire system. This technique finds widespread use across various domains, including structural analysis [3], heat transfer [7], fluid dynamics [2], and electromagnetics [13]. Its versatility and flexibility render it a popular choice for solving complex problems in engineering and science. Moreover, it has been effectively employed in addressing problems of nonlinear behavior, including the bioheat model [34], two-sided space-fractional evolution equation [22], Allen–Cahn type phase-field model [38], and other diverse applications.

In this paper, we are interested in solving the nonlinear Kudryashov–Sinelshchikov equation in the following form:

$$U_t + \alpha U U_x + \beta U_{xxx} + \gamma (U U_{xx})_x + \rho U_x U_{xx} = 0, \tag{1}$$

where $U(x,t)$ is a field variable, and it is the unknown that is to be determined. In addition, $\alpha, \beta, \gamma, \rho$ are real parameters, and the subscripts shown are represented as differentiation concerning time and space respectively. Depending on the context of the simulated model, the variable $U(x,t)$ in this equation can represent different physical quantities of interest. In the field of fluid dynamics, $U(x,t)$ may describe the velocity, pressure, or free surface displacement of water waves, allowing researchers to analyze the formation and evolution of solitary waves and solitons. In plasma physics, $U(x,t)$ can be interpreted as the electric potential or the density of charged particles, providing insights into the nonlinear behavior of plasma waves, including ion-acoustic solitons. Moreover, in the domain of solid mechanics, $U(x,t)$ may correspond to the displacement or deformation of a solid material under the influence of nonlinear effects, with implications for the design of advanced nonlinear mechanical systems and wave-based nondestructive testing. By investigating the analytical and numerical solutions of the Kudryashov–Sinelshchikov equation, as presented in this study, researchers can deepen their understanding of the rich nonlinear phenomena underlying these diverse physical applications. The solution to (1) will be discussed using the known B-spline finite element method. The B-splines, or basis splines, are mathematical functions widely used for approximating complex curves and surfaces.

They are constructed by combining a set of piecewise polynomial functions, each defined over a small interval. The resulting function is continuous and smooth, with adjustable levels of curvature or smoothness. B-splines are crucial in computer-aided design, computer graphics, and computer vision, where they are used to represent and manipulate complex geometric shapes.

The novelty of the presented work lies within the following points:

1. The Kudryashov–Sinelshchikov equation is investigated in this work.
2. The solution is based on the Petrov–Galerkin finite element technique accompanied by the quintic B-splines as basis functions.
3. A Crank–Nicolson approach is employed for time and the B-spline for spatial discretization.
4. A detailed stability analysis using Von-Neumann stability is illustrated to prove that the scheme is unconditionally stable.
5. Numerical experiments are undertaken to verify that the method yields precise solutions, offering insights into the dynamics of the novel model.

The paper is organized as follows: Section 2 presents a detailed description of the proposed technique. In Section 3, the Crank–Nicolson technique is illustrated to reduce the proposed model (1), whereas Section 4 presents the initial state of the solution. Section 5 focuses on the stability analysis of the proposed algorithm. The main results of the collocation technique are presented in Section 6, and the final and concluding remarks, along with potential future work, are addressed in Section 7.

2 Petrov–Galerkin finite element technique

In this section, we will illustrate the main steps for solving model (1) using the Petrov–Galerkin finite element method. We first define model (1); throughout $[a, b]$ is a finite region with boundary conditions. Next, let a partition of $[a, b]$ be $a = x_0 < x_1 < \cdots < x_N = b$ by the equally spaced knots x_i and let quintic B-splines with knots at the points x_i, $0 < i < N$ be $\phi_i(x)$, where the set $\phi_{i-2}, \phi_{i-1}, \phi_i, \phi_{i+1}, \phi_{i+2}, \phi_{i+3}$ constitutes a series of splines, serving as the basis for functions sought within the finite region $[a, b]$. The solution approximation $U_N(x,t)$ for $U(x,t)$ is defined as follows:

$$U_N(x,t) = \sum_{i=-2}^{N+2} \phi_i(x) u_i(t), \tag{2}$$

where u_i represent time-dependent parameters that can be computed from boundary conditions,

$$U(a,t) = U(b,t) = 0, \qquad U_x(a,t) = U_x(b,t) = 0. \tag{3}$$

The intervals $[x_i, x_{i+1}]$ are utilized to define finite elements with nodes positioned at x_i and x_{i+1}. Each element $[x_i, x_{i+1}]$ is spanned by six splines $(\phi_{i-2}, \phi_{i-1}\ , \phi_i, \phi_{i+1},\ \phi_{i+2}, \phi_{i+3})$, which are expressed within a local coordinate system denoted by ζ, defined as $h\zeta = (x - x_i)$ where $h = x_{i+1} - x_i$ and $0 \leq \zeta \leq 1$. The approximate solution in the quintic B-spline collocation method can be expressed as a combination of quintic B-spline basis functions for the approximation of the space variables under consideration. The formulations for all these splines across the element $[x_i, x_{i+1}]$ are given by

$$\begin{aligned}
\phi_{i-2} &= 1 - 5\zeta + 10\zeta^2 - 10\zeta^3 + 5\zeta^4 - \zeta^5, \\
\phi_{i-1} &= 26 - 50\zeta + 20\zeta^2 + 20\zeta^3 - 20\zeta^4 + 5\zeta^5, \\
\phi_i &= 66 - 60\zeta^2 + 30\zeta^4 - 10\zeta^5, \\
\phi_{i+1} &= 26 + 50\zeta + 20\zeta^2 - 20\zeta^3 - 20\zeta^4 + 10\zeta^5, \\
\phi_{i+2} &= 1 + 5\zeta + 10\zeta^2 + 10\zeta^3 + 5\zeta^4 - 5\zeta^5, \\
\phi_{i+3} &= \zeta^5.
\end{aligned} \tag{4}$$

Outside the interval $[x_{i-3}, x_{i+3}]$, the spline $\phi_i(x)$ and its fifth derivatives are zero. When we utilize (4) to formulate equations based on the element parameters u_i^e, these curves serve as “shape” functions for the element. The $U_N(x,t)$ variation across the element $[x_{i-3}, x_{i+3}]$ is provided by

$$u^e(x,t) = \sum_{j=i-2}^{i+3} \phi_j(x) u_i(t). \tag{5}$$

The derivatives at the knots and the nodal value of $U_N(x,t)$ are represented in terms of the element parameters as demonstrated below:

$$\begin{aligned} U_i &= u_{i-2} + 26u_{i-1} + 66u_i + 26u_{i+1} + u_{i+2}, \\ hU_i^{'} &= 5(u_{i+2} + 10u_{i+1} - 10u_{i-1} - u_{i-2}), \\ h^2U_i^{''} &= 20(u_{i-2} + 2u_{i-1} - 6u_i + 2u_{i+1} + u_{i+2}), \\ h^3U_i^{'''} &= 60(u_{i+2} - 2u_{i+1} + 2u_{i-1} - u_{i-2}), \\ h^4U_i^{''''} &= 120(u_{i-2} - 4u_{i-1} + 6u_i - 4u_{i+1} + u_{i+2}). \end{aligned} \tag{6}$$

The dashes denote differentiation with respect to x. When the Petrov–Galerkin method is employed in (1) and by utilizing the weight functions $W(x)$, the outcome is

$$\int_b^a W\left(U_t + \alpha UU_x + \beta U_{xxx} + \gamma(UU_{xx})_x + \rho U_x U_{xx}\right) dx = 0. \tag{7}$$

Now, let us establish the corresponding element matrices. For the contribution of the standard element $[x_i, x_{i+1}]$, we acquire

$$\int_e W\left(u_t^e + \alpha u^e u_x + \beta u_{xxx}^e + \gamma(u^e u_{xx}^e)_x + \rho u_x^e u_{xx}^e\right) dx = 0, \tag{8}$$

$$\begin{aligned} &\sum_{i=l-2}^{l+3}\left(\int_{x_l}^{x_{l+1}} \phi_k\phi_i dx\right)\dot{u}_i^e + \beta\sum_{i=l-2}^{l+3}\left(\int_{x_l}^{x_{l+1}} \phi_k\phi_i^{'''} dx\right)u_i^e \\ &+ \alpha\sum_{j=l-2}^{l+3}\sum_{i=l-2}^{l+3}\left(\left(\int_{x_l}^{x_{l+1}} \phi_i\phi_j^{'}\phi_k dx\right)u_i^e\right)u_j^e \\ &+ (\gamma+\rho)\sum_{j=l-2}^{l+3}\sum_{i=l-2}^{l+3}\left(\left(\int_{x_l}^{x_{l+1}} \phi_k\phi_i^{'}\phi_j^{''} dx\right)u_i^e\right)u_j^e \\ &+ \rho\sum_{j=l-2}^{l+3}\sum_{i=l-2}^{l+3}\left(\left(\int_{x_l}^{x_{l+1}} \phi_k\phi_i\phi_j^{'''} dx\right)u_i^e\right)u_j^e = 0. \end{aligned} \tag{9}$$

The matrix form is constructed as

$$A^e\dot{u}^e + \beta B^e u^e + \alpha C^e u^{eT} u^e + (\gamma+\rho)u^{eT} D^e u^e + \gamma u^{eT} E^e u^e = 0, \tag{10}$$

where the dot represents differentiation with respect to time t, and

$$u^e = (u_{l-2}, u_{l-1}, u_l, u_{l+1}, u_{l+2}, u_{l+3})^T. \tag{11}$$

The element matrices are given by

$$\begin{aligned} A^e_{ij} &= \int_{x_l}^{x_{l+1}} \phi_k \phi_i dx, \qquad B^e_{ij} = \int_{x_l}^{x_{l+1}} \phi_k \phi_i^{'''} dx, \qquad C^e_{ij} = \int_{x_l}^{x_{l+1}} \phi_i \phi_j^{'} \phi_k dx, \\ D^e_{ij} &= \int_{x_l}^{x_{l+1}} \phi_k \phi_i^{'} \phi_j^{''} dx, \quad E^e_{ij} = \int_{x_l}^{x_{l+1}} \phi_k \phi_i \phi_j^{'''} dx, \end{aligned} \tag{12}$$

where i, j, and k range from $l-2$ to $l+3$ for the element $[x_l, x_{l+1}]$. Consequently, the matrices A^e and B^e are of size 6×6, while C^e, D^e, and E^e are of size $6 \times 6 \times 6$. In our algorithm, instead of C^e, D^e, and E^e, we employ the corresponding 6×6 matrices c^e, d^e, and e^e.

$$C^e_{ij} = \sum_{k=l-2}^{l+3} c^e_{ijk} u^e_k, \qquad D^e_{ij} = \sum_{k=l-2}^{l+3} d^e_{ijk} u^e_k, \qquad E^e_{ij} = \sum_{k=l-2}^{l+3} e^e_{ijk} u^e_k. \tag{13}$$

This is contingent on the variables u^e_k. The matrices of elements A^e and B^e are algebraically determined from (12), where u^e_k is provided by (11). The equation below is derived by assembling the elements from (10).

$$A\dot{u} + (\beta B + \alpha C + (\gamma + \rho) D + \gamma E) u = 0, \tag{14}$$

where the matrices A, B, C, D, E are constructed from the element matrices A^e, B^e, C^e, D^e, E^e, respectively, in the usual way and

$$u = (u_{-2}, u_{-1}, u_0, \ldots, u_{N+1}, u_{N+2})^T. \tag{15}$$

In the next section, we will illustrate the Crank–Nicolson approach for simulating the solution for the main model.

3 Crank–Nicolson approach

In this section, we will demonstrate the time discretization for the system given by (14) using the Crank–Nicolson approach. The Crank–Nicolson technique is a finite difference method utilized for numerically solving partial differential equations such as the heat equation. Time is centered around $(n + \frac{1}{2})\Delta t$, where Δt denotes the time step. Subsequently, we employ the Crank–Nicolson method, in the following form

$$u = \frac{1}{2}(u^n + u^{n+1}), \qquad \dot{u} = \frac{1}{\Delta t}(u^{n+1} - u^n). \tag{16}$$

Substituting (16) into (14), we obtain the recurrence relationship

$$\frac{A}{\Delta t}(u^{n+1} - u^n) + \frac{1}{2}(\beta B + \alpha C + (\gamma + \rho)D + \gamma E)(u^{n+1} + u^n) = 0, \tag{17}$$

and then

$$\begin{aligned} &\Big(A + \frac{\Delta t}{2}(\beta B + \alpha C + (\gamma + \rho)D + \gamma E)\Big)u^{n+1} \\ &= \Big(A - \frac{\Delta t}{2}(\beta B + \alpha C + (\gamma + \rho)D + \gamma E)\Big)u^n. \end{aligned} \tag{18}$$

In the system (18), the time indices are denoted by the superscripts n and $n + 1$. This system comprises $N + 1$ linear equations with $N + 5$ variables. To ensure a unique solution, an additional four conditions derived from the boundary conditions must be satisfied. These conditions can be used to eliminate $u_{-2}, u_{-1}, u_0, \ldots, u_{N+1}$, and u_{N+2} from the recurrence relationships (18), resulting in an 11-banded $(N + 5) \times (N + 5)$ matrix equation.

During each time step, an inner iteration is executed to verify the convergence of the nonlinear term. The following outlines the iteration algorithm:

Initially, u^0 is known. The first approximation, u^1_1 to u, is computed using $u = u^0$ as per (16). Subsequently, the second approximation u^1_2 is obtained with $u = \frac{1}{2}(u^0 + u^1_1)$, followed by the third u^1_3 with $u = \frac{1}{2}(u^0 + u^1_2)$. Typically, we find that ten iterations are sufficient to obtain a reasonable approximation for u^1 in this initial stage.

To obtain a first approximation, denoted as u^{n+1}_1, for u^{n+1} in general, we use $u = u^n + \frac{1}{2}(u^n + u^{n-1})$. Subsequently, a second approximation u^{n+1}_2 is obtained with $u = \frac{1}{2}(u^n + u^{n+1})$, and so forth. Typically, convergence is attained after two or three iterations. The time evolution of u^n is determined by a system of the decadiagonal and as a result after the initial vector of the parameters u^0 is obtained, $U_N(x,t)$ can be begun.

The initial state for the main model is illustrated in the next section.

4 The initial state

In this section, the initial state for solving system (18) is outlined. We initiate the time assessment of u^{n+1} by computing the vector u^0 from the initial condition using the recurrence relation (18). Referring to (5), if we rewrite the global trial functions as follows:

$$U_N(x,0)=\sum_{i=-2}^{N+2}\phi_i(x)u_i^0. \tag{19}$$

The parameters u_i^0 represent unknowns that need to be determined. To establish the initial vector u_i^0, U_N must satisfy the following conditions: At the knots x_j, it agrees with the analytical initial condition; applying (5), it leads to $N+1$ conditions. The solution of matrix equations is then used to get the start up vector u^0,

$$Mu^0=b, \tag{20}$$

where

$$M=\begin{bmatrix} 3 & 30 & 27 & & & & & & & \\ 1 & 18 & 33 & 8 & & & & & & \\ 1 & 26 & 66 & 26 & 1 & & & & & \\ & 1 & 26 & 66 & 26 & 1 & & & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & & & \\ & & & & 1 & 26 & 66 & 26 & 1 & \\ & & & & & 1 & 26 & 66 & 26 & 1 \\ & & & & & & 8 & 33 & 18 & 1 \\ & & & & & & & 27 & 30 & 3 \end{bmatrix}, \tag{21}$$

$$\begin{aligned} b&=(U^{''}(x_0),U^{'}(x_0),U(x_0),U(x_1),\ldots,U(x_N),U^{'}(x_N),U^{''}(x_N))^T,\\ u^0&=(u_{-2}^0,u_{-1}^0,u_0^0,\ldots,u_N^0,u_{N+1}^0,u_{N+2}^0)^T. \end{aligned} \tag{22}$$

After determining the initial vector u^0 as the solution of the undecadiagonal matrix (18), the system is solved using a Thomas algorithm [37].

5 Stability analysis

In this section, the Kudryashov–Sinelshchikov equation represented by model (1) is amenable to stability analysis through Fourier transform techniques. Specifically, one can express the solution of the equation as a Fourier series expansion and substitute it back into the equation to obtain a system of algebraic equations for the Fourier coefficients [25]. By analyzing the solutions of these algebraic equations, one can determine the stability properties of the system in question [41]. We begin by applying the Fourier transform for the main problem, which results in the following form:

$$U_m^n = z^n e^{imkh}, \tag{23}$$

Here, z denotes the growth factor of the error in a typical mode of amplitude z_n, h represents the element size, and k indicates the mode number. By substituting the Fourier mode given in (23) into the system (18), we obtain the following equality:

$$\begin{aligned}&\Big(A + \frac{\Delta t}{2}(\beta B + \alpha C + (\gamma+\rho)D + \gamma E)\Big) z^{n+1} e^{i(m+1)kh} \\ &= \Big(A - \frac{\Delta t}{2}(\beta B + \alpha C + (\gamma+\rho)D + \gamma E)\Big) z^n e^{imkh}.\end{aligned} \tag{24}$$

Let $a = \Big(A + \frac{\Delta t}{2}(\beta B + \alpha C + (\gamma+\rho)D + \gamma E)\Big)$ and $b = \Big(A - \frac{\Delta t}{2}(\beta B + \alpha C + (\gamma+\rho)D + \gamma E)\Big)$. Then (24) took the form

$$a z^{n+1} e^{i(m+1)kh} = b z^n e^{imkh}, \tag{25}$$

for more simplicity, we divide both sides by $z^n e^{im\theta}$, where $\theta = kh$ so that the equation would be

$$aze^{i\theta} = b. \tag{26}$$

We obtain growth factor as follows:

$$z = \frac{b}{ae^{i\theta}}, \tag{27}$$

where $e^{i\theta} = cos(\theta) + isin(\theta)$. As $|z| < 1$, then the scheme is unconditionally stable.

6 Computational results

In this section, we present the computational results to validate the theoretical findings. The numerical algorithm developed in Section 3 will be validated through the examination of test problems involving the migration and interaction of solitons. We utilize the L_2 and L_∞ error norms to quantify the disparity between the numerical and analytical solutions, thereby demonstrating the predictive accuracy of the scheme regarding the position and amplitude of the solution as the simulation progresses. The L_2 and L_∞ norms of the solution are defined as follows:

$$
\begin{aligned}
L_2 &= \|U^{exact} - U^n\|_2 = [h \sum_{i=1}^{N} |U_i^{exact} - U_i^n|^2]^{\frac{1}{2}}, \\
L_\infty &= \|U^{exact} - U^n\|_\infty = \max_i |U_i^{exact} - U_i^n|.
\end{aligned} \tag{28}
$$

The Kudryashov–Sinelshchikov (1) has two conserved quantities, which are as follows:

$$
\begin{aligned}
C_1 &= \int_a^b \grave{u} dx \simeq h \sum_{i=1}^{N} U_j^n, \\
C_2 &= \int_a^b \left[\frac{1}{(\gamma + \sigma)} \Big(\big(\gamma u + \beta^{\sigma\gamma+1} - \beta^{\sigma\gamma+1} \gamma^{-\sigma\gamma} \big) \Big) \right] dx \\
&\simeq h \sum_{i=1}^{N} \left[\frac{1}{(\gamma + \sigma)} \Big(\big(\gamma u_j^n + \beta^{\sigma\gamma+1} - \beta^{\sigma\gamma+1} \gamma^{-\sigma\gamma} \big) \Big) \right].
\end{aligned} \tag{29}
$$

First, we give the case of a single solitary wave.

6.1 Single solitary wave

The exact solution to the Kudryashov–Sinelshchikov equation is of the form

$$
U(x,t) = A\mathrm{sech}^2 \left[B \Big(x - x_0 - ct \Big) \right], \tag{30}
$$

and the initial condition can take the following equation:

$$U(x,0) = A\text{sech}^2\left[B\Big(x - x_0\Big)\right], \tag{31}$$

where $A = \frac{3\beta k^2\mu^2}{\alpha-\gamma k^2\mu^2}$, $B = \frac{k\mu}{2}$, and $c = \beta k^2\mu^2$. The constants are in the form $\alpha = 1.5, \beta = 1.7, \gamma = 0.8, \sigma = 2.4, c = -2.4, k = 1, \mu = 0.5$. Three experiments were carried out: the first at $\Delta t = h = 0.1$,the second at $\Delta t = h = 0.025$, and the third at $\Delta t = 0.025, h = 0.1$. The error norms and conserved quantities are determined for various h and Δt values up to time $t = 20$.

For the first example, the algorithm is executed within the calculation range $[-100, 100]$ up to time $t = 20$ to demonstrate the proper functioning of our numerical technique. During the simulation calculations, typical values of $\Delta t = 0.1$ and $\Delta t = 0.025$ are employed, along with $h = 0.1$ and $h = 0.025$. Tables 1 and 2 present the values of the error norms and invariants at various time levels, enabling us to promptly discern the impact of the number of grid points on the numerical technique. As observed from the tables, the two conserved quantities remain nearly constant over time. Moreover, the computed values of the error norms are determined to be sufficiently small, with these errors scarcely altering over time.

Table 1: Error norms and conservative quantities single solitary wave at $\Delta t = h = 0.1$

t	L_2	L_∞	C_1	C_2
5.0	3.52322809e-04	1.114142546e-03	7.84615390357	-12.36234575240
10.0	3.52312669e-04	1.114034293e-03	7.84615401202	-12.36234573439
15.0	3.52288576e-04	1.114034292e-03	7.846154326030	-12.36234568224
20.0	3.52300820e-04	1.114034292e-03	7.84615523469	-12.36234553134

Additionally, Table 3 presents a comparison with earlier methods, specifically method one (finite element method with quintic B-spline, FEMQB-spline) [1] and method two (finite element method with septic B-spline, FEMSB-spline) [17] at time $t = 20$. It can be observed from the table that the error norms obtained using the current method align well with those obtained using the earlier methods.

Table 2: Error norms for the Kudryashov–Sinelshchikov equation's single solitary wave when $\Delta t = h = 0.025$

t	L_2	L_∞	C_1	C_2
5.0	2.3295e-08	4.5041e-10	1.4657e-07	-3.345103498
10.0	1.7861e-08	5.8745e-10	1.1238e-07	-3.34510350
15.0	1.3694e-08	7.6617e-10	8.6166e-08	-3.345103508
20.0	1.0500e-08	9.9928e-10	6.6066e-08	-3.345103511

Table 3: Error norms for the Kudryashov–Sinelshchikov equation's single solitary wave when $\Delta t = h = 0.025$

	Present Method		FEMQB-spline [1]		FEMSB-spline [17]	
t	L_2	L_∞	L_2	L_∞	L_2	L_∞
5.0	2.3295e-08	4.5041e-10	3.4853e10-6	1.4600e10-6	3.1100e-06	1.3154e-06
10.0	1.7861e-08	5.8745e-10	5.7917e10-6	2.1598e10-6	5.7796e-06	2.0899e-06
15.0	1.3694e-08	7.6617e-10	1.2314e10-5	5.8342e10-6	1.2013e-05	5.4335e-06
20.0	1.0500e-08	9.9928e-10	2.2657e10-5	1.0635e10-5	2.2673e-05	1.0848e-05

Table 4 demonstrates that the error norms (L_2 and L_∞) and the two conserved quantities remain stable and consistent across different time values (t), which suggests that the numerical method used for solving the Kudryashov–Sinelshchikov equation's single solitary wave is reliable and maintains accuracy over time.

Table 4: Error norms for the Kudryashov–Sinelshchikov equation's single solitary wave when $\Delta t = 0.025, h = 0.1$

t	L_2	L_∞	C_1	C_2
5.0	3.80470850877e-04	12.03154472e-04	7.84615382769	-12.36234576810
10.0	3.80458787872e-04	12.03116325e-04	7.84615382159	-12.36234576911
15.0	3.80358154750e-04	12.0279809e-04	7.84615381376	-12.36234577041
20.0	3.80461773590e-04	12.0312576e-04	7.84615380362	-12.3623457721

Figure 1 illustrates the two-dimensional states of the bell-shaped solitary wave solutions at time $t = 30$ seconds, along with the distribution of numerical errors at the same time for $h = \Delta t = 0.1$. Figure 2 presents the three-dimensional states of the bell-shaped solitary wave solutions obtained from $t = 0$ to $t = 20$. Additionally, the contour lines depicting the movement of each individual wave are visible in Figure 2. Figure 3 illustrates how the wave has changed over time. These visualizations demonstrate that the method under examination accurately captures the propagation movement of a single wave while preserving its amplitude and shape.

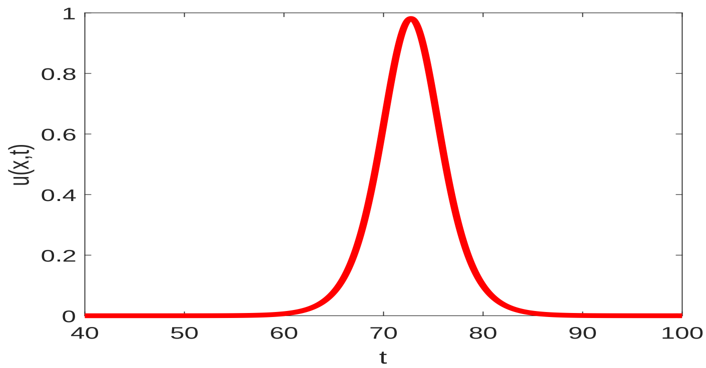


Figure 1: Behavior of numerical solution for $t = 30$.

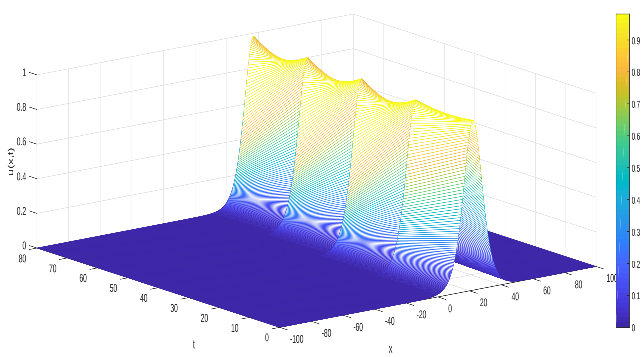


Figure 2: Space plot of one soliton solution for $t \in [0, 80]$.

6.2 Interaction of two solitary waves

The interaction of two solitary waves moving in the same direction is investigated by utilizing an initial condition obtained from the linear sum of two well-separated solitary waves with distinct amplitudes. This initial condition can take the following form

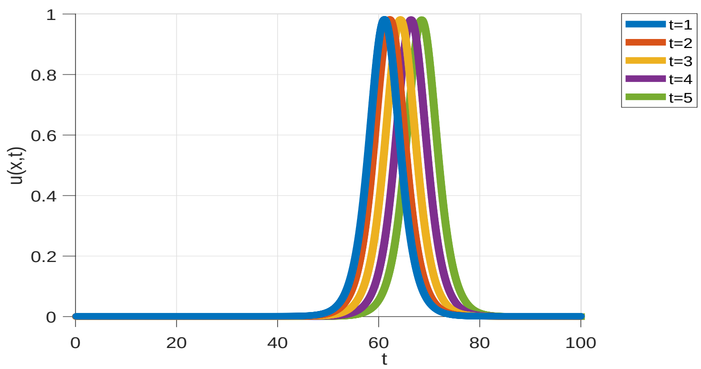


Figure 3: Behavior of numerical solution for $t \in [1, 5]$.

$$U(x,0) = \sum_{i=1}^{2} A_i \mathrm{sech}^2 \left[B_i \Big(x_i - x_0 \Big) \right]. \tag{32}$$

The collision of two solitary waves is analyzed in this problem using the following exact solution:

$$U(x,t) = \sum_{i=1}^{2} A_i \mathrm{sech}^2 \left[B_i \Big(x_i - x_0 - c_i t \Big) \right], \tag{33}$$

where $A_i = \frac{3\beta k^2 \mu_i^2}{\alpha - \gamma k^2 \mu_i^2}$, $B_i = \frac{k\mu_i}{2}$, and $c_i = \beta k^2 \mu_i^2$. The constants are in the form $\alpha = 1.5, \beta = 1.7, \gamma = 0.8, \sigma = -2.4, c_1 = 0.833, c_2 = 0.425, k = 1, \mu_1 = 0.7, \mu_2 = 0.5$. Two experiments were carried out: the first at $\Delta t = h = 0.1$ and the second at $\Delta t = h = 0.025$.

To validate the proper functioning of our numerical technique, the algorithm is executed within the calculation range $[-100, 100]$ up to time $t = 20$ in this initial example. The values of the error norms and invariants for various time levels are presented in Table 5. This allows us to promptly assess the impact of the number of grid points on the numerical technique. The table indicates that the two conserved quantities remained nearly consistent over time. Moreover, the computed error norm values are determined to be appropriately small, with these errors hardly ever changing over time. Additionally, Table 6 compares this method with the older method finite element method with septic B-spline (FEMSB-spline) [17] at time $t = 20$. Furthermore, Tables 7 and 8 illustrate how the double solitary wave has evolved as have changed $\Delta t = h = 0.025$ then $\Delta t = h = 0.1$. It can be observed from the table that the error norms calculated using the current method are consistent with those obtained using the older method.

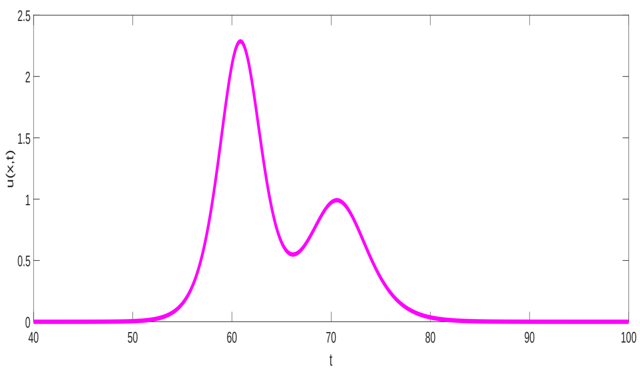


Figure 4: Behavior of numerical solution at $t = 25$ seconds.

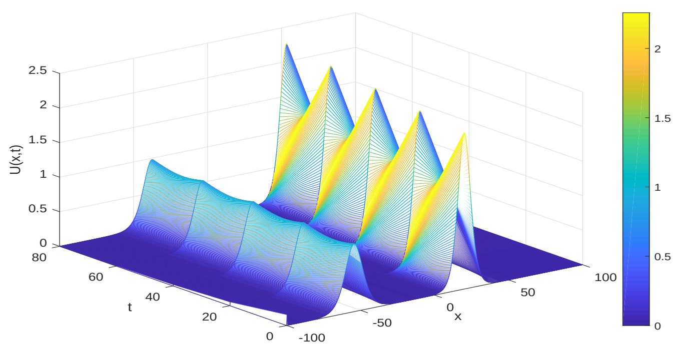


Figure 5: Numerical solution for $x \in [-100, 100]$ and $t \in [0, 80]$.

Figure 4 depicts the two-dimensional states of the bell-shaped solitary wave solutions at time $t = 25$ seconds, as well as the numerical error distribution for $h = \Delta t = 0.1$ at time $t = 30$ seconds. Figure 5 shows that at time zero, the higher energy single wave lags behind the lower energy second wave. Depending on their magnitude, both waves flow to the right. More energy equals faster, according to solitary wave theory. Figure 6 illustrates how the wave has changed over time. As a result, the larger wave eventually reaches the smaller one, causing interference. Over time, the wave with the

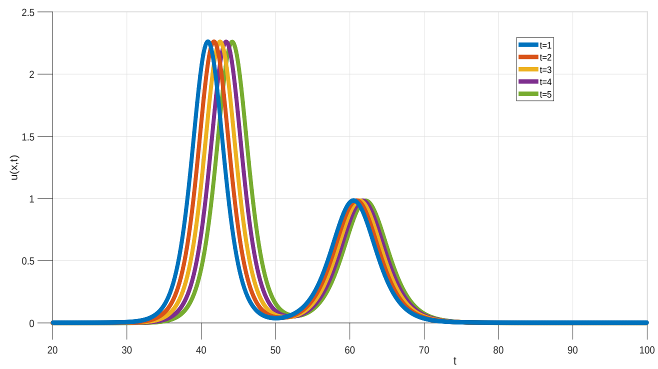


Figure 6: Behavior of numerical solution for $t \in [1, 5]$.

most energy leaves the second wave with the least energy, and solo waves finally revert to their initial form.

Table 5: Error norms for the Kudryashov–Sinelshchikov equation's two solitary waves when $\Delta t = h = 0.1$

t	L_2	L_∞	C_1	C_2
5.0	11.49337207e-04	3.63452337e-03	20.7342405464	-11.0358676030
10.0	11.49282935e-04	3.63435175e-03	20.7342406546	-11.0387332828
15.0	11.50661291e-04	3.6387104977e-03	20.7342409686	-11.0458076846
20.0	11.52842423e-04	3.6456078414e-03	20.7342418772	-11.06220804894

Table 6: Error norms for the Kudryashov–Sinelshchikov equation's two solitary waves when $\Delta t = h = 0.1$

	Proposed technique		FEMSB-spline [17]	
t	C_1	C_2	C_1	C_2
5.0	20.7342405464	-11.0358676030	20.7343904	-24.400625
10.0	20.7342406546	-11.0387332828	20.734406	-24.400625
15.0	20.7342409686	-11.0458076846	20.734405	-24.400625
20.0	20.7342418772	-11.06220804894	20.734415	-24.400624

Table 7: Error norms for the Kudryashov–Sinelshchikov equation's two solitary waves when $\Delta t = h = 0.025$

t	L_2	L_∞	C_1	C_2
5.0	1.7674008765e-05	5.5890123084e-05	7.9329863569e-05	-3.34509034835
10.0	1.76740087654e-05	5.5890123084e-05	7.9329863569e-05	-3.34509034835
15.0	8.5319829343e-06	2.6980499030e-05	3.8304877651e-05	-3.34509716131
20.0	4.1201731071e-06	1.302913136e-05	1.8504767045e-05	-3.34510044952

Table 8: Error norms and conservative quantities single solitary wave at $\Delta t = h = 0.1$ and $\alpha = 111.5$

t	L_2	L_∞	C_1	C_2
5.0	4.1151810516e-06	1.301334510e-05	0.091644205	-13.36524688650
10.0	4.1147812099e-06	1.3012080696e-05	0.0916442067	-13.36524688629
15.0	4.1140013153e-06	1.3009614453e-05	0.09164421	-13.36524688568
20.0	4.1149242293e-06	1.301253296e-05	0.091644221	-13.36524688398

7 Conclusion

In this study, we present a detailed and comprehensive investigation of the numerical solution of the Kudryashov–Sinelshchikov equation using a numerical collocation approach through simulations to elucidate its behavior. The Kudryashov–Sinelshchikov equation is a nonlinear partial differential equation that can model the dynamics of various physical field variables, such as fluid velocity, wave amplitude, plasma density, or displacement/deformation in solid mechanics. Through the finite element approach, we meticulously present a diverse array of accurate solutions for this equation, which has direct relevance to the study of nonlinear wave propagation phenomena in these physical systems.

The validation of our method's effectiveness is substantiated through the utilization of a collocation finite element approach based on quintic B-spline functions. Error norms, including L_2 and L_∞, serve as robust metrics for evaluating the precision of our numerical solutions, affirming the reliability of our approach. Valuable insights into the dynamic evolution of the equation over time and under varying parameters, enriching our understanding of its behavior, have been presented. Furthermore, our study delves into the computation of conservation quantities of motion and the investigation of numerical scheme stability using the Von Neumann theory, further bolstering the credibility of our findings.

Based on these findings, the congruence between our analytical and numerical results underscores the robustness and efficacy of our approach in comprehensively addressing the Kudryashov–Sinelshchikov equation. These findings not only contribute to the broader understanding of nonlinear partial differential equations but also offer promising avenues for future research endeavors, particularly in extending our techniques to tackle more complex models with direct physical applications in the areas of fluid dynamics, wave propagation, plasma physics, and solid mechanics.

Acknowledgements

The authors would like to convey their thanks to the Editor and Reviewers for the helpful comments and suggestions which improved the work.

References

[1] Ak, T., Osman, M.S., and Kara, A.H. *Polynomial and rational wave solutions of Kudryashov–Sinelshchikov equation and numerical simulations for its dynamic motions,* J. Appl. Anal. Comput. 10(5), (2020), 2145–2162.

[2] Ali, L., Liu, X., Ali, B., Mujeed, S., Abdal, S., and Mutahir, A. *The impact of nanoparticles due to applied magnetic dipole in micropolar fluid flow using the finite element method,* Symmetry, 12(4), (2020), 520.

[3] Belhocine, A., and Abdullah, O.I. *Thermomechanical model for the analysis of disc brake using the finite element method in frictional contact,* Multiscale Science and Engineering, 2(1), (2020), 27–41.

[4] Bildik, N., and Deniz, S. *New approximate solutions to the nonlinear Klein–Gordon equations using perturbation iteration techniques,* Discrete Continuous Dyn. Syst. Ser S, 13(3), (2020), 503–518.

[5] Chettri, K., Tamang, J., Chatterjee, P., and Saha, A. *Dynamics of nonlinear ion-acoustic waves in Venus' lower ionosphere,* Astrophys. Space Sci. 369(5), (2024).

[6] El-shamy, O., El-barkoki, R., Ahmed, H.M., Abbas, W., and Samir, I. *Exploration of new solitons in optical medium with higher-order dispersive and nonlinear effects via improved modified extended tanh function method,* Alex. Eng. J. 68, (2023), 611–618.

[7] Feng, Y.-Y., and Wang, C.-H. *Discontinuous finite element method applied to transient pure and coupled radiative heat transfer,* Int. Commun. Heat Mass Transf. 122, (2021), 105156.

[8] Gepreel, K. A. *Analytical methods for nonlinear evolution equations in mathematical physics,* Mathematics, 8(12), (2020), 2211.

[9] Ghanbari, B., Kumar, S., Niwas, M., and Baleanu, D. *The Lie symmetry analysis and exact Jacobi elliptic solutions for the Kawahara–KdV type equations,* Results Phys. 23, (2021), 104006.

[10] Gohar, M., Li, C., and Li, Z. *Finite difference methods for Caputo–Hadamard fractional differential equations,* Mediterr. J. Math. 17(6), (2020), 194.

[11] He, J.-H., Qie, N., and He, C.-H. *Solitary waves travelling along an unsmooth boundary,* Results Phys. 24, (2021), 104104.

[12] Hyder, A.-A., and Barakat, M.A. *General improved Kudryashov method for exact solutions of nonlinear evolution equations in mathematical physics,* Phys. Scr. 95(4), (2020), 045212.

[13] Jamshed, W. *Finite element method in thermal characterization and streamline flow analysis of electromagnetic silver-magnesium oxide nanofluid inside grooved enclosure,* Int. Commun. Heat Mass Transf. 130, (2022), 105795.

[14] Jin, Y.-T., and Chen, A.-H. *Resonant solitary wave and resonant periodic wave solutions of the Kudryashov–Sinelshchikov equation,* Phys. Scr. 95(8), (2020), 085208.

[15] Jisha, C.R., Dubey, R.K., Benton, D., and Rashid, A. *The exact solutions for Kudryashov and Sinelshchikov equation with variable coefficients,* Phys. Scr. 97(9), (2022), 095212.

[16] Johnson, R.S. *Water waves and Korteweg–de Vries equations,* J. Fluid Mech. 97(04), (1980), 701.

[17] Karakoc, S.B.G., Saha, A., Bhowmik, S.K., and Sucu, D.Y. *Numerical and dynamical behaviors of nonlinear traveling wave solutions of the Kudryashov–Sinelshchikov equation,* Wave Motion, 118, (2023), 103121.

[18] Karakoc, S.B.G., Saha, A., and Sucu, D. *A novel implementation of Petrov–Galerkin method to shallow water solitary wave pattern and superperiodic traveling wave and its multistability: Generalized Korteweg-de Vries equation,* Chin. J. Phys. 68, (2020), 605–617.

[19] Kruglov, V.I., and Triki, H. *Propagation of coupled quartic and dipole multi-solitons in optical fibers medium with higher-order dispersions,* Chaos Soliton. Fract. 172, (2023), 113526.

[20] Kumar, S., Niwas, M., and Dhiman, S.K. *Abundant analytical soliton solutions and different wave profiles to the Kudryashov–Sinelshchikov equation in mathematical physics,* J. Ocean Eng. Sci. 7(6), (2022), 565–577.

[21] Leissa, A. *The historical bases of the Rayleigh and Ritz methods,* J. Sound Vib. 287(4-5), (2005), 961–978.

[22] Liu, H., Zheng, X., Wang, H., and Fu, H. *Error estimate of finite element approximation for two-sided space-fractional evolution equation with variable coefficient,* J. Sci. Comput. 90(1), (2022), 15.

[23] Lu, J. (2018). *New exact solutions for Kudryashov–Sinelshchikov equation,* Adv. Diff. Equ. 2018(1), 374.

[24] Ma, Y.-L., Wazwaz, A.-M., and Li, B.-Q. *Soliton resonances, soliton molecules, soliton oscillations and heterotypic solitons for the nonlinear Maccari system,* Nonlinear Dyn. 111(19), (2023), 18331–18344.

[25] Mohebbi, A., and Dehghan, M. *High-order compact solution of the one-dimensional heat and advection–diffusion equations,* Appl. Math. Model. 34(10), (2010), 3071–3084.

[26] Nakazawa, S. *Computational Galerkin methods,* Comput. Methods Appl. Mech. Eng. 50(2), (1985), 199–200.

[27] Ozisik, M., Secer, A., Bayram, M., Cinar, M., Ozdemir, N., Esen, H., and Onder, I. *Investigation of optical soliton solutions of higher-order nonlinear Schrödinger equation having Kudryashov nonlinear refractive index,* Optik, 274, (2023), 170548.

[28] Pal, N.K., Chatterjee, P., and Saha, A. *Solitons, multi-solitons and multi-periodic solutions of the generalized Lax equation by Darboux transformation and its quasiperiodic motions,* Int. J. Mod. Phys. B, (2023).

[29] Rabinowitz, P.H. *On a class of nonlinear Schrodinger equations,* Z. fur Angew. Math. Phys., 43(2), (1992), 270–291.

[30] Ramadan, M., and Aly, H. *New approach for solving of extended KdV equation,* Alfarama Journal of Basic & Applied Sciences, (2022).

[31] Rezaei, S., Harandi, A., Moeineddin, A., Xu, B.-X., and Reese, S. *A mixed formulation for physics-informed neural networks as a potential solver for engineering problems in heterogeneous domains: Comparison with finite element method,* Comput. Methods Appl. Mech. Eng. 401, (2022), 115616.

[32] Rubinstein, J. *Sine-Gordon equation,* J. Math. Phys. 11(1), (1970), 258–266.

[33] Ryabov, P.N. *Exact solutions of the Kudryashov–Sinelshchikov equation,* Appl. Math. Comput. 217(7), (2010), 3585–3590.

[34] Saeed, T., and Abbas, I. *Finite element analyses of nonlinear DPL bio-heat model in spherical tissues using experimental data,* Mechanics Based Design of Structures and Machines, 50(4), (2022), 1287–1297.

[35] Shaikhova, G., Kutum, B., Altaybaeva, A., and Rakhimzhanov, B. *Exact solutions for the (3+1)-dimensional Kudryashov–Sinelshchikov equation,* J. Phys. Conf. Ser. 1416(1), (2019), 012030.

[36] Sultana, M., Arshad, U., Abdel-Aty, A.-H., Akgül, A., Mahmoud, M., and Eleuch, H. *New numerical approach of solving highly nonlinear fractional partial differential equations via fractional novel analytical method,* Fract. Fract. 6(9), (2022), 512.

[37] Weickert, J., Romeny, B., and Viergever, M. *Efficient and reliable schemes for nonlinear diffusion filtering,* IEEE Trans. Image Process. 7(3), (1998), 398–410.

[38] Yang, X., and He, X. *A fully-discrete decoupled finite element method for the conserved Allen–Cahn type phase-field model of three-phase fluid flow system,* Comput. Methods Appl. Mech. Eng. 389, (2022), 114376.

[39] Yin, X., Xu, L., and Yang, L. *Evolution and interaction of soliton solutions of Rossby waves in geophysical fluid mechanics,* Nonlinear Dyn. 111(13), (2023), 12433–12445.

[40] Yong, W., Zhang, W., Nguyen, H., Bui, X.-N., Choi, Y., Nguyen-Thoi, T., Zhou, J., and Tran, T.T. *Analysis and prediction of diaphragm wall deflection induced by deep braced excavations using finite element method and artificial neural network optimized by metaheuristic algorithms,* Reliab. Eng. Syst. Saf., 221, (2022), 108335.

[41] Zahra, W.K., Ouf, W.A., and El-Azab, M.S. *An effective scheme based on quartic B-spline for the solution of Gardner equation and Harry Dym equation,* Communications on Advanced Computational Science with Applications, 2016(2), (2016), 82–94.

Aims and scope

Iranian Journal of Numerical Analysis and Optimization (IJNAO) is published twice a year by the Department of Applied Mathematics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad. Papers dealing with different aspects of numerical analysis and optimization, theories and their applications in engineering and industry are considered for publication.

Journal Policy

All submissions to IJNAO are first evaluated by the journal's Editor-in-Chief or one of the journal's Associate Editors for their appropriateness to the scope and objectives of IJNAO. If deemed appropriate, the paper is sent out for review using a single blind process. Manuscripts are reviewed simultaneously by reviewers who are experts in their respective fields. The first review of every manuscript is performed by at least two anonymous referees. Upon the receipt of the referee's reports, the paper is accepted, rejected, or sent back to the author(s) for revision. Revised papers are assigned to an Associate Editor who makes an evaluation of the acceptability of the revision. Based upon the Associate Editor's evaluation, the paper is accepted, rejected, or returned to the author(s) for another revision. The second revision is then evaluated by the Editor-in-Chief, possibly in consultation with the Associate Editor who handled the original paper and the first revision, for a usually final resolution.

The authors can track their submissions and the process of peer review via: http://ijnao.um.ac.ir

All manuscripts submitted to IJNAO are tracked by using ”iThenticate” for possible plagiarism before acceptance.

Instruction for Authors

The Journal publishes all papers in the fields of numerical analysis and optimization. Articles must be written in English.

All submitted papers will be refereed and the authors may be asked to revise their manuscripts according to the referee's reports. The Editorial Board of the Journal keeps the right to accept or reject the papers for publication.

The papers with more than one authors, should determine the corresponding author. The e-mail address of the corresponding author must appear at the end of the manuscript or as a footnote of the first page.

It is strongly recommended to set up the manuscript by Latex or Tex, using the template provided in the web site of the Journal. Manuscripts should be typed double-spaced with wide margins to provide enough room for editorial remarks.

References should be arranged in alphabetical order by the surname of the first author as examples below:

[1] Brunner, H. *A survey of recent advances in the numerical treatment of Volterra integral and integro-differential equations*, J. Comput. Appl. Math. 8 (1982), 213-229.

[2] Stoer, J. and Bulirsch, R. *Introduction to Numerical Analysis*, Springer-verlag, New York, 2002.

Iranian Journal of Numerical Analysis and Optimization

CONTENTS Vol. 14, No. 4, pp 991-1333

H. Samy, W. Adel, I. Hanafy and M. Ramadan

web site: **https://ijnao.um.ac.ir**
Email: **ijnao@um.ac.ir**
ISSN-Print: **2423-6977**
ISSN-Online: **2423-6969**