Degree of Approximation by Kantorovich-Shilkret quasi-interpolation neural network operators revisited

George A. Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

Abstract

In this article we exhibit multivariate basic approximation by a Kantorovich-Shilkret type quasi-interpolation neural network operators with respect to supremum norm. This is done with rates using the multivariate modulus of continuity. We approximate continuous and bounded functions on $\mathbb{R}^N$, $N \in \mathbb{N}$. When they are additionally uniformly continuous we derive pointwise and uniform convergences. We include also the related Complex approximation. Our activation functions are induced by the arctangent, algebraic, Gudermannian and generalized symmetrical sigmoid functions.

2020 AMS Mathematics Subject Classification: 41A17, 41A25, 41A30, 41A35.

Keywords and Phrases: arctangent-algebraic-Gudermannian-generalized symmetrical activation functions, multivariate quasi-interpolation neural network approximation, Kantorovich-Shilkret type operators.

1 Introduction

Here we are motivated by [1].

The author here performs multivariate arctangent-algebraic-Gudermannian-generalized symmetrical activation functions based neural network approximation to continuous functions over $\mathbb{R}^N$, $N \in \mathbb{N}$, and then he extends his results to complex valued functions. The convergences here are with rates expressed via the multivariate modulus of continuity of the involved function and given by very tight Jackson type inequalities.

RGMIA Res. Rep. Coll. **25** (2022), Art. 46, 18 pp. Received 13/04/22

The author comes up with the "right" precisely defined flexible quasi-interpolation Kantorovich-Shilkret type integral coefficient neural network operators associated to the arctangent-algebraic-Gudermannian-generalized symmetrical activation functions. This is a continuation of [3], Chapter 11.

Feed-forward neural network (FNNs) with one hidden layer with deal with, are expressed mathematicaly as

$$N_n(x) = \sum_{j=0}^{n} c_j \sigma\left(\langle a_j \cdot x\rangle + b_j\right), \quad x \in \mathbb{R}^s,\ s \in \mathbb{N},$$

where for $0 \leq j \leq n$, $b_j \in \mathbb{R}$ are the thresholds, $a_j \in \mathbb{R}^s$ are the connection weights, $c_j \in \mathbb{R}$ are the coefficients, $\langle a_j \cdot x\rangle$ is the inner product of a_j and x, and σ is the activation function of the network. In many fundamental neural network models the activation functions are the arctangent- algebraic- Gudermannian-generalized symmetrical activation functions.

About neural networks in general you may read [12], [13], [14]. In recent years non-additive integrals, like the N. Shilkret one [15], have become fashionable and more useful in Economic theory, etc.

2 Background

Here we follow [15].

Let $\mathcal{F}$ be a σ-field of subsets of an arbitrary set Ω. An extended non-negative real valued function μ on $\mathcal{F}$ is called maxitive if $\mu(\varnothing) = 0$ and

$$\mu\left(\cup_{i\in I} E_i\right) = \sup_{i\in I} \mu\left(E_i\right), \tag{1}$$

where the set I is of cardinality at most countable. We also call μ a maxitive measure. Here f stands for a non-negative measurable function on Ω. In [15], Niel Shilkret developed his non-additive integral defined as follows:

$$(N^*)\int_D f d\mu := \sup_{y\in Y}\left\{y \cdot \mu\left(D \cap \{f \geq y\}\right)\right\}, \tag{2}$$

where $Y = [0, m]$ or $Y = [0, m)$ with $0 < m \leq \infty$, and $D \in \mathcal{F}$. Here we take $Y = [0, \infty)$.

It is easily proved that

$$(N^*)\int_D f d\mu = \sup_{y>0}\left\{y \cdot \mu\left(D \cap \{f > y\}\right)\right\}. \tag{3}$$

The Shilkret integral takes values in $[0, \infty]$.

The Shilkret integral ([15]) has the following properties:

$$(N^*)\int_\Omega \chi_E d\mu = \mu(E), \tag{4}$$

where χ_E is the indicator function on $E \in \mathcal{F}$,

$$(N*) \int_D cf d\mu = c(N^*) \int_D f d\mu, \quad c \geq 0, \tag{5}$$

$$(N^*) \int_D \sup_{n \in \mathbb{N}} f_n d\mu = \sup_{n \in \mathbb{N}} (N^*) \int_D f_n d\mu, \tag{6}$$

where f_n, $n \in \mathbb{N}$, is an increasing sequence of elementary (countably valued) functions converging uniformly to f. Furthermore we have

$$(N^*) \int_D f d\mu \geq 0, \tag{7}$$

$$f \geq g \text{ implies } (N^*) \int_D f d\mu \geq (N^*) \int_D g d\mu, \tag{8}$$

where $f, g : \Omega \to [0, \infty]$ are measurable.

Let $a \leq f(\omega) \leq b$ for almost every $\omega \in E$, then

$$a\mu(E) \leq (N^*) \int_E f d\mu \leq b\mu(E);$$

$$(N^*) \int_E 1 d\mu = \mu(E);$$

$f > 0$ almost everywhere and $(N^*) \int_E f d\mu = 0$ imply $\mu(E) = 0$;

$(N^*) \int_\Omega f d\mu = 0$ if and only $f = 0$ almost everywhere;

$(N^*) \int_\Omega f d\mu < \infty$ implies that

$$\overline{N}(f) := \{\omega \in \Omega | f(\omega) \neq 0\} \text{ has } \sigma\text{-finite measure}; \tag{9}$$

$$(N^*) \int_D (f+g) d\mu \leq (N^*) \int_D f d\mu + (N^*) \int_D g d\mu;$$

and

$$\left| (N^*) \int_D f d\mu - (N^*) \int_D g d\mu \right| \leq (N^*) \int_D |f - g| d\mu. \tag{10}$$

From now on in this article we assume that $\mu : \mathcal{F} \to [0, +\infty)$.

2.1 About the arctangent activation function

We consider the

$$\arctan x = \int_0^x \frac{dz}{1+z^2}, \quad x \in \mathbb{R}. \tag{11}$$

We will be using

$$h(x) := \frac{2}{\pi} \arctan\left(\frac{\pi}{2} x\right) = \frac{2}{\pi} \int_0^{\frac{\pi x}{2}} \frac{dz}{1+z^2}, \quad x \in \mathbb{R}, \tag{12}$$

which is a sigmoid type function and it is strictly increasing. We have that

$$h(0)=0,\ \ h(-x)=-h(x),\ \ h(+\infty)=1,\ \ h(-\infty)=-1,$$

and

$$h'(x)=\frac{4}{4+\pi^2x^2}>0,\ \ \text{all } x\in\mathbb{R}. \tag{13}$$

We consider the activation function

$$\psi_1(x):=\frac{1}{4}\left(h(x+1)-h(x-1)\right),\ \ x\in\mathbb{R}, \tag{14}$$

and we notice that

$$\psi_1(-x)=\psi_1(x), \tag{15}$$

it is an even function.

Since $x+1>x-1$, then $h(x+1)>h(x-1)$, and $\psi_1(x)>0$, all $x\in\mathbb{R}$.

We see that

$$\psi_1(0)=\frac{1}{\pi}\arctan\frac{\pi}{2}\cong 18.31. \tag{16}$$

Let $x>0$, we have that

$$\psi_1'(x)=\frac{1}{4}\left(h'(x+1)-h'(x-1)\right)=$$

$$\frac{-4\pi^2x}{\left(4+\pi^2(x+1)^2\right)\left(4+\pi^2(x-1)^2\right)}<0. \tag{17}$$

That is

$$\psi_1'(x)<0,\ \text{for } x>0. \tag{18}$$

That is ψ_1 is strictly decreasing on $[0,\infty)$ and clearly is strictly increasing on $(-\infty,0]$, and $\psi_1'(0)=0$.

Observe that

$$\begin{aligned}&\lim_{x\to+\infty}\psi_1(x)=\tfrac{1}{4}\left(h(+\infty)-h(+\infty)\right)=0,\\ &\text{and}\\ &\lim_{x\to-\infty}\psi_1(x)=\tfrac{1}{4}\left(h(-\infty)-h(-\infty)\right)=0.\end{aligned} \tag{19}$$

That is the x-axis is the horizontal asymptote on ψ_1.

All in all, ψ_1 is a bell symmetric function with maximum $\psi_1(0)\cong 18.31$.

We need

Theorem 1 *([2], p. 286) We have that*

$$\sum_{i=-\infty}^{\infty}\psi_1(x-i)=1,\ \ \forall\ x\in\mathbb{R}. \tag{20}$$

Theorem 2 *([2], p. 287) It holds*

$$\int_{-\infty}^{\infty} \psi_1(x)\, dx = 1. \tag{21}$$

So that $\psi_1(x)$ is a density function on $\mathbb{R}$.

We mention

Theorem 3 *([2], p. 288) Let $0 < \alpha < 1$, and $n \in \mathbb{N}$ with $n^{1-\alpha} > 2$. It holds*

$$\sum_{\substack{\begin{cases} k = -\infty \\ : |nx - k| \geq n^{1-\alpha} \end{cases}}}^{\infty} \psi_1(nx - k) < \frac{2}{\pi^2 (n^{1-\alpha} - 2)}. \tag{22}$$

We introduce (see [8])

$$Z_1(x_1, ..., x_N) := Z_1(x) := \prod_{i=1}^{N} \psi_1(x_i), \quad x = (x_1, ..., x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{23}$$

It has the properties:

(i) $Z_1(x) > 0, \ \forall\, x \in \mathbb{R}^N$,

(ii)

$$\sum_{k=-\infty}^{\infty} Z_1(x - k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z_1(x_1 - k_1, ..., x_N - k_N) = 1, \tag{24}$$

where $k := (k_1, ..., k_n) \in \mathbb{Z}^N$, $\forall\, x \in \mathbb{R}^N$,

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_1(nx - k) = 1, \tag{25}$$

$\forall\, x \in \mathbb{R}^N$; $n \in \mathbb{N}$,

and

(iv)

$$\int_{\mathbb{R}^N} Z_1(x)\, dx = 1, \tag{26}$$

that is Z_1 is a multivariate density function.

(v) It is clear that

$$\sum_{\substack{\begin{cases} k = -\infty \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta} \end{cases}}}^{\infty} Z_1(nx - k) < \frac{2}{\pi^2 (n^{1-\beta} - 2)} =: c_1(\beta, n), \tag{27}$$

$0 < \beta < 1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$.

Above it is $\|x\|_\infty := \max\{|x_1|, ..., |x_N|\}$, $x \in \mathbb{R}^N$, also set $\infty := (\infty, ..., \infty)$, $-\infty = (-\infty, ... - \infty)$ upon the multivariate context.

2.2 About the algebraic activation function

Here see also [4].

We consider the generator algebraic function

$$\varphi(x) = \frac{x}{\sqrt[2m]{1 + x^{2m}}}, \quad m \in \mathbb{N}, \ x \in \mathbb{R}, \tag{28}$$

which is a sigmoidal type of function and is a strictly increasing function.

We see that $\varphi(-x) = -\varphi(x)$ with $\varphi(0) = 0$. We get that

$$\varphi'(x) = \frac{1}{(1 + x^{2m})^{\frac{2m+1}{2m}}} > 0, \ \ \forall \ x \in \mathbb{R}, \tag{29}$$

proving φ as strictly increasing over $\mathbb{R}, \varphi'(x) = \varphi'(-x)$. We easily find that $\lim_{x \to +\infty} \varphi(x) = 1$, $\varphi(+\infty) = 1$, and $\lim_{x \to -\infty} \varphi(x) = -1$, $\varphi(-\infty) = -1$.

We consider the activation function

$$\psi_2(x) = \frac{1}{4}[\varphi(x+1) - \varphi(x-1)]. \tag{30}$$

Clearly it is $\psi_2(x) = \psi_2(-x)$, $\forall \ x \in \mathbb{R}$, so that ψ_2 is an even function and symmetric with respect to the y-axis. Clealry $\psi_2(x) > 0$, $\forall \ x \in \mathbb{R}$.

Also it is

$$\psi_2(0) = \frac{1}{2 \sqrt[2m]{2}}. \tag{31}$$

By [4], we have that $\psi_2'(x) < 0$ for $x > 0$. That is ψ_2 is strictly decreasing over $(0, +\infty)$.

Clearly, ψ_2 is strictly increasing over $(-\infty, 0)$ and $\psi_2'(0) = 0$.

Furthermore we obtain that

$$\lim_{x \to +\infty} \psi_2(x) = \frac{1}{4}[\varphi(+\infty) - \varphi(+\infty)] = 0, \tag{32}$$

and

$$\lim_{x \to -\infty} \psi_2(x) = \frac{1}{4}[\varphi(-\infty) - \varphi(-\infty)] = 0. \tag{33}$$

That is the x-axis is the horizontal asymptote of ψ_2.

Conclusion, ψ_2 is a bell shape symmetric function with maximum

$$\psi_2(0) = \frac{1}{2 \sqrt[2m]{2}}, \quad m \in \mathbb{N}. \tag{34}$$

We need

Theorem 4 *([4]) We have that*

$$\sum_{i=-\infty}^{\infty} \psi_2(x-i) = 1, \ \ \forall \ x \in \mathbb{R}. \tag{35}$$

Theorem 5 *([4]) It holds*

$$\int_{-\infty}^{\infty} \psi_2(x)\,dx = 1. \tag{36}$$

Theorem 6 *([4]) Let* $0<\alpha<1$*, and* $n \in \mathbb{N}$ *with* $n^{1-\alpha} > 2$*. It holds*

$$\sum_{\begin{cases} k=-\infty \\ : |nx-k| \geq n^{1-\alpha} \end{cases}}^{\infty} \psi_2(nx-k) < \frac{1}{4m\left(n^{1-\alpha}-2\right)^{2m}}, \quad m \in \mathbb{N}. \tag{37}$$

We introduce (see also [9])

$$Z_2(x_1,...,x_N) := Z_2(x) := \prod_{i=1}^{N} \psi_2(x_i), \quad x = (x_1,...,x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{38}$$

It has the properties:

(i) $Z_2(x) > 0, \ \forall \ x \in \mathbb{R}^N$,

(ii)

$$\sum_{k=-\infty}^{\infty} Z_2(x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z_2(x_1-k_1,...,x_N-k_N) = 1, \tag{39}$$

where $k := (k_1,...,k_n) \in \mathbb{Z}^N$, $\forall \ x \in \mathbb{R}^N$,

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_2(nx-k) = 1, \tag{40}$$

$\forall \ x \in \mathbb{R}^N$; $n \in \mathbb{N}$,

and

(iv)

$$\int_{\mathbb{R}^N} Z_2(x)\,dx = 1, \tag{41}$$

that is Z_2 is a multivariate density function.

(v) It is clear that

$$\sum_{\begin{cases} k=-\infty \\ \left\| \frac{k}{n}-x \right\|_{\infty} > \frac{1}{n^{\beta}} \end{cases}}^{\infty} Z_2(nx-k) < \frac{1}{4m\left(n^{1-\beta}-2\right)^{2m}} =: c_2(\beta,n), \tag{42}$$

$0<\beta<1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$, $m \in \mathbb{N}$.

2.3 About the Gudermannian activation function

See also [5], [16].

Here we consider $gd(x)$ the Gudermannian function [16], which is a sigmoid function, as a generator function:

$$\sigma(x) = 2\arctan\left(\tanh\left(\frac{x}{2}\right)\right) = \int_0^x \frac{dt}{\cosh t} =: gd(x), \ x \in \mathbb{R}. \tag{43}$$

Let the normalized generator sigmoid function

$$f(x) := \frac{4}{\pi}\sigma(x) = \frac{4}{\pi}\int_0^x \frac{dt}{\cosh t} = \frac{8}{\pi}\int_0^x \frac{1}{e^t + e^{-t}}dt, \quad x \in \mathbb{R}. \tag{44}$$

Here

$$f'(x) = \frac{4}{\pi \cosh x} > 0, \quad \forall\ x \in \mathbb{R},$$

hence f is strictly increasing on $\mathbb{R}$.

Notice that $\tanh(-x) = -\tanh x$ and $\arctan(-x) = -\arctan x$, $x \in \mathbb{R}$.

So, here the neural network activation function will be:

$$\psi_3(x) = \frac{1}{4}\left[f(x+1) - f(x-1)\right], \ x \in \mathbb{R}. \tag{45}$$

By [5], we get that

$$\psi_3(x) = \psi_3(-x), \quad \forall\ x \in \mathbb{R}, \tag{46}$$

i.e. it is even and symmetric with respect to the y-axis. Here we have $f(+\infty) = 1$, $f(-\infty) = -1$ and $f(0) = 0$. Clearly it is

$$f(-x) = -f(x), \ \forall\ x \in \mathbb{R}, \tag{47}$$

an odd function, symmetric with respect to the origin. Since $x+1 > x-1$, and $f(x+1) > f(x-1)$, we obtain $\psi_3(x) > 0$, $\forall\ x \in \mathbb{R}$.

By [5], we have that

$$\psi_3(0) = \frac{2}{\pi} gd(1) \cong 0.551. \tag{48}$$

By [5] ψ_3 is strictly decreasing on $(0, +\infty)$, and strictly increasing on $(-\infty, 0)$, and $\psi_3'(0) = 0$.

Also we have that

$$\lim_{x \to +\infty} \psi_3(x) = \lim_{x \to -\infty} \psi_3(x) = 0, \tag{49}$$

that is the x-axis is the horizontal asymptote for ψ_3.

Conclusion, ψ_3 is a bell shaped symmetric function with maximum $\psi_3(0) \cong 0.551$.

We need

Theorem 7 *([5]) It holds that*

$$\sum_{i=-\infty}^{\infty} \psi_3 (x - i) = 1, \;\; \forall \; x \in \mathbb{R}. \tag{50}$$

Theorem 8 *([5]) We have that*

$$\int_{-\infty}^{\infty} \psi_3 (x) \, dx = 1. \tag{51}$$

So $\psi_3 (x)$ is a density function.

Theorem 9 *([5]) Let $0 < \alpha < 1$, and $n \in \mathbb{N}$ with $n^{1-\alpha} > 2$. It holds*

$$\sum_{\begin{cases} k = -\infty \\ : |nx - k| \geq n^{1-\alpha} \end{cases}}^{\infty} \psi_3 (nx - k) < \frac{1}{\pi e^{(n^{1-\alpha} - 2)}} = \frac{4e^2}{\pi e^{n^{1-\alpha}}}. \tag{52}$$

We introduce (see also [7])

$$Z_3 (x_1, ..., x_N) := Z_3 (x) := \prod_{i=1}^{N} \psi_3 (x_i), \quad x = (x_1, ..., x_N) \in \mathbb{R}^N, \; N \in \mathbb{N}. \tag{53}$$

It has the properties:

(i) $Z_3 (x) > 0, \;\; \forall \; x \in \mathbb{R}^N$,

(ii)

$$\sum_{k=-\infty}^{\infty} Z_3 (x - k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z_3 (x_1 - k_1, ..., x_N - k_N) = 1, \tag{54}$$

where $k := (k_1, ..., k_n) \in \mathbb{Z}^N$, $\forall \; x \in \mathbb{R}^N$,

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_3 (nx - k) = 1, \tag{55}$$

$\forall \; x \in \mathbb{R}^N; \, n \in \mathbb{N}$,

and

(iv)

$$\int_{\mathbb{R}^N} Z_3 (x) \, dx = 1, \tag{56}$$

that is Z_3 is a multivariate density function.

(v) It is also clear that

$$\sum_{\substack{k=-\infty \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\infty} Z_3(nx-k) < \frac{4e^2}{\pi e^{n^{1-\beta}}} = c_3(\beta, n), \tag{57}$$

$0 < \beta < 1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$, $m \in \mathbb{N}$.

2.4 About the generalized symmetrical activation function

Here we consider the generalized symmetrical sigmoid function ([6], [11])

$$f_1(x) = \frac{x}{\left(1 + |x|^\mu\right)^{\frac{1}{\mu}}}, \quad \mu > 0,\ x \in \mathbb{R}. \tag{58}$$

This has applications in immunology and protection from disease together with probability theory. It is also called a symmetrical protection curve.

The parameter μ is a shape parameter controling how fast the curve approaches the asymptotes for a given slope at the inflection point. When $\mu = 1$ f_1 is the absolute sigmoid function, and when $\mu = 2$, f_1 is the square root sigmoid function. When $\mu = 1.5$ the function approximates the arctangent function, when $\mu = 2.9$ it approximates the logistic function, and when $\mu = 3.4$ it approximates the error function. Parameter μ is estimated in the likelihood maximization ([11]). For more see [11].

Next we study the particular generator sigmoid function

$$f_2(x) = \frac{x}{\left(1 + |x|^\lambda\right)^{\frac{1}{\lambda}}}, \quad \lambda \text{ is an odd number, } x \in \mathbb{R}. \tag{59}$$

We have that $f_2(0) = 0$, and

$$f_2(-x) = -f_2(x), \tag{60}$$

so f_2 is symmetric with respect to zero.

When $x \geq 0$, we get that ([6])

$$f_2'(x) = \frac{1}{\left(1 + x^\lambda\right)^{\frac{\lambda+1}{\lambda}}} > 0, \tag{61}$$

that is f_2 is strictly increasing on $[0, +\infty)$ and f_2 is strictly increasing on $(-\infty, 0]$. Hence f_2 is strictly increasing on $\mathbb{R}$.

We also have $f_2(+\infty) = f_2(-\infty) = 1$.

Let us consider the activation function ([6]):

$$\psi_4(x) = \frac{1}{4}\left[f_2(x+1) - f_2(x-1)\right] =$$

$$\frac{1}{4}\left[\frac{(x+1)}{\left(1+|x+1|^{\lambda}\right)^{\frac{1}{\lambda}}} - \frac{(x-1)}{\left(1+|x-1|^{\lambda}\right)^{\frac{1}{\lambda}}}\right]. \tag{62}$$

Clearly it holds ([6])

$$\psi_4(x) = \psi_4(-x), \quad \forall\, x \in \mathbb{R}. \tag{63}$$

and

$$\psi_4(0) = \frac{1}{2\sqrt[\lambda]{2}}, \tag{64}$$

and $\psi_4(x) > 0$, $\forall\, x \in \mathbb{R}$.

Following [6], we have that ψ_4 is strictly decreasing over $[0, +\infty)$, and ψ_4 is strictly increasing on $(-\infty, 0]$, by ψ_4-symmetry with respect to y-axis, and $\psi_4'(0) = 0$.

Clearly it is

$$\lim_{x\to+\infty}\psi_4(x) = \lim_{x\to-\infty}\psi_4(x) = 0, \tag{65}$$

therefore the x-axis is the horizontal asymptote of $\psi_4(x)$.

The value

$$\psi_4(0) = \frac{1}{2\sqrt[\lambda]{2}}, \ \lambda \text{ is an odd number}, \tag{66}$$

is the maximum of ψ_4, which is a bell shaped function.

We need

Theorem 10 *([6]) It holds*

$$\sum_{i=-\infty}^{\infty}\psi_4(x-i) = 1, \ \ \forall\, x \in \mathbb{R}. \tag{67}$$

Theorem 11 *([6]) We have that*

$$\int_{-\infty}^{\infty}\psi_4(x)\,dx = 1. \tag{68}$$

So that $\psi_4(x)$ is a density function on $\mathbb{R}$.

We need

Theorem 12 *([6]) Let $0 < \alpha < 1$, and $n \in \mathbb{N}$ with $n^{1-\alpha} > 2$. It holds*

$$\sum_{\substack{j=-\infty \\ : |nx-j| \geq n^{1-\alpha}}}^{\infty}\psi_4(nx-j) < \frac{1}{2\lambda\left(n^{1-\alpha}-2\right)^{\lambda}}, \tag{69}$$

where $\lambda \in \mathbb{N}$ is an odd number.

We introduce (see also [10])

$$Z_4(x_1,...,x_N) := Z_4(x) := \prod_{i=1}^{N} \psi_4(x_i), \quad x = (x_1,...,x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{70}$$

It has the properties:

(i) $Z_4(x) > 0, \ \forall\, x \in \mathbb{R}^N$,

(ii)

$$\sum_{k=-\infty}^{\infty} Z_4(x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z_4(x_1-k_1,...,x_N-k_N) = 1, \tag{71}$$

where $k := (k_1,...,k_n) \in \mathbb{Z}^N$, $\forall\, x \in \mathbb{R}^N$,

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_4(nx-k) = 1, \tag{72}$$

$\forall\, x \in \mathbb{R}^N$; $n \in \mathbb{N}$,

and

(iv)

$$\int_{\mathbb{R}^N} Z_4(x)\, dx = 1, \tag{73}$$

that is Z_4 is a multivariate density function.

(v) It is clear that

$$\sum_{\begin{cases} k=-\infty \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\infty} Z_4(nx-k) < \frac{1}{2\lambda \left(n^{1-\beta}-2\right)^{\lambda}} =: c_4(\beta,n), \tag{74}$$

$0 < \beta < 1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$, λ is odd.

For $f \in C_B^+\left(\mathbb{R}^N\right)$ (continuous and bounded functions from $\mathbb{R}^N$ into $\mathbb{R}_+$), we define the first modulus of continuity

$$\omega_1(f,\delta) := \sup_{\substack{x,y \in \mathbb{R}^N \\ \|x-y\|_\infty \leq h}} |f(x) - f(y)|, \quad h > 0. \tag{75}$$

Given that $f \in C_U^+\left(\mathbb{R}^N\right)$ (uniformly continuous from $\mathbb{R}^N$ into $\mathbb{R}_+$, same definition for ω_1), we have that

$$\lim_{h\to 0} \omega_1(f,h) = 0. \tag{76}$$

When $N = 1$, ω_1 is defined as in (75) with $\|\cdot\|_\infty$ collapsing to $|\cdot|$ and has the property (76).

3 Main Results

We need

Definition 13 *Let $\mathcal{L}$ be the Lebesgue σ-algebra on $\mathbb{R}^N$, $N \in \mathbb{N}$, and the maxitive measure $\mu : \mathcal{L} \to [0, +\infty)$, such that for any $A \in \mathcal{L}$ with $A \neq \varnothing$, we get $\mu(A) > 0$.*

For $f \in C_B^+(\mathbb{R}^N)$, we define the multivariate Kantorovich-Shilkret type neural network operators for any $x \in \mathbb{R}^N$:

$$
{}_jT_n^{\mu}(f, x) = {}_jT_n^{\mu}(f, x_1, ..., x_N) :=
$$

$$
\sum_{k=-\infty}^{\infty} \left(\frac{(N^*) \int_{\left[0, \frac{1}{N}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t)}{\mu\left(\left[0, \frac{1}{n}\right]^N\right)} \right) Z_j(nx - k) =
$$

$$
\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \left(\frac{(N^*) \int_0^{\frac{1}{n}} \cdots \int_0^{\frac{1}{n}} f\left(t_1 + \frac{k_1}{n}, t_2 + \frac{k_2}{n}, ..., t_N + \frac{k_N}{n}\right) d\mu(t_1, ..., t_N)}{\mu\left(\left[0, \frac{1}{n}\right]^N\right)} \right) \tag{77}
$$

$$
\cdot \left(\prod_{i=1}^{N} \psi_j(nx_i - k_i) \right),
$$

where $x = (x_1, ..., x_N) \in \mathbb{R}^N$, $k = (k_1, ..., k_N)$, $t = (t_1, ..., t_N)$, $n \in \mathbb{N}$; $j = 1, 2, 3, 4$.

Clearly here $\mu\left(\left[0, \frac{1}{n}\right]^N\right) > 0$, $\forall\, n \in \mathbb{N}$.

Above we notice that

$$
\left\| {}_jT_n^{\mu}(f) \right\|_{\infty} \leq \|f\|_{\infty}, \tag{78}
$$

so that ${}_jT_n^{\mu}(f, x)$ is well-defined, $j = 1, 2, 3, 4$.

We make

Remark 14 *Let $t \in \left[0, \frac{1}{n}\right]^N$ and $x \in \mathbb{R}^N$, then*

$$
f\left(t + \frac{k}{n}\right) = f\left(t + \frac{k}{n}\right) - f(x) + f(x) \leq \left| f\left(t + \frac{k}{n}\right) - f(x) \right| + f(x), \tag{79}
$$

hence

$$
(N^*) \int_{\left[0, \frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t) \leq
$$

$$
(N^*) \int_{\left[0, \frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t) + f(x)\, \mu\left(\left[0, \frac{1}{n}\right]^N\right). \tag{80}
$$

That is

$$
(N^*) \int_{\left[0, \frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t) - f(x)\, \mu\left(\left[0, \frac{1}{n}\right]^N\right) \leq \tag{81}
$$

$$(N^*) \int_{\left[0, \frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t).$$

Similarly, we have

$$f(x) = f(x) - f\left(t + \frac{k}{n}\right) + f\left(t + \frac{k}{n}\right) \leq \left| f\left(t + \frac{k}{n}\right) - f(x) \right| + f\left(t + \frac{k}{n}\right),$$

hence

$$(N^*) \int_{\left[0, \frac{1}{n}\right]^N} f(x) \, d\mu(t) \leq (N^*) \int_{\left[0, \frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t)$$

$$+ (N^*) \int_{\left[0, \frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t).$$

That is

$$f(x) \, \mu\left(\left[0, \frac{1}{n}\right]^N\right) - (N^*) \int_{\left[0, \frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t) \leq \tag{82}$$

$$(N^*) \int_{\left[0, \frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t).$$

By (81) and (82) we derive

$$\left| (N^*) \int_{\left[0, \frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t) - f(x) \, \mu\left(\left[0, \frac{1}{n}\right]^N\right) \right| \leq$$

$$(N^*) \int_{\left[0, \frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t). \tag{83}$$

In particular it holds

$$\left| \frac{(N^*) \int_{\left[0, \frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t)}{\mu\left(\left[0, \frac{1}{n}\right]^N\right)} - f(x) \right| \leq$$

$$\frac{(N^*) \int_{\left[0, \frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t)}{\mu\left(\left[0, \frac{1}{n}\right]^N\right)}. \tag{84}$$

We present the following approximation result.

Theorem 15 *Let* $f \in C_B^+\left(\mathbb{R}^N\right)$, $0 < \beta < 1$, $x \in \mathbb{R}^N$; $N, n \in \mathbb{N}$ *with* $n^{1-\beta} > 2$; $j = 1, 2, 3, 4$. *Then*

i)

$$\sup_{\mu} \left| {}_jT_n^{\mu}(f, x) - f(x) \right| \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + 2\|f\|_\infty c_j(\beta, n) =: \lambda_{jn}, \tag{85}$$

and

ii)

$$\sup_{\mu} \left\| {}_{j}T_{n}^{\mu}(f) - f \right\|_{\infty} \leq \lambda_{jn}. \tag{86}$$

Given that $f \in \left(C_U^{+}\left(\mathbb{R}^N\right) \cap C_B^{+}\left(\mathbb{R}^N\right)\right)$, we obtain $\lim_{n\to\infty} {}_{j}T_{n}^{\mu}(f) = f$, uniformly. Above $c_j(\beta, n)$ are as in (27), (42), (57) and (74), respectively.

Proof. We observe that

$$\left| {}_{j}T_{n}^{\mu}(f,x) - f(x) \right| =$$

$$\left| \sum_{k=-\infty}^{\infty} \left(\frac{(N^*) \int_{\left[0,\frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)} \right) Z_j(nx-k) - \sum_{k=-\infty}^{\infty} f(x)\, Z_j(nx-k) \right| =$$

$$\left| \sum_{k=-\infty}^{\infty} \left(\left(\frac{(N^*) \int_{\left[0,\frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)} \right) - f(x) \right) Z_j(nx-k) \right| \leq$$

$$\sum_{k=-\infty}^{\infty} \left| \left(\frac{(N^*) \int_{\left[0,\frac{1}{n}\right]^N} f\left(t + \frac{k}{n}\right) d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)} \right) - f(x) \right| Z_j(nx-k) \overset{(84)}{\leq}$$

$$\sum_{k=-\infty}^{\infty} \left(\frac{(N^*) \int_{\left[0,\frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)} \right) Z_j(nx-k) =$$

$$\sum_{\begin{cases} k=-\infty \\ : \left\| \frac{k}{n} - x \right\|_{\infty} \leq \frac{1}{n^{\beta}} \end{cases}}^{\infty} \left(\frac{(N^*) \int_{\left[0,\frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)} \right) Z_j(nx-k) +$$

$$\sum_{\begin{cases} k=-\infty \\ : \left\| \frac{k}{n} - x \right\|_{\infty} > \frac{1}{n^{\beta}} \end{cases}}^{\infty} \left(\frac{(N^*) \int_{\left[0,\frac{1}{n}\right]^N} \left| f\left(t + \frac{k}{n}\right) - f(x) \right| d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)} \right) Z_j(nx-k) \leq \tag{87}$$

$$\sum_{\begin{cases} k=-\infty \\ : \left\| \frac{k}{n} - x \right\|_{\infty} \leq \frac{1}{n^{\beta}} \end{cases}}^{\infty} \left(\frac{(N^*) \int_{\left[0,\frac{1}{n}\right]^N} \omega_1\left(f, \|t\|_{\infty} + \left\| \frac{k}{n} - x \right\|_{\infty}\right) d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)} \right) Z_j(nx-k)$$

$$+2\|f\|_{\infty} \left(\sum_{\begin{cases} k=-\infty \\ : \left\| \frac{k}{n} - x \right\|_{\infty} > \frac{1}{n^{\beta}} \end{cases}}^{\infty} Z_j(nx-k) \right) \text{ (by (27), (42), (57), (74))}$$

$$\leq \omega_1\left(f, \frac{1}{n}+\frac{1}{n^{\beta}}\right)+2\|f\|_{\infty} c_j(\beta, n), \tag{88}$$

proving the claim. ■

Additionally we give

Definition 16 *Denote by* $C_B^{+}\left(\mathbb{R}^N, \mathbb{C}\right)=\{f: \mathbb{R}^N \rightarrow \mathbb{C} \mid f=f_1+i f_2$, *where* $f_1, f_2 \in C_B^{+}\left(\mathbb{R}^N\right)$, $N \in \mathbb{N}\}$. *We set for* $f \in C_B^{+}\left(\mathbb{R}^N, \mathbb{C}\right)$ *that*

$${}_jT_n^{\mu}(f, x):={}_jT_n^{\mu}\left(f_1, x\right)+i\ {}_jT_n^{\mu}\left(f_2, x\right), \quad j=1,2,3,4, \tag{89}$$

$\forall\ n \in \mathbb{N}$, $x \in \mathbb{R}^N$, $i=\sqrt{-1}$.

Theorem 17 *Let* $f \in C_B^{+}\left(\mathbb{R}^N, \mathbb{C}\right)$, $f=f_1+i f_2$, $N \in \mathbb{N}$, $0<\beta<1$, $x \in \mathbb{R}^N$; $n \in \mathbb{N}$ *with* $n^{1-\beta}>2$; $j=1,2,3,4$. *Then*

i)

$$\sup_{\mu}\left|{}_jT_n^{\mu}(f, x)-f(x)\right| \leq\left[\omega_1\left(f_1, \frac{1}{n}+\frac{1}{n^{\beta}}\right)+\omega_1\left(f_2, \frac{1}{n}+\frac{1}{n^{\beta}}\right)\right]$$
$$+2\left(\left\|f_1\right\|_{\infty}+\left\|f_2\right\|_{\infty}\right) c_j(\beta, n)=: l_{jn}, \tag{90}$$

and

ii)

$$\sup_{\mu}\left\|{}_jT_n^{\mu}(f)-f\right\| \leq l_{jn}. \tag{91}$$

Proof.

$$\left|{}_jT_n^{\mu}(f, x)-f(x)\right|=\left|{}_jT_n^{\mu}\left(f_1, x\right)+i\ {}_jT_n^{\mu}\left(f_2, x\right)-f_1(x)-i f_2(x)\right|=$$
$$\left|\left({}_jT_n^{\mu}\left(f_1, x\right)-f_1(x)\right)+i\left({}_jT_n^{\mu}\left(f_2, x\right)-f_2(x)\right)\right| \leq$$
$$\left|{}_jT_n^{\mu}\left(f_1, x\right)-f_1(x)\right|+\left|{}_jT_n^{\mu}\left(f_2, x\right)-f_2(x)\right| \overset{(85)}{\leq} \tag{92}$$
$$\left(\omega_1\left(f_1, \frac{1}{n}+\frac{1}{n^{\beta}}\right)+2\left\|f_1\right\|_{\infty} c_j(\beta, n)\right)+$$
$$\left(\omega_1\left(f_2, \frac{1}{n}+\frac{1}{n^{\beta}}\right)+2\left\|f_2\right\|_{\infty} c_j(\beta, n)\right)=$$
$$\left[\omega_1\left(f_1, \frac{1}{n}+\frac{1}{n^{\beta}}\right)+\omega_1\left(f_2, \frac{1}{n}+\frac{1}{n^{\beta}}\right)\right]+$$
$$2\left(\left\|f_1\right\|_{\infty}+\left\|f_2\right\|_{\infty}\right) c_j(\beta, n), \tag{93}$$

proving the claim. ■

References

[1] G.A. Anastassiou, *Intelligent Systems II: Complete Approximation by Neural Network Operators*, Springer, Heidelberg, New York, 2016.

[2] G.A. Anastassiou, *Intelligent Computations: Abstract Fractional Calculus, Inequalities, Approximations*, Springer, Heidelberg, New York, 2018.

[3] G.A. Anastassiou, *Ordinary and Fractional Approximation by Non-additive Integrals: Choquet, Shilkret and Sugeno Integral Approximators*, Springer, Heidelberg, New York, 2019.

[4] G.A. Anastassiou, *Algebraic function based Banach space valued ordinary and fractional neural network approximations*, New Trends in Mathematical Sciences, 10 special issues (1) (2022), 100-125.

[5] G.A. Anastassiou, *Gudermannian function activated Banach space valued ordinary and fractional neural network approximation*, Advances in Non-linear Variational Inequalities, 25 (2) (2022), 27-64.

[6] G.A. Anastassiou, *Generalized symmetrical sigmoid function activated Banach space valued ordinary and fractional neural network approximation*, Analele Universităţii Oradea, Fasc. Matematica, accepted for publication, 2022.

[7] G.A. Anastassiou, *Abstract multivariate Gudermannian function activated neural network approximations*, Panamerican Mathematical Journal, accepted, 2022.

[8] G.A. Anastassiou, *General multivariate arctangent function activated neural network approximations*, submitted, 2022.

[9] G.A. Anastassiou, *Abstract multivariate algebraic function activated neural network approximations*, submitted, 2022.

[10] G.A. Anastassiou, *Generalized symmetrical sigmoid function activated neural network multivariate approximation*, submitted, 2022.

[11] A.J. Dunning, J. Kensler, L. Goudeville, F. Bailleux, *Some extensions in continuous methods for immunological correlates of protection*, BMC Medical Research Methodology 15 (107) (28 Dec. 2015), doi:10.1186/s12874-015-0096-9.

[12] I.S. Haykin, *Neural Networks: A Comprehensive Foundation* (2 ed.), Prentice Hall, New York, 1998.

[13] W. McCulloch and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, Bulletin of Mathematical Biophysics, 7 (1943), 115-133.

[14] T.M. Mitchell, *Machine Learning*, WCB-McGraw-Hill, New York, 1997.

[15] Niel Shilkret, *Maxitive measure and integration*, Indagationes Mathematicae, 33 (1971), 109-116.

[16] E.W. Weisstein, *Gudermannian*, MathWorld.