Available online at http://ijim.srbiau.ac.ir/
Int. J. Industrial Mathematics (ISSN 2008-5621)
Vol. 12, No. 3, 2020 Article ID IJIM-1396, 11 pages
DOR: http://dorl.net/dor/20.1001.1.20085621.2020.12.3.10.0
Research Article

Improved Neural Network and the Pontryagin's minimum Principle for Solve Fuzzy Optimal Control Problems

S. Askari *, S. Abbasbandy †‡

Received Date: 2019-12-01 Revised Date: 2020-02-06 Accepted Date: 2020-02-22

Abstract

In this paper, a novel and practical approach is proposed to solve the fuzzy optimal control (FOC) using an improved multi-layer perceptron (IMLP) network along with the Pontryagin minimum principle (PMP).

Keywords : FOC problem; Pontryagin minimum principle; IMLP networks; BFGs method.

1 Introduction

Optimal control issues appear widely in various sciences [21, 27]. The numerical solution of these issues is of great importance. As such, several methods have been proposed to solve such problems such as [2, 4, 5, 9, 10, 11, 12, 19, 23, 25] which are often, based on PMP [14], Hamilton-Jacobi-Bellman (HJB), partial differential equation (PDE) [14, 17], and the optimization control problems (OCP). On the other hand, the MLP network is a powerful tool to estimate many functions [3, 6]. Effati and Pakdaman (2013) utilized the neural networks to approximate the solution for a OCP [6]. In many systems, uncertainty is common while the fuzzy sets are useful to handle this uncertainty [15]. In the past few decades, the concept of the fuzzy set [28] has recently expanded in various research fields such as optimization, differential equations, and OCP. Therefore, fuzzy optimal control theory generates a suitable condition to formulate the real world problems under uncertainty [11]. The interested readers can refer to some applications of the fuzzy control systems. In [12] authors described the Nowak-May model using fuzzy variables and then proposed the fuzzy control model to maximize the uninfected cells of HIV disease. Zarei et al. [29] developed a fuzzy mathematical model of HIV infection including the linear fuzzy differential equations. Mazandarani, and Kamyad [23] provided a fuzzy model of diabetes mellitus type 2. Although the notion of fuzzy sets is widely spread for several control optimization problems, establishing necessary optimality conditions for fuzzy optimal control problems is seldom available in literature [15, 16, 22, 24, 26].
In this research, a novel and practical method is developed based on the fuzzy multi-layer percep-

*Department of Applied Mathematics, Hamedan Branch, Islamic Azad University, Hamedan, Iran.
†Corresponding author. abbasbandy@ikiu.ac.ir, Tel:+98(912)1305326.
‡Department of Applied Mathematics, Faculty of Science, Imam Khomeini International University, Qazvin, Iran.

tion (FMLP) network to solve the fuzzy optimal control problems (FOCP). The current research consists of following sections: In Section 2, the basic and required concepts are expressed. In Section 3, the estimate of control optimization problems has been stated using FMLP network. In Section 4, the numerical examples are presented. Finally, in Section 5, the conclusion is presented.

2 Definitions and required theorems

In this section we introduce notations, definitions, some preliminary notions, and required theorems, which is used to propose the model.

Definition 2.1 *The fuzzy number* $\tilde{a}:\mathbb{R}\to[0,1]$ *is a mapping with the properties:*

(i) $\tilde{a}$ *is normal, i.e. there exists* $x\in\mathbb{R}$ *such that* $\tilde{a}(x)=1$.

(ii) $\tilde{a}$ *is a convex, i.e.*
$\forall x,y\in\mathbb{R}$ *and* $\lambda\in[0,1]$,
$\tilde{a}(\lambda x+(1-\lambda)y)\geq\min\{\tilde{a}(x),\tilde{a}(y)\}$.

(iii) $\tilde{a}$ *is upper semi continuous.*

(iv) $cl\{s\in\mathbb{R}\mid u(s)>0\}$. *is compact in* $x\in\mathbb{R}$.

The α-level set of a fuzzy number $u\in E^1, 0\leq\alpha\leq 1$, denoted by u_α, is defined as:

$$|\tilde{a}|_\alpha=\begin{cases}\{s\in\mathbb{R}\mid\tilde{a}(s)\geq\alpha\} & 0<\alpha\leq 1\\ cl\{s\in\mathbb{R}\mid u(s)>0\} & \alpha=0\end{cases}\tag{2.1}$$

If $\tilde{a}:\mathbb{R}\to[0,1]$, then $\tilde{a}$ is fuzzy convex, so $[\tilde{a}]_\alpha$ is closed and bounded in $\mathbb{R}$, i.e. $[\tilde{a}]_\alpha=[\underline{a}(\alpha),\overline{a}(\alpha)]$, where $\underline{a}(\alpha)=\inf\{x\in\mathbb{R}:\tilde{a}(x)\geq\alpha\}>-\infty$ and $\overline{a}(\alpha)=\sup\{x\in\mathbb{R}:\tilde{a}(x)\geq\alpha\}<\infty$. For more detail see [3].

Lemma 2.1 *(see [20]) Denote* $I=[0;1]$. *Assumed that* $\underline{a}:I\to\mathbb{R}$ *and* $\overline{a}:I\to\mathbb{R}$ *satisfy the following conditions:*

(1) $\underline{a}:I\to\mathbb{R}$ is a bounded increasing function;

(2) $\overline{a}:I\to\mathbb{R}$ is a bounded decreasing function;

(3) $\underline{a}(1)\leq\overline{a}(1)$,

(4) for $0<k\leq 1$, $\lim_{\alpha\to k^-}\underline{a}(\alpha)=\underline{a}(k)$ and $\lim_{\alpha\to k^-}\overline{a}(\alpha)=\overline{a}(k)$

(5) $\lim_{\alpha\to 0^+}\underline{a}(\alpha)=\underline{a}(0)$ and $\lim_{\alpha\to 0^+}\overline{a}(\alpha)=\overline{a}(0)$.

Then $\tilde{a}:[0,1]\to\mathbb{R}$ characterized by $\tilde{a}(x)=\sup\{\alpha:\underline{a}(\alpha)\leq x\leq\overline{a}(\alpha)\}$ is a fuzzy number. Also if $\tilde{a}:\mathbb{R}\to[0,1]$ is a fuzzy number with $[\tilde{a}]_\alpha=[\underline{a}(\alpha),\overline{a}(\alpha)]$, then functions $\underline{a}(\alpha)$ and $\overline{a}(\alpha)$ satisfy conditions $(1)-(5)$ in Lemma (2.1).

Definition 2.2 *If* A *and* B *be fuzzy numbers with* $[A]_\alpha=[\underline{a}(\alpha),\overline{a}(\alpha)]$, $[\tilde{B}]_\alpha=[\underline{b}(\alpha),\overline{b}(\alpha)]$ *and* $\alpha\in[0,1]$, *then fuzzy operation between them are defined as follows [30].*

$$\begin{aligned}&[A+B]_\alpha=[\underline{a}(\alpha)+\underline{b}(\alpha),\overline{a}(\alpha)+\overline{b}(\alpha)],\\&[-A]_\alpha=[-\underline{a}(\alpha),-\overline{a}(\alpha)],\\&[A-B]_\alpha=[\underline{a}(\alpha)-\overline{b}(\alpha),\overline{a}(\alpha)-\underline{b}(\alpha)],\\&[\lambda A]_\alpha=[\lambda\underline{a}(\alpha),\lambda\overline{a}(\alpha)],\quad\lambda>0,\\&[\lambda A]_\alpha=[\lambda\overline{a}(\alpha),\lambda\underline{a}(\alpha)],\quad\lambda<0.\end{aligned}$$

So, the fuzzy number $\tilde{a}$ is triangular if $\underline{a}(1)=\overline{a}(1)$, $\underline{a}(\alpha)=\underline{a}(1)-(1-\alpha)(\underline{a}(1)-\overline{a}(0))$ and $\overline{a}(\alpha)=\underline{a}(1)+(1-\alpha)(\overline{a}(0)-\underline{a}(1))$. The triangular fuzzy number $\tilde{a}$ is generally denoted by $\tilde{a}=\langle\underline{a}(0),\underline{a}(1),\overline{a}(0)\rangle$.

Definition 2.3 *Let* $\tilde{a},\tilde{b}\in F$. *We write [8]*

1) $\tilde{a}\leq\tilde{b}$ *if* $\underline{a}(\alpha)\leq\underline{b}(\alpha)$ *and* $\overline{a}(\alpha)\leq\overline{b}(\alpha)$ *for all* $\alpha\in[0,1]$,

2) $\tilde{a}<\tilde{b}$ *if* $\tilde{a}\leq\tilde{b}$ *and there exists an* $\alpha\in[0,1]$ *so that* $\underline{a}(\alpha)\leq\underline{b}(\alpha)$ *and* $\overline{a}(\alpha)\leq\overline{b}(\alpha)$,

3) $\tilde{a}=\tilde{b}$ *if* $\tilde{a}\leq\tilde{b}$ *and* $\tilde{a}\geq\tilde{b}$. *In the other words,* $\tilde{a}=\tilde{b}$, *if* $\tilde{a}(\alpha)=\tilde{b}(\alpha)$ *for all* $\alpha\in[0,1]$,

4) $\tilde{a},\tilde{b}\in F$ *are comparable if either* $\tilde{a}\leq\tilde{b}$ *or* $\tilde{a}\geq\tilde{b}$, *and non-comparable otherwise.*

Definition 2.4 *Let* $\tilde{x}:T\to F(\mathbb{R})$ *is Hukuhara differentiable at* $t_0\in T\subseteq\mathbb{R}$ *if for some* $h_0>0$ *the Hukuhara difference* $\tilde{x}(t_0+\triangle t)\sim_h\tilde{x}(t_0)$, $\tilde{x}(t_0)\sim_h\tilde{x}(t_0+\triangle t)$, *exist in* E *for all* $0<\triangle<h_0$ *and if there exist an element* $\dot{\tilde{x}}(t_0)\in F(\mathbb{R})$ *such that*

$$\lim_{\triangle t\to 0^+}d_\infty\Big(\frac{\tilde{x}(t_0+\triangle t)\sim_h\tilde{x}(t_0)}{\triangle t},\dot{\tilde{x}}(t_0)\Big)=0.\tag{2.2}$$

And

$$\lim_{\triangle t \to 0^+} d_\infty\Big(\frac{\tilde{x}(t_0) \sim_h \tilde{x}(t_0 + \triangle t)}{\triangle t}, \dot{\tilde{x}}(t_0)\Big) = 0. \quad (2.3)$$

The fuzzy set $\dot{\tilde{x}}(t_0)$ is called the Hukuhara derivative of $\tilde{x}$ at t_0 [1]. Recall that $U \sim_h V = W$ are defined on level sets, where $[U]_\alpha \sim_h [V]_\alpha = [W]_\alpha$ for all $0 \le \alpha \le 1$. By consideration of definition of the metric d_∞, all the level set mappings $[\tilde{x}(.)]_\alpha$ are Hukuhara differentiable at t_0 with Hukuhara derivatives $[\dot{\tilde{x}}(t_0)]_\alpha$ for each $0 \le \alpha \le 1$, when $\tilde{x} : T \to F(\mathbb{R})$ is Hukuhara differentiable at with Hukuhara derivative $[\dot{\tilde{x}}(t_0)]_\alpha$.

Definition 2.5 *A mapping $\tilde{x} : T \to F(\mathbb{R})$ is called a fuzzy process. We denote*

$$[\tilde{x}(t)]_\alpha = [\underline{x}_\alpha(t), \overline{x}_\alpha(t)],\ t \in I,\ \alpha \in (0, 1]. \quad (2.4)$$

The Seikkala derivative $x'(t)$ of a fuzzy process y is defined by

$$[\dot{\tilde{x}}(t)]_\alpha = [\dot{\underline{x}}_\alpha(t), \dot{\overline{x}}_\alpha(t)],\ \alpha \in (0, 1]. \quad (2.5)$$

Provided the equation defines a fuzzy number $\dot{x}(t) \in F(\mathbb{R})$.

Definition 2.6 *(see [13]). Let $\tilde{f}$ be a function defined by $\tilde{f} : V \to F(\mathbb{R})$, where V is a real vector space. Then $\tilde{f}$ is called fuzzy-valued function on V. For any $x \in V, \tilde{f}(x)$, is a fuzzy number (since $\tilde{f}(x) \in F(\mathbb{R})$). Defined two real-valued functions $\underline{f}(\alpha)$ and $\overline{f}(\alpha)$ on V for each $\alpha \in [0, 1]$. Therefore, Obtained the real numbers $\underline{f}(x, \alpha)$ and $\overline{f}(x, \alpha)$ for each $\alpha \in [0, 1]$. When writed*

$$\underline{f}(x, \alpha) = \Big(\underline{f}(x)\Big) and \overline{f}(x, \alpha) = \Big(\overline{f}(x)\Big). \quad (2.6)$$

Therefore, the fuzzy valued function $\tilde{f}$ defined on the real vector space V can induce a family of real-valued functions $\underline{f}(x, \alpha)$ and $\overline{f}(x, \alpha)$ for each $\alpha \in [0, 1]$ which are given in equation (2.7). Also it holds $\underline{f}(x, \alpha) \le \overline{f}(x, \alpha)$ for any $\alpha \in [0, 1]$.

Definition 2.7 *(see [8]) We say that $\tilde{f} : V \subseteq \mathbb{R} \to F$ is continuous at $x \in V$, if both $\underline{f}(x, \alpha)$ and $\overline{f}(x, \alpha)$ are continuous functions of $x \in V$, for all $\alpha \in [0, 1]$.*

Definition 2.8 *(see [8]) Suppose that $\tilde{f} : V \subseteq \mathbb{R} \to F$ is fuzzy-valued function with $[\tilde{f}(x)]_\alpha = (\underline{f}(x, \alpha), \overline{f}(x, \alpha))$. If the partial derivatives of $\underline{f}(x, \alpha)$ and $\overline{f}(x, \alpha)$ with respect to $x \in \mathbb{R}$ exist and the interval $(\underline{f}'(x, \alpha), \overline{f}'(x, \alpha))$ defines the α-level set of a fuzzy number for $x \in \mathbb{R}$, $\alpha \in [0, 1]$. Then For $x \in \mathbb{R}$, $\alpha \in [0, 1]$ is called differentiable and we write*

$$\tilde{f}'(x, \alpha) = \Big(\underline{f}'(x, \alpha), \overline{f}'(x, \alpha)\Big). \quad (2.7)$$

Definition 2.9 *If $\Big(\frac{\partial \underline{f}(x, \alpha)}{\partial x_i}, \frac{\partial \overline{f}(x, \alpha)}{\partial x_i}\Big)$, $i = 1, 2, \ldots n$. Then $\tilde{f} : V \subseteq \mathbb{R}^n \to F$ is the gradient of fuzzy-valued function. Defines the α-level set of a fuzzy number, then the gradient of $\tilde{f}$ at x is*

$$\triangledown[\tilde{f}(x)]^\alpha = \Big(\frac{\partial[\tilde{f}(x)]_\alpha}{\partial x_1}, \frac{\partial[\tilde{f}(x)]_\alpha}{\partial x_2}, \ldots, \frac{\partial[\tilde{f}(x)]_\alpha}{\partial x_n}\Big). \quad (2.8)$$

Usind Lemma (2.1), the sufficient conditions that the gradient of $\tilde{f}$ at x exist are: For $\alpha \in [0, 1]$ the partial derivatives of $\underline{f}(x, \alpha)$ and $\overline{f}(x, \alpha)$ exist with respect to x_i, [8].

$\frac{\partial[\underline{f}(x)]_\alpha}{\partial x_i}$ is a continuous increasing function of α (it provide condition (1)),

$\frac{\partial[\overline{f}(x)]_\alpha}{\partial x_i}$ is a continuous decreasing function of (it provide condition (2)),

$\frac{\partial[\underline{f}(x)]_\alpha}{\partial x_i} \le \frac{\partial[\overline{f}(x)]_\alpha}{\partial x_i}$ provide condition (3).

Definition 2.10 *(see [8]) Suppose that $\tilde{f} : V \subseteq \mathbb{R} \to F$ is integrable with respect to x, if both $\underline{f}(x, \alpha)$ and $\overline{f}(x, \alpha)$ are Lebesgue integrable functions of $x \in \mathbb{R}$, for all $\alpha \in [0, 1]$ and $(\int \underline{f}(x, \alpha)dx, \int \overline{f}(x, \alpha)dx)$, defines the α-level set of a fuzzy number. Denoted the integral of fuzzy function $\tilde{f}$ with respect to for $\alpha \in [0, 1]$ by*

$$\int [\tilde{f}(x)]_\alpha dx = \int \underline{f}(x, \alpha)dx, \int \overline{f}(x, \alpha)dx. \quad (2.9)$$

Using Lemma (2.1), the suffcient conditions that $\int \underline{f}(x, \alpha)dx$ and $\int \overline{f}(x, \alpha)dx$ provide condition (1) and condition (2), respectively. Also $\int \underline{f}(x, 1)dx \le \int \partial\overline{f}(x, 1)dx$ provide condition (3).

Definition 2.11 *(Distance measure between fuzzy functions [8]). Suppose that $\tilde{f} : V \subseteq \mathbb{R} \to F$ and $\tilde{g} : V \subseteq \mathbb{R} \to F$ are two fuzzy functions. The distance measure between $\tilde{f}$ and $\tilde{g}$ is defined by*

$$D_F(\tilde{f},\tilde{g}) = \sup_{0\leq\alpha\leq 1} H([\tilde{f}(x)]_\alpha, [\tilde{g}(x)]_\alpha) = \max\left\{ \sup_{z\in[\tilde{f}(x)]_\alpha} H(z,[\tilde{g}(x)]_\alpha)), \sup_{y\in[\tilde{g}(x)]_\alpha)} H([\tilde{f}(x)]_\alpha), y)\right\} \quad (2.10)$$

where H is the well-known Hausdorff metric on the family of all nonempty compact subsets of $\mathbb{R}$, and $d(a,B) = \inf_{b\in B} d(a,b)$. For notational convenience, defined for any $\tilde{f} : V \subseteq \mathbb{R} \to F$

$$\|\tilde{f}(x)\|_F^2 = D_F(\tilde{f}(x), \tilde{f}(x)),\ \forall x \in S. \quad (2.11)$$

Definition 2.12 *Let $\tilde{x}(.)$ and $\tilde{x}(.)+\delta\tilde{x}(.)$ be fuzzy functions for which the fuzzy functional $\tilde{J}$ is efined. The increment of $\tilde{J}$, denoted by $\triangle\tilde{J}$, is defined as $\triangle\tilde{J} := \tilde{J}(\tilde{x}+\delta\tilde{x}) \ominus \tilde{J}(\tilde{x})$, where $\delta\tilde{x}(.)$ is known as the variation of $\delta\tilde{x}(.)$. In order to emphasize that the increment $\triangle\tilde{J}$ depends on the fuzzy functions $\tilde{x}$ and $\delta\tilde{x}$, we may denote $\triangle\tilde{J}$ by $\triangle\tilde{J}(\tilde{x},\delta\tilde{x})$ [8].*

Definition 2.13 *Let the increment of $\tilde{J}$ can be written as [8]*

$$\triangle\tilde{J}(\tilde{x},\delta\tilde{x}) := \delta\tilde{J}(\tilde{x},\delta\tilde{x}) + \eta(\tilde{x},\delta\tilde{x}).\|\delta\tilde{x}\|_F \quad (2.12)$$

where $\delta\tilde{J}$ is linear in $\delta\tilde{x}$. If for any $\epsilon > 0$,

$$D_F(\eta(\tilde{x},\delta\tilde{x}),\tilde{0}) < \epsilon, \text{ as } \|\delta\tilde{x}\|_F \to 0. \quad (2.13)$$

Then, we say that $\tilde{J}$ is differentiable on $\tilde{x}$.

Definition 2.14 *A fuzzy functional $\tilde{J}$ with domain $\tilde{C}[t_0,t_f]$, the class of all fuzzy continuous functions on $\tilde{C}[t_0,t_f]$, has a fuzzy relative minimizer $\tilde{x}^* = \tilde{x}^*(t)$, if the increment of $\tilde{J}$ is fuzzy non-negative, that is, $\tilde{J}(\tilde{x}) \geq \tilde{J}(\tilde{x}^*)$ for all fuzzy functions $\tilde{x}$ in $\tilde{C}[t_0,t_f]$. Notice that the The above-mentioned inequality holds if and only if $\underline{\tilde{x},\alpha} \geq \underline{\tilde{x}^*,\alpha}$ and $\overline{\tilde{x},\alpha} \geq \overline{\tilde{x}^*,\alpha}$ for all $\alpha \in [0,1]$.*

Theorem 2.1 *(Fuzzy fundamental theorem) Suppose that $\tilde{x},\delta\tilde{x} \in \tilde{C}[t_0,t_f]$ are fuzzy functions of $t \in [t_0,t_f]$ and $\tilde{J}(x)$ is differentiable fuzzy functional of $\tilde{x}$. If $\tilde{x}^*$ is a fuzzy minimizer of $\tilde{J}$, then the variation of $\tilde{J}$ regardless of any boundary conditions must vanish on $\tilde{x}^*$, that is,*

$$\delta\tilde{J}(\tilde{x},\delta\tilde{x}) = \tilde{0}, \quad (2.14)$$

for all admissible $\delta\tilde{x}$ having the property $\tilde{x},\delta\tilde{x} \in \tilde{C}[t_0,t_f]$.

Proof. See [8].

Now we consider the fuzzy initial value problem

$$\dot{\tilde{x}}(t) = \tilde{f}(t,\tilde{x}(t)),\ \tilde{x}(0) = 0 \quad (2.15)$$

where $\tilde{f} : [0,T] \times F(\mathbb{R})$ is obtained by Zadeh's extension principle from a continuous function $\tilde{f} : [0,T] \times \mathbb{R} \to \mathbb{R}$ Note that $\tilde{f}$ is continuous because f is continuous (see [24]), and if $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is a continuous function then according to Zadehs extension principle one can extend f to $\tilde{f} : \mathcal{F}(\mathbb{R})\times\mathcal{F}(\mathbb{R}) \to \mathcal{F}(\mathbb{R})$ by the equation $\tilde{f}(\tilde{u},\tilde{v})(z) = \sup_{z=f(s,t)} \min(\tilde{\mu}(s),\tilde{v}(t))$. It is well known that $[\tilde{f}(\tilde{u},\tilde{v})]_\alpha = \tilde{f}([\tilde{u}]_\alpha,[\tilde{v}]_\alpha)$, $\alpha \in [0,1]$, $\tilde{\mu} \in F(\mathbb{R})$, $\tilde{v} \in F(\mathbb{R})$ we have $[\tilde{f}(t,\tilde{x})]_\alpha = \tilde{f}(t,[\tilde{x}]_\alpha)$ where $f(t,A) = \{f(t,a) \mid a \in A\}$. Associated with (2.15) we can consider the following crisp differential equation

$$\dot{x}(t) = f(t,x(t)),\ x(0) = x_0 \quad (2.16)$$

where $\dot{x}(t)$ is the derivative of a crisp function $x : [0,T] \to \mathbb{R}$.

Theorem 2.2 *(see [7]) let $\tilde{x} \in \mathcal{F}(\mathbb{R})$. Suppose that f is a continuous function and for each $x_0 \in \mathbb{R}$ there exists a unique solution $x(.,x_0)$ for (2.16) and that $x(t,.)$ is continuous in $\mathbb{R}$ for each $t \in [0,T]$. Then: If $\tilde{f}$ is nondecreasing with respect to the second argument, then the fuzzy solution of (2.15) and the solution of (2.16).*

2.1 Multi-layer perceptron Network and it training

A multi-layer perceptron (MLP) is a class of feed forward artificial neural network. An MLP consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called back propagation for training. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron. It can distinguish data that is not linearly separable. Multi-layer networks use a variety of learning techniques, the most popular being back-propagation (BP), which is based on

The error correction learning rule. So, to calculate sensitivities for the different layers of neurons in the MLP network the Derivative of conversion neurons functions is required. So functions used that have derivative. One of these functions is the sigmoid function defined. it is A real, bounded, and derivative function. The sigmoid function has a positive derivative, and has the following general relationship $S(n) = \frac{1}{1-e^{-n}}$.

Theorem 2.3 *(The World approximation Builder) The multi-layer perceptron network with one hidden layer with a sigmoid function in the middle layer and linear transformation functions in output layer are able to approximate All functions in any degree of the integral of the square (see [12]).*

2.2 Quasi-Newton BFGS method

For compute the value of some predefined error-function the output values are compared with the correct answer. which it by various techniques, the error is then fed back through the network. Using this information, the algorithm adjusts the weights of each connection in order to reduce the value of the error function by some small amount. After repeating this process for a sufficiently large number of training cycles, the network will usually converge to some state where the error of the calculations is small. In this case, one would say that the network has learned a certain target function. To adjust weights properly, must minimize unconstrained optimization problem. For that purpose, minimization techniques such as the steepest descent method and the conjugate gradient or Quasi-Newton methods can be employed. The Newton method is one of the important algorithms in nonlinear optimization. The main disadvantage of the Newton method is that it is necessary to evaluate the second derivative matrix (Hessian matrix). Quasi-Newton methods were originally proposed by Davidon in 1959 and were later developed by Fletcher and Powell (1963). The most fundamental idea in quasi-Newton methods is the requirement to calculate an approximation of the Hessian matrix. Here the Quasi-Newton BFGS (Broyden Fletcher Goldfarb Shanno) method is used. This method is quadratically (see [18]).

3 Solution of Fuzzy optimization control problem (FOCP) using Artificial Neural Network

In this paper, the following type of FOCP is considered as follows:

$$Minimize\, \tilde{J}(\tilde{u}) : \int_{t_0}^{t_f} \tilde{g}(\tilde{x}(t), \tilde{u}(t), t)dt$$

$$subject\ to\, \dot{\tilde{x}}(t) = \tilde{h}(\tilde{x}(t), \tilde{u}(t), t), \tag{3.17}$$

$$\tilde{x}(t_0) = \tilde{x}_0, \quad \tilde{x}(t_f) = \tilde{x}_f.$$

Where $\tilde{g}$ and $\tilde{h}$ assign a fuzzy number to the fuzzy point $(\tilde{x}(t), \tilde{u}(t), t) \in \mathbb{F}^2 \times \mathbb{R}$, where the fuzzy state $\tilde{x}(t)$ and the fuzzy control $\tilde{u}(t)$ are fuzzy functions of belonging to the specified interval $[t_0, t_f]$. Assuming that the integrand $\tilde{g}$ and fuzzy function $\tilde{h}$ have continuous first and second partial derivatives with respect to all of their arguments. Suppose $\tilde{x} = \tilde{x}(t)$ is admissible, if it satisfies the endpoints conditions and also is twice continuously differentiable with respect to $t \in [t_0, t_f]$. Also an admissible fuzzy control $\tilde{u} = \tilde{u}(t)$ is that is not bounded.
In order to model this problem, the fuzzy Lagrange multiplier is adopted. To begin with, the fuzzy augmented functional is formulated as follows:

$$\tilde{J}_a(\tilde{u}) = \int_{t_0}^{t_f} \tilde{g}_a(\tilde{x}(t), \tilde{u}(t), \tilde{P}(t), \dot{\tilde{x}}(t), t)dt. \tag{3.18}$$

Where

$$\begin{aligned}\tilde{g}_a(\tilde{x}(t),&\tilde{u}(t), \tilde{P}(t), \dot{\tilde{x}}(t), t) := \tilde{g}(\tilde{x}(t), \tilde{u}(t), t) \\ &+ \tilde{P}(t)\tilde{h}(\tilde{x}(t), \tilde{u}(t), t) \ominus \dot{\tilde{x}}(t).\end{aligned} \tag{3.19}$$

In order to simplify the result presentations, the special case is stated in the following assumption.

Remark 3.1 *To simplify the variation equations, it is assumed that $\underline{J}_\alpha(\tilde{u})$ or $\overline{J}_\alpha(\tilde{u})$ is stated in terms containing only $\underline{x}(t,\alpha)$ and $\overline{x}(t,\alpha)$ or $(\overline{x}(t,\alpha)$ and $\dot{\overline{x}}(t,\alpha))$. In this case, $\underline{J}_\alpha(\underline{u},\alpha)$ and $\overline{J}_\alpha(\overline{u},\alpha)$ are considered instead of $\underline{J}_\alpha(\tilde{u},\alpha)$ and $\overline{J}_\alpha(\tilde{u},\alpha)$, respectively.*

On the $\tilde{u}^*$, the variation of $\tilde{J}_\alpha$ must be zero. That is, $\delta\tilde{J}_\alpha(\tilde{u}^*) = \tilde{0}$. This admits for all $\alpha \in [0,1]$,

$$\delta\underline{J}_\alpha(\underline{u}^*, \alpha) = \tilde{0}, \tag{3.20}$$

and

$$\delta\overline{J}_\alpha(\overline{u}^*, \alpha) = \tilde{0}. \tag{3.21}$$

Originally, for relation (3.20), the following formulation is considered:

$$\begin{aligned}\delta\underline{J}_\alpha(\underline{u}^*, \alpha) = &\int_{t_0}^{t_f} \Big\{\Big[\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t) \\ &, \underline{P}^*(t), \underline{\dot{x}}^*(t), t, \alpha) \\ &- \frac{d}{dt}\Big(\frac{\partial \underline{g}_\alpha}{\partial \underline{\dot{x}}}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}^*(t), \underline{\dot{x}}^*(t), t, \alpha)\Big)\Big]\delta\underline{x} \\ &+ \Big[\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}^*(t), \underline{\dot{x}}^*(t), t, \alpha)\Big]\delta\underline{u} \\ &+ \Big[\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}^*(t), \underline{\dot{x}}^*(t), t, \alpha)\Big]\Big\}\delta\underline{P}dt \end{aligned} \tag{3.22}$$

$$= 0$$

Regarding the fuzzy augmented integrand function $\tilde{g}_\alpha$ defined in (3.19), the latter equation can be written as follows:

$$\begin{aligned}&\int_{t_0}^{t_f} \Big\{\Big[\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}^*(t), \underline{\dot{x}}^*(t), t, \alpha) \\ &+ \underline{P}^*(t,\alpha)\Big(\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), t, \alpha\Big) - \\ &\frac{d}{dt}(-\underline{P}^*(t,\alpha))\Big]\delta\underline{x} + \Big[\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), t, \alpha) \\ &+ \underline{P}^*(t,\alpha)\Big(\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), t, \alpha\Big)\Big]\delta\underline{u} + \\ &\underline{h}[\underline{h}(\underline{x}^*(t), \underline{u}^*(t), t, \alpha) - \underline{\dot{x}}^*(t,\alpha)]\delta\underline{P}\Big\}dt = 0 \end{aligned} \tag{3.23}$$

As this constraint must be satisfied by the extremal $\tilde{u}^*$, it is found out that

$$\underline{\dot{x}}^*(t,\alpha) = \underline{h}(\underline{h}(\underline{x}^*(t), \underline{u}^*(t), t) \tag{3.24}$$

and hence the coefficient of $\delta\underline{P}$ in (3.23) is zero. Moreover, the arbitrary fuzzy Lagrange multiplier $\underline{P}^*$ can be chosen such that the coefficient of $\delta\underline{x}$ does not appear in the above integral. Thus:

$$\begin{aligned}\underline{P}^*(t,\alpha) = &\Big[\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}^*(t), \underline{\dot{x}}^*(t), t, \alpha) \\ &+ \underline{P}^*(t,\alpha)\Big(\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), t, \alpha\Big)\Big]. \end{aligned} \tag{3.25}$$

There exists nevertheless a term inside the integral (3.23) to deal with. As the equality (3.23) must be satisfied, it is obtained that:

$$\begin{aligned}&\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}^*(t), \underline{\dot{x}}^*(t), t, \alpha) + \underline{P}^*(t,\alpha) \\ &\Big(\frac{\partial \underline{g}_\alpha}{\partial \underline{x}}(\underline{x}^*(t), \underline{u}^*(t), t, \alpha\Big) = 0. \end{aligned} \tag{3.26}$$

As such, the well-known Hamiltonian function is constructed as follows:

$$\begin{aligned}&H(\tilde{x}(t), \tilde{u}(t), \tilde{P}(t), t) := \\ &\tilde{g}(\tilde{x}(t), \tilde{u}(t), t) + \tilde{P}(t), \tilde{h}(\tilde{x}(t), \tilde{u}(t), t) \end{aligned} \tag{3.27}$$

where $\tilde{P}$ is the so-called fuzzy Lagrange multiplier with the α-level set

$$\begin{aligned}&\underline{H}(\tilde{x}(t), \tilde{u}(t), \tilde{P}(t), t)[\alpha] = \Big(\\ &\underline{H}(\underline{x}(t), \underline{u}(t), \underline{P}(t), t, \alpha), \overline{H}(\overline{x}(t), \overline{u}(t), \overline{P}(t), t, \alpha)\Big) \end{aligned} \tag{3.28}$$

where

$$\begin{aligned}&\underline{H}(\underline{x}(t), \underline{u}(t), \underline{P}(t), t, \alpha) = \\ &\underline{g}(\underline{x}(t), \underline{u}(t), t, \alpha) + \underline{P}(t), \underline{h}(\underline{x}(t), \underline{u}(t), t, \alpha) \end{aligned} \tag{3.29}$$

$$\begin{aligned}&\overline{H}(\overline{x}(t), \overline{u}(t), \overline{P}(t), t, \alpha) = \\ &\overline{g}(\overline{x}(t), \overline{u}(t), t) + \overline{P}(t), \overline{h}(\overline{x}(t), \overline{u}(t), t, \alpha) \end{aligned} \tag{3.30}$$

Using the left-hand Hamiltonian function $\underline{H}(\underline{x}(t), \underline{u}(t), \underline{P}(t), t, \alpha)$ defined in (3.29), the equations (3.24)-(3.26) can be written more compactly as follows:

$$\underline{\dot{x}}^*(t) = \frac{\partial \underline{H}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}(t)^*, t)}{\partial \underline{P}^*} \tag{3.31}$$

$$\underline{\dot{P}}^*(t) = -\frac{\partial \underline{H}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}(t)^*, t)}{\partial \underline{x}^*} \tag{3.32}$$

$$0 = \frac{\partial \underline{H}(\underline{x}^*(t), \underline{u}^*(t), \underline{P}(t)^*, t)}{\partial \underline{u}^*} \tag{3.33}$$

for all $\alpha \in [0,1]$, $t \in [t_0, t_f]$. According to the above, it can be said that:

Theorem 3.1 *Let $\underline{x}^*_\alpha = \underline{x}^*(t,\alpha)$ be an admissible fuzzy function, i.e., it is twice continuously differentiable fuzzy function. Then, in order that $\underline{x}^*_\alpha$ give a relative (local) minimum to the fuzzy functional $\underline{J}_\alpha$ in (FOCP), it is necessary that for all $\alpha \in [0,1]$, $t \in [t_0, t_f]$.*

Again, following the scheme of obtaining (3.31)-(3.33) and adapting it to the case under consideration involving (3.21), one may show that for all $\alpha \in [0,1]$, $t \in [t_0, t_f]$.

$$\dot{\overline{x}}^*(t) = \frac{\partial \overline{H}(\overline{x}^*(t), \overline{u}^*(t), \overline{P}(t)^*, t)}{\partial \overline{P}^*} \tag{3.34}$$

$$\dot{\overline{P}}^*(t) = -\frac{\partial \overline{H}(\overline{x}^*(t), \overline{u}^*(t), \overline{P}(t)^*, t)}{\partial \overline{x}^*} \tag{3.35}$$

$$0 = \frac{\partial \overline{H}(\overline{x}^*(t), \overline{u}^*(t), \overline{P}(t)^*, t)}{\partial \overline{u}^*} \tag{3.36}$$

Theorem 3.2 *Let $\overline{x}^*_\alpha = \overline{x}^*(t,\alpha)$ be an admissible fuzzy function, i.e. it is twice continuously differentiable fuzzy function. Then, in order that $\overline{x}^*_\alpha$ give a relative (local) minimum to the fuzzy functional $\overline{J}_\alpha$ in (FOCP), it is necessary that for all $\alpha \in [0,1]$, $t \in [t_0, t_f]$.*

Now, using Theorem (3.1) and Theorem (3.2), let $\tilde{x}^*_\alpha = \tilde{x}^*(t,\alpha)$ be an admissible fuzzy function, i.e. it is twice continuously differentiable fuzzy function. Then, in order that $\tilde{x}^*_\alpha$ give a relative (local) minimum to the fuzzy functional $\tilde{J}_\alpha$ in (FOCP), it is necessary that for all $\alpha \in [0,1]$, $t \in [t_0, t_f]$. Indeed, for the Hamiltonian function in (3.27) and using (3.28) necessary conditions for the binary $(\underline{x}^*_\alpha, \overline{x}^*_\alpha)$ the optimal solution of (3.17) is the existence of a costate vector function $(\underline{u}^*_\alpha, \overline{u}^*_\alpha)$ that satisfies the following differential equations (3.31)-(3.36), where $(\underline{x}^*_\alpha, \overline{x}^*_\alpha)$ is the same as $\tilde{x}^*_\alpha$ while $(\underline{u}^*_\alpha, \overline{u}^*_\alpha)$ is the same as $\tilde{u}^*_\alpha$.

Now, it is tried to develop an approximation method to solve the equations arising in PMP. This proposed method is a three-equation system with initial conditions and boundary conditions. The proposed equation system is defined as follows:

$$\begin{cases} \tilde{x}_T = f_1(t) + f_2(t, \widetilde{Net}_x(t,\alpha,\tilde{\varphi}_x)) \\ \tilde{p}_T = f_3(t)\widetilde{Net}_p(t,\alpha,\tilde{\varphi}_p) \\ \tilde{u}_T = \widetilde{Net}_u(t,\alpha,\tilde{\varphi}_u) \end{cases} \tag{3.37}$$

Where, $\forall i = 1,2,3$, $f_i(t)$, is an arbitrary function in terms of the variable t, so that the conditions of the issue are true. includes all weights of relation (3.37). Three fuzzy neural networks are introduced for each equation mentioned above as follows:

$$\widetilde{Net}_x(t,\alpha,\tilde{\varphi}_x) = \sum_{i=1}^{k} \frac{\tilde{v}_{ix}}{1 + e^{-\tilde{w}_{1ix}t - \tilde{w}_{2ix}\alpha - \tilde{b}_{ix}}} \tag{3.38}$$

$$\widetilde{Net}_p(t,\alpha,\tilde{\varphi}_p) = \sum_{i=1}^{k} \frac{\tilde{v}_{ip}}{1 + e^{-\tilde{w}_{1ip}t - \tilde{w}_{2ip}\alpha - \tilde{b}_{ip}}} \tag{3.39}$$

$$\widetilde{Net}_u(t,\alpha,\tilde{\varphi}_u) = \sum_{i=1}^{k} \frac{\tilde{v}_{iu}}{1 + e^{-\tilde{w}_{1iu}t - \tilde{w}_{2iu}\alpha - \tilde{b}_{iu}}}. \tag{3.40}$$

By replacing the approximate solutions into the fuzzy Hamiltonian function, an approximate fuzzy Hamiltonian $\tilde{H}_T$ is defined where the functions $\tilde{x}, \tilde{p}$ and $\tilde{u}$ are replaced by their corresponding approximate format

$$\tilde{H}(\tilde{x}_T(t), \tilde{u}_T(t), \tilde{p}_T(t), t)[\alpha] = (\underline{H}(\underline{x}_T(t), \underline{u}_T(t), \underline{p}_T(t), t, \alpha), \overline{H}(\overline{x}_T(t), \overline{u}_T(t), \overline{p}_T(t), t, \alpha)) \tag{3.41}$$

with the α-level set, we have, $\tilde{x}_T = (\underline{x}_T, \overline{x}_T)$, $\tilde{p}_T = (\underline{p}_T, \overline{p}_T)$, $\tilde{u}_T = (\underline{u}_T, \overline{u}_T)$ where

$$\begin{cases} \underline{x}_T = f_1(t) + f_2(t)\underline{Net}_x(t,\alpha,\underline{\varphi}_x) \\ \underline{p}_T = f(t)\underline{Net}_p(t,\alpha,\underline{\varphi}_p) \\ \underline{u}_T = \underline{Net}_u(t,\alpha,\underline{\varphi}_p) \\ \overline{x}_T = f_1(t) + f_2(t)\overline{Net}_x(t,\alpha,\overline{\varphi}_x) \\ \overline{p}_T = f(t)\overline{Net}_p(t,\alpha,\overline{\varphi}_p) \\ \overline{u}_T = \overline{Net}_u(t,\alpha,\overline{\varphi}_p) \end{cases} \tag{3.42}$$

Also, with the α-level set, we have, $\widetilde{Net}_x = (\underline{Net}_x, \overline{Net}_x)$, $\widetilde{Net}_p = (\underline{Net}_p, \overline{Net}_p)$, $\widetilde{Net}_u =$

$(\underline{Net}_u, \overline{Net}_u)$, where

$$\begin{cases} \underline{Net}_x(t,\alpha,\underline{\varphi}_x) = \sum_{i=1}^{k} \dfrac{\underline{v}_{ix}}{1+e^{-\underline{w}_{1ix}t-\underline{w}_{2ix}\alpha-\underline{b}_{ix}}} \\ \underline{Net}_p(t,\alpha,\underline{\varphi}_p) = \sum_{i=1}^{k} \dfrac{\underline{v}_{ip}}{1+e^{-\underline{w}_{1ip}t-\underline{w}_{2ip}\alpha-\underline{b}_{ip}}} \\ \underline{Net}_u(t,\alpha,\underline{\varphi}_u) = \sum_{i=1}^{k} \dfrac{\underline{v}_{iu}}{1+e^{-\underline{w}_{1iu}t-\underline{w}_{2iu}\alpha-\underline{b}_{iu}}} \\ \overline{Net}_x(t,\alpha,\overline{\varphi}_x) = \sum_{i=1}^{k} \dfrac{\overline{v}_{ix}}{1+e^{-\overline{w}_{1ix}t-\overline{w}_{2ix}\alpha-\overline{b}_{ix}}} \\ \overline{Net}_p(t,\alpha,\overline{\varphi}_p) = \sum_{i=1}^{k} \dfrac{\overline{v}_{ip}}{1+e^{-\overline{w}_{1ip}t-\overline{w}_{2ip}\alpha-\overline{b}_{ip}}} \\ \overline{Net}_u(t,\alpha,\overline{\varphi}_u) = \sum_{i=1}^{k} \dfrac{\overline{v}_{iu}}{1+e^{-\overline{w}_{1iu}t-\overline{w}_{2iu}\alpha-\overline{b}_{iu}}} \end{cases} \tag{3.43}$$

On the other hand the trial solutions (3.42) must satisfy conditions (3.34)-(3.36), so we replace them in (3.34)-(3.36):

$$\dot{\tilde{P}}_T(t) + \frac{\partial \tilde{H}(\tilde{x}_T(t),\tilde{u}_T(t),\tilde{p}_T(t),t)}{\partial \tilde{x}_T} = 0 \tag{3.44}$$

$$\dot{\tilde{x}}_T(t) - \frac{\partial \tilde{H}(\tilde{x}_T(t),\tilde{u}_T(t),\tilde{p}_T(t),t)}{\partial \tilde{p}_T} = 0 \tag{3.45}$$

$$\frac{\partial \tilde{H}(\tilde{x}_T(t),\tilde{u}_T(t),\tilde{p}_T(t),t)}{\partial \tilde{u}_T} = 0 \tag{3.46}$$

To solve the equations (3.44)-(3.46), three error functions are considered corresponding to each equation as follows:

$$\tilde{e}_1(\varphi,t) = \left[\dot{\tilde{P}}_T(t) + \frac{\partial \tilde{H}(\tilde{x}_T(t),\tilde{u}_T(t),\tilde{p}_T(t),t)}{\partial \tilde{x}_T}\right]^2 \tag{3.47}$$

$$\tilde{e}_2(\varphi,t) = \left[\dot{\tilde{x}}_T(t) - \frac{\partial \tilde{H}(\tilde{x}_T(t),\tilde{u}_T(t),\tilde{p}_T(t),t)}{\partial \tilde{p}_T}\right]^2 \tag{3.48}$$

$$\tilde{e}_3(\varphi,t) = \left[\frac{\partial \tilde{H}(\tilde{x}_T(t),\tilde{u}_T(t),\tilde{p}_T(t),t)}{\partial \tilde{u}_T}\right]^2 \tag{3.49}$$

Where φ is a vector containing all weights of there MLP networks (3.43). Note that φ contains all weights $\tilde{v}_x, \tilde{w}_{1x}, \tilde{w}_{2x}, \tilde{v}_p, \tilde{w}_{1p}, \tilde{w}_{2p}, \tilde{v}_u, \tilde{w}_{1u}$ and $\tilde{w}_{2u}$. With the α-level set, for relations (3.47)-(3.49) and by deriving the functions contained in relation (3.42) the following equation is obtained as follows:

$$\begin{cases} \underline{e}_1(\varphi,t) = \sum\limits_{i=1}^{k}\Big[\Big(f_3'(t)\underline{Net}_p + f_3(t)\dfrac{\partial \underline{Net}_p}{\partial t}\Big) \\ \quad + \dfrac{\partial \underline{H}(\underline{x}_T(t),\underline{u}_T(t),\underline{p}_T(t),t,\alpha)}{\partial \underline{x}_T}\Big]^2 \\ \underline{e}_2(\varphi,t) = \sum\limits_{i=1}^{k}\Big[\Big(f_1'(t) + f_2'(t)\underline{Net}_x + f_2(t) \\ \quad \dfrac{\partial \underline{Net}_x}{\partial t}\Big) - \dfrac{\partial \underline{H}(\underline{x}_T(t),\underline{u}_T(t),\underline{p}_T(t),t,\alpha)}{\partial \underline{p}_T}\Big]^2 \\ \underline{e}_3(\varphi,t) = \sum\limits_{i=1}^{k}\Big[\dfrac{\partial \underline{H}(\underline{x}_T(t),\underline{u}_T(t),\underline{p}_T(t),t,\alpha)}{\partial \underline{u}_T}\Big]^2 \\ \overline{e}_1(\varphi,t) = \sum\limits_{i=1}^{k}\Big[\Big(f_3'(t)\overline{Net}_p + f_3(t)\dfrac{\partial \overline{Net}_p}{\partial t}\Big) \\ \quad + \dfrac{\partial \overline{H}(\overline{x}_T(t),\overline{u}_T(t),\overline{p}_T(t),t,\alpha)}{\partial \overline{x}_T}\Big]^2 \\ \overline{e}_2(\varphi,t) = \sum\limits_{i=1}^{k}\Big[\Big(f_1'(t) + f_2'(t)\overline{Net}_x + f_2(t) \\ \quad \dfrac{\partial \overline{Net}_x}{\partial t}\Big) - \dfrac{\partial \overline{H}(\overline{x}_T(t),\overline{u}_T(t),\overline{p}_T(t),t,\alpha)}{\partial \overline{p}_T}\Big]^2 \\ \overline{e}_3(\varphi,t) = \sum\limits_{i=1}^{k}\Big[\dfrac{\partial \overline{H}(\overline{x}_T(t),\overline{u}_T(t),\overline{p}_T(t),t,\alpha)}{\partial \overline{u}_T}\Big]^2 \end{cases} \tag{3.50}$$

Finally, a total error function $e(\varphi,t) = \underline{e}_1(\varphi,t) + \underline{e}_2(\varphi,t) + \underline{e}_3(\varphi,t) + \overline{e}_1(\varphi,t) + \overline{e}_2(\varphi,t) + \overline{e}_3(\varphi,t)$ is obtained.

So, instead of solving equations (3.44), (3.45) and (3.46), the interval $[t_0, t_f]$ (by m points) is discredited and then solved using following unconstrained fuzzy optimization problem:

$$E(p) = \min \sum_{j=1}^{m} e(\varphi, t_j) \tag{3.51}$$

In this work, the relation (3.51) is minimized using the BFGS Quasi-Newton method (see [15] for more details).

4 Numerical Examples

In this section, two examples that illustrate the efficiency, simplicity and precision of the proposed method are presented.

Example 4.1 *Find the fuzzy control that minimize*

$$\tilde{J}(\tilde{u}) = \int_0^1 \tilde{u}^2(t)dt, \tag{4.52}$$

subject to

$$\dot{\tilde{x}}(\tilde{u}) = \tilde{u}(t) - (0,1,3)\tilde{x}(t), \quad t \in [0,1],$$

with boundary conditions

$$\tilde{x}(0) = 1 = (1,1,1), \quad \tilde{x}(1) = 0 = (0,0,0).$$

In First step we must construct the Hamiltonian function

$$\tilde{H}(\tilde{x}(t), \tilde{u}(t), \tilde{P}(t), t, \alpha) := \tilde{u}^2(t) + \tilde{P}(t)[\tilde{u}(t) - (0,1,3)\tilde{x}(t)] \quad (4.53)$$

and hence the α-level set of $\tilde{H}$ is characterized by

$$\underline{H}(\underline{x}(t), \underline{u}(t), \underline{P}(t), t, \alpha) = \underline{u}^2(t,\alpha) + \underline{p}(t,\alpha)[\underline{u}(t,\alpha) - (3-2\alpha)\underline{u}(t,\alpha)] \quad (4.54)$$

$$\overline{H}(\overline{x}(t), \overline{u}(t), \overline{P}(t), t, \alpha) = \overline{u}^2(t,\alpha) + \overline{p}(t,\alpha)[\overline{u}(t,\alpha) - (3-2\alpha)\overline{u}(t,\alpha)] \quad (4.55)$$

According to (3.31)-(3.33), we must have:

$$\begin{cases} \dot{\underline{x}}(t,\alpha) = \underline{u}(t,\alpha) - (3-2\alpha)\underline{x}(t,\alpha) \\ \dot{\overline{x}}(t,\alpha) = \overline{u}(t,\alpha) - (\alpha)\overline{x}(t,\alpha) \\ \dot{\underline{p}}(t,\alpha) = (3-2\alpha)\underline{p}(t,\alpha) \\ \dot{\overline{p}}(t,\alpha) = (\alpha)\overline{p}(t,\alpha) \\ 0 = 2\underline{u}(t,\alpha) + \underline{p}(t,\alpha) \\ 0 = 2\overline{u}(t,\alpha) + \overline{p}(t,\alpha) \end{cases} \quad (4.56)$$

We can select the approximate solution according to the initial conditions $x(0) = 1 = (1,1,1)$ and $x(1) = 0 = (0,0,0)$, as follows:

$$\begin{cases} \tilde{x}_T = (1-t)(1 + t\widetilde{Net}_x(t,\alpha,\tilde{\varphi}_x)) \\ \tilde{p}_T = \widetilde{Net}_x(t,\alpha,\tilde{\varphi}_p)) \\ \tilde{u}_T = \widetilde{Net}_u(t,\alpha,\tilde{\varphi}_u)) \end{cases} \quad (4.57)$$

In this example, error function for two neurons units with sigmoid function in the hidden layer and for $m = 50$ equally spaced points inside the interval $[0,1]$ is trained. The control and state functions and Lagrange multipliers are shown for each α-cut in Fig. 1.
$E(P) = 1.618451668844628e-04$

Example 4.2 *Find the fuzzy control that minimize*

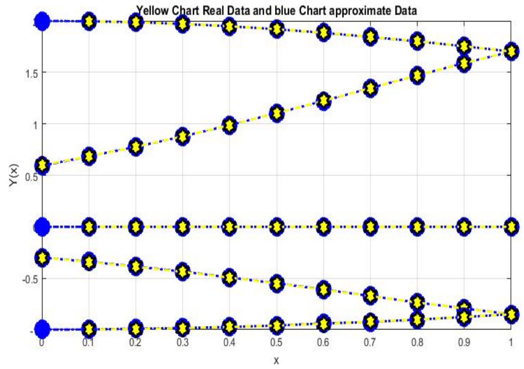

Figure 1: The control and state functions and Lagrange multipliers for Example 4.1

$$\tilde{J}(\tilde{u}) = \int_0^1 [(2 - \tilde{x}(x))^2 + \tilde{u}^2(t)]dt, \quad (4.58)$$

subject to

$$\dot{\tilde{x}}(\tilde{u}) = -(0,0.25,0.5)\tilde{x}(t) + \tilde{u}(t), \quad t \in [0,1],$$

with boundary conditions

$$\tilde{x}(0) = 0 = (0,0,0), \quad \tilde{x}(1) = 2 = (2,2,2).$$

In First step we must construct the Hamiltonian function

$$\tilde{H}(\tilde{x}(t), \tilde{u}(t), \tilde{p}(t), t) := [(2 - \tilde{x}(t))^2 \tilde{u}^2(t)] + \tilde{P}(t)[-(0,0.25,0.5)\tilde{x}(t) + \tilde{u}(t)] \quad (4.59)$$

and hence the α-level set of $\tilde{H}$ is characterized by

$$\underline{H}(\underline{x}(t), \underline{u}(t), \underline{p}(t), t, \alpha) := [(2 - \underline{x}(t,\alpha))^2 + \underline{u}^2(t,\alpha)] + \underline{p}(t,\alpha)[-(0.5 - 0.25\alpha)\underline{x}(t,\alpha) + \underline{u}(t,\alpha)] \quad (4.60)$$

$$\overline{H}(\overline{x}(t), \overline{u}(t), \overline{p}(t), t, \alpha) := [(2 - \overline{x}(t,\alpha))^2 + \overline{u}^2(t,\alpha)] + \overline{p}(t,\alpha)[-(0.25\alpha)\overline{x}(t,\alpha) + \overline{u}(t,\alpha)] \quad (4.61)$$

According to (3.31)-(3.33), we must have:

$$\begin{cases} \dot{\underline{x}}(t,\alpha) = -(0.5 - 0.25\alpha)\underline{x}(t,\alpha) + \underline{u}(t,\alpha) \\ \dot{\overline{x}}(t,\alpha) = -(0.25\alpha)\overline{x}(t,\alpha) + \overline{u}(t,\alpha) \\ \dot{\underline{p}}(t,\alpha) = -2(2 - \underline{x}(t,\alpha)) + \underline{p}(t,\alpha) \\ \qquad (-(0.5 - 0.25\alpha)) \\ \dot{\overline{p}}(t,\alpha) = -2(2 - \overline{x}(t,\alpha)) + \overline{p}(t,\alpha)(-0.25\alpha) \\ 0 = 2\underline{u}(t,\alpha) + \underline{p}(t,\alpha) \\ 0 = 2\overline{u}(t,\alpha) + \overline{p}(t,\alpha) \end{cases} \quad (4.62)$$

We can select the approximate solution according to the initial conditions $x(0) = 0 = (0, 0, 0)$ and $x(1) = 2 = (2, 2, 2)$, as follows:

$$\begin{cases} \tilde{x}_T = 2t + t(t-1)\widetilde{Net}_x(t, \alpha, \tilde{\varphi}_x) \\ \tilde{p}_T = \widetilde{Net}_x(t, \alpha, \tilde{\varphi}_p) \\ \tilde{u}_T = \widetilde{Net}_u(t, \alpha, \tilde{\varphi}_u) \end{cases} \tag{4.63}$$

In this example, error function for two neurons units with sigmoid function in the hidden layer and for $m = 50$ equally spaced points inside the interval $[0, 1]$ is trained. The state function is shown for each α-cut in Fig. 2 The control function is shown for each α-cut in Fig. 3.
$E(P) = 0.005391814261496$

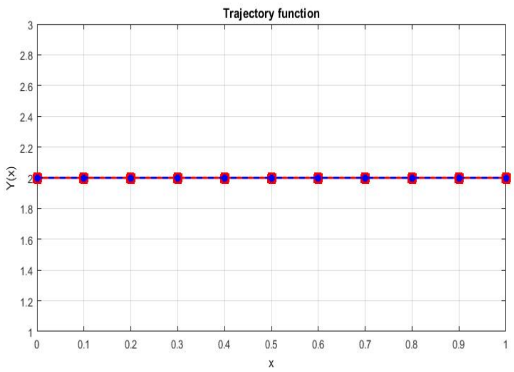

Figure 2: State function for Example 4.2

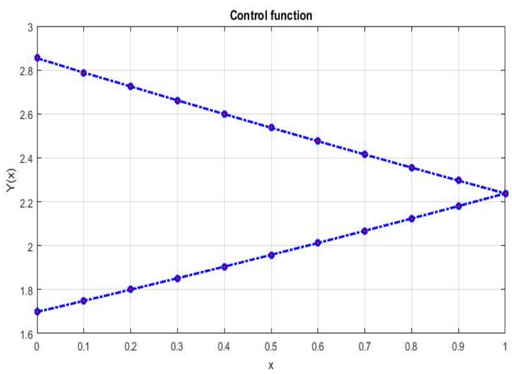

Figure 3: Control function for Example 4.2

5 Conclusion

In this paper, it is tried to provide an approximate solution to FOCP. This approach allows transforming a fuzzy problem to a set of crisp problems via cut- , that can be solved with MLP network method. The MLP network training is based on minimizing the sum of squares errors of all Crisp problems. One of the advantages of the proposed method is that it provides a general form of algorithms that can be designed based on the neural network to approximate the solution of FHJB equation and also for problems arising in calculus of variations, formula of each of the functions used in this method is determined by the type of problem. Another advantage of the proposed method is that as it obtains more solutions that are precise, the number of hidden layers and training points can be set depending on the type of problem. In some cases, it can be seen that the number of points or hidden layer of the neural network can increased while in some cases, they can be decreased. The proposed solution is an appropriate function for state, co-state, and control functions. Using some examples, the main novelties of this paper are highlighted by solving some examples.

References

[1] B. Bede, S. G. Gal, Generalizations of the differentiability of fuzzy-number-valued functions with applications to fuzzy differential equations, *Fuzzy Sets and Systems* 151 (2005) 581-599.

[2] S. Berkani, F. Manseur, A. Maidi, Optimal control based on the variational iteration method, *Comput Math Appl* 64 (2012) 604-610.

[3] T. Cheng, H. Sun, Z. Qu, FL. Lewis, Neural network solution for suboptimal control of non-holonomic chained form system, *Trans Inst Measurement Control* 31 (2009) 475-494.

[4] T. Cheng, FL. Lewis, Neural network solution for finite horizon H-infinity constrained optimal control of nonlinear systems, *J Control Theory Appl* 5 (2007) 1-11.

[5] D. Clever, J. Lang, S. Ulbrich, JC. Ziems, Combination of an adaptive multilevel SQP method and a space-time adaptive PDAE solver for optimal control problems, *Procedia Comput Sci.* 2 (2010) 1435-1443.

[6] S. Effati, M. Pakdaman, Optimal control problem via neural networks, *Neural Computing and Applications* 23 (2013) 7-8.

[7] M. H. Farahi, M. Keshtegar, M. Najariyan, Fuzzy optimal control of a poisoning-pest model by using α-cuts, *Journal of engineering research and technology, volume* 1 (2014) 79-82.

[8] B. Farhadinia, Pontryagins minimum principle for fuzzy optimal control problems, *Iranian Journal of Fuzzy Systems* 11 (2014) 27-43.

[9] D. Garg D, M. Patterson, WW. Hagera, AV. Raoa, DA. Bensonb, GT. Huntington, A unified framework for the numerical solution of optimal control problems using pseudo-spectral methods, *Automatica* 46 (2010) 1843-1851.

[10] M. Gerdts, A non-smooth Newton's method for control-state constrained optimal control problems, *Math Comput Simul* 79 (2008) 925-936.

[11] Y. Gao, Y, J, Liu, Adaptive fuzzy optimal control using direct heuristic dynamic programming for chaotic discrete-time system, *J. Vib. Control* 22 (2016) 595-603.

[12] K. Hornick, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* 2 (1989) 359-366.

[13] Wu. Hsien-Chung, Duality Theory in Fuzzy Optimization Problems, *Fuzzy Optimization and Decision Making* 3 (2005) 345-365.

[14] DE. Kirk, Optimal control theory an introduction, *Dover Publications, Mineola, NY* (2004).

[15] C. C. Ku, P. H. Huang, W. J. Chang, Passive fuzzy controller design for nonlinear systems with multiplicative noises, *J. Franklin Institute* 347 (2010) 732-750.

[16] Y. C. Kwun, J. S. Kim, M. J. H. Park, J. Park, Nonlocal controllability for the semi linear fuzzy integro–differential equations in n-dimensional fuzzy vector space, *Advances in Difference Equations* (2009).

[17] F. Lewis, VL. Syrmos, Optimal control. Wiley, *New York* (1995).

[18] D. G. Luenberger, Linear and Nonlinear Programming, second ed, *Addison Wesley* (1984).

[19] H. Modares, MB. Naghibi Sistani, Solving nonlinear optimal control problems using a hybrid IPSOSQP algorithm, *Eng Appl Artif Intell* 24 (2011) 476-484.

[20] M. Najariyan, M. H. Farahi, Optimal control of fuzzy liner controlled system with fuzzy initial conditions, *Iranian Journal of Fuzzy Systems* 10 (2013) 21-35.

[21] Hu. Pengwei, Y. Xiaohui, L. Xin, Y. Zhu-Hong, Incorporating the Coevolving Information of Substrates in Predicting HIV-1 Protease Cleavage Sites, *IEEE/ACM transactions on computational biology and bioinformatics* (2019).

[22] N. D. Phu, T. T. Tung, Existence of solutions of fuzzy control differential equations, *J. Science and Technology Development* 10 (2007) 5-12.

[23] R. Pinch, Optimal control and the calculus of variations, Oxford Science Publications, *Oxford Univ. Press, New York* (1993).

[24] F. Roman, L. Barros, R. Bassanezi, A note on the Zadehs extensions, *Fuzzy sets on Banach spaces, Inform.Sci* 144 (2002) 227-247.

[25] H. Saberi Nik, S. Effati, M. Shirazian, An approximate-analytical solution for the Hamilton Jacobi Bellman equation via homotopy perturbation method, *Appl Math Model* 36 (2012) 5614-5623.

[26] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. on SMC.* 15 (1985) 116-132.

[27] L. Xin, Zh. MengChu, L. Shuai, H. Lun, and Mingsheng. Sh, Non-Negativity Constrained Missing Data Estimation for High-Dimensional and Sparse Matrices from Industrial Applications, *IEEE transactions on cybernetics* (2019).

[28] D. Yang, K. Y. Cai, Finite-time quantized guaranteed cost fuzzy control for continuous-time nonlinear systems, *Expert Systems with Applications* http://dx.doi.org/10.1016/j.eswa(2010)03.024/.

[29] H. Zarei, A. V. Kamyad, A. A. Heydari, Fuzzy modeling and control of HIV infection, *Computational and Mathematical Methods in Medicine Art. no. 893474.* (2012).

[30] H. J. Zimmermann, Fuzzy set theory and its applications, *Kluwer Academic, Boston* (1991).

Sahar Askari is a PhD. Student in Mathematics at IAU, Iran, Hamedan. Her recent research interests numerical solution of Fuzzy Optimal Control, Mathematical Education and Numerical Linear Algebra.

Saeid Abbasbandy is Professor at the Department of Mathematics, Imam Khomeini International University, Qazvin, Iran. He received a Master of Science degree and a PhD from Kharazmi University. His researches encompass numerical analysis, homotopy analysis method, reproducing kernel Hilbert space method.