Research Paper
Computational Methods for Differential Equations
http://cmde.tabrizu.ac.ir
Vol. *, No. *, *, pp. 1-25
DOI:10.22034/cmde.2025.66698.3144

Strong and weak solutions of fuzzy nonlinear optimal control problems via a Jacobi-based neural network scheme

Aneseh Kazemi and Alireza Nazemi*
Faculty of Mathematical Sciences, Shahrood University of Technology, Shahrood, Iran.

Abstract

In this paper, an artificial neural network architecture is proposed for solving a class of fuzzy optimal control problems. At the first step, we consider the Pontryagin minimum principle for the mentioned problems. The necessary optimality conditions for these problems are stated in the form of two-point boundary value problems. Then, for the first time, a neural network solution method is introduced in which Jacobi functions are employed as activation functions in one of the hidden layers to approximate solutions to two-point boundary value problems. This neural network uses roots of Jacobi polynomials as the training dataset, and the Levenberg-Marquardt algorithm is chosen as the optimizer. By relying on the ability of the generalized fuzzy hyperbolic models as function approximator, the trial solutions of variables are substituted in the related two-point boundary value problem. The obtained algebraic nonlinear equations system is then reduced into an error function minimization problem. A learning scheme based on the Levenberg-Marquardt algorithm is employed as the optimizer to derive the adjustable parameters of fuzzy solutions. To show the effectiveness of the presented neural network, some numerical results are provided.

Keywords. Fuzzy optimal control, Necessary optimality conditions, Jacobi neural network, Levenberg-Marquardt algorithm.
2010 Mathematics Subject Classification. 03-XX, 34A07, 49-XX, 92B20,65Kxx.

1. Introduction

Classical optimal control problems play a crucial role in science and engineering. These problems aim to find a pair of control-state variables that satisfy a dynamic system and minimize an objective functional. Typically, optimal control problems are solved using the Pontryagin minimum principle (PMP), a generalization of the classic Euler–Lagrange and Weierstrass necessary optimality conditions for the calculus of variations. Many authors have approached optimal control problems from various perspectives, with some of the latest research found in [3, 4, 12, 42, 43, 47, 50].

In recent decades, fuzzy optimal control problems have gained considerable attention. Various approaches to fuzzy optimal control for nonlinear systems have been studied, often based on the Takagi–Sugeno (T–S) fuzzy model [49]. Filev and Angelov [18] used fuzzy mathematical programming to formulate and solve the fuzzy optimal control problem. Zhao and Zhu [52] explored the concept of fuzzy processes and addressed the existence and uniqueness of a linear quadratic fuzzy optimal control problem. Farhadinia [17] applied the fuzzy variational approach from [16] to fuzzy optimal control problems, deriving necessary optimality conditions based on the Buckley and Feuring derivative [7]. However, the main theorems in [16, 17] were found to be invalid, and the results deduced were incorrect. Specifically, in Example 7.1 of [17], the fuzzy Lagrange multiplier $\tilde{\lambda}(\alpha) = [24(3-\alpha), 48\alpha]$ is not a fuzzy number, and the solution $\tilde{x}(t,\alpha)$ is not differentiable when $t \in (\frac{1}{2}, 1]$, meaning $[\tilde{\dot{x}}(t)]^{\alpha}$ did not represent an α-level set of a fuzzy number. Similarly, in Example 7.1 of [16], the α-level set of optimal fuzzy control

$$[\tilde{u}^*(t)]^{\alpha} = [\frac{-(3-2\alpha)e^{-(1-t)(3-2\alpha)}}{\sinh(3-2\alpha)}, \frac{-\alpha e^{-(1-t)\alpha}}{\sinh(\alpha)}],$$

Received: 09 April 2025; Accepted: 09 May 2026.
∗ Corresponding author. Email: nazemi20042003@gmail.com.

did not represent an α-level set of a fuzzy number.

Najariyan and Farahi explored fuzzy optimal control problems of a linear dynamical system with uncertain initial conditions [33] and uncertain parameters [34]. However, Example 4.2 by Solaki et al. [46] highlighted that the solutions derived were not extremals for a specific instance of fuzzy optimal control problems, indicating a discrepancy in [34]. Moreover, in [33, 34], the authors examined the notion of Strongly Generalized Hukuhara (SGH) differentiability in a fuzzy dynamical system and for designing fuzzy optimal control. However, the SGH derivative has some shortcomings: (i) it does not always exist, and (ii) it requires the monotony of the uncertainty. To address these issues, Najariyan and Farahi evaluated a class of fuzzy optimal control problems using the generalized Hukuhara (gH) differentiability concept [35]. Although the gH differentiability notion does not require the monotony of solution fuzziness, the existence of this derivative is not guaranteed. Additionally, approaches based on SGH and gH differentiability face a significant limitation called the Unnatural Behavior in Modeling (UBM) phenomenon [29]. The UBM phenomenon arises from the fact that, based on fuzzy standard interval arithmetic (FSIA), the solutions of different forms of a fuzzy differential equation with the same structure may differ. For example, consider the fuzzy differential equations

$$\dot{\tilde{x}}(t) = \tilde{x}(t) + \tilde{\kappa}, \dot{\tilde{x}}(t) - \tilde{x}(t) = \tilde{\kappa}, \text{ and } \dot{\tilde{x}}(t) - \tilde{d} = \tilde{\kappa},$$

where the initial conditions are the same and equal to a fuzzy number, $\tilde{\kappa}$ is a fuzzy number, and the derivative $\dot{\tilde{x}}$ is SGH or gH. As a result, the solutions obtained by solving these fuzzy differential equations may differ, exemplifying the UBM phenomenon. Additionally, [29] discusses other limitations associated with these derivatives, such as doubling property and multiplicity of solutions.

These limitations motivated some authors to explore fuzzy optimal control problems using a method that not only defines fuzzy derivatives but also addresses these issues. To overcome the difficulties of the gH derivative, Mustafa et al. [31, 32] defined the fuzzy optimal control problem with state conditions at the final time and derived the necessary conditions for this problem as the main goal of their paper. They introduced two types of fuzzy Hamiltonian functions, used alongside the concepts of parameterizing the fuzzy-valued function and its differentiation and integration by the left- and right-hand functions of its α-level set and fuzzy variational approaches to prove the necessary conditions for the fuzzy optimal control problem with state conditions at the final time. Additionally, they introduced the concepts of strong and weak solutions to ensure that the optimal solutions are fuzzy functions.

Artificial neural networks (ANNs) are abstract computational schemes inspired by biological systems. They resemble the behavior and functionality of the human brain and neural system by acquiring and storing knowledge. ANNs consist of neurons and their interconnections, representing synaptic connections and neuron activation functions. These functions model possible processes occurring inside each neuron in the human brain, enabling ANNs to learn and produce desired outputs for various problems. ANNs are typically represented as directed graphs, with nodes performing simple mathematical operations (activation functions) on data.

Numerical weights associated with directed edges between nodes represent information within these links and allow ANNs to learn and adapt during several iterations of inputting sample datasets. The process of updating the weights based on the type of network inputs is called learning. ANNs are categorized based on their topologies. Single-layer ANNs are directly connected, while multilayer ANNs have additional neurons that act as signal bridges, connecting neurons in previous and next layers, forming hidden layers. The number of layers in multilayer ANNs depends on various criteria, such as the number of input signals, input and output neurons, and learning paradigms. The outputs of neurons are multiplied by their related weights, and the products are fed through the activation function to generate a result.

Recently, ANN methods for solving various types of differential equations have been rapidly developing. Different ANN frameworks, such as radial basis function networks [22, 26], orthogonal neural networks [27, 28], and deep neural networks [38], simulate dynamical systems using nonlinear optimization methodologies. Orthogonal neural networks, in particular, utilize orthogonal functions, which have suitable properties for approximating unknown functions. One important class of orthogonal functions is the Jacobi polynomials, which have been used to solve various problems in science and engineering. Recent studies on this topic can be found in [1, 2, 5, 30, 40].

Motivated by these discussions, this paper uses Jacobi polynomials as the basis of ANNs for solving fuzzy variational and fuzzy optimal control problems. We first consider PMP for fuzzy optimal control problems based on the concepts of differentiability and integrability of a fuzzy mapping parameterized by the left- and right-hand functions of its α-level sets. This approach leads to related two-point boundary value problems (TPBVPs). Next, we use a Jacobi neural network architecture to approximate the trial solutions of the necessary optimality conditions. These neural networks use the roots of Jacobi polynomials as the training dataset, and the Levenberg-Marquardt algorithm is selected to optimize the unknown weight parameters. It should be noted that, until now, fuzzy variational and fuzzy optimal control problems have not been solved using any neural network scheme.

The main contributions of this paper are reflected in the following aspects:

• The primary aim of this paper is to study the necessary conditions of the fuzzy optimal control problem of multiple variables, based on the concepts of differentiability and integrability of a fuzzy-valued function parameterized by the left- and right-hand functions of its α-level set and variational approaches, in order to provide solutions. However, the solutions of the fuzzy optimal control problem, optimal controls, and corresponding optimal states are not always fuzzy functions. To ensure that the solutions of the fuzzy optimal control problem of several variables are always fuzzy functions, we introduce the concepts of strong (fuzzy) and weak solutions.

• To address fuzzy optimal control problems, we propose a novel Jacobi-based neural network approach using Jacobi polynomials. To the best of our knowledge, this method has not been explored in the existing literature.

• The solution of the fuzzy optimal control via the Jacobi neural network is a smooth solution that can be easily used in subsequent calculations.

• A significant advantage of the Jacobi neural network is that it produces differentiable solutions expressed in a closed analytical form.

• Employing the Jacobi neural network provides a solution for fuzzy optimal control with excellent interpolation properties. The comparative results with the method outlined in [31, 32] clearly confirm this point.

The manuscript is organized as follows:

In Section 2, we recall and generalize some fundamental concepts of differentiability and integrability of a fuzzy mapping. In Section 3, we establish our main results concerning the necessary conditions of the fuzzy optimal control problem of multiple variables. Additionally, we state the definitions of strong (fuzzy) and weak solutions for the problem. Section 4 contains a description of Jacobi Polynomials and introduce a Jacobi neural network framework. Section 5 explains the method of updating the neural network weights via the Levenberg-Marquardt algorithm. Section 6 presents several numerical simulation results. Finally, section 7 concludes this paper.

2. Preliminaries

In this section, we recall some fundamental concepts of fuzzy analysis [17, 19, 51].

Definition 2.1. [19] A fuzzy set $\tilde{a} : \mathbb{R} \longrightarrow [0,1]$ is called a fuzzy number if $\tilde{a}$ is normal, convex fuzzy set, upper semi-continuous and $\text{supp}(a) = \overline{\{x \in \mathbb{R} : \tilde{a}(x) > 0\}}$ is compact, where $\overline{M}$ denotes the closure of M. n the rest of this paper we use E^1 to denote the fuzzy number space.

Definition 2.2. [19] The α–level set of $\tilde{a} \in E^1$ denoted by $\tilde{a}[\alpha] = [\underline{a}(\alpha), \overline{a}(\alpha)]$, is defined by

$$\tilde{a}[\alpha] = \begin{cases} \{x \in \mathbb{R} : \tilde{a}(x) \geq \alpha\}, & \text{if } 0 < \alpha \leq 1, \\ \\ \overline{\{x \in \mathbb{R} : \tilde{a}(x) > 0\}}, & \text{if } \alpha = 0. \end{cases}$$

Obviously, the α-level set $\tilde{a}[\alpha] = [\underline{a}(\alpha), \overline{a}(\alpha)]$ is bounded closed interval in $\mathbb{R}$ for all $\alpha \in [0,1]$, where $\underline{a}(\alpha)$ and $\overline{a}(\alpha)$ denote the left-hand and right-hand endpoints of $\tilde{a}[\alpha]$, respectively.

Definition 2.3. [19] The opposite of a fuzzy number $\tilde{a}$ is $-\tilde{a}$. It means $-\tilde{a}[\alpha] = [-\overline{a}(\alpha), -\underline{a}(\alpha)]$. Also, for $\tilde{a}, \tilde{b} \in E^1$, with $\tilde{a}[\alpha] = [\underline{a}(\alpha), \overline{a}(\alpha)]$, $\tilde{b}[\alpha] = [\underline{b}(\alpha), \overline{b}(\alpha)]$ and $k \in \mathbb{R}$,

- $(\tilde{a} \oplus \tilde{b})[\alpha] = \tilde{a}[\alpha] + \tilde{b}[\alpha] = \{x + y : x \in \tilde{a}[\alpha], y \in \tilde{b}[\alpha]\}$,
- $(k \odot \tilde{a})[\alpha] = k\tilde{a}[\alpha] = \{kx : x \in \tilde{a}[\alpha]\}$,
- $(\tilde{a} \odot \tilde{b})[\alpha] = [\min\{\underline{a}(\alpha)\underline{b}(\alpha), \underline{a}(\alpha)\overline{b}(\alpha), \overline{a}(\alpha)\underline{b}(\alpha), \overline{a}(\alpha)\overline{b}(\alpha)\},$
  $\max\{\underline{a}(\alpha)\underline{b}(\alpha), \underline{a}(\alpha)\overline{b}(\alpha), \overline{a}(\alpha)\underline{b}(\alpha), \overline{a}(\alpha)\overline{b}(\alpha)\}]$.

Any crisp number with value k can be regarded as a fuzzy number $\tilde{a}$ if its membership function is defined by,

$$\tilde{a}(x) = \begin{cases} 1, & x = k, \\ 0, & x \neq k. \end{cases}$$

Definition 2.4. [19] The metric structure is defined by the Hausdorff distance $D : E^1 \times E^1 \times \mathbb{R} \longrightarrow \mathbb{R}_+ \cup \{0\}$,

$$D(\tilde{a}, \tilde{b}) = \sup_{\alpha \in [0,1]} \max\{\mid \underline{a}(\alpha) - \underline{b}(\alpha) \mid, \mid \overline{a}(\alpha) - \overline{b}(\alpha) \mid\}.$$

D satisfies the following properties:

- (E^n, D) is a complete metric space,
- $D(\tilde{a} \oplus \tilde{c}, \tilde{b} \oplus \tilde{c}) = D(\tilde{a}, \tilde{b})$,
- $D(k \odot \tilde{a}, k \odot \tilde{b}) = \mid k \mid D(\tilde{a}, \tilde{b})$, where $\tilde{a}, \tilde{b}, \tilde{c} \in E^1$ and $k \in \mathbb{R}$.

Definition 2.5. [19] A triangular fuzzy number is denoted by $\tilde{a} = (u, v, w)$, and its α-cuts is denoted by $\tilde{a}[\alpha] = [u + \alpha(v - u), w - \alpha(w - v)]$.

Lemma 2.6. *[19] Let $\underline{a} : [0; 1] \longrightarrow \mathbb{R}$ and $\overline{a} : [0; 1] \longrightarrow \mathbb{R}$ satisfy the following conditions:*

C_1**:** *$\underline{a}$ is a bounded increasing function;*

C_2**:** *$\overline{a}$ is a bounded decreasing function;*

C_3**:** *$\underline{a}(1) \leq \overline{a}(1)$;*

C_4**:** $\lim_{\alpha \longrightarrow k^-} \underline{a}(\alpha) = \underline{a}(k)$, and $\lim_{\alpha \longrightarrow k^-} \overline{a}(\alpha) = \overline{a}(k)$, for all $0 < k \leq 1$;

C_5**:** $\lim_{\alpha \longrightarrow 0^+} \underline{a}(\alpha) = \underline{a}(0)$, and $\lim_{\alpha \longrightarrow 0^+} \overline{a}(\alpha) = \overline{a}(0)$.

Then $\tilde{a} : [0, 1] \longrightarrow \mathbb{R}$ characterized by $\tilde{a}(x) = \sup\{\alpha : \underline{a}(\alpha) \leq x \leq \overline{a}(\alpha)\}$ is a fuzzy number with with $[\tilde{a}]_\alpha = [\underline{a}(\alpha), \overline{a}(\alpha)]$. Moreover $\tilde{a} : [0, 1] \longrightarrow \mathbb{R}$ is a fuzzy number with $[\tilde{a}]_\alpha = [\underline{a}(\alpha), \overline{a}(\alpha)]$, then $\underline{a}(\alpha), \overline{a}(\alpha)$ satisfy conditions C_1 - C_5.

Definition 2.7. [17] Suppose that $\tilde{a}, \tilde{b} \in E^1$, where $\tilde{a}[\alpha] = [\underline{a}(\alpha), \overline{a}(\alpha)]$ and $\tilde{b}[\alpha] = [\underline{b}(\alpha), \overline{b}(\alpha)]$ for all $\alpha \in [0, 1]$, the H–difference of $\tilde{a}$ and $\tilde{b}$ is defined by

$$\tilde{a} \ominus_H \tilde{b} = \tilde{c} \Longleftrightarrow \tilde{a} = \tilde{b} \oplus \tilde{c}.$$

Obviously, $\tilde{a} \ominus_H \tilde{a} = \tilde{0}$, and the α–level set of H–difference is

$$(\tilde{a} \ominus_H \tilde{b})[\alpha] = [\underline{a}(\alpha) - \underline{b}(\alpha), \overline{a}(\alpha) - \overline{b}(\alpha)], \qquad \forall \alpha \in [0, 1].$$

Definition 2.8. [17] Let $\tilde{a}, \tilde{b} \in E^1$, we write $\tilde{a} \preceq \tilde{b}$, if $\underline{a}(\alpha) \leq \underline{b}(\alpha)$ and $\overline{a}(\alpha) \geq \overline{b}(\alpha)$ for all $\alpha \in [0, 1]$. We also write $\tilde{a} \prec \tilde{b}$, if $\tilde{a} \preceq \tilde{b}$ and there exists $\alpha_0 \in [0, 1]$ such that $\underline{a}(\alpha_0) < \underline{b}(\alpha_0)$ or $\overline{a}(\alpha_0) < \overline{b}(\alpha_0)$. Furthermore,$\tilde{a} = \tilde{b}$, if $\tilde{a} \preceq \tilde{b}$ and $\tilde{a} \succeq \tilde{b}$. In other words, $\tilde{a} = \tilde{b}$, if $\tilde{a}[\alpha] = \tilde{b}[\alpha]$ for all $\alpha \in [0, 1]$. $\tilde{a}, \tilde{b} \in E^1$ are said to be comparable if either $\tilde{a} \preceq \tilde{b}$ or $\tilde{a} \succeq \tilde{b}$, and non–comparable otherwise.

Definition 2.9. [51] The function $\tilde{F} : S \subseteq \mathbb{R} \longrightarrow E^1$ is called a fuzzy valued function if $\tilde{F}(t)$ is assign a fuzzy number for any $t \in S$. We also denote $\tilde{F}(t)[\alpha] = [\underline{F}(t, \alpha), \overline{F}(t, \alpha)]$, where $\underline{F}(t, \alpha) = \underline{(\tilde{F}(t))}(\alpha) = \min\{\tilde{F}(t)[\alpha]\}$ and $\overline{F}(t, \alpha) = \overline{(\tilde{F}(t))}(\alpha) = \max\{\tilde{F}(t)[\alpha]\}$. Therefore any fuzzy valued function $\tilde{F}$ may be understood by $\underline{F}(t, \alpha)$ and $\overline{F}(t, \alpha)$ being respectively a bounded increasing function of α and a bounded decreasing function of $\alpha$for $\alpha \in [0, 1]$. Also it holds $\underline{F}(t, \alpha) \leq \overline{F}(t, \alpha)$ for any $\alpha \in [0, 1]$.

Definition 2.10. [17] The function $\tilde{F} : S \subseteq \mathbb{R} \longrightarrow E^1$ is continuous at $t \in S$, if both $\underline{F}(t, \alpha)$ and $\overline{F}(t, \alpha)$ are continuous functions at $t \in S$ for all $\alpha \in [0, 1]$.

Definition 2.11. [17] Let $S \subseteq \mathbb{R}$ and $t_0 \in S$ and h be such that $t_0 + h \in S$. A fuzzy valued function $\tilde{F} : S \longrightarrow E^1$ is said to be H-differentiable at $t_0 \in S$ if and only if there exists a fuzzy number $\dot{\tilde{F}}(t_0)$ such that the limits (with respect to metric D)

$$\lim_{h \longrightarrow 0^+} \frac{\tilde{F}(t_0 + h) \ominus_H \tilde{F}(t_0)}{h}, \quad \lim_{h \longrightarrow 0^+} \frac{\tilde{F}(t_0) \ominus_H \tilde{F}(t_0 - h)}{h},$$

both exist and are equal to $\dot{\tilde{F}}(t_0) \in E^1$ In this case $\dot{\tilde{F}}(t_0)$ is called the H-derivative of $\tilde{F}$ at t_0. If $\tilde{F}$ is H-differentiable at any $t \in S$, we call $\tilde{F}$ is H-differentiable over S.

Moreover, if a fuzzy valued function $\tilde{F} : S \longrightarrow E^1$ is H-differentiable at $t_0 \in S$, then $\dot{\underline{F}}(t, \alpha)$ and $\dot{\overline{F}}(t, \alpha)$ are differentiable at $t_0 \in S$ for all $\alpha \in [0, 1]$, and we have

$$\dot{\tilde{F}}(t_0)[\alpha] = [\dot{\underline{F}}(t_0, \alpha), \dot{\overline{F}}(t_0, \alpha)].$$

Definition 2.12. [17] Let $\tilde{F} : [t_0, t_1] \longrightarrow E^1$. We say that $\tilde{F}$ is Fuzzy–Riemann integrable to $I \in E^1$ if for any $\epsilon > 0$, there exists $\delta > 0$ such that for any division $P = \{[i, j]; \xi\}$ of $[t_0, t_1]$ with the norms $\Delta(p) < \delta$, we have

$$D\big(\sum_{p}^{*}(i - j) \odot \tilde{F}(\xi); I\big) < \epsilon,$$

where $\sum^*$ denotes the fuzzy summation. We choose to write $I := (FR) \int_{t_0}^{t_1} \tilde{F}(t)dt$. Furthermore, for any $\alpha \in [0, 1]$

$$(FR) \int_{t_0}^{t_1} \tilde{F}(t)[\alpha]dt = \big[(R) \int_{t_0}^{t_1} \underline{F}(t, \alpha)dt, (R) \int_{t_0}^{t_1} \overline{F}(t, \alpha)dt\big].$$

Definition 2.13. [17] Suppose that $\tilde{F}, \tilde{G} : S \subseteq \mathbb{R} \longrightarrow E^1$ are two fuzzy functions. The distance measure between $\tilde{F}$ and $\tilde{G}$ is defined by

$$\begin{aligned} D_{E^1}(\tilde{F}(t), \tilde{G}(t)) &= \sup_{\alpha\in[0,1]} \mathbf{H}(\tilde{F}(t)[\alpha], \tilde{G}(t)[\alpha]) \\ &= \max\{ \sup_{z\in\tilde{F}(t)[\alpha]} d(z, \tilde{G}(t)[\alpha]), \sup_{y\in\tilde{G}(t)[\alpha]} d(\tilde{F}(t)[\alpha], y)\}, \quad \forall t \in S, \end{aligned}$$

where $\mathbf{H}$ is the Hausdorff metric on the family of all nonempty compact subsets of $\mathbb{R}$, and

$$d(a, B) = \inf_{b\in B} d(a, b).$$

3. Fuzzy optimal control problems

In this section, we begin by addressing the fuzzy optimal control problem, which includes state constraints at the final time, along with the associated necessary conditions. Subsequently, we explore both strong and weak solutions of the fuzzy optimal control problem.

3.1. **Statement of the problem.** Let $\tilde{x} = \tilde{x}(t)$ be a fuzzy function of $t \in [t_0, t_f] \subseteq \mathbb{R}$ and belonging to the class of fuzzy functions with continuous first derivatives with respect to $t \in [t_0, t_f]$. The fuzzy optimal control problem is stated as follows:

$$\begin{aligned} &\min_{\tilde{u}} \tilde{J}(\tilde{u}) = \tilde{\psi}(\tilde{x}(t_f), t_f) \oplus (FR) \int_{t_0}^{t_f} \tilde{f}(\tilde{x}(t), \tilde{u}(t), t)dt, \\ &\text{s.t. } \dot{\tilde{x}}(t) = \tilde{g}(\tilde{x}(t), \tilde{u}(t), t), \\ &\tilde{x}(t_0) = \tilde{x}_0, \qquad \tilde{x}(t_f) \text{ is free,} \end{aligned} \tag{3.1}$$

where $\tilde{f}, \tilde{g} : E^1 \times E^1 \times \mathbb{R} \longrightarrow E^1$ are fuzzy valued functions, the fuzzy state $\tilde{x}(t)$ and the fuzzy control $\tilde{u}(t)$ are functions of $t \in [t_0, t_f] \subseteq \mathbb{R}$, meanwhile, the fuzzy state function $\tilde{x}(t)$ is assumed to be H–differentiable with respect to $t \in [t_0, t_f]$. The integrand $\tilde{f}$ and the right hand side of the fuzzy differential equation $\tilde{g}$ are considered to have continuous first and second partial derivatives with respect to all of their arguments.

Definition 3.1. (Admissible fuzzy state)[32] We say that $\tilde{x}(t)$ is admissible, if it satisfies appropriate boundary condition and also is twice continuously differentiable with respect to $t \in [t_0, t_f]$. Throughout this paper, we assume that an admissible fuzzy control $\tilde{u}(t)$ is not bounded.

Remark 3.2. The fuzzy Hamiltonian function for Problem (3.1) is defined as follows:

$$\tilde{H}(\tilde{x}(t),\tilde{u}(t),\tilde{p}(t),t) = \tilde{f}(\tilde{x}(t),\tilde{u}(t),t) \oplus \tilde{p}(t) \odot \tilde{g}(\tilde{x}(t),\tilde{u}(t),t). \tag{2.4}$$

Note that, according to the definition of the multiplication of two fuzzy numbers presented in Definition 2.3, the α-cut corresponding to the fuzzy Hamiltonian function defined in (2.4) has 16 possible cases. For simplicity, we consider only one of these 16 cases as follows:

$$\tilde{H}(\tilde{x}(t),\tilde{u}(t),\tilde{p}(t),t)[\alpha] = [\underline{H}(\underline{x},\underline{u},\underline{p},t,\alpha), \overline{H}(\overline{x},\overline{u},\overline{p},t,\alpha)], \tag{3.2}$$

with the α- level set

$$\underline{H}(\underline{x},\underline{u},\underline{p},t,\alpha) = \underline{f}(\underline{x},\underline{u},\underline{p},t,\alpha) + \underline{p}^\top \underline{g}(\underline{x},\underline{u},\underline{p},t,\alpha),$$
$$\overline{H}(\overline{x},\overline{u},\overline{p},t,\alpha) = \overline{f}(\overline{x},\overline{u},\overline{p},t,\alpha) + \overline{p}^\top \overline{g}(\overline{x},\overline{u},\overline{p},t,\alpha).$$

Theorem 3.3. *[32] (Necessary optimality conditions). Assume that $\tilde{x}^*(t)$ be an admissible fuzzy state and $\tilde{u}^*(t)$ be an admissible fuzzy control. If $(\tilde{x}^*,\tilde{u}^*)$ is an optimal solution to the problem (3.1), then there exists costate function $\tilde{p}^*(t)$ such that $([\tilde{x}^*(t),\tilde{u}^*(t),\tilde{p}^*(t),t,\alpha])$ satisfies the following system for all $\alpha \in [0,1]$, $t \in [t_0,t_f]$:*

$$\begin{cases} \dot{\underline{x}}^*(t,\alpha) = \dfrac{\partial \underline{H}}{\partial \underline{p}}(\underline{x}^*(t,\alpha),\underline{u}^*(t,\alpha),\underline{p}^*(t,\alpha),t,\alpha), \\ \dot{\overline{x}}^*(t,\alpha) = \dfrac{\partial \overline{H}}{\partial \overline{p}}(\overline{x}^*(t,\alpha),\overline{u}^*(t,\alpha),\overline{p}^*(t,\alpha),t,\alpha), \\ \dot{\underline{p}}^*(t,\alpha) = -\dfrac{\partial \underline{H}}{\partial \underline{x}}(\underline{x}^*(t,\alpha),\underline{u}^*(t,\alpha),\underline{p}^*(t,\alpha),t,\alpha), \\ \dot{\overline{p}}^*(t,\alpha) = -\dfrac{\partial \overline{H}}{\partial \overline{x}}(\overline{x}^*(t,\alpha),\overline{u}^*(t,\alpha),\overline{p}^*(t,\alpha),t,\alpha), \\ 0 = \dfrac{\partial \underline{H}}{\partial \underline{u}}(\underline{x}^*(t,\alpha),\underline{u}^*(t,\alpha),\underline{p}^*(t,\alpha),t,\alpha), \\ 0 = \dfrac{\partial \overline{H}}{\partial \overline{u}}(\overline{x}^*(t,\alpha),\overline{u}^*(t,\alpha),\overline{p}^*(t,\alpha),t,\alpha), \\ \underline{p}(t_f,\alpha) = \dfrac{\partial \underline{\psi}}{\partial \underline{x}}\,|_{t=t_f}, \\ \overline{p}(t_f,\alpha) = \dfrac{\partial \overline{\psi}}{\partial \overline{x}}\,|_{t=t_f}, \end{cases} \tag{3.3}$$

where the fuzzy Hamiltonian function $\tilde{H}$ is defined as (3.2).

3.2. **The strong and weak solutions.** As demonstrated above, Theorem 3.3 provided the necessary conditions for $(\tilde{x}^*,\tilde{u}^*)$ to be optimal solutions of problem (3.1). From Lemma 2.6, we know that the left and right-hand sides of the generated functions of the α-level set of any fuzzy number must satisfy certain properties. Consequently, the left and right-hand side functions of the α-level set of optimal fuzzy control-state $(\tilde{x}^*,\tilde{u}^*)$ must adhere to these properties. To ensure that the solutions of the problem (3.1) and the associated problem are fuzzy functions, Mustafa et al. [31, 32] suggested the concepts of strong and weak solutions. Furthermore, $\tilde{u}^*(t)$ and $\tilde{x}^*(t)$ are fuzzy numbers, where $\tilde{u}^*(t)[\alpha] = [\underline{u}^*(t,\alpha),\overline{u}^*(t,\alpha)]$ and $\tilde{x}^*(t)[\alpha] = [\underline{x}^*(t,\alpha),\overline{x}^*(t,\alpha)]$ if $\underline{u}^*(t,\alpha)$, $\overline{u}^*(t,\alpha)$, $\underline{x}^*(t,\alpha)$, and $\overline{x}^*(t,\alpha)$ satisfy the properties stated in Lemma 2.6.

In the following, based on conditions C_1 and C_2 of Lemma 2.6, the definitions of strong and weak solutions are introduced for the fuzzy optimal control problem with state conditions at the final time, as defined in (3.1).

Definition 3.4. • (Strong solution): $\tilde{u}^*(t)[\alpha]$ and $\tilde{x}^*(t)[\alpha]$ are defined as strong solutions of the problem (3.1), if $\underline{u}^*(t,\alpha),\overline{u}^*(t,\alpha),\underline{x}^*(t,\alpha)$ and $\overline{x}^*(t,\alpha)$ obtained from (3.3) satisfy conditions C_1 and C_2 of Lemma 2.6, for all $t \in [t_0,t_f]$ and $\alpha \in [0,1]$.

• (Weak solution): $\tilde{u}^*(t)[\alpha]$ and $\tilde{x}^*(t)[\alpha]$ are defined as weak solutions of the problem (3.1), if $\underline{u}^*(t,\alpha),\overline{u}^*(t,\alpha),\underline{x}^*(t,\alpha)$ and $\overline{x}^*(t,\alpha)$ obtained from (3.3) do not satisfy conditions C_1 and C_2 of Lemma 2.6, meanwhile, $\tilde{u}^*(t)[\alpha]$ and $\tilde{x}^*(t)[\alpha]$

are defined as

$$\tilde{u}^*(t)[\alpha] = \begin{cases} [2\overline{u}^*(t,1) - \underline{u}^*(t,\alpha), \overline{u}^*(t,\alpha)], \underline{u}^*, \overline{u}^* \text{are decreasing functions of } \alpha, \\ [\underline{u}^*(t,\alpha), 2\underline{u}^*(t,1) - \overline{u}^*(t,\alpha)], \underline{u}^*, \overline{u}^* \text{are increasing functions of } \alpha, \\ [\overline{u}^*(t,\alpha), \underline{u}^*(t,\alpha)], \underline{u}^* \text{is decreasing and } \quad \overline{u}^* \text{is an increasing of } \ \alpha, \end{cases} \tag{3.4}$$

and

$$\tilde{x}^*(t)[\alpha] = \begin{cases} [2\overline{x}^*(t,1) - \underline{x}^*(t,\alpha), \overline{x}^*(t,\alpha)], \underline{x}^*, \overline{x}^* \text{are decreasing functions of } \alpha, \\ [\underline{x}^*(t,\alpha), 2\underline{x}^*(t,1) - \overline{x}^*(t,\alpha)], \underline{x}^*, \overline{x}^* \text{are increasing functions of } \alpha, \\ [\overline{x}^*(t,\alpha), \underline{x}^*(t,\alpha)], \underline{x}^* \text{is decreasing and } \quad \overline{x}^* \text{is an increasing of } \ \alpha, \end{cases} \tag{3.5}$$

for all $t \in [t_0, t_f]$ and $\alpha \in [0,1]$.

4. Jacobi neural network

In this section, we first consider some properties of Jacobi polynomials. Then, a Jacobi based neural network architecture is constructed for solving the optimality conditions (3.3).

4.1. Jacobi polynomials.

Orthogonal functions have many applications in approximation theory as well as in different numerical algorithms such as spectral schemes. Here an important theorem about the orthogonal functions is remarked.

Theorem 4.1. *[23] Let $\{\varsigma_n(x)\}$ be the sequence of orthogonal polynomials determined by the positive integrable function $w(x)$. These polynomials are assumed to be orthogonal over the closed finite interval $[a,b]$, that is*

$$\int_a^b \varsigma_n(x)\varsigma_m(x)w(x)dx = c_n\delta_{m,n},$$

where c_n is a constant and $\delta_{m,n}$ is Kronecker delta function. Then the set $\{\varsigma_n(x)\}$ is complete with respect to continuous functions over $[a,b]$.

One of the most valuable and significant types of orthogonal functions is Jacobi polynomials. There are some special cases of this orthogonal polynomial family such as Legendre, the four kinds of Chebyshev, and Gegenbauer polynomials [11, 14]. The Jacobi polynomials that are denoted by $J_M^{\beta,\gamma}(Z)$ are orthogonal over $[-1,1]$ interval with $W^{\beta,\gamma}(Z) = (1-Z)^\beta(1+Z)^\gamma$ weight function. Jacobi polynomials can be defined as eigenfunctions of a singular Sturm-Liouville differential equation as follows:

$$\frac{d}{dZ}\left((1-Z)^{\beta+1}(1+Z)^{\gamma+1}\frac{d}{dZ}J_M^{\beta,\gamma}(Z)\right) + (1-Z)^\beta(1+Z)^\gamma\lambda_M J_M^{\beta,\gamma}(Z) = 0, \tag{4.1}$$

where $\lambda_M = M(M+\beta+\gamma+1)$. Besides, these polynomials can be calculated by a recursive formula as

$$\begin{aligned} &J_0^{\beta,\gamma}(Z) = 1, \qquad J_1^{\beta,\gamma}(Z) = \frac{1}{2}(\beta+\gamma+2)Z + \frac{1}{2}(\beta-\gamma), \\ &J_{M+1}^{\beta,\gamma}(Z) = (a_M Z - b_M)J_M^{\beta,\gamma}(Z) - C_M J_{M-1}^{\beta,\gamma}(Z), \qquad M \geq 1, \end{aligned} \tag{4.2}$$

where $a_M = \dfrac{(2M+\beta+\gamma+1)(2M+\beta+\gamma+2)}{2(M+1)(M+\beta+\gamma+1)}$, $b_M = \dfrac{(\gamma^2-\beta^2)(2M+\beta+\gamma+1)}{2(M+1)(M+\beta+\gamma+1)(2M+\beta+\gamma)}$, and $c_M = \dfrac{(M+\beta)(M+\gamma)(2M+\beta+\gamma+2)}{(M+1)(M+\beta+\gamma+1)(2M+\beta+\gamma)}$. On top of that, these polynomials have some special properties, which are brought as follows:

$$J_M^{\beta,\gamma}(-Z) = (-1)^M J_M^{\gamma,\beta}(Z), \tag{4.3}$$

$$J_M^{\beta,\gamma}(1) = \frac{\Gamma(M+\beta+1)}{M! \times \Gamma(\beta+1)}, \tag{4.4}$$

$$\frac{d^m}{dZ^m}(J_M^{\beta,\gamma}(Z)) = \frac{\Gamma(m+M+\beta+\gamma+1)}{2^m \times \Gamma(M+\beta+\gamma+1)} J_{M-m}^{\beta+m,\gamma+m}(Z). \tag{4.5}$$

Theorem 4.2. *[48] The Jacobi polynomial $J_M^{\beta,\gamma}(Z)$ has exactly M real zeros on the interval $(-1,1)$.*

In order to obtain Jacobi functions over $[0,\eta]$, we use a transformation $Z = 2(\frac{t}{\eta}) - 1$,while $\eta > 0$. Utilizing this transformation, Jacobi functions are obtained, denoted by $J_M^{\beta,\gamma}(t)$ where these functions are orthogonal over the interval $[0,\eta]$ with weight function $W^{\beta,\gamma}(t) = (\eta - t)^\beta t^\gamma$. Similarly, there exists a Sturm-Liouville differential equation for fractional order of Jacobi functions. This Sturm-Liouville equation is as follows:

$$\frac{d}{dt}\left((\eta-t)^{\beta+1}t^{(\gamma+1)}\frac{d}{dt}J_M^{\beta,\gamma}(t)\right) + (\eta-t)^\beta t^\gamma \lambda_M J_M^{\beta,\gamma}(t) = 0, \tag{4.6}$$

where $\lambda_M = M(M+\beta+\gamma+1)$. In addition, the Jacobi functions over the interval $[0,\eta]$ can be calculated from the following recursive formula:

$$\begin{aligned} &J_0^{\beta,\gamma}(t) = 1, \qquad J_1^{\beta,\gamma}(t) = (\beta+\gamma+2)(\frac{t}{\eta}) - (\gamma+1), \\ &J_{M+1}^{\beta,\gamma}(t) = \left(2a_M(\frac{t}{\eta}) - (a_M+b_M)\right) J_M^{\beta,\gamma}(t) - c_M J_{M-1}^{\beta,\gamma}(t), M \geq 1. \end{aligned} \tag{4.7}$$

Theorem 4.3. *[39] The Jacobi function $J_M^{\beta,\gamma}(t)$ has precisely M real zeros on the interval $(0,\eta)$ with the following form:*

$$t_k = \eta(\frac{1+Z_k}{2}), \qquad k = 1,2,\ldots,M, \tag{4.8}$$

where Z_k is a root of $J_M^{\beta,\gamma}(Z)$.

Definition 4.4. Consider $\Lambda = \{t \mid 0 < t < \eta\}$ and $L^p_{W^{\beta,\gamma}}(\Lambda) = \{f : \Lambda \longrightarrow \mathbb{R} | f$ is measurable and $\| f \|_{W^{\beta,\gamma}} < \infty\}$, where $\| f \|_{W^{\beta,\gamma}} = \left(\int_0^\eta f^p(t) W^{\beta,\gamma}(t)dt\right)^{\frac{1}{p}}$. Moreover, define $\mathcal{F}_N = span\{J_M^{\beta,\gamma}(t), 0 \leq M \leq N\}$.

Theorem 4.5. *[6] For any function $\varpi \in \mathcal{F}_N$ and $1 \leq p \leq q \leq \infty$, we have*

$$\| \varpi \|_{L^q_{W^{\beta,\gamma}}}(\Lambda) \leq cN^{\sigma(\beta,\gamma)(\frac{1}{p}-\frac{1}{q})} \| \varpi \|_{L^p_{W^{\beta,\gamma}}}(\Lambda),$$

in which

$$\sigma(\beta,\gamma) = \begin{cases} 2\max(\beta,\gamma)+2, & \max(\beta,\gamma) \geq -\frac{1}{2}, \\ 1, & \textit{Otherwise.} \end{cases}$$

Theorem 4.6. *[6] Suppose that $\frac{d^k g}{dt^k} \in C[0,1]$ for $k = 0,1,\cdots,N-1$, and $(3+2N+\gamma) > 0$. If $g_N(t)$ is the best approximation to $g(x)$ from $\mathcal{F}_N$, then the error bound is presented as follows:*

$$\| g(t) - g_N(t) \|_{W^{\beta,\gamma}} \leq \frac{E}{\Gamma(N+2)}\sqrt{\frac{\Gamma(1+\beta)\Gamma(3+2N+\gamma)}{\Gamma(4+2N+\beta+\gamma)}},$$

in which $E > \mid \frac{d^{N+1}g}{dt^{N+1}} \mid$ and $t \in [0,1]$.

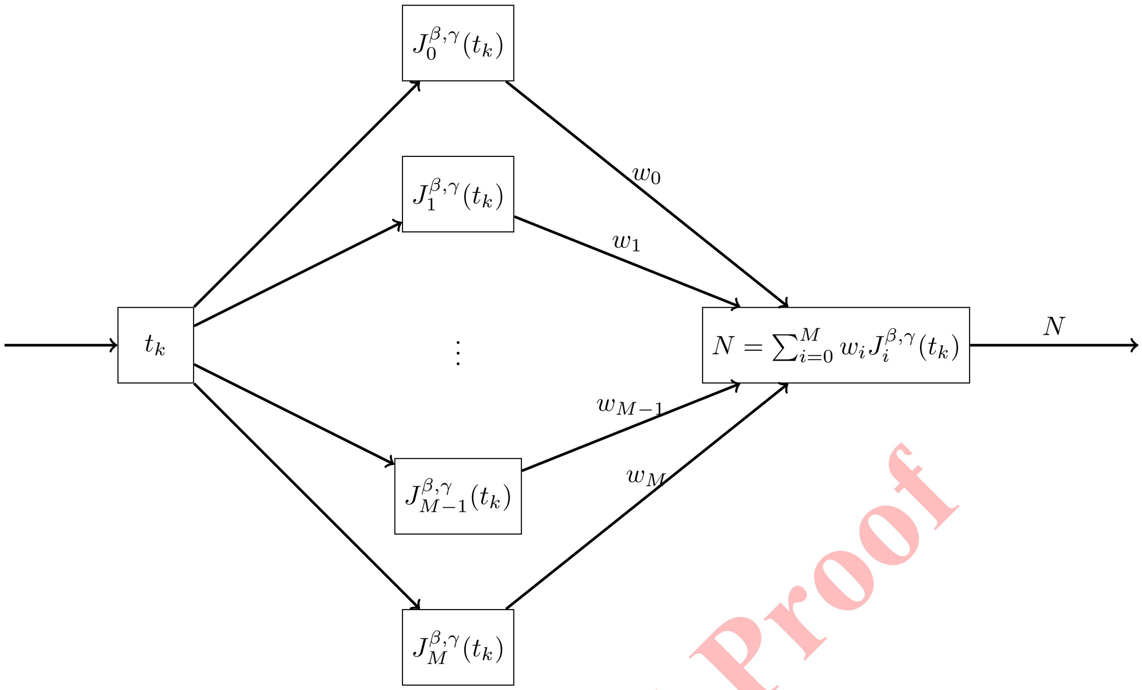


FIGURE 1. The topology of Jacobi neural network.

4.2. **Structure of Jacobi neural network.** Among the numerical methods, ANN schemes have demonstrated broad capabilities, proving to be flexible and robust tools for mathematical modeling and solving natural phenomena. Orthogonal neural networks are an advanced type of neural network based on the properties of orthogonal functions.

The approximation theory of orthogonal functions for building neural networks was initiated by Huang and Cheng [20]. According to this theory, any arbitrary function can be approximated as an orthogonal function. Additionally, the expansion terms of orthogonal functions are independent of each other, and the coefficients of these expansion terms are unique. Figure 1 shows a fundamental representation of an orthogonal neural network with one input layer, an orthogonal layer as the hidden layer (considered a functional expansion block), and an output layer with a single node. This topology, as depicted in Figure 1, includes $J_i^{\beta,\gamma}(t)$ as the Jacobi functions on $[0,\eta]$ and w_i as the weights of the network. Thus, considering the network as shown in Figure 1, the output of this neural network is as follows:

$$N(t,w) = \sum_{i=0}^{M} w_i J_i^{\beta,\gamma}(t), \tag{4.9}$$

where $w = (w_0, w_1, \dots, w_M)^\top$. The remaining vital steps to be taken are satisfying the optimality conditions set and training the network in such a way that it approximates the solution on the whole problem domain accurately and efficiently. In order to solve (3.3) by a Jacobi orthogonal neural network, first, it is necessary to approximate state variables $(\underline{x}(t)^\top, \overline{x}(t)^\top)^\top$, control functions $(\underline{u}(t)^\top, \overline{u}(t)^\top)^\top$ Lagrangian multipliers $(\underline{p}(t)^\top, \overline{p}(t)^\top)^\top$ using Jacobi orthogonal neural network. Then by constructing a suitable cost function and minimizing it, and, finally, training the network such that the approximations of $(\underline{x}(t)^\top, \underline{u}(t)^\top, \underline{p}(t)^\top)^\top$ and $(\overline{x}(t)^\top, \overline{u}(t)^\top, \overline{p}(t)^\top)^\top$ satisfy (3.3). Based on these facts and using (4.9), the trial solutions may be written as the sum of two terms in the following forms

$$\begin{cases} \underline{x}_T(t,\alpha) = \underline{x}(t_0) + (t-t_0)N_{\underline{x}}(t,\alpha,w^{\underline{x}}),\\ \overline{x}_T(t,\alpha) = \overline{x}(t_0) + (t-t_0)N_{\overline{x}}(t,\alpha,w^{\overline{x}}),\\ \underline{p}_T(t,\alpha) = (t-t_f)N_{\underline{p}}(t,\alpha,w^{\underline{p}}),\\ \overline{p}_T(t,\alpha) = (t-t_f)N_{\overline{p}}(t,\alpha,w^{\overline{p}}),\\ \underline{u}_T(t,\alpha) = N_{\underline{u}}(t,\alpha,w^{\underline{u}}),\\ \overline{u}_T(t,\alpha) = N_{\overline{u}}(t,\alpha,w^{\overline{u}}),\\ \underline{x}_T(t_0) = \underline{x}_0,\\ \overline{x}_T(t_0) = \overline{x}_0,\\ \underline{p}_T(t_f) = 0,\\ \overline{p}_T(t_f) = 0, \end{cases} \tag{4.10}$$

where the first term satisfies initial or boundary conditions and contains no adjustable parameters. The second term makes no contribution to initial or boundary conditions but this is used to a neural network whose weights are adjusted to minimize the error function.

Approximate solutions in (4.10) should be satisfied in system (3.3), we then get

$$\begin{cases} \dot{\underline{x}}(t,\alpha) = \dfrac{\partial \underline{H}_T}{\partial \underline{p}_T},\\ \dot{\overline{x}}(t,\alpha) = \dfrac{\partial \overline{H}_T}{\partial \overline{p}_T},\\ \dot{\underline{p}}(t,\alpha) = -\dfrac{\partial \underline{H}_T}{\partial \underline{x}_T},\\ \dot{\overline{p}}(t,\alpha) = -\dfrac{\partial \overline{H}_T}{\partial \overline{x}_T},\\ 0 = \dfrac{\partial \underline{H}_T}{\partial \underline{u}_T},\\ 0 = \dfrac{\partial \overline{H}_T}{\partial \overline{u}_T},\\ \underline{x}(t_0) = \underline{x}_0,\\ \overline{x}(t_0) = \overline{x}_0,\\ \underline{p}(t_f) = 0,\\ \overline{p}(t_f) = 0, \end{cases} \tag{4.11}$$

where

$$\underline{H}_T = H(\underline{x}_T, \underline{u}_T, \underline{p}_T, t), \qquad \overline{H}_T = H(\overline{x}_T, \overline{u}_T, \overline{p}_T, t).$$

We use the rth shifted Jacobi polynomial. The next step of this method is finding the best training dataset on which the network is taught to find the solution with the minimum error. To achieve this purpose, the roots of shifted Jacobi polynomials are used because not only do they maximize the convergence rate, but they are also well-scattered on the problem domain and will approximate the satisfactory solution on the interval rather than only on the neighborhood of training points. Let us assume $D = \{t_1, t_2, \ldots, t_r\}$ as the training data set, which contains the roots of rth-shifted Jacobi polynomial. The following cost function is formed in a way that $\zeta = (\underline{x}(t)^\top, \underline{u}(t)^\top, \underline{p}(t)^\top, \overline{x}(t)^\top, \overline{u}(t)^\top, \overline{p}(t)^\top)^\top$ will be the solution of (4.11):

$$E(\zeta,\vartheta) = \frac{1}{2}\sum_{i=1}^{6}\sum_{k=1}^{r}\{E_i(t_k,\vartheta)\}, \tag{4.12}$$

where $\vartheta = (w^{\underline{x}}, w^{\overline{x}}, w^{\underline{p}}, w^{\overline{p}}, w^{\underline{u}}, w^{\overline{u}})^\top$ and

$$\begin{cases} E_1(t_k,\vartheta) = \|\dot{\underline{x}}(t,\alpha) - \dfrac{\partial \underline{H}_T}{\partial \underline{p}_T}\|^2_{|t=t_k}, \\ E_2(t_k,\vartheta) = \|\dot{\overline{x}}(t,\alpha) - \dfrac{\partial \overline{H}_T}{\partial \overline{p}_T}\|^2_{|t=t_k}, \\ E_3(t_k,\vartheta) = \|\dot{\underline{p}}(t,\alpha) + \dfrac{\partial \underline{H}_T}{\partial \underline{x}_T}\|^2_{|t=t_k}, \\ E_4(t_k,\vartheta) = \|\dot{\overline{p}}(t,\alpha) + \dfrac{\partial \overline{H}_T}{\partial \overline{x}_T}\|^2_{|t=t_k}, \\ E_5(t_k,\vartheta) = \|\dfrac{\partial \underline{H}_T}{\partial \underline{u}_T}\|^2_{|t=t_k}, \\ E_6(t_k,\vartheta) = \|\dfrac{\partial \overline{H}_T}{\partial \overline{u}_T}\|^2_{|t=t_k}. \end{cases} \tag{4.13}$$

For leading ϑ^* to become an optimal solution of (4.11), the cost function should be minimized. Hence, the problem reduces to the following nonlinear least square problem:

$$\text{argmin}_{\vartheta} \quad E(\zeta,\vartheta). \tag{4.14}$$

In the next section, we illustrate the approach used for solving this nonlinear least square problem. For this purpose, the Levenberg-Marquardt algorithm is established, which is used to minimize the cost function value in (4.14).

5. Levenberg-Marquardt algorithm

The Levenberg-Marquardt algorithm is an iterative optimization technique used to solve nonlinear least squares problems, particularly in curve fitting and parameter estimation [24]. The Levenberg-Marquardt algorithm is widely used in nonlinear optimization, machine learning, and engineering applications. It combines the Gauss-Newton method, which provides fast convergence near a solution, and the gradient descent method, which enhances robustness when far from the solution, using a damping factor

Consider a vector function $f:\mathbb{R}^n \to \mathbb{R}^m$, $m \ge n$, and the objective

$$F(W) = \frac{1}{2}\sum_{i=1}^{m} f_i^2(W) = \frac{1}{2}\|f(W)\|^2.$$

Goal is to find $W^* \in \mathbb{R}^n$ such that

$$W^* = \arg\min_W \{F(W)\}, \quad \text{where } F(W^*) = 0.$$

(1) Initialize:
   - Choose an initial guess W_0.
   - Set initial damping parameter $\mu_0 > 0$ (small value).
   - Set convergence tolerances ϵ_1, ϵ_2 and maximum iterations $N_{\max}$.

(2) Compute Residuals and Jacobian:
   - Evaluate $f(W_n)$.
   - Compute the Jacobian $J_n = \frac{\partial f}{\partial W}(W_n) \in \mathbb{R}^{m\times n}$.
   - Compute gradient vector $g_n = J_n^\top f(W_n)$.

(3) Formulate the Damped Update: Solve

$$(J_n^\top J_n + \mu_n I)h_n = -g_n,$$

where I is the $n \times n$ identity matrix. - If μ_n is large, $h_n \approx -\frac{1}{\mu_n} g_n$ (steepest descent). - If μ_n is small, $h_n \approx -(J_n^\top J_n)^{-1} g_n$ (Gauss–Newton step).

(4) Update Parameters:

$$W_{n+1} = W_n + h_n.$$

(5) Evaluate Cost Function:

$$F(W_{n+1}) = \frac{1}{2}\|f(W_{n+1})\|^2.$$

(6) Adjust Damping Parameter μ_n:
- If $F(W_{n+1}) < F(W_n)$ (step is successful), decrease μ_n (move toward Gauss–Newton):

$\mu_{n+1} = \mu_n/\nu, \quad \nu > 1.$

- Otherwise (step unsuccessful), increase μ_n (move toward gradient descent):

$\mu_{n+1} = \mu_n \cdot \nu.$

(7) Check Convergence:
- Stop if $\|h_n\| < \epsilon_1$ or $|F(W_{n+1}) - F(W_n)| < \epsilon_2$ or $n > N_{\max}$.
- Otherwise, set $n \leftarrow n + 1$ and go to Step 2.

6. Numerical examples

In this section, we are going to illustrate the accuracy of the suggested neural network by applying it on various types of fuzzy unconstrained and constrained. In this part, four test cases are solved by employing the presented network. Selected test cases have the exact solutions in order to compute the accuracy of the method. By considering the interval $[t_0, t_f]$ as the spatial domain of the problem, the following errors are used to compare the exact strong solution with an approximate solution:

$$\begin{cases} \underline{e}(\alpha) = ||\underline{x}_T(t,\alpha) - \underline{x}^*(t,\alpha)||_2^2 = \int_{t_0}^{t_f} (\underline{x}_T(t,\alpha) - \underline{x}^*(t,\alpha))^2 dt, \\ \overline{e}(\alpha) = ||\overline{x}_T(t,\alpha) - \overline{x}^*(t,\alpha)||_2^2 = \int_{t_0}^{t_f} (\overline{x}_T(t,\alpha) - \overline{x}^*(t,\alpha))^2 dt, \end{cases} \tag{6.1}$$

where $(\underline{x}_T(t,\alpha)^\top, \overline{x}_T(t,\alpha)^\top)^\top$ are the trial solutions (approximate solutions) of the exact strong solutions

$$(\underline{x}^*(t,\alpha)^\top, \overline{x}(t,\alpha)^{*\top})^\top.$$

The defined errors in (6.1) are modified when the problem has the weak solutions. It should be noted that we perform our computations on a 2.7 GHz Intel Core i5 CPU machine with 8 GB of memory.

Example 6.1. [36] Consider the following fuzzy control problem

$$\begin{aligned} &\min\ (FR) \int_0^2 \frac{1}{2}(\tilde{x}^2(t) \oplus_H \tilde{u}^2(t))dt, \\ &\text{s.t.} \quad \dot{\tilde{x}}(t) = (0,1,3)\tilde{u}(t), \\ &\tilde{x}(0) = (1,1,1), \qquad \tilde{x}(2) \text{ is free.} \end{aligned}$$

The fuzzy Hamiltonian function is

$$\tilde{H}(\tilde{x}(t), \tilde{u}(t), \tilde{p}(t), t)[\alpha] = [\underline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha), \overline{H}(\overline{x}, \overline{u}, \overline{p}, t, \alpha)], \tag{6.2}$$

where

$$\begin{aligned} \underline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha) &= \frac{1}{2}(\underline{x}^2 + \underline{u}^2) + \alpha \underline{p}\underline{u}, \\ \overline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha) &= \frac{1}{2}(\overline{x}^2 + \overline{u}^2) + (3 - 2\alpha)\overline{p}\overline{u}. \end{aligned}$$

The fuzzy Hamiltonian must be satisfied in the following conditions:

$$\begin{cases} \dfrac{\partial \overline{H}}{\partial \overline{x}} = -\dot{\overline{p}} \longrightarrow \overline{x} = -\dot{\overline{p}}, \\ \dfrac{\partial \underline{H}}{\partial \underline{x}} = -\dot{\underline{p}} \longrightarrow \underline{x} = -\dot{\underline{p}}, \\ \dfrac{\partial \overline{H}}{\partial \overline{u}} = 0 \longrightarrow \overline{u} = -(3-2\alpha)\overline{p}, \\ \dfrac{\partial \underline{H}}{\partial \underline{u}} = 0 \longrightarrow \underline{u} = -\alpha \underline{p}, \\ \dfrac{\partial \overline{H}}{\partial \overline{p}} = \dot{\overline{x}} \longrightarrow \dot{\overline{x}} = (3-2\alpha)\overline{u}, \\ \dfrac{\partial \underline{H}}{\partial \underline{p}} = \dot{\underline{x}} \longrightarrow \dot{\underline{x}} = \alpha \underline{u}, \\ \overline{x}(0) = 1, \\ \underline{x}(0) = 1, \\ \overline{p}(2) = 0, \\ \underline{p}(2) = 0. \end{cases} \tag{6.3}$$

The exact solutions of (6.3) are

$$\begin{cases} \underline{x}^*(t,\alpha) = \dfrac{1}{e^{4\alpha}+1} e^{\alpha t} + \dfrac{e^{4\alpha}}{e^{4\alpha}+1} e^{-\alpha t}, \\ \underline{u}^*(t,\alpha) = \dfrac{1}{e^{4\alpha}+1} e^{\alpha t} - \dfrac{e^{4\alpha}}{e^{4\alpha}+1} e^{-\alpha t}, \\ \overline{x}^*(t,\alpha) = \dfrac{1}{e^{4(3-2\alpha)}+1} e^{(3-2\alpha)t} + \dfrac{e^{4(3-2\alpha)}}{e^{4(3-2\alpha)}+1} e^{-(3-2\alpha)t}, \\ \overline{u}^*(t,\alpha) = \dfrac{1}{e^{4(3-2\alpha)}+1} e^{(3-2\alpha)t} - \dfrac{e^{4(3-2\alpha)}}{e^{4(3-2\alpha)}+1} e^{-(3-2\alpha)t}. \end{cases}$$

For $t \in [0,2]$ and $\alpha \in [0,1]$, the monotonicity with respect to α are as: $\underline{x}^*(t,\alpha)$ is decreasing in α, $\overline{x}^*(t,\alpha)$ is increasing in α, $\underline{u}^*(t,\alpha)$ is decreasing in α, $\overline{u}^*(t,\alpha)$ is increasing in α. Using Definition 3.4, the weak solution of the problem is given by

$$\begin{cases} \tilde{u}^*(t)[\alpha] = [\overline{u}^*(t,\alpha), \underline{u}^*(t,\alpha)], \\ \tilde{x}^*(t)[\alpha] = [\overline{x}^*(t,\alpha), \underline{x}^*(t,\alpha)], \end{cases}$$

for all $t \in [0,2]$ and $\alpha \in [0,1]$. The trial solutions can be considered as follows

$$\begin{cases} \underline{x}_T(t,\alpha) = 1 + tN_{\underline{x}}(t,\alpha,w), \\ \overline{x}_T(t,\alpha) = 1 + tN_{\overline{x}}(t,\alpha,w), \\ \underline{p}_T(t,\alpha) = (t-2)N_{\underline{p}}(t,\alpha,w), \\ \overline{p}_T(t,\alpha) = (t-2)N_{\overline{p}}(t,\alpha,w), \\ \underline{u}_T(t,\alpha) = N_{\underline{u}}(t,\alpha,w), \\ \overline{u}_T(t,\alpha) = N_{\overline{u}}(t,\alpha,w). \end{cases}$$

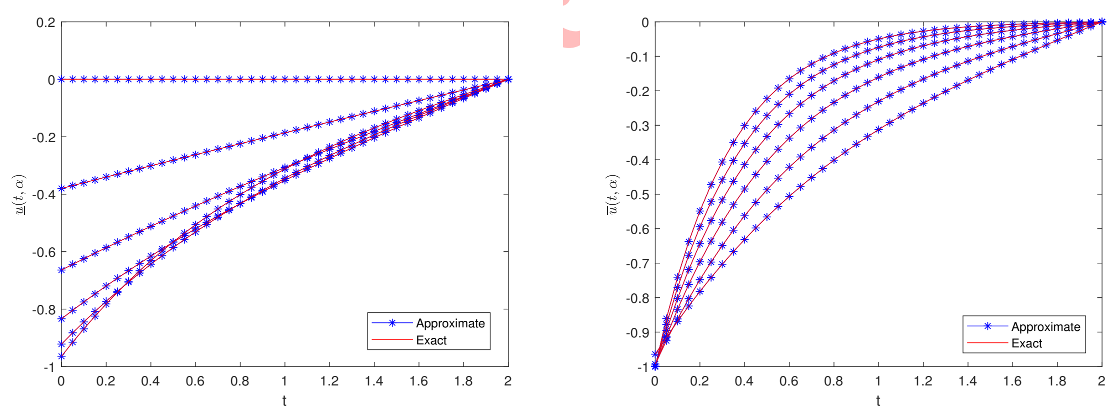
TABLE 1. The absolute errors for different values of $\alpha-$ cuts in Example 6.1.

α	$\underline{e}(\alpha)$	$\overline{e}(\alpha)$
0	3.0343×10^{-41}	5.5075×10^{-11}
0.3	1.23×10^{-20}	3.1081×10^{-11}
0.6	7.0267×10^{-17}	1.5955×10^{-11}
0.9	1.0379×10^{-12}	7.3502×10^{-12}

The computations for this example are done by the presented neural network using 10 neurons in the hidden layer and Jacobi activation functions, so the output of the network is

$$\begin{cases} N_{\underline{x}}(t,\alpha,w^{\underline{x}}) = \sum_{i=0}^{9} w_i^{\underline{x}} J_i^{\beta,\gamma}(t), \\ N_{\overline{x}}(t,\alpha,w^{\overline{x}}) = \sum_{i=0}^{9} w_i^{\overline{x}} J_i^{\beta,\gamma}(t), \\ N_{\underline{u}}(t,\alpha,w^{\underline{u}}) = \sum_{i=0}^{9} w_i^{\underline{u}} J_i^{\beta,\gamma}(t), \\ N_{\overline{u}}(t,\alpha,w^{\overline{u}}) = \sum_{i=0}^{9} w_i^{\overline{u}} J_i^{\beta,\gamma}(t), \\ N_{\underline{p}}(t,\alpha,w^{\underline{p}}) = \sum_{i=0}^{9} w_i^{\underline{p}} J_i^{\beta,\gamma}(t), \\ N_{\overline{p}}(t,\alpha,w^{\overline{p}}) = \sum_{i=0}^{9} w_i^{\overline{p}} J_i^{\beta,\gamma}(t). \end{cases}$$

Also, a training dataset containing 10 roots of the 10th Jacobi polynomial is used to train the network. The Levenberg-Marquardt algorithm does this optimization in an iterative way. The obtained results of applying the Levenberg-Marquardt algorithm to the related optimization problem is reported in Table 1. In this table, the absolute errors of the approximated solution per different values of $\alpha \in [0,1]$ are reported. In Figures 2 and 3, the exact and approximate solutions of the control-state functions for different values of α are plotted. The numerical simulations guarantee that the state and control functions for different values of $\alpha \in [0,1]$ are also in good agreement.

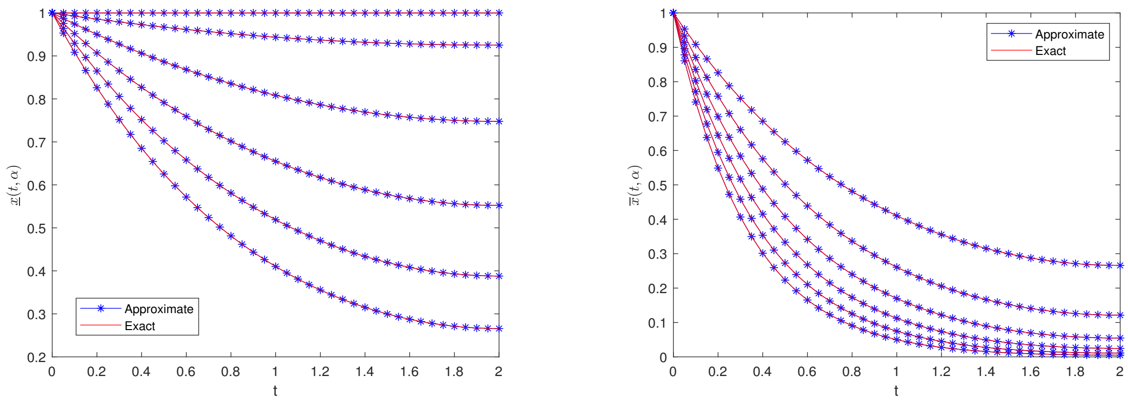


(a) (b)

FIGURE 2. The exact and approximate solutions of $(\underline{u}(t,\alpha), \overline{u}(t,\alpha))$ for different values of α in Example 6.1.

Example 6.2. [36] Consider the following fuzzy optimal control problem

$$\begin{aligned} &\min \ \tilde{J}(\tilde{u}(t)) = (FR)\int_0^2 \ominus_H \frac{\tilde{x}(2)}{2} dt, \\ &\text{s.t.} \qquad \dot{\tilde{x}}(t) = (0, 2.5, 3)(\ominus\tilde{x}(t) \oplus \tilde{x}(t) \odot \tilde{u}(t) \ominus \tilde{u}(t) \odot \tilde{u}(t)), \end{aligned}$$



(a) (b)

FIGURE 3. The exact and approximate solutions of $(\underline{x}(t,\alpha), \overline{x}(t,\alpha))$ for different values of α in Example 6.1.

$$\tilde{x}(0) = (1,1,1), \qquad \tilde{x}(2) \text{ is free.}$$

The fuzzy Hamiltonian function is

$$\tilde{H}(\tilde{x}(t), \tilde{u}(t), \tilde{p}(t), t)[\alpha] = [\underline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha), \overline{H}(\overline{x}, \overline{u}, \overline{p}, t, \alpha)], \tag{6.4}$$

where

$$\underline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha) = -\frac{\underline{x}(2)}{2} + \frac{5}{2}\alpha(-\underline{x} + \underline{x}\underline{u} - \underline{u}^2)\underline{p},$$
$$\overline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha) = -\frac{\overline{x}(2)}{2} + (3 - \frac{\alpha}{2})(-\overline{x} + \overline{x}\overline{u} - \overline{u}^2)\overline{p}.$$

The necessary conditions is

$$\begin{cases} \dfrac{\partial \overline{H}}{\partial \overline{x}} = -\dot{\overline{p}} \longrightarrow \overline{p}(3 - \dfrac{\alpha}{2})(-1 + \overline{u}) = -\dot{\overline{p}}, \\ \dfrac{\partial \underline{H}}{\partial \underline{x}} = -\dot{\underline{p}} \longrightarrow \underline{p}(\dfrac{5}{2}\alpha)(-1 + \underline{u}) = -\dot{\underline{p}}, \\ \dfrac{\partial \overline{H}}{\partial \overline{u}} = 0 \longrightarrow \overline{p}(3 - \dfrac{\alpha}{2})(\overline{x} - 2\overline{u}) = 0 \longrightarrow \overline{u} = \dfrac{1}{2}\overline{x}, \\ \dfrac{\partial \underline{H}}{\partial \underline{u}} = 0 \longrightarrow \underline{p}(\dfrac{5}{2}\alpha)(\underline{x} - 2\underline{u}) = 0 \longrightarrow \underline{u} = \dfrac{1}{2}\underline{x}, \\ \dfrac{\partial \overline{H}}{\partial \overline{p}} = \dot{\overline{x}} \longrightarrow \dot{\overline{x}} = (3 - \dfrac{\alpha}{2})(-\overline{x} + \overline{x}\overline{u} - \overline{u}^2), \\ \dfrac{\partial \underline{H}}{\partial \underline{p}} = \dot{\underline{x}} \longrightarrow \dot{\underline{x}} = \dfrac{5}{2}\alpha(-\underline{x} + \underline{x}\underline{u} - \underline{u}^2), \\ \underline{x}(0) = 1, \\ \overline{x}(0) = 1, \\ \underline{p}(2) = 0, \\ \overline{p}(2) = 0. \end{cases} \tag{6.5}$$

The exact solutions of (6.5) are

$$\begin{cases} \underline{x}^*(t,\alpha) = \dfrac{4}{1+3e^{\frac{5}{2}\alpha t}}, \\ \underline{u}^*(t,\alpha) = \dfrac{2}{1+3e^{\frac{5}{2}\alpha t}}, \\ \overline{x}^*(t,\alpha) = \dfrac{4}{1+3e^{(3-\frac{\alpha}{2})t}}, \\ \overline{u}^*(t,\alpha) = \dfrac{2}{1+3e^{(3-\frac{\alpha}{2})t}}. \end{cases}$$

For $t \in [0,2]$ and $\alpha \in [0,1]$, the monotonicity with respect to α are as: $\underline{x}^*(t,\alpha)$ is decreasing in α, $\overline{x}^*(t,\alpha)$ is increasing in α, $\underline{u}^*(t,\alpha)$ is decreasing in α, $\overline{u}^*(t,\alpha)$ is increasing in α. Using Definition 3.4, the weak solution of the problem is given by

$$\begin{cases} \tilde{u}^*(t)[\alpha] = [\overline{u}^*(t,\alpha), \underline{u}^*(t,\alpha)], \\ \tilde{x}^*(t)[\alpha] = [\overline{x}^*(t,\alpha), \underline{x}^*(t,\alpha)], \end{cases}$$

for all $t \in [0,2]$ and $\alpha \in [0,1]$. We can choose the trial solutions as

$$\begin{cases} \underline{x}_T(t,\alpha) = 1 + tN_{\underline{x}}(t,\alpha,w), \\ \overline{x}_T(t,\alpha) = 1 + tN_{\overline{x}}(t,\alpha,w), \\ \underline{p}_T(t,\alpha) = (t-2)N_{\underline{p}}(t,\alpha,w), \\ \overline{p}_T(t,\alpha) = (t-2)N_{\overline{p}}(t,\alpha,w), \\ \underline{u}_T(t,\alpha) = N_{\underline{u}}(t,\alpha,w), \\ \overline{u}_T(t,\alpha) = N_{\overline{u}}(t,\alpha,w), \end{cases}$$

The simulations are performed using the method proposed in this manuscript with 10 neurons in the first hidden layer and five neurons for the second hidden layer. This network is trained by 20 training points that are the roots of 20th-shifted Jacobi polynomial. The obtained results is reported in Table 2 which shows the accuracy of the proposed method in this paper is acceptable. In Figure 4, the exact and approximate solutions of the state variables for different values of α are plotted.

TABLE 2. The absolute errors for different values of $\alpha-$ cuts in Example 6.2.

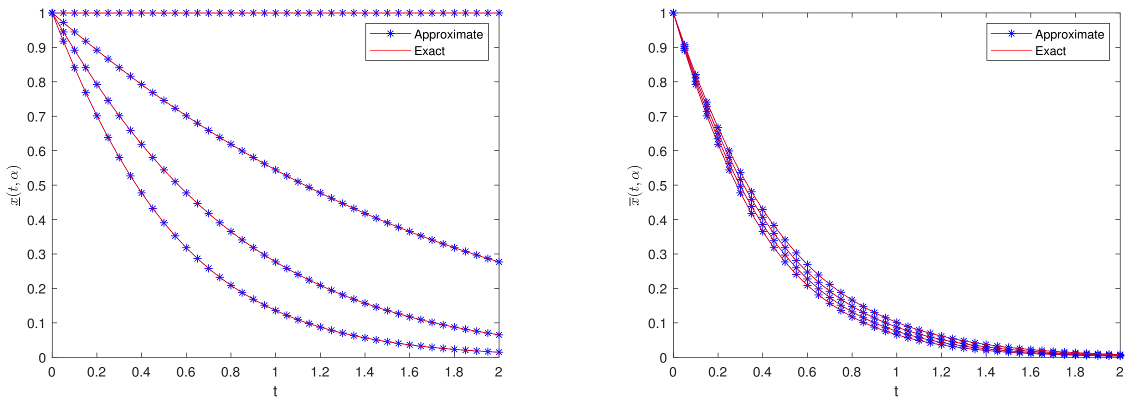
α	$\underline{e}(\alpha)$	$\overline{e}(\alpha)$
0	1.4194×10^{-17}	4.6032×10^{-10}
0.4	3.3666×10^{-17}	3.4781×10^{-12}
0.6	7.8471×10^{-17}	7.519×10^{-14}
0.8	1.8725×10^{-16}	4.2291×10^{-15}
1	4.7449×10^{-16}	4.7445×10^{-16}

Example 6.3. [32] Find the fuzzy control such that

$$\begin{aligned} &\min \quad \tilde{J}(\tilde{u}(t)) = \tilde{x}(1) \oplus (FR)\int_0^1 \tilde{u}^2(t)dt, \\ &\text{s.t. } \dot{\tilde{x}}(t) = \tilde{u}(t) \ominus_H (0,1,3)\tilde{x}(t), \qquad t \in [0,1], \\ &\tilde{x}(0) = \tilde{1} = (1,1,1), \qquad \tilde{x}(1) \text{ is free.} \end{aligned}$$

The fuzzy Hamiltonian function is

$$\tilde{H}(\tilde{x}(t), \tilde{u}(t), \tilde{p}(t), t)[\alpha] = [\underline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha), \overline{H}(\overline{x}, \overline{u}, \overline{p}, t, \alpha)], \tag{6.6}$$



(a) (b)

FIGURE 4. The exact and approximate solutions of $(\underline{x}(t,\alpha),\overline{x}(t,\alpha))$ for different values of α in Example 6.2.

where

$$\begin{cases}\underline{H}(\underline{x},\underline{u},\underline{p},t,\alpha)=\underline{u}^2(t,\alpha)+\underline{p}(t,\alpha)[\underline{u}(t,\alpha)-(3-2\alpha)\underline{x}(t,\alpha)],\\ \overline{H}(\overline{x},\overline{u},\overline{p},t,\alpha)=\overline{u}^2(t,\alpha)+\overline{p}(t,\alpha)[\overline{u}(t,\alpha)-\alpha\overline{x}(t,\alpha)],\\ [\underline{\psi}(\underline{x}(1),1),\overline{\psi}(\overline{x}(1),1)]=[\underline{x}(1),\overline{x}(1)],\\ \left[\dfrac{\partial\underline{\psi}}{\partial\underline{x}}\,|_{t=1},\dfrac{\partial\overline{\psi}}{\partial\overline{x}}\,|_{t=1}\right]=[1,1].\end{cases}$$

The necessary conditions for optimality is

$$\begin{cases}\dot{\underline{x}}(t,\alpha)=\dfrac{\partial\underline{H}}{\partial\underline{p}}=\underline{u}(t,\alpha)-(3-2\alpha)\underline{x}(t,\alpha),\\ \dot{\underline{p}}(t,\alpha)=-\dfrac{\partial\underline{H}}{\partial\underline{x}}=(3-2\alpha)\underline{p}(t,\alpha),\\ 0=\dfrac{\partial\underline{H}}{\partial\underline{u}}=2\underline{u}(t,\alpha)+\underline{p}(t,\alpha),\\ \underline{p}(1,\alpha)=\dfrac{\partial\underline{\psi}}{\partial\underline{x}}|_{t=1}=1,\\ \dot{\overline{x}}(t,\alpha)=\dfrac{\partial\overline{H}}{\partial\overline{p}}=\overline{u}(t,\alpha)-\alpha\overline{x}(t,\alpha),\\ \dot{\overline{p}}(t,\alpha)=-\dfrac{\partial\overline{H}}{\partial\overline{x}}=\alpha\overline{p}(t,\alpha),\\ 0=\dfrac{\partial\overline{H}}{\partial\overline{u}}=2\overline{u}(t,\alpha)+\overline{p}(t,\alpha),\\ \overline{p}(1,\alpha)=\dfrac{\partial\overline{\psi}}{\partial\overline{x}}|_{t=1}=1.\end{cases}\tag{6.7}$$

The exact solutions of (6.7) are

$$\begin{cases} \underline{u}^*(t,\alpha) = -\dfrac{1}{2}e^{(3-2\alpha)(t-1)}, \\ \overline{u}^*(t,\alpha) = -\dfrac{1}{2}e^{\alpha(t-1)}, \\ \underline{x}^*(t,\alpha) = \dfrac{-e^{(3-2\alpha)(t-1)} + 4(3-2\alpha)e^{-(3-2\alpha)t} + e^{-(3-2\alpha)(t+1)}}{4(3-2\alpha)}, \\ \overline{x}^*(t,\alpha) = \dfrac{-e^{\alpha(t-1)} + 4\alpha e^{-\alpha t} + e^{-\alpha(t+1)}}{4\alpha}. \end{cases}$$

It is easily seen that $\underline{u}^*(t,\alpha)$ is a decreasing function of α, and $\overline{u}^*(t,\alpha)$ is an increasing function of α. Therefore, $\underline{u}^*(t,\alpha)$ and $\overline{u}^*(t,\alpha)$ don't satisfy the conditions C_1 and C_2 of Lemma 2.6. Also we observe that $\underline{x}^*(t,\alpha)$ is an increasing function of α and $\overline{x}^*(t,\alpha)$ is a decreasing function of α. Therefore, $\underline{x}^*(t,\alpha)$ and $\overline{x}^*(t,\alpha)$ satisfy the conditions C_1 and C_2 of Lemma 2.6. Moreover, the condition C_3 of Lemma 2.6 for all $t \in [0,1]$ is satisfied. Using (3.4) and (3.5) of Definition 3.4, the weak solution $\tilde{u}^*(t)[\alpha]$ and the strong solution $\tilde{x}^*(t)[\alpha]$ of the problem are respectively given by

$$\begin{cases} \tilde{u}^*(t)[\alpha] = \big[\overline{u}^*(t,\alpha), \underline{u}^*(t,\alpha)\big] = \big[-\dfrac{1}{2}e^{\alpha(t-1)}, -\dfrac{1}{2}e^{(3-2\alpha)(t-1)}\big]. \\ \tilde{x}^*(t)[\alpha] = \big[\underline{x}^*(t,\alpha), \overline{x}^*(t,\alpha)\big] \\ = \big[\dfrac{-e^{(3-2\alpha)(t-1)} + 4(3-2\alpha)e^{-(3-2\alpha)t} + e^{-(3-2\alpha)(t+1)}}{4(3-2\alpha)} \dfrac{-e^{\alpha(t-1)} + 4\alpha e^{-\alpha t} + e^{-\alpha(t+1)}}{4\alpha}\big]. \end{cases} \tag{6.8}$$

We can choose the trial solutions as

$$\begin{cases} \underline{x}_T(t,\alpha) = 1 + tN_{\underline{x}}(t,\alpha,w), \\ \overline{x}_T(t,\alpha) = 1 + tN_{\overline{x}}(t,\alpha,w), \\ \underline{p}_T(t,\alpha) = 1 + (1-t)N_{\underline{p}}(t,\alpha,w), \\ \overline{p}_T(t,\alpha) = 1 + (1-t)N_{\overline{p}}(t,\alpha,w), \\ \underline{u}_T(t,\alpha) = N_{\underline{u}}(t,\alpha,w), \\ \overline{u}_T(t,\alpha) = N_{\overline{u}}(t,\alpha,w), \end{cases}$$

The simulations are performed using the method proposed in this manuscript with 10 neurons in the first hidden layer and five neurons for the second hidden layer. This network is trained by 20 training points that are the roots of 20th-shifted Jacobi polynomial. The obtained results is reported in Table 3. Figures 5 and 6 show the exact and approximate solutions of $(\underline{u}(t,\alpha), \overline{u}(t,\alpha))$ and $(\underline{x}(t,\alpha), \overline{x}(t,\alpha))$ for different values of $\alpha-$ cuts. Thus, our results are in good agreement. Figure 7 also shows the triangular optimal fuzzy solutions $\tilde{u}(0.5)$ and $\tilde{x}(0.5)$ for different values of $\alpha-$cuts.

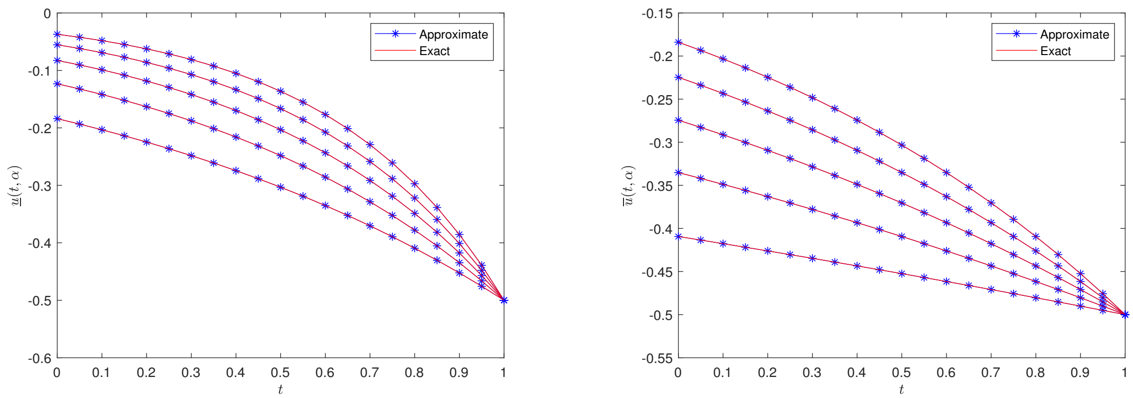
TABLE 3. The absolute errors for different values of $\alpha-$ cuts in Example 6.3.

α	$\underline{e}(\alpha)$	$\overline{e}(\alpha)$
0.2	1.4296×10^{-18}	2.9777×10^{-18}
0.4	2.3801×10^{-18}	2.4512×10^{-18}
0.6	2.5731×10^{-18}	2.25×10^{-18}
0.8	2.4402×10^{-18}	2.2118×10^{-18}
1	2.258×10^{-18}	2.2584×10^{-18}

Example 6.4. [32] Find the fuzzy control such that

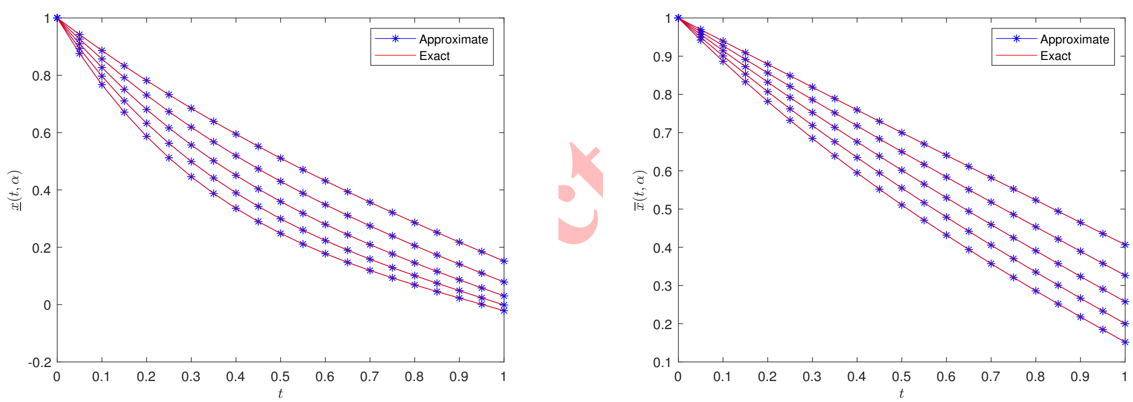
$$\min\ \tilde{J}(\tilde{u}(t)) = \tilde{x}(1) \oplus (FR)\int_0^1 \tilde{u}^2(t)dt,$$

$$\text{s.t. } \dot{\tilde{x}}(t) = \tilde{u}(t) \oplus (0,1,3)\tilde{x}(t), \qquad t \in [0,1],$$



(a) (b)

FIGURE 5. The exact and approximate solutions of $(\underline{u}(t,\alpha), \overline{u}(t,\alpha))$ for different values of α in Example 6.3.



(a) (b)

FIGURE 6. The exact and approximate solutions of $(\underline{x}(t,\alpha), \overline{x}(t,\alpha))$ for different values of α in Example 6.3.

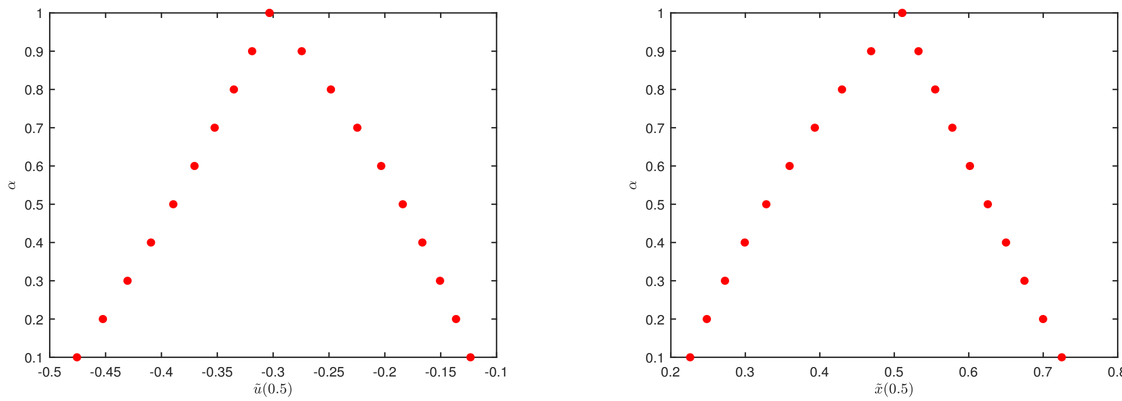
$$\tilde{x}(0) = \tilde{1} = (1,1,1), \qquad \tilde{x}(t_1) \text{ is free.}$$

The fuzzy Hamiltonian function is

$$\tilde{H}(\tilde{x}(t), \tilde{u}(t), \tilde{p}(t), t)[\alpha] = [\underline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha), \overline{H}(\overline{x}, \overline{u}, \overline{p}, t, \alpha)], \tag{6.9}$$

where

$$\begin{cases} \underline{H}(\underline{x}, \underline{u}, \underline{p}, t, \alpha) = \underline{u}^2(t,\alpha) + \underline{p}(t,\alpha)[\underline{u}(t,\alpha) + \alpha \underline{x}(t,\alpha)], \\ \overline{H}(\overline{x}, \overline{u}, \overline{p}, t, \alpha) = \overline{u}^2(t,\alpha) + \overline{p}(t,\alpha)[\overline{u}(t,\alpha) + (3-2\alpha)\overline{x}(t,\alpha)], \\ [\underline{\psi}(\underline{x}(1), 1), \overline{\psi}(\overline{x}(1), 1)] = [\underline{x}(1), \overline{x}(1)], \\ [\dfrac{\partial \underline{\psi}}{\partial \underline{x}} \mid_{t=1}, \dfrac{\partial \overline{\psi}}{\partial \overline{x}} \mid_{t=1}] = [1,1]. \end{cases}$$



(a) (b)

FIGURE 7. The triangular optimal fuzzy solutions for $t = 0.5$ in Example 6.3.

The necessary conditions for optimality is

$$\begin{cases} \dot{\underline{x}}(t,\alpha) = \dfrac{\partial \underline{H}}{\partial \underline{p}} = \underline{u}(t,\alpha) + \alpha \underline{x}(t,\alpha), \\ \dot{\underline{p}}(t,\alpha) = -\dfrac{\partial \underline{H}}{\partial \underline{x}} = -\alpha \underline{p}(t,\alpha), \\ 0 = \dfrac{\partial \underline{H}}{\partial \underline{u}} = 2\underline{u}(t,\alpha) + \underline{p}(t,\alpha), \\ \underline{p}(1,\alpha) = \dfrac{\partial \underline{\psi}}{\partial \underline{x}}|_{t=1} \Longrightarrow \underline{p}(1,\alpha) = 1, \\ \dot{\overline{x}}(t,\alpha) = \dfrac{\partial \overline{H}}{\partial \overline{p}} = \overline{u}(t,\alpha) + (3-2\alpha)\overline{x}(t,\alpha), \\ \dot{\overline{p}}(t,\alpha) = -\dfrac{\partial \overline{H}}{\partial \overline{x}} = -(3-2\alpha)\overline{p}(t,\alpha), \\ 0 = \dfrac{\partial \overline{H}}{\partial \overline{u}} = 2\overline{u}(t,\alpha) + \overline{p}(t,\alpha), \\ \overline{p}(1,\alpha) = \dfrac{\partial \overline{\psi}}{\partial \overline{x}}|_{t=1} = 1. \end{cases} \tag{6.10}$$

The exact solutions of (6.10) are

$$\begin{cases} \underline{u}^*(t,\alpha) = -\dfrac{1}{2}e^{\alpha(1-t)}, \\ \overline{u}^*(t,\alpha) = -\dfrac{1}{2}e^{(3-2\alpha)(1-t)}, \\ \underline{x}^*(t,\alpha) = \dfrac{e^{\alpha(1-t)} + 4\alpha e^{\alpha t} - e^{\alpha(t+1)}}{4\alpha}. \\ \overline{x}^*(t,\alpha) = \dfrac{e^{(3-2\alpha)(1-t)} + 4(3-2\alpha)e^{(3-2\alpha)t} - e^{(3-2\alpha)(t+1)}}{4(3-2\alpha)}. \end{cases}$$

It is easily seen that $\underline{u}^*(t,\alpha)$ is a decreasing function of α, and $\overline{u}^*(t,\alpha)$ is an increasing function of α. Therefore, $\underline{u}^*(t,\alpha)$ and $\overline{u}^*(t,\alpha)$ don't satisfy the conditions C_1 and C_2 of Lemma 2.6. Also we observe that $\underline{x}^*(t,\alpha)$ and $\overline{x}^*(t,\alpha)$ are increasing functions of α. Moreover, the condition C_3 of Lemma 2.6 for all $t \in [0,1]$ is satisfied. Using (3.4) and

(3.5) of Definition 3.4, the weak solution of the problem is given by

$$\begin{aligned}\tilde{u}^*(t)[\alpha] &= [\overline{u}^*(t,\alpha), \underline{u}^*(t,\alpha)] = [-\frac{1}{2}e^{(3-2\alpha)(1-t)}, -\frac{1}{2}e^{\alpha(1-t)}],\\ \tilde{x}^*(t)[\alpha] &= [\underline{x}^*(t,\alpha), -\overline{x}^*(t,\alpha) + 2\underline{x}^*(t,1)]\\ &= [\frac{e^{\alpha(1-t)} + 4\alpha e^{\alpha t} - e^{\alpha(1+t)}}{4\alpha}, \frac{e^{(1-t)} + 4e^t - e^{(1+t)}}{2} - (\frac{e^{(3-2\alpha)(1-t)} + 4(3-2\alpha)e^{(3-2\alpha)t} - e^{(3-2\alpha)(1+t)}}{4(3-2\alpha)})].\end{aligned}$$

We can choose the trial solutions as

$$\begin{cases}\underline{x}_T(t,\alpha) = 1 + tN_{\underline{x}}(t,\alpha,w),\\ \overline{x}_T(t,\alpha) = 1 + tN_{\overline{x}}(t,\alpha,w),\\ \underline{p}_T(t,\alpha) = 1 + (1-t)N_{\underline{p}}(t,\alpha,w),\\ \overline{p}_T(t,\alpha) = 1 + (1-t)N_{\overline{p}}(t,\alpha,w),\\ \underline{u}_T(t,\alpha) = N_{\underline{u}}(t,\alpha,w),\\ \overline{u}_T(t,\alpha) = N_{\overline{u}}(t,\alpha,w),\end{cases}$$

The simulations are performed using the method proposed in this manuscript with 10 neurons in the first hidden layer and five neurons for the second hidden layer. This network is trained by 20 training points that are the roots of 20th-shifted Jacobi polynomial. The obtained results is reported in Table 4. As Table 4 shows, the accuracy of the proposed method in this paper is acceptable. Figures 8 and 9 show the exact and approximate solutions of $(\underline{u}(t,\alpha), \overline{u}(t,\alpha))$ and $(\underline{x}(t,\alpha), \overline{x}(t,\alpha))$ for different values of $\alpha-$ cuts. Thus, our results are in good agreement. Figure 10 also shows the triangular optimal fuzzy solutions $\tilde{u}(0.5)$ and $\tilde{x}(0.5)$ for different values of $\alpha-$cuts.

TABLE 4. The absolute errors for different values of $\alpha-$ cuts in Example 6.4.

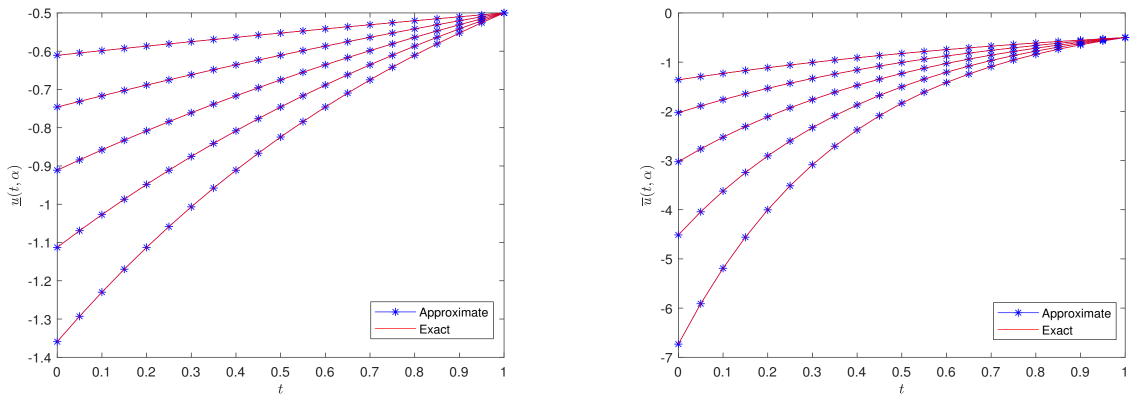
α	$\underline{e}(\alpha)$	$\overline{e}(\alpha)$
0.2	6.9969×10^{-18}	5.9777×10^{-18}
0.4	1.379×10^{-17}	2.4512×10^{-17}
0.6	3.1331×10^{-17}	2.25×10^{-18}
0.8	7.9944×10^{-17}	6.2118×10^{-18}
1	2.254×10^{-16}	2.3284×10^{-16}

To end this section, we summarize some advantages of the proposed method as follows:

- Unlike the multi-layer neural networks, the proposed method has no defects such as slow convergence speed, local minimum, initial values of weights and number of processing elements.
- The solution of neural network scheme is continuous over all the domain of the problem. In contrast, the other numerical methods provide solutions only over discrete points, and the solution between these points must be interpolated.
- It is worth mentioning that, in this algorithm, the linearization method is not needed. So, the efficiency of this approach is more than those of algorithms that have a linearization part.

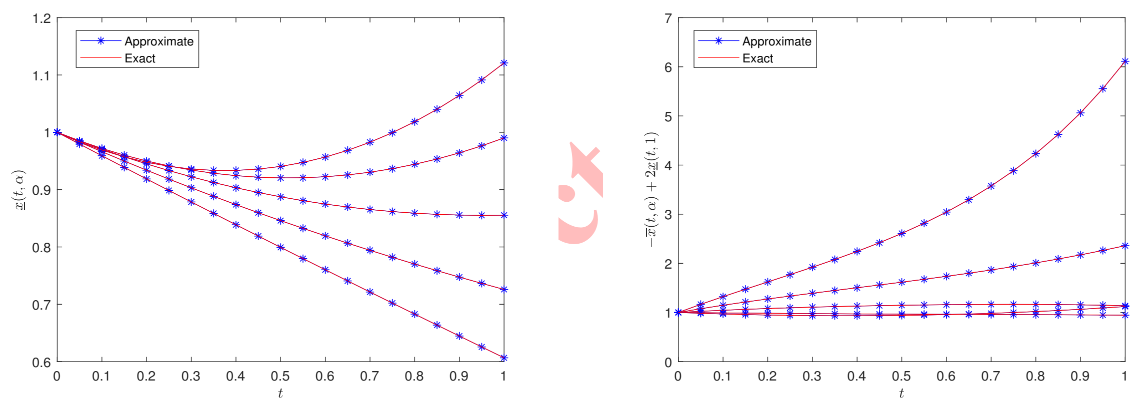
7. Conclusion

In this paper, we addressed the fuzzy optimal control problem with state conditions at the final time. To propose a solution method, we used the necessary conditions, as derived in [32], and introduced a neural network scheme employing Jacobi functions as the activation functions and the Levenberg-Marquardt algorithm as the optimizer to perform numerical simulations of these conditions in fuzzy nonlinear optimal control problems. To ensure that the solutions are fuzzy functions, we applied the concepts of strong and weak solutions. This approach eliminates the need for linearization methods for nonlinear problems. Jacobi functions serve as the activation functions for the hidden layer, while the identity function is used as the activation function for the output layer, enhancing accuracy.



(a) (b)

FIGURE 8. The exact and approximate solutions of $(\underline{u}(t,\alpha), \overline{u}(t,\alpha))$ for different values of α in Example 6.4.



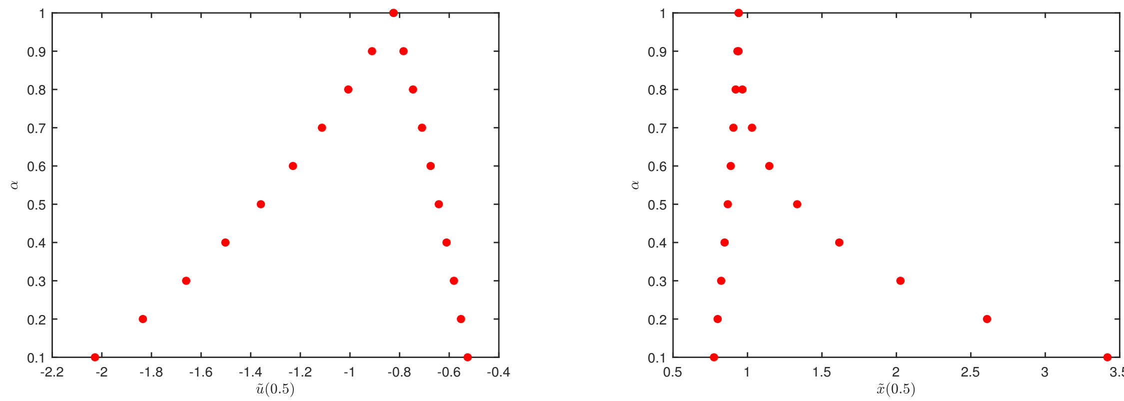
(a) (b)

FIGURE 9. The exact and approximate solutions of $(\underline{x}(t,\alpha), -\overline{x}(t,\alpha)) + 2\underline{x}(t,1))$ for different values of α in Example 6.4.

The reported results show that this method can perform computations with a few neurons and Levenberg-Marquardt iterations with high precision, leading to high efficiency. However, there are challenges not addressed in this article, which could be explored in future research. For instance, parallel implementation has technical aspects that warrant further discussion. Additionally, solving high-dimensional fractional fuzzy problems using neural networks has not been adequately explored, making it a valuable area for future work to introduce an efficient neural network framework for such problems.

8. Declarations

Availability of supporting data: Enquiries about data availability should be directed to the authors.
Competing interests: Not Applicable
Funding: Not Applicable



(a) (b)

FIGURE 10. The triangular optimal fuzzy solutions for $t = 0.5$ in Example 6.4.

Consent to participate: All authors consent to participate
Consent for publication: All authors consent for publication
Human and Animal Ethics: Not Applicable
Authors' contributions: Contributions of all authors are equal.
Acknowledgments: Not Applicable
Conflict of interest Both of the authors declare that they have no conflict of interest.
Ethical approval This article does not contain any studies with animals performed by any of the authors.

REFERENCES

[1] M. A. Abdelkawy, A. Z. M. Amin, A. H. Bhrawy, J. A. T. Machado, and A. M. Lopes, *Jacobi collocation approximation for solving multi-dimensional Volterra integral equations*, International Journal of Nonlinear Sciences and Numerical Simulation, *18*(5) (2017), 411–425.

[2] M. A. Abdelkawy, A. M. Lopes, and M. M. Babatin, *Shifted fractional Jacobi collocation method for solving fractional functional differential equations of variable order*, Applied Mathematics and Computation, *134*(4) (2020), 109721.

[3] S. D. Agashe, B. K. Lande, V. Jain, G. Chaudhary, and F. Al-turjman, *Numerical solution of optimal control problem*, Soft Computing, *27*(3) (2023), 3389–3401.

[4] M. Azizi, M. Amirfakhrian, and M. A. Fariborzi Araghi, *Application of fuzzy systems on the numerical solution of the elliptic PDE-constrained optimal control problems*, Computational Methods for Differential Equations, *10*(2) (2022), 351–371.

[5] A. H. Bhrawy, *A Jacobi spectral collocation method for solving multi-dimensional nonlinear fractional sub-diffusion equations*, Numerical Algorithms, *73*(1) (2016), 91–113.

[6] A. H. Bhrawy and M. A. Zaky, *Shifted fractional-order Jacobi orthogonal functions: application to a system of fractional differential equations*, Applied Mathematical Modelling, *40*(2) (2016), 832–845.

[7] J. J. Buckley and T. Feuring, *Introduction to fuzzy partial differential equations*, Fuzzy Sets and Systems, *105*(2) (1999), 241–248.

[8] T. Cheng, H. Sun, Z. Qu, and F. L. Lewis, *Neural network solution for suboptimal control of non-holonomic chained form system*, Transactions of the Institute of Measurement and Control, *31*(6) (2009), 475–494.

[9] M. Delkhosh and K. Parand, *Generalized pseudospectral method: Theory and applications*, Journal of Computational Science, *34*(4) (2019), 11–32.

[10] M. Delkhosh, K. Parand, and A. H. Hadian-Rasanan, *A development of Lagrange interpolation, part I: theory*, Journal of Computational and Applied Mathematics, *356*(1) (2019), 44–63.

[11] E. H. Doha, A. H. Bhrawy, and S. S. Ezz-Eldien, *Efficient Chebyshev spectral methods for solving multi-term fractional orders differential equations*, Applied Mathematical Modelling, *35*(12) (2011), 5662–5672.

[12] A. Ekrami Kivaj, A. Basohbat Novinzadeh, and F. Pazooki, *Spacecraft reentry trajectory optimization by heuristic optimization methods and optimal control theory*, International Journal of Dynamics and Control, *11*(3) (2023), 1132–1141.

[13] S. S. Ezz-Eldien, *On solving systems of multi-pantograph equations via spectral tau method*, Applied Mathematics and Computation, *321*(1) (2018), 63–73.

[14] S. S. Ezz-Eldien and E. H. Doha, *Fast and precise spectral method for solving pantograph type Volterra integro-differential equations*, Numerical Algorithms, *81*(1) (2019), 57–77.

[15] S. S. Ezz-Eldien, R. M. Hafez, A. H. Bhrawy, D. Baleanu, and A. A. El-Kalaawy, *New numerical approach for fractional variational problems using shifted Legendre orthonormal polynomials*, Journal of Optimization Theory and Applications, *174*(1) (2017), 295–320.

[16] B. Farhadinia, *Necessary optimality conditions for fuzzy variational problems*, Information Sciences, *181*(7) (2011), 1348–1357.

[17] B. Farhadinia, *Pontryagin's minimum principle for fuzzy optimal control problems*, Iranian Journal of Fuzzy Systems, *11*(4) (2014), 27–43.

[18] D. Filev and P. Angelov, *Fuzzy optimal control*, Fuzzy Sets and Systems, *47*(2) (1992), 151–156.

[19] M. Friedman, M. Ma, and A. Kandel, *Numerical solution of fuzzy differential and integral equations*, Fuzzy Sets and Systems, *106*(1) (1999), 35–48.

[20] H. P. Huang and J. J. Cheng, *Unstable backpropagation method in neural networks: A remedy*, in 2nd International Conference on Automation Technology, *1* (1992), 249–255.

[21] M. Hukuhara, *Integration des applications mesurables dont la valeur est un compact convexe*, Funkcialaj Ekvacioj, *10*(3) (1967), 205–223.

[22] L. Jianyu, L. Siwei, Q. Yingjian, and H. Yaping, *Numerical solution of elliptic partial differential equation using radial basis function neural networks*, Neural Networks, *16*(5-6) (2003), 729–734.

[23] S. Kazem, S. Abbasbandy, and S. Kumar, *Fractional-order Legendre functions for solving fractional-order differential equations*, Applied Mathematical Modelling, *37*(7) (2013), 5498–5510.

[24] K. Levenberg, *A method for the solution of certain non-linear problems in least squares*, Quarterly of Applied Mathematics, *2*(2) (1944), 164–168.

[25] A. Lotfi, S. A. Yousefi, and M. Dehghan, *Numerical solution of a class of fractional optimal control problems via the Legendre orthonormal basis combined with the operational matrix and the Gauss quadrature rule*, Journal of Computational and Applied Mathematics, *250*(1) (2013), 143–160.

[26] N. Mai-Duy and T. Tran-Cong, *Numerical solution of differential equations using multiquadric radial basis function networks*, Neural Networks, *14*(2) (2001), 185–199.

[27] S. Mall and S. Chakraverty, *Chebyshev neural network based model for solving Lane-Emden type equations*, Applied Mathematics and Computation, *247*(1) (2014), 100–114.

[28] S. Mall and S. Chakraverty, *Single layer Chebyshev neural network model for solving elliptic partial differential equations*, Neural Processing Letters, *45*(3) (2017), 825–840.

[29] M. Mazandarani, N. Pariz, and A. V. Kamyad, *Granular differentiability of fuzzy number-valued functions*, IEEE Transactions on Fuzzy Systems, *26*(1) (2017), 310–323.

[30] A. K. Mittal, *Error analysis and approximation of Jacobi pseudospectral method for the integer and fractional order integro-differential equation*, Applied Mathematics and Computation, *432*(1) (2022), 128322.

[31] A. M. Mustafa, Z. Gong, and M. Osman, *Fuzzy optimal control problem of several variables*, Advances in Mathematical Physics, *2019*(1) (2019), 9872836.

[32] A. M. Mustafa, Z. Gong, and M. Osman, *The strong and weak solutions of fuzzy optimal control problem with state conditions at the final time*, Canadian Journal of Applied Mathematics, *2*(1) (2020), 45–67.

[33] M. Najariyan and M. H. Farahi, *Optimal control of fuzzy linear controlled system with fuzzy initial conditions*, Iranian Journal of Fuzzy Systems, *10*(3) (2013), 21–35.

[34] M. Najariyan and M. H. Farahi, *A new approach for the optimal fuzzy linear time invariant controlled system with fuzzy coefficients*, Journal of Computational and Applied Mathematics, *259*(1) (2014), 682–694.

[35] M. Najariyan and M. H. Farahi, *A new approach for solving a class of fuzzy optimal control systems under generalized Hukuhara differentiability*, Journal of the Franklin Institute, *352*(5) (2015), 1836–1848.

[36] S. Panahian Fard, R. Pourabbas, and J. Pouramini, *Approximate solutions of fuzzy optimal control problems using sigmoid-weighted neural networks*, Soft Computing, *25*(7) (2021), 5355–5364.

[37] J. A. Rad, S. Kazem, M. Shaban, K. Parand, and A. Yildirim, *Numerical solution of fractional differential equations with a Tau method based on Legendre and Bernstein polynomials*, Mathematical Methods in the Applied Sciences, *37*(3) (2014), 329–342.

[38] M. Raissi, P. Perdikaris, and G. E. Karniadakis, *Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, Journal of Computational Physics, *378*(1) (2019), 686–707.

[39] A. H. Hadian Rasanan, N. Bajalan, K. Parand, and J. Amani Rad, *Simulation of nonlinear fractional dynamics arising in the modeling of cognitive decision making using a new fractional neural network*, Mathematical Methods in the Applied Sciences, *43*(4) (2020), 1997–2014.

[40] T. Rezazadeh and E. Najafi, *Jacobi collocation method and smoothing transformation for numerical solution of neutral nonlinear weakly singular Fredholm integro-differential equations*, Applied Mathematics and Computation, *432*(1) (2022), 127388.

[41] A. Saadatmandi and M. Dehghan, *A Legendre collocation method for fractional integro-differential equations*, Journal of Vibration and Control, *17*(13) (2011), 2050–2058.

[42] M. A. Sadygov and H. S. Akhundov, *Optimal control problem described by the Goursat-Darboux equations*, Applied and Computational Mathematics, An International Journal, *24*(1) (2025), 146–161.

[43] N. Saldi, *Common information approach for static team problems with polish spaces and existence of optimal policies*, Applied and Computational Mathematics, An International Journal, *23*(3) (2024), 307–324.

[44] M. A. Snyder, *Chebyshev Methods in Numerical Approximation*, Prentice-Hall, (1966).

[45] J. Soolaki, O. S. Fard, and A. H. Borzabadi, *Generalized Euler-Lagrange equations for fuzzy variational problems*, SeMA Journal, *73*(2) (2016), 131–148.

[46] J. Soolaki, O. S. Fard, and A. H. Borzabadi, *A necessary condition of Pontryagin type for fuzzy control systems*, Journal of Computational and Applied Mathematics, *37*(1) (2018), 1263–1278.

[47] S. Soradi-Zeid and M. Alipour, *A collocation method using generalized Laguerre polynomials for solving nonlinear optimal control problems governed by integro-differential equations*, Journal of Computational and Applied Mathematics, *436*(1) (2024), 115410.

[48] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer Science, (2013).

[49] T. Takagi and M. Sugeno, *Fuzzy identification of systems and its applications to modeling and control*, IEEE Transactions on Systems, Man, and Cybernetics, *15*(1) (1985), 116–132.

[50] X. Wang, J. Liu, H. Peng, and X. Zhao, *An iterative framework to solve nonlinear optimal control with proportional delay using successive convexification and symplectic multi-interval pseudospectral scheme*, Applied Mathematics and Computation, *435*(1) (2022), 127448.

[51] H. C. Wu, *Duality theory in fuzzy optimization problems*, Fuzzy Optimization and Decision Making, *3*(4) (2005), 345–365.

[52] Y. Z. Zhao and Y. G. Zhu, *Fuzzy optimal control of linear quadratic models*, Computers and Mathematics with Applications, *60*(1) (2010), 67–73.

[53] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*, Kluwer Academic Publishers, (1991).