# Quantitative Approximation by Multiple sigmoids Kantorovich-Shilkret quasi-interpolation neural network operators

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

**Abstract**

In this article we derive multivariate quantitative approximation by Kantorovich-Shilkret type quasi-interpolation neural network operators with respect to supremum and $L_p$ norms. 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 also uniformly continuous we have pointwise and uniform convergences, plus $L_p$ estimates. We include also the related Complex approximation. Our activation functions are induced by multiple general sigmoid functions.

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

**Keywords and Phrases:** multiple general sigmoid activation functions, univariate and multivariate quasi-interpolation neural network approximation, Kantorovich-Shilkret type operators.

## 1 Introduction

The author in [1] and [2], see Chapters 2-5, was the first to establish neural network approximations to continuous functions with rates by very specifically defined neural network operators of Cardaliaguet-Euvrard and "Squashing" types, by employing the modulus of continuity of the engaged function or its high order derivative, and producing very tight Jackson type inequalities. He treats there both the univariate and multivariate cases. The defining these operators "bell-shaped" and "squashing" functions are assumed to be compact support. Also



*RGMIA Res. Rep. Coll.* **25** (2022), Art. 108, 13 pp. Received 16/10/22

in [2] he gives the $N$th order asymptotic expansion for the error of weak approximation of these two operators to a special natural class of smooth functions, see Chaptes 4-5 there.

The author inspired by [16], continued his studies on neural networks approximation by introducing and using the proper quasi-interpolation operators of sigmoidal and hyperbolic tangent type which resulted into [3] - [7], by treating both the univariate and multivariate cases. He did also the corresponding fractional case [8]. For recent works see [9] - [15].

The author here performs multivariate multiple general sigmoid activation functions based neural network approximation to continuous functions over the whole $\mathbb{R}^N$, $N \in \mathbb{N}$, then he extends his results to complex valued functions. $L_p$ approximations are included. All convergences here are with rates expressed via the modulus of continuity of the involved function and given by very tight Jackson type inequalities.

The author comes up with the "right" precisely defined flexible quasi-interpolation, Kantorovich-Shilkret type integral coefficient neural networks operators associated with: multiple general sigmoid activation functions. In preparation to prove our results we present important properties of the general density functions defining our operators.

Feed-forward neural networks (FNNs) with one hidden layer, the only type of networks we deal with in this article, are mathematically expressed 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 $\sigma$ is the activation function of the network. In many fundamental neural network models, the activation functions are based on multiple general sigmoid activation functions. About neural networks in general read [17], [18], [19].

In recent years non-additive integrals, like the N. Shilkret one [20], have become fashionable and more useful in Economics, etc.

# 2 Background

## 2.1 About Shilkret integral

Here we follow [20].

Let $\mathcal{F}$ be a $\sigma$-field of subsets of an arbitrary set $\Omega$. An extended non-negative real valued function $\mu$ 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 $\mu$ a maxitive measure. Here $f$ stands for a non-negative measurable function on $\Omega$. In [20], 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 ([20]) has the following properties:

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

where $\chi_E$ 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 $\mu : \mathcal{F} \to [0, +\infty)$.

## 2.2 On activation functions

Let $i = 1, ..., N \in \mathbb{N}$ and $h_i : \mathbb{R} \to [-1, 1]$ be a general sigmoid function, such that it is strictly increasing, $h_i(0) = 0$, $h_i(-x) = -h_i(x)$, $h_i(+\infty) = 1$, $h_i(-\infty) = -1$. Also $h_i$ is strictly convex over $(-\infty, 0]$ and striclty concave over $[0, +\infty)$, with $h_i^{(2)} \in C(\mathbb{R}, [-1, 1])$.

Some examples of related sigmoid functions follow: $\frac{1}{1+e^{-x}}$; $\tanh x$; $\frac{2}{\pi}\arctan\left(\frac{\pi}{2}x\right)$; $\frac{x}{\sqrt[2m]{1+x^{2m}}}$, $m \in \mathbb{N}$; $\frac{4}{\pi} gd(x)$; $\frac{x}{(1+|x|)^{\frac{1}{\lambda}}}$, $\lambda$ is odd; $erf\left(\frac{\sqrt{\pi}}{2}x\right)$; $\frac{1}{1+e^{-\mu x}}$, $\tanh \mu x$, $\mu > 0$; for all $x \in \mathbb{R}$.

We consider the activation function

$$\psi_i(x) := \frac{1}{4}\left(h_i(x+1) - h_i(x-1)\right), \ \ x \in \mathbb{R}, \ i = 1, ..., N. \tag{11}$$

As in [11], p. 285, we get that $\psi_i(-x) = \psi_i(x)$, thus $\psi_i$ is an even function. Since $x+1 > x-1$, then $h_i(x+1) > h_i(x-1)$, and $\psi_i(x) > 0$, all $x \in \mathbb{R}$.

We see that

$$\psi_i(0) = \frac{h_i(1)}{2}, \quad i = 1, ..., N. \tag{12}$$

Let $x > 1$, we have that

$$\psi_i'(x) = \frac{1}{4}\left(h_i'(x+1) - h_i'(x-1)\right) < 0,$$

by $h_i'$ being strictly decreasing over $[0, +\infty)$.

Let now $0 < x < 1$, then $1 - x > 0$ and $0 < 1 - x < 1 + x$. It holds $h_i'(x-1) = h_i'(1-x) > h_i'(x+1)$, so that again $\psi_i'(x) < 0$. Consequently $\psi_i$ is stritly decreasing on $(0, +\infty)$.

Clearly, $\psi_i$ is strictly increasing on $(-\infty, 0)$, and $\psi_i'(0) = 0$.

See that

$$\lim_{x\to+\infty}\psi_i(x) = \frac{1}{4}\left(h_i(+\infty) - h_i(+\infty)\right) = 0, \tag{13}$$

and

$$\lim_{x\to-\infty}\psi_i(x) = \frac{1}{4}\left(h_i(-\infty) - h_i(-\infty)\right) = 0. \tag{14}$$

That is the $x$-axis is the horizontal asymptote on $\psi_i$.

Conclusion, $\psi$ is a bell symmetric function with maximum

$$\psi_i(0) = \frac{h_i(1)}{2}.$$

We need

**Theorem 1** *We have that*

$$\sum_{i=-\infty}^{\infty} \psi_i (x - i) = 1, \ \ \forall \ x \in \mathbb{R}, \ \ i = 1, ..., N. \tag{15}$$

**Proof.** As exactly the same as in [11], p. 286 is omitted. ■

**Theorem 2** *It holds*

$$\int_{-\infty}^{\infty} \psi_i (x) \, dx = 1, \quad i = 1, ..., N. \tag{16}$$

**Proof.** Similar to [11], p. 287. It is omitted. ■

Thus $\psi_i (x)$ is a density function on $\mathbb{R}$, $i = 1, ..., N$.

We need also

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

$$\sum_{\begin{cases} k = -\infty \\ : |nx - k| \geq n^{1-\alpha} \end{cases}}^{\infty} \psi_i (nx - k) < \frac{\left(1 - h_i \left(n^{1-\alpha} - 2\right)\right)}{2}, \ i = 1, ..., N. \tag{17}$$

*Notice that*

$$\lim_{n \to +\infty} \frac{\left(1 - h_i \left(n^{1-\alpha} - 2\right)\right)}{2} = 0, \ \ i = 1, ..., N.$$

**Proof.** Similar to [13], as such is omitted. ■

We make

**Remark 4** *We define*

$$Z (x_1, ..., x_N) := Z (x) := \prod_{i=1}^{N} \psi_i (x_i), \quad x = (x_1, ..., x_N) \in \mathbb{R}^N, \ \ N \in \mathbb{N}. \tag{18}$$

*It has the properties:*

*(i)*

$$Z (x) > 0, \ \ \forall x \in \mathbb{R}^N, \tag{19}$$

*(ii)*

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

$$\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} \prod_{i=1}^{N} \psi_i (x_i - k_i) = \prod_{i=1}^{N} \left( \sum_{k_i=-\infty}^{\infty} \psi_i (x_i - k_i) \right) \overset{(5)}{=} 1.$$

*Hence*
$$\sum_{k=-\infty}^{\infty} Z(x-k) = 1. \tag{20}$$
*That is*

*(iii)*
$$\sum_{k=-\infty}^{\infty} Z(nx-k) = 1, \quad \forall x \in \mathbb{R}^N;\ n \in \mathbb{N}. \tag{21}$$
*And*

*(iv)*
$$\int_{\mathbb{R}^N} Z(x)\,dx = \int_{\mathbb{R}^N} \left(\prod_{i=1}^{N} \psi_i(x_i)\right) dx_1 ... dx_N = \prod_{i=1}^{N} \left(\int_{-\infty}^{\infty} \psi_i(x_i)\,dx_i\right) \overset{(16)}{=} 1, \tag{22}$$
*thus*
$$\int_{\mathbb{R}^N} Z(x)\,dx = 1, \tag{23}$$
*that is* $Z$ *is a multivariate density function.*

*Here denote* $x = (x_1, ..., x_N)$, $\|x\|_\infty := \max\{|x_1|, ..., |x_N|\}$, $x \in \mathbb{R}^N$, *also set* $\infty := (\infty, ..., \infty)$, $-\infty := (-\infty, ..., -\infty)$ *upon the multivariate context,* $0 < \beta < 1$,

*(v) We have*
$$\sum_{\begin{cases} k=-\infty \\ \left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\infty} Z(nx-k) = \sum_{k_1=-\infty}^{\infty} ... \sum_{\substack{k_N=-\infty \\ \left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta}}}^{\infty} \left(\prod_{i=1}^{N} \psi_i(nx_i-k_i)\right) =$$
$$\prod_{i=1}^{N} \left( \sum_{\begin{cases} k_i=-\infty \\ \left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\infty} \psi_i(nx_i-k_i) \right) \leq \quad \textit{(for some } r \in \{1, ..., N\} \textit{)}$$
$$\left(\prod_{\substack{i=1 \\ i\neq r}}^{N} \left(\sum_{k_i=-\infty}^{\infty} \psi_i(nx_i-k_i)\right)\right) \left( \sum_{\begin{cases} k_r=-\infty \\ \left|\frac{k_r}{n}-x_r\right| > \frac{1}{n^\beta} \end{cases}}^{\infty} \psi_r(nx_r-k_r) \right) = \tag{24}$$
$$\sum_{\begin{cases} k_r=-\infty \\ \left|\frac{k_r}{n}-x_r\right| > \frac{1}{n^\beta} \end{cases}}^{\infty} \psi_r(nx_r-k_r) = \sum_{\begin{cases} k_r=-\infty \\ |nx_r-k_r| > n^{1-\beta} \end{cases}}^{\infty} \psi_r(nx_r-k_r) \overset{(17)}{<}$$

$$\frac{1-h_r\left(n^{1-\beta}-2\right)}{2} \leq \max_{i\in\{1,...,N\}}\left(\frac{1-h_i\left(n^{1-\beta}-2\right)}{2}\right).$$

*That is*

$$\sum_{\begin{cases} k=-\infty \\ \left\|\frac{k}{n}-x\right\|_{\infty} > \frac{1}{n^{\beta}} \end{cases}}^{\infty} Z(nx-k) < \max_{i\in\{1,...,N\}}\left(\frac{1-h_i\left(n^{1-\beta}-2\right)}{2}\right), \tag{25}$$

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

Denote by

$$\delta_N(\beta,n) := \max_{i\in\{1,...,N\}}\left(\frac{1-h_i\left(n^{1-\beta}-2\right)}{2}\right), \tag{26}$$

$0<\beta<1$.

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

$$\omega_1(f,\delta) := \sup_{\substack{x,y\in\mathbb{R}^N \\ \|x-y\|_{\infty}\leq h}} |f(x)-f(y)|, \quad h>0. \tag{27}$$

Given that $f\in C_U^+\left(\mathbb{R}^N\right)$ (uniformly continuous from $\mathbb{R}^N$ into $\mathbb{R}_+$, same definition for $\omega_1$), we have that

$$\lim_{h\to 0}\omega_1(f,h)=0. \tag{28}$$

When $N=1$, $\omega_1$ is defined as in (27) with $\|\cdot\|_{\infty}$ collapsing to $|\cdot|$ and has the property (28).

# 3 Main Results

We need

**Definition 5** *Let $\mathcal{L}$ be the Lebesgue $\sigma$-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 $f\in C_B^+\left(\mathbb{R}^N\right)$, we define the multivariate Kantorovich-Shilkret type neural network operators for any $x\in\mathbb{R}^N$ :*

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

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

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

$$\cdot\left(\prod_{i=1}^{N}\psi_i\left(nx_i-k_i\right)\right), \tag{29}$$

*where* $x=(x_1,...,x_N)\in\mathbb{R}^N$, $k=(k_1,...,k_N)$, $t=(t_1,...,t_N)$, $n\in\mathbb{N}$.

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

*Above we notice that*

$$\left\|T_n^{\mu}(f)\right\|_{\infty}\leq\|f\|_{\infty}, \tag{30}$$

*so that* $T_n^{\mu}(f,x)$ *is well-defined.*

We make

**Remark 6** *Let* $t\in\left[0,\frac{1}{n}\right]^N$ *and* $x\in\mathbb{R}^N$, *then*

$$f\left(t+\frac{k}{n}\right)=f\left(t+\frac{k}{n}\right)-f(x)+f(x)\leq\left|f\left(t+\frac{k}{n}\right)-f(x)\right|+f(x), \tag{31}$$

*hence*

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

$$(N^*)\int_{\left[0,\frac{1}{n}\right]^N}\left|f\left(t+\frac{k}{n}\right)-f(x)\right|d\mu(t)+f(x)\,\mu\left(\left[0,\frac{1}{n}\right]^N\right). \tag{32}$$

*That is*

$$(N^*)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)d\mu(t)-f(x)\,\mu\left(\left[0,\frac{1}{n}\right]^N\right)\leq \tag{33}$$

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

*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*

$$f(x)\,\mu\left(\left[0,\frac{1}{n}\right]^N\right)-(N^*)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)d\mu(t)\leq \tag{34}$$

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

*By (33) and (34) we derive*

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

$$(N^*)\int_{\left[0,\frac{1}{n}\right]^N}\left|f\left(t+\frac{k}{n}\right)-f(x)\right|d\mu(t). \tag{35}$$

*In particular it holds*

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

$$\frac{(N^*)\int_{\left[0,\frac{1}{n}\right]^N}\left|f\left(t+\frac{k}{n}\right)-f(x)\right|d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)}. \tag{36}$$

We present the following approximation result.

**Theorem 7** *Let $f\in C_B^+\left(\mathbb{R}^N\right)$, $0<\beta<1$, $x\in\mathbb{R}^N$; $N,n\in\mathbb{N}$ with $n^{1-\beta}>2$. Then*

*i)*

$$\sup_{\mu}\left|T_n^{\mu}(f,x)-f(x)\right|\leq\omega_1\left(f,\frac{1}{n}+\frac{1}{n^{\beta}}\right)+2\|f\|_{\infty}\,\delta_N(\beta,n)=:\lambda_n, \tag{37}$$

*and*

*ii)*

$$\sup_{\mu}\left\|T_n^{\mu}(f)-f\right\|_{\infty}\leq\lambda_n. \tag{38}$$

*Given that $f\in\left(C_U^+\left(\mathbb{R}^N\right)\cap C_B^+\left(\mathbb{R}^N\right)\right)$, we obtain $\lim_{n\to\infty}T_n^{\mu}(f)=f$, uniformly. Above $\delta_N(\beta,n)$ is as in (26).*

**Proof.** We observe that

$$\left|T_n^{\mu}(f,x)-f(x)\right|=$$

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

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

$$\sum_{k=-\infty}^{\infty}\left|\left(\frac{(N^*)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)}\right)-f(x)\right| Z(nx-k) \overset{(36)}{\leq}$$

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

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

$$\sum_{\substack{k=-\infty \\ :\left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta}}}^{\infty}\left(\frac{(N^*)\int_{\left[0,\frac{1}{n}\right]^N}\left|f\left(t+\frac{k}{n}\right)-f(x)\right|d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)}\right)Z(nx-k) \leq \tag{39}$$

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

$$+2\|f\|_\infty\left(\sum_{\substack{k=-\infty \\ :\left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta}}}^{\infty} Z(nx-k)\right) \text{ (by (25))}$$

$$\leq \omega_1\left(f,\frac{1}{n}+\frac{1}{n^\beta}\right)+2\|f\|_\infty\,\delta_N(\beta,n), \tag{40}$$

proving the claim. ■

Additionally we give

**Definition 8** *Denote by* $C_B^+\left(\mathbb{R}^N,\mathbb{C}\right)=\{f:\mathbb{R}^N\to\mathbb{C}|f=f_1+if_2,$ *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*

$$T_n^\mu(f,x) := T_n^\mu(f_1,x)+i\,T_n^\mu(f_2,x), \tag{41}$$

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

**Theorem 9** *Let $f \in C_B^+\left(\mathbb{R}^N, \mathbb{C}\right)$, $f = f_1 + if_2$, $N \in \mathbb{N}$, $0 < \beta < 1$, $x \in \mathbb{R}^N$; $n \in \mathbb{N}$ with $n^{1-\beta} > 2$. Then*

*i)*

$$\sup_{\mu} \left| T_n^{\mu}(f,x) - f(x) \right| \leq \left[ \omega_1\left(f_1, \frac{1}{n} + \frac{1}{n^{\beta}}\right) + \omega_1\left(f_2, \frac{1}{n} + \frac{1}{n^{\beta}}\right) \right]$$

$$+2\left(\|f_1\|_{\infty} + \|f_2\|_{\infty}\right) \delta_N(\beta, n) =: l_n, \tag{42}$$

*and*

*ii)*

$$\sup_{\mu} \left\| T_n^{\mu}(f) - f \right\| \leq l_n. \tag{43}$$

**Proof.**

$$\left| T_n^{\mu}(f,x) - f(x) \right| = \left| T_n^{\mu}(f_1,x) + i\, T_n^{\mu}(f_2,x) - f_1(x) - if_2(x) \right| =$$

$$\left| \left(T_n^{\mu}(f_1,x) - f_1(x)\right) + i\left(T_n^{\mu}(f_2,x) - f_2(x)\right) \right| \leq$$

$$\left| T_n^{\mu}(f_1,x) - f_1(x) \right| + \left| T_n^{\mu}(f_2,x) - f_2(x) \right| \overset{(37)}{\leq} \tag{44}$$

$$\left( \omega_1\left(f_1, \frac{1}{n} + \frac{1}{n^{\beta}}\right) + 2\|f_1\|_{\infty} \delta_N(\beta, n) \right) +$$

$$\left( \omega_1\left(f_2, \frac{1}{n} + \frac{1}{n^{\beta}}\right) + 2\|f_2\|_{\infty} \delta_N(\beta, n) \right) =$$

$$\left[ \omega_1\left(f_1, \frac{1}{n} + \frac{1}{n^{\beta}}\right) + \omega_1\left(f_2, \frac{1}{n} + \frac{1}{n^{\beta}}\right) \right] +$$

$$2\left(\|f_1\|_{\infty} + \|f_2\|_{\infty}\right) \delta_N(\beta, n). \tag{45}$$

proving the claim. ■

We finish with an $L_{p_1}$, $p_1 \geq 1$, estimate.

**Theorem 10** *Let $f \in C_B^+\left(\mathbb{R}^N, \mathbb{C}\right)$, $f = f_1 + if_2$, $N \in \mathbb{N}$, $0 < \beta < 1$, $n \in \mathbb{N}$ with $n^{1-\beta} > 2$, and $p_1 \geq 1$. Then*

$$\left\| T_n^{\mu}(f) - f \right\|_{p_1, \Lambda} \leq l_n \left| \Lambda \right|^{\frac{1}{p_1}}, \tag{46}$$

*where $|\Lambda| < \infty$, is the Lebesgue measure of compact $\Lambda \subset \mathbb{R}^N$, and $l_n$ as in (42).*

**Proof.** By integrating (42), etc. ■

## References


[1] G.A. Anastassiou, *Rate of convergence of some neural network operators to the unit-univariate case*, J. Math. Anal. Appli. 212 (1997), 237-262.

[2] G.A. Anastassiou, *Quantitative Approximations*, Chapman & Hall/CRC, Boca Raton, New York, 2001.

[3] G.A. Anastassiou, *Intelligent Systems: Approximation by Artificial Neural Networks*, Intelligent Systems Reference Library, Vol. 19, Springer, Heidelberg, 2011.

[4] G.A. Anastassiou, *Univariate hyperbolic tangent neural network approximation*, Mathematics and Computer Modelling, 53 (2011), 1111-1132.

[5] G.A. Anastassiou, *Multivariate hyperbolic tangent neural network approximation*, Computers and Mathematics 61 (2011), 809-821.

[6] G.A. Anastassiou, *Multivariate sigmoidal neural network approximation*, Neural Networks 24 (2011), 378-386.

[7] G.A. Anastassiou, *Univariate sigmoidal neural network approximation*, J. of Computational Analysis and Applications, Vol. 14, No. 4, 2012, 659-690.

[8] G.A. Anastassiou, *Fractional neural network approximation*, Computers and Mathematics with Applications, 64 (2012), 1655-1676.

[9] G.A. Anastassiou, *Univariate error function based neural network approximation*, Indian J. of Math., Vol. 57, No. 2 (2015), 243-291.

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

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

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

[13] 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.

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

[15] G.A. Anastassiou, *Generalized symmetrical sigmoid function activated Banach space valued ordinary and fractional neural network approximation*, Analele Universităţii Oradea, Fasc. Matematica, XXX(1) (2022), 117-134.

[16] Z. Chen and F. Cao, *The approximation operators with sigmoidal functions*, Computers and Mathematics with Applications, 58 (2009), 758-765.

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

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

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

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