# Journal of Nonlinear Sciences and Applications

# A version of the Stone-Weierstrass theorem in fuzzy analysis

Juan J. Font*, Delia Sanchis, Manuel Sanchis

*Departament de Matemàtiques and Institut Universitari de Matemàtiques i Aplicacions de Castelló (IMAC), Universitat Jaume I, Avda. Sos Baynat s/n, 12071, Castelló (Spain).*

Communicated by C. Zaharia

## Abstract

Let $C(K, \mathbb{E}^1)$ be the space of continuous functions defined between a compact Hausdorff space $K$ and the space of fuzzy numbers $\mathbb{E}^1$ endowed with the supremum metric. We provide a set of sufficient conditions on a subspace of $C(K, \mathbb{E}^1)$ in order that it be dense. We also obtain a similar result for interpolating families of $C(K, \mathbb{E}^1)$. As a corollary of the above results we prove that certain fuzzy-number-valued neural networks can approximate any continuous fuzzy-number-valued function defined on a compact subspace of $\mathbb{R}^n$. ©2017 All rights reserved.

Keywords: Stone-Weierstrass theorem, fuzzy numbers, fuzzy-number-valued continuous functions.
*2010 MSC:* 41A30, 26E50.

## 1. Introduction

Fuzzy numbers provide formalized tools to deal with non-precise quantities. They are indeed fuzzy sets in the real line and were introduced in 1978 by Dubois and Prade [3], who also defined their basic operations. Since then, fuzzy analysis has developed based on the notion of fuzzy number just as much as classical real analysis did based on the concept of real number. Such development was eased by a characterization of fuzzy numbers provided in 1986 by Goetschel and Voxman [4] leaning on their level sets.

As real-valued functions do in the classical setting, fuzzy-number-valued functions, that is, functions defined on a topological space taking values in the space of fuzzy numbers, play a central role in fuzzy analysis. Namely, fuzzy-number-valued functions have become the main tool in several fuzzy contexts, such as fuzzy differential equations [1], fuzzy integrals [11] or fuzzy optimization [5]. However the main difficulty of dealing with these functions is the fact that the space they form is not a linear space. Indeed it is not a group with respect to addition.

In this paper we focus on the conditions under which continuous (with respect to the supremum metric) fuzzy-number-valued functions defined on a compact Hausdorff space can be (uniformly) approximated to any degree of accuracy. More precisely and based on ideas of Jewett [7] and Prolla [10],

---

*Corresponding author
Email addresses: font@uji.es (Juan J. Font), dsanchis@uji.es (Delia Sanchis), sanchis@uji.es (Manuel Sanchis)

we provide a set of sufficient conditions on a subspace of the space of fuzzy-number-valued functions in order that it be dense, which is to say a Stone-Weierstrass type result. The celebrated Stone-Weierstrass theorem is one of the most important results in classical analysis, plays a key role in the development of general approximation theory and, particularly, is in the essence of the approximation capabilities of neural networks. We also obtain a similar result for interpolating families of continuous fuzzy-number-valued functions in the sense that the uniform approximation can also demand exact agreement at any finite number of points.

As a corollary of the above results we prove that fuzzy-number-valued four-layer regular neural networks can approximate any continuous fuzzy-number-valued function defined on a compact subspace of $\mathbb{R}^n$.

## 2. Preliminaries

Let $F(\mathbb{R})$ denote the family of all fuzzy subsets on the real numbers $\mathbb{R}$. For $u \in F(\mathbb{R})$ and $\lambda \in [0, 1]$, the $\lambda$-level set of $u$ is defined by

$$[u]^\lambda := \{x \in \mathbb{R} : u(x) \geqslant \lambda\}, \quad \lambda \in ]0, 1], \quad [u]^0 := cl_{\mathbb{R}}\{x \in \mathbb{R} : u(x) > 0\}.$$

The fuzzy number space $\mathbb{E}^1$ is the set of elements $u$ of $F(\mathbb{R})$ satisfying the following properties:

(1) $u$ is normal, i.e., there exists an $x_0 \in \mathbb{R}$ with $u(x_0) = 1$;
(2) $u$ is convex, i.e., $u(\lambda x + (1 - \lambda)y) \geqslant \min\{u(x), u(y)\}$ for all $x, y \in \mathbb{R}, \lambda \in [0, 1]$;
(3) $u$ is upper-semicontinuous;
(4) $[u]^0$ is a compact set in $\mathbb{R}$.

Notice that if $u \in \mathbb{E}^1$, then the $\lambda$-level set $[u]^\lambda$ of $u$ is a compact interval for each $\lambda \in [0, 1]$. We denote $[u]^\lambda = [u^-(\lambda), u^+(\lambda)]$. Every real number $r$ can be considered a fuzzy number since $r$ can be identified with the fuzzy number $\tilde{r}$ defined as

$$\tilde{r}(t) := \begin{cases} 1 & \text{if } t = r, \\ 0 & \text{if } t \neq r. \end{cases}$$

We can now state the characterization of fuzzy numbers provided by Goetschel and Voxman [4].

**Theorem 2.1.** *Let $u \in \mathbb{E}^1$ and $[u]^\lambda = [u^-(\lambda), u^+(\lambda)], \lambda \in [0, 1]$. Then the pair of functions $u^-(\lambda)$ and $u^+(\lambda)$ has the following properties:*

  (i) *$u^-(\lambda)$ is a bounded left continuous nondecreasing function on $(0, 1]$;*
 (ii) *$u^+(\lambda)$ is a bounded left continuous nonincreasing function on $(0, 1]$;*
(iii) *$u^-(\lambda)$ and $u^+(\lambda)$ are right continuous at $\lambda = 0$;*
(iv) *$u^-(1) \leqslant u^+(1)$.*

*Conversely, if a pair of functions $\alpha(\lambda)$ and $\beta(\lambda)$ satisfy the above conditions (i)–(iv), then there exists a unique $u \in \mathbb{E}^1$ such that $[u]^\lambda = [\alpha(\lambda), \beta(\lambda)]$ for each $\lambda \in [0, 1]$.*

Given $u, v \in \mathbb{E}^1$ and $k \in \mathbb{R}$, we can define

$$u + v := [u^-(\lambda), u^+(\lambda)] + [v^-(\lambda), v^+(\lambda)],$$

and

$$ku := k[u^-(\lambda), u^+(\lambda)].$$

It is well-known that $\mathbb{E}^1$ endowed with these two natural operations is not a vector space. Indeed $(\mathbb{E}^1, +)$ is not a group.

On the other hand, we can endow $\mathbb{E}^1$ with the following metric.

**Definition 2.2** ([2, 4]). For $u, v \in \mathbb{E}^1$, we can define

$$d_\infty(u, v) := \sup_{\lambda \in [0,1]} \max \left\{ |u^-(\lambda) - v^-(\lambda)|, |u^+(\lambda) - v^+(\lambda)| \right\}.$$

It is called the supremum metric on $\mathbb{E}^1$, and $(\mathbb{E}^1, d_\infty)$ is well-known to be a complete metric space. Notice that, by the definition of $d_\infty$, $\mathbb{R}$ endowed with the Euclidean topology can be topologically identified with the closed subspace $\tilde{\mathbb{R}} = \{\tilde{x} : x \in \mathbb{R}\}$ of $(\mathbb{E}^1, d_\infty)$ where $\tilde{x}^+(\lambda) = \tilde{x}^-(\lambda) = x$ for all $\lambda \in [0,1]$. As a metric space, we shall always consider $\mathbb{E}^1$ equipped with the metric $d_\infty$.

**Proposition 2.3.** *The metric $d_\infty$ satisfies the following properties:*

1. $d_\infty(\sum_{i=1}^m u_i, \sum_{i=1}^m v_i) \leqslant \sum_{i=1}^m d_\infty(u_i, v_i)$ *where $u_i, v_i \in \mathbb{E}^1$ for $i = 1, \cdots, m$.*
2. $d_\infty(ku, kv) = k d_\infty(u, v)$ *where $u, v \in \mathbb{E}^1$ and $k > 0$.*
3. $d_\infty(ku, \mu u) = |k - \mu| \, d_\infty(u, 0)$, *where $u \in \mathbb{E}^1$, $k \geqslant 0$ and $\mu \geqslant 0$.*
4. $d_\infty(ku, \mu v) \leqslant |k - \mu| \, d_\infty(u, 0) + \mu d_\infty(u, v)$, *where $u, v \in \mathbb{E}^1$, $k \geqslant 0$ and $\mu \geqslant 0$.*

*Proof.* The proofs of (1) and (2) can be found, for example, in [2].

In order to prove (3), let us assume that $\mu < k$ and rewrite $ku$ and $\mu u$ as

$$ku = (\mu + (k - \mu))u = \mu u + (k - \mu)u,$$

and

$$\mu u = \mu u + (k - \mu)0.$$

By (1), we know that

$$d_\infty(ku, \mu u) \leqslant d_\infty(\mu u, \mu u) + d_\infty((k - \mu)u, (k - \mu)0) = |k - \mu| \, d_\infty(u, 0).$$

Consequently,

$$d_\infty(ku, \mu v) \leqslant d_\infty(ku, \mu u) + d_\infty(\mu u, \mu v) \leqslant |k - \mu| \, d_\infty(u, 0) + \mu d_\infty(u, v).$$

$\square$

We shall denote by $C(K, \mathbb{E}^1)$ the space of continuous functions defined between the compact Hausdorff space $K$ and the metric space $(\mathbb{E}^1, d_\infty)$. In $C(K, \mathbb{E}^1)$ we shall consider the following metric:

$$D(f, g) = \sup_{t \in K} d_\infty(f(t), g(t)),$$

which induces the uniform convergence topology on $C(K, \mathbb{E}^1)$.

**Proposition 2.4.** *Let $\phi \in C(K, \mathbb{R}^+)$ and $f \in C(K, \mathbb{E}^1)$. Then the function $k \mapsto \phi(k)f(k)$, $k \in K$, belongs to $C(K, \mathbb{E}^1)$.*

*Proof.* It is apparent that the pointwise product $\phi(k)f(k)$ is well-defined since $\phi(k) > 0$ for all $k \in K$. Let us see that it is continuous. Choose $s, t \in K$. By Proposition 2.3 (4), we know that

$$d_\infty(\phi(t) \cdot f(t), \phi(s) \cdot f(s)) \leqslant |\phi(t) - \phi(s)| \, d_\infty(f(t), 0) + \phi(s) d_\infty(f(t), f(s)).$$

Since $\phi$ and $f$ are continuous, the proof follows from the well-known fact that every continuous function defined between a compact space and a metric space is bounded. $\square$

## 3. A version of the Stone-Weierstrass theorem in fuzzy analysis

Let us first introduce a basic tool to obtain our main theorem (Theorem 3.7).

**Definition 3.1.** Let $W$ be a nonempty subset of $C(K, \mathbb{E}^1)$. We define

$$\text{Conv}(W) = \{\varphi \in C(K, [0,1]) : \varphi f + (1-\varphi)g \in W \text{ for all } f, g \in W\}.$$

**Proposition 3.2.** *Let $W$ be a nonempty subset of $C(K, \mathbb{E}^1)$. Then we have:*

1. $\phi \in \text{Conv}(W)$ *implies that* $1 - \phi \in \text{Conv}(W)$.
2. *If* $\phi, \varphi \in \text{Conv}(W)$, *then* $\phi \cdot \varphi \in \text{Conv}(W)$.
3. *If* $\phi$ *belongs to the uniform closure of* $\text{Conv}(W)$, *then so does* $1 - \phi$.
4. *If* $\phi, \varphi$ *belong to the uniform closure of* $\text{Conv}(W)$, *then so does* $\phi \cdot \varphi$.

*Proof.* (1) is clear. To see (2), let us suppose that $\phi, \varphi \in \text{Conv}(W)$. The identity

$$1 - \phi \cdot \varphi = (1 - \phi) + \phi(1 - \varphi),$$

implies, for every pair $f, g \in W$, that

$$(\phi \cdot \varphi)f + (1 - \phi \cdot \varphi)g = \phi[\varphi f + (1 - \varphi)g] + (1 - \phi)g,$$

and, consequently, $\phi \cdot \varphi \in \text{Conv}(W)$.

In order to prove (3), let us suppose that $\phi$ belongs to the uniform closure of $\text{Conv}(W)$. Then there exists a sequence $\{\phi_n\} \subset \text{Conv}(W)$ converging to $\phi$. Hence, $\{1 - \phi_n\}$, which is contained in $\text{Conv}(W)$ by (1), converges uniformly to $1 - \phi$.

Finally, (4) can be proved similarly.

$\square$

**Definition 3.3.** It is said that $M \subset C(K, [0,1])$ separates the points of $K$ if given $s, t \in K$, there exists $\phi \in M$ such that $\phi(s) \neq \phi(t)$.

Next we state two technical lemmas which will be used in the sequel.

**Lemma 3.4** ([7, Lemma 2]). *Let $0 < a < b < 1$ and $0 < \delta < \frac{1}{2}$. There exists a polynomial $p(x) = (1 - x^m)^n$ such that*

1. $p(x) > 1 - \delta$ *for all* $0 \leqslant x \leqslant a$;
2. $p(x) < \delta$ *for all* $b \leqslant x \leqslant 1$.

**Lemma 3.5** ([7, Theorem 1]). *Let $W \subset C(K, \mathbb{E}^1)$. The maximum of two elements of $\text{Conv}(W)$ belongs to the uniform closure of $\text{Conv}(W)$.*

**Lemma 3.6.** *Let $W \subseteq C(K, \mathbb{E}^1)$. If $\text{Conv}(W)$ separates the points of $K$, then, given $x_0 \in K$ and an open neighborhood $N$ of $x_0$, there exists a neighborhood $U$ of $x_0$ contained in $N$ such that for all $0 < \delta < \frac{1}{2}$, there is $\varphi \in \text{Conv}(W)$ such that*

1. $\varphi(t) > 1 - \delta$, *for all* $t \in U$;
2. $\varphi(t) < \delta$, *for all* $t \notin N$.

*Proof.* Let $H$ be the complement of $N$ in $K$. For each $t \in H$ and since $\text{Conv}(W)$ separates the points of $K$ by hypothesis, we can assume, with no loss of generality, that there is a $\varphi_t \in \text{Conv}(W)$ such that $\varphi_t(t) < \varphi_t(x_0)$.

Let us choose two real numbers $a$ and $b$ such that $\varphi_t(t) < a < b < \varphi_t(x_0)$. Then, taking $\delta = \frac{1}{4}$ in Lemma 3.4, we have a polynomial $p_t(x) = (1 - x^m)^n$ such that $p_t(x) < \frac{1}{4}$ for $b \leqslant x \leqslant 1$, and $p_t(x) > \frac{3}{4}$ for $0 \leqslant x \leqslant a$. Hence, $p_t(\varphi_t(x_0)) < \frac{1}{4}$ and $p_t(\varphi_t(t)) > \frac{3}{4}$.

Hence, for every $t \in H$, we can define $U(t) = \{s \in K : p_t(\varphi_t(s)) > \frac{3}{4}\}$, which is clearly an open neighborhood of $t$. Since $H$ is compact, there exist $t_1, \cdots, t_m \in H$ such that $H \subset U(t_1) \cup U(t_2) \cup \cdots \cup U(t_m)$. For each $i = 1, \cdots, m$ and all $s \in K$ we can define

$$\varphi_i(s) = p_{t_i}(\varphi_{t_i}(s)).$$

Let us mention that $\varphi_i \in \text{Conv}(W)$ for all $i = 1, \cdots, m$. Indeed, we have $p_{t_i}(\varphi_{t_i}(s)) = (1 - [\varphi_{t_i}(s)]^m)^n$ and since $\varphi_{t_i}(s) \in \text{Conv}(W)$, we infer that so is $p_{t_i}(\varphi_{t_i})$ by Proposition 3.2, (1) and (2).

Let us define $\psi(s) = \max(\varphi_1(s), \cdots, \varphi_m(s))$, $s \in K$. By Lemma 3.5, we know that $\psi$ belongs to the uniform closure of $\text{Conv}(W)$. We remark that $\psi(x_0) < \frac{1}{4}$ and $\psi(t) > \frac{3}{4}$, for all $t \in H$ due to the properties of the polynomials $p_t(x)$. Let us define $U = \{s \in S; \psi(s) < \frac{1}{4}\}$. Clearly, $U$ is an open neighborhood of $x_0$ in $K$. We claim that $U$ is contained in $N$. Indeed, if $t \in U$ and $t \notin N$, then $t \in H$ and, consequently, $\psi(t) > \frac{3}{4}$, that is, it cannot be in $U$.

Let us choose $\delta$ satisfying $0 < \delta < \frac{1}{2}$ and let $p$ be the polynomial defined by Lemma 3.4, applied to $a = \frac{1}{4}$, $b = \frac{3}{4}$ and $\delta/2$. Define $\eta(s) = p(\psi(s))$, for $s \in K$. By Proposition 3.2, (3) and (4), the function $\eta$ belongs also to the uniform closure of $\text{Conv}(W)$.

If $t \in U$, then $\eta(t) > 1 - \delta/2$ by construction. If $t \notin N$, then $t \in H$ and $\eta(t) < \delta/2$ since $\psi(t) > \frac{3}{4}$.

As $\eta$ belongs to the uniform closure of $\text{Conv}(W)$, given $\frac{\delta}{2}$, there exists $\varphi \in \text{Conv}(W)$ such that $\| \varphi - \eta \|_\infty = \sup_{t \in K} | \varphi(t) - \eta(t) | < \delta/2$. Consequently, $\varphi$ satisfies the desired properties. $\qquad\square$

Gathering the information obtained so far, we can now state and prove a version of the Stone-Weierstrass theorem for fuzzy-number-valued continuous functions.

**Theorem 3.7.** *Let $W$ be a nonempty subset of $C(K, \mathbb{E}^1)$ and assume that $\text{Conv}(W)$ separates points. If given $f \in C(K, \mathbb{E}^1)$ and $\varepsilon > 0$, there exists, for each $x \in K$, $g_x \in W$ such that $d_\infty(f(x), g_x(x)) < \varepsilon$, then $W$ is dense in $(C(K, \mathbb{E}^1), D)$.*

*Proof.* Fix $\varepsilon > 0$ and $f \in C(K, \mathbb{E}^1)$. Our goal is to find $g \in W$ such that $D(f, g) < \varepsilon$. For each $x \in K$ and $0 < \varepsilon(x) < \varepsilon$, let us define $N(x) = \{t \in K : d_\infty(f(t), g_x(t)) < \varepsilon(x) < \varepsilon\}$, which is an open neighborhood of $x$. Let $U(x)$ be an open neighborhood of $x$ satisfying the properties in Lemma 3.6. Fix $x_1 \in K$ and let $S$ be the complement of $N(x_1)$ in $K$. By the compacity of $S$, there is a finite set $\{x_2, \cdots, x_m\} \subset S$ such that $S \subset U(x_2) \cup \cdots \cup U(x_m)$. Let us choose $0 < \delta < \frac{1}{2}$ such that $\delta Qm < \varepsilon - \varepsilon'$, where

$$\varepsilon' = \max\{\varepsilon(x_i) : 1 \leqslant i \leqslant m\},$$

and

$$Q = \max\{D(f, g_{x_i}) : 1 \leqslant i \leqslant m\}.$$

By Lemma 3.6, we know that there exist $\phi_2, \cdots, \phi_m \in \text{Conv}(W)$ such that for all $i = 2, \cdots, m$,

(i) $\phi_i(t) > 1 - \delta$, for all $t \in U(x_i)$;
(ii) $0 \leqslant \phi_i(t) < \delta$, if $t \notin N(x_i)$.

Let us define the functions

$$\psi_2 := \phi_2,$$
$$\psi_3 := (1 - \phi_2)\phi_3,$$
$$\vdots$$
$$\psi_m := (1 - \phi_2)(1 - \phi_3) \cdots (1 - \phi_{m-1})\phi_m.$$

It is apparent that $\psi_i \in \text{Conv}(W)$ for all $i = 2, \cdots, m$. Next we claim that

$$\psi_2 + \cdots + \psi_j = 1 - (1 - \phi_2)(1 - \phi_3) \cdots (1 - \phi_j),$$

$j = 2, \cdots, m$. Indeed, it is clear that

$$\psi_2 + \psi_3 = \phi_2 + (1 - \phi_2)\phi_3 = 1 - (1 - \phi_2) \cdot (1 - \phi_3).$$

We proceed by induction. Assume that the result is true for a certain $j \in \{4, \cdots, m-1\}$ and let us check

$$\psi_2 + \cdots + \psi_j + \psi_{j+1} = 1 - (1 - \phi_2)(1 - \phi_3) \cdots (1 - \phi_j)(1 - \phi_{j+1}).$$

Namely,

$$\psi_2 + \cdots + \psi_j + \psi_{j+1} = 1 - (1 - \phi_2)(1 - \phi_3) \cdots (1 - \phi_j) + (1 - \phi_2)(1 - \phi_3) \cdots (1 - \phi_j)\phi_{j+1}$$
$$= 1 - (1 - \phi_2)(1 - \phi_3) \cdots (1 - \phi_j)(1 - \phi_{j+1}),$$

as was to be checked. Finally, we can define $\psi_1 := (1 - \phi_2) \cdots (1 - \phi_m)$, which belongs to $\text{Conv}(W)$ and satisfies $\psi_1 + \psi_2 + \cdots + \psi_m = 1$.

On the other hand, we claim that

$$\psi_i(t) < \delta, \quad \forall\, t \notin N(x_i), \quad i = 1, \cdots, m. \tag{3.1}$$

Indeed, if $i \geqslant 2$, then $\psi_i(t) \leqslant \phi_i(t)$ by definition, which is to say, $\psi_i(t) < \delta$ for all $t \notin N(x_i)$ from (ii) above. If $t \notin N(x_1)$, we have $t \in S$. Hence, $t \in U(x_j)$ for some $j = 2, \cdots, m$. From (i), we have $\psi_1(t) = 1 - \phi_j(t) < \delta$ and then

$$\psi_i(t) = (1 - \phi_j(t)) \prod_{i \neq j}(1 - \phi_i(t)) < \delta.$$

Let

$$g := \psi_1 g_1 + \psi_2 g_2 + \cdots + \psi_m g_m, \tag{3.2}$$

where $g_i$ stands for $g_{x_i}$ for $i = 1, \cdots, m$. A routine manipulation yields

$$g = \phi_2 g_2 + (1 - \phi_2)[\phi_3 g_3 + (1 - \phi_3)[\phi_4 g_4 + \cdots + (1 - \phi_{m-1})[\phi_m g_m + (1 - \phi_m)g_1] \cdots]].$$

Therefore, $g \in W$ since $\phi_i \in \text{Conv}(W)$ for $i = 2, \cdots, m$ (see Definition 3.1). From Proposition 2.3, we know that, given $x_0 \in K$,

$$d_\infty(f(x_0), g(x_0)) = d_\infty\left(\sum_{i=1}^{m} \psi_i(x_0)f(x_0), \sum_{i=1}^{m} \psi_i(x_0)g_i(x_0)\right) \leqslant \sum_{i=1}^{m} \psi_i(x_0)d_\infty(f(x_0), g_i(x_0)).$$

Let $I = \{1 \leqslant i \leqslant m : x_0 \in N(x_i)\}$ and $J = \{1 \leqslant i \leqslant m : x_0 \notin N(x_i)\}$. Then, for all $i \in I$, we have

$$\psi_i(x_0)d_\infty(f(x_0), g_i(x_0)) \leqslant \psi_i(x_0)\varepsilon',$$

and, for all $i \in J$, the inequality (3.1) yields

$$\psi_i(x_0)d_\infty(f(x_0), g_i(x_0)) \leqslant \delta Q.$$

From these two inequalities, we deduce

$$\sum_{i=1}^{m} \psi_i(x_0)d_\infty(f(x_0), g_i(x_0)) \leqslant \sum_{i \in I} \psi_i(x_0)\varepsilon' + \sum_{i \in J} \delta Q \leqslant \varepsilon' + \delta Q m < \varepsilon.$$

Finally, gathering all the information and since $x_0$ is arbitrary in $K$, we infer $d_\infty(f(x), g(x)) < \varepsilon$ for all $x \in K$, which yields $D(f, g) \leqslant \varepsilon$ and we are done.

$\square$

Given $u \in \mathbb{E}^1$, we shall write $\widehat{U}$ to denote the function in $C(K, \mathbb{E}^1)$ which takes the constant value $u$. If we assume that $W$ contains the constant functions, then we can obtain the following improved version of Theorem 3.7.

**Corollary 3.8.** *Let $W$ be a nonempty subset $C(K, \mathbb{E}^1)$ that contains the constant functions. If we further assume that $\mathrm{Conv}(W)$ separates points, then $W$ is dense in $(C(K, \mathbb{E}^1), D)$.*

*Proof.* Given $f \in C(K, \mathbb{E}^1)$ and $\varepsilon > 0$, we can take, for each $x \in K$, $g_x := \widehat{f(x)}$. In this case $d_\infty(f(x), g_x(x)) = 0$ and the result follows from Theorem 3.7. $\square$

**Corollary 3.9.** *Let $f \in C(K, \mathbb{E}^1)$. Then there exist finitely many functions $\psi_i \in C(K, [0,1])$ and $u_i \in \mathbb{E}^1$, $i = 1, \cdots, m$, such that*
$$D(f, \psi_1 \widehat{U_1} + \cdots + \psi_m \widehat{U_m}) < \varepsilon.$$

*Proof.* Let us first remark that $\mathrm{Conv}(C(K, \mathbb{E}^1)) = C(K, [0,1])$, which clearly separates the points of $K$. Furthermore, in the definition of the function $g$ in the proof of Theorem 3.7 (see (3.2)), we can take $g_{x_i} := \widehat{f(x_i)}$ for $i = 1, \cdots, m$. $\square$

**Definition 3.10.** A subset $A \subset C(K, \mathbb{E}^1)$ is said to be an interpolating family, if given any finite subset $K' \subset K$ and any $f \in C(K, \mathbb{E}^1)$, there exists $g \in A$ such that $f(x) = g(x)$ for all $x \in K'$.

**Theorem 3.11.** *Let $A \subset C(K, \mathbb{E}^1)$ be an interpolating family such that $\mathrm{Conv}(A)$ separates the points of $K$. Then, for every $f \in C(K, \mathbb{E}^1)$, every $\varepsilon > 0$ and every finite subset $K' \subset K$, there exists $g \in A$ such that $D(f, g) < \varepsilon$ and $f(x) = g(x)$ for all $x \in K'$. In particular, $A$ is uniformly dense in $(C(K, \mathbb{E}^1), D)$.*

*Proof.* Fix $f_0 \in C(K, \mathbb{E}^1)$ and define $W = \{g \in A : f_0(x) = g(x) \text{ for all } x \in K'\}$, which is not empty since $A$ is a interpolating family.

Let us see that $\mathrm{Conv}(A) \subseteq \mathrm{Conv}(W)$. Let $f, g \in W$. We need to check that $\varphi f + (1 - \varphi)g \in W$ for any $\varphi \in \mathrm{Conv}(A)$. It is apparent that $\varphi f + (1 - \varphi)g \in A$. Furthermore

$$\varphi(x)f(x) + (1 - \varphi(x))g(x) = \varphi(x)f_0(x) + (1 - \varphi)f_0(x) = f_0(x)$$

for all $x \in K'$ and consequently, $\varphi f + (1 - \varphi)g \in W$.

For each $x \in K$, consider the finite set $K' \cup \{x\}$. Since $A$ is an interpolating family for $C(K, \mathbb{E}^1)$, there exists $g_x \in A$ such that $f(t) = g_x(t)$ for all $t \in K' \cup \{x\}$. In particular, $f(t) = g_x(t)$ for all $t \in K'$. Hence $g_x \in W$. On the other hand $f(x) = g_x(x)$ implies $d_\infty(f(x), g_x(x)) < \varepsilon$. By Theorem 3.7, there exists $g \in W \subseteq A$ such that $D(f, g) < \varepsilon$ and $g(t) = f(t)$ for all $t \in K'$. $\square$

## 4. Approximation by means of fuzzy-number-valued regular neural networks

Let us first recall (see, e.g., [6, 9]) that fuzzy-number-valued four-layer regular neural networks (four-layer RFNNs) are defined by

$$H(x) = \sum_{i=1}^{m} u_i \left( \sum_{j=1}^{s} w_{ij} \cdot \sigma(x \cdot a_j + \theta_j) \right)$$

for each $x \in \mathbb{R}$, where $u_i \in \mathbb{E}^1$, the coefficients $w_{ij}$, the weights $a_j$ and the thresholds $\theta_j$ are real numbers and $\sigma : \mathbb{R} \to \mathbb{R}$ stands for the activation function in the hidden layer. That is, $H$ is a $\mathbb{R}$-based fuzzy-number-valued function. In [9], the author proved that, although three-layer RFNNs cannot approximate the set of all continuous $\mathbb{R}$-based fuzzy-number-valued functions, four-layer RFNNs can. He used sigmoid or bounded continuous nonconstant activation functions to achieve such approximation property of four-layer RFNNs.

In this section we deal with $\mathbb{R}^n$-based fuzzy-number-valued four-layer RFNNs and, therefore, both the variable $x$ and the weights $a_j$ belong to $\mathbb{R}^n$. Given an activation function $\sigma$, we denote the set of all $\mathbb{R}^n$-based fuzzy-number-valued four-layer RFNNs by $\mathcal{H}(\sigma)$.

We will show, based on the results in the previous section, that the elements of $\mathcal{H}(\sigma)$ can approximate any continuous fuzzy-number-valued function defined on a compact subspace of $\mathbb{R}^n$ provided the activation function $\sigma$ be a non-polynomial continuous function.

**Theorem 4.1.** *Assume that* $\sigma : \mathbb{R} \to \mathbb{R}$ *is a non-polynomial continuous function and that* $K$ *is a compact subset of* $\mathbb{R}^n$. *Then, given any* $f \in C(K, \mathbb{E}^1)$ *and* $\varepsilon > 0$, *there exists a neural network* $H \in \mathcal{H}(\sigma)$ *such that* $D(f, H) < \varepsilon$.

*Proof.* By Corollary 3.9, there exist finitely many functions $\psi_i \in C(K, [0,1])$ and $u_i \in \mathbb{E}^1$, $i = 1, \cdots, m$, such that

$$D(f, \psi_1 \widehat{U_1} + \cdots + \psi_m \widehat{U_m}) < \frac{\varepsilon}{2}.$$

On the other hand, by [8], we know that for each $\psi_i$, $i = 1, \cdots, m$, there exist $w_{ij}, \theta_j \in \mathbb{R}$ and $a_j \in \mathbb{R}^n$ such that

$$\left| \psi_i(x) - \sum_{j=1}^{s} w_{ij} \cdot \sigma(x \cdot a_j + \theta_j) \right| < \frac{\varepsilon}{2m \cdot d_\infty(u_i, 0)}$$

for all $x \in K$. Hence

$$D\left( \left( \sum_{j=1}^{s} w_{ij} \cdot \sigma(x \cdot a_j + \theta_j) \right) \widehat{U_i}, \ \psi_i(x) \cdot \widehat{U_i} \right)$$

$$= \sup_{x \in K} d_\infty \left( \left( \sum_{j=1}^{s} w_{ij} \cdot \sigma(x \cdot a_j + \theta_j) \right) u_i, \ \psi_i(x) \cdot u_i \right)$$

$$= \sup_{x \in K} \left| \psi_i(x) - \left( \sum_{j=1}^{s} w_{ij} \cdot \sigma(x \cdot a_j + \theta_j) \right) \right| \sup_{\lambda \in [0,1]} \max\{| u_i^-(\lambda) |, | u_i^+(\lambda) |\}$$

$$= \sup_{x \in K} \left| \psi_i(x) - \left( \sum_{j=1}^{s} w_{ij} \cdot \sigma(x \cdot a_j + \theta_j) \right) \right| d_\infty(u_i, 0) < \frac{\varepsilon}{2m},$$

which clearly yields

$$D\left( f, \ \sum_{i=1}^{m} u_i \left( \sum_{j=1}^{s} w_{ij} \cdot \sigma((\cdot) a_j + \theta_j) \right) \right) < \varepsilon.$$

□

## 5. Conclusion

We have proved that, under certain natural assumptions, continuous (with respect to the supremum metric) fuzzy-number-valued functions defined on a compact Hausdorff space can be (uniformly) approximated to any degree of accuracy, which yields a Stone-Weierstrass type result in this setting. A similar result for interpolating families of continuous fuzzy-number-valued functions in the sense that the uniform approximation can also demand exact agreement at any finite number of points. As a corollary, we infer that fuzzy-number-valued four-layer regular neural networks can approximate any continuous fuzzy-number-valued function defined on a compact subspace of $\mathbb{R}^n$.

## Acknowledgment

# References

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

[2] P. Diamond, P. Kloeden, *Metric Spaces of Fuzzy Sets: Theory and Applications*, World Scientific, Singapore, (1994). 2.2, 2

[3] D. Dubois, H. Prade, *Operations on fuzzy numbers*, Internat. J. Systems Sci., **9** (1978), 613–626. 1

[4] R. Goetschel, W. Voxman, *Elementary fuzzy calculus*, Fuzzy Sets and Systems, **18** (1986), 31–42. 1, 2, 2.2

[5] S. Hai, Z. Gong, H. Li, *Generalized differentiability for n-dimensional fuzzy-number-valued functions and fuzzy optimization*, Inf. Sci., **374** (2016), 151–163. 1

[6] H. Huang, C. Wu, *Approximation of level fuzzy-valued functions by multilayer regular fuzzy neural networks*, Math. Comput. Modelling, **49** (2009), 1311–1318. 4

[7] R. I. Jewett, *A variation on the Stone-Weierstrass theorem*, Proc. Amer. Math. Soc., **14** (1963), 690–693. 1, 3.4, 3.5

[8] M. Leshno, V. Y. Lin, A. Pinkus, S. Schocken, *Multilayer feedforward networks with a nonpolynomial activation function can approximate any function*, Neural Networks, **6** (1993), 861–867. 4

[9] P. Y. Liu, *Universal approximations of continuous fuzzy-valued function by multi-layer regular fuzzy neural networks*, Fuzzy Sets and Systems, **119** (2001), 313–320. 4

[10] J. B. Prolla, *On the Weierstrass-Stone Theorem*, J. Approx. Theory, **78** (1994), 299–313. 1

[11] C. Wu, Z. Gong, *On Henstock integral of fuzzy-number-valued functions (I)*, Fuzzy Sets and Systems, **120** (2001), 523–532. 1