Parametrized arctangent based Banach space valued multi layer neural network multivariate approximations

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

Abstract

Here we examine the multivariate quantitative approximations of Banach space valued continuous multivariate functions on a box or $\mathbb{R}^N$, $N \in \mathbb{N}$, by the multivariate normalized, quasi-interpolation, Kantorovich type and quadrature type neural network operators. We research also the case of approximation by iterated operators of the last four types, that is multi hidden layer approximations. These approximations are achieved by establishing multidimensional Jackson type inequalities involving the multivariate modulus of continuity of the engaged function or its high order Fréchet derivatives. Our multivariate operators are defined by using a multidimensional density function induced by a parametrized arctangent sigmoid function. The approximations are pointwise and uniform. The related feed-forward neural networks are with one or multi hidden layers.

2020 AMS Mathematics Subject Classification: 41A17, 41A25, 41A30, 46B36.

Keywords and Phrases: parametrized arctangent sigmoid function, multivariate neural network approximation, quasi-interpolation operator, Kantorovich type operator, quadrature type operator, multivariate modulus of continuity, iterated multi layer approximation.

1 Introduction

The author in [2] and [3], 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 Cardaliagnet-Euvrard and "Squashing" types,

RGMIA Res. Rep. Coll. **25** (2022), Art. 123, 26 pp. Received 17/11/22

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 of compact support. Also in [3] he gives the Nth order asymptotic expansion for the error of weak approximation of these two operators to a special natural class of smooth functions, see chapters 4-5 there.

For this article the author is motivated by the article [16] of Z. Chen and F. Cao, also by [4]-[12], [17], [18].

The author here performs multivariate parametrized arctangent sigmoid function based neural network approximations to continuous functions over boxes or over the whole $\mathbb{R}^N$, $N \in \mathbb{N}$. Also he does the iterated multlayer approximation. All convergences here are with rates expressed via the multivariate modulus of continuity of the involved function or its high order Fréchet derivative and given by very tight multidimensional Jackson type inequalities.

The author here comes up with the "right" precisely defined multivariate normalized, quasi-interpolation neural network operators related to boxes or $\mathbb{R}^N$, as well as Kantorovich type and quadrature type related operators on $\mathbb{R}^N$. Our boxes are not necessarily symmetric to the origin. In preparation to prove our results we establish important properties of the basic multivariate density function induced by a parametrized arctangent sigmoid function and 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

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

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 network models the activation function is the arctangent sigmoid function. About neural networks read [19]-[21].

2 Background

We consider the function

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

We will be using the following parametrized function with a parameter $\lambda > 0$:

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

We have that

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

and

$$h'_\lambda(x) = \frac{2}{\pi}\left(\frac{1}{1+\frac{\pi^2\lambda^2x^2}{4}}\right)\frac{\pi\lambda}{2} = \frac{4\lambda}{4+\pi^2\lambda^2x^2} > 0 \tag{3}$$

all $x \in \mathbb{R}$.

So that h_λ is a strictly increasing function from $\mathbb{R}$ into $[-1,1]$, with horizontal asymptotes $y = \pm 1$.

Furthermore we get that

$$h''_\lambda(x) = -\left(\frac{8\pi^2\lambda^3}{\left(4+\pi^2\lambda^2x^2\right)^2}\right)x,\quad x \in \mathbb{R}. \tag{4}$$

Clearly then

$$h''_\lambda(x) < 0,\ \text{ for } x \in (0,+\infty),$$
$$\text{and}$$
$$h''_\lambda(x) > 0,\ \text{ for } x \in (-\infty,0),$$

with $h''_\lambda(0) = 0$.

That is h_λ is strictly concave over $[0,+\infty)$ and h_λ is strictly convex over $(-\infty,0]$. Obviosly $h''_\lambda \in C(\mathbb{R})$.

Therefore h_λ is a sigmoid function fulfilling exactly all the properties of the general sigmoid function described in [13].

When $0 < \lambda < 1$, h_λ is expected to outperform the *ReLu* and *Leaky ReLu* activation functions.

We consider the activation function

$$\psi_\lambda(x) := \frac{1}{4}\left(h_\lambda(x+1) - h_\lambda(x-1)\right),\ \ x \in \mathbb{R}, \tag{5}$$

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

We see that

$$\psi_\lambda(0) = \frac{h_\lambda(1)}{2} = \frac{\arctan\left(\frac{\pi}{2}\lambda\right)}{\pi}. \tag{6}$$

Let $x > 1$, we have that

$$\psi'_\lambda(x) = \frac{1}{4}\left(h'_\lambda(x+1) - h'_\lambda(x-1)\right) < 0,$$

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

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

Clearly, ψ_λ is strictly increasing on $(-\infty, 0)$, and $\psi'_\lambda(0) = 0$.

See that

$$\lim_{x \to +\infty} \psi_\lambda(x) = \frac{1}{4}\left(h_\lambda(+\infty) - h_\lambda(+\infty)\right) = 0, \tag{7}$$

and

$$\lim_{x \to -\infty} \psi_\lambda(x) = \frac{1}{4}\left(h_\lambda(-\infty) - h_\lambda(-\infty)\right) = 0. \tag{8}$$

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

Conclusion, ψ_λ is a bell symmetric function with maximum

$$\psi_\lambda(0) = \frac{h_\lambda(1)}{2} = \frac{\arctan\left(\frac{\pi\lambda}{2}\right)}{\pi}.$$

We need

Theorem 1 *We have that*

$$\sum_{i=-\infty}^{\infty} \psi_\lambda(x - i) = 1, \quad \forall\, x \in \mathbb{R}. \tag{9}$$

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

Theorem 2 *It holds*

$$\int_{-\infty}^{\infty} \psi_\lambda(x)\, dx = 1. \tag{10}$$

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

Thus $\psi_\lambda(x)$ is a density function on $\mathbb{R}$.

We give

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

$$\sum_{\begin{cases} k = -\infty \\ : |nx - k| \geq n^{1-\alpha} \end{cases}}^{\infty} \psi_\lambda(nx - k) < \frac{\left(1 - h_\lambda\left(n^{1-\alpha} - 2\right)\right)}{2}. \tag{11}$$

Notice that

$$\lim_{n \to +\infty} \frac{\left(1 - h_\lambda\left(n^{1-\alpha} - 2\right)\right)}{2} = 0.$$

Proof. By [13]. ■

Denote by $\lfloor \cdot \rfloor$ the integral part of the number and by $\lceil \cdot \rceil$ the ceiling of the number.

We further give

Theorem 4 *Let $x \in [a,b] \subset \mathbb{R}$ and $n \in \mathbb{N}$ so that $\lceil na \rceil \leq \lfloor nb \rfloor$. It holds*

$$\frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi_\lambda (nx-k)} < \frac{1}{\psi_\lambda (1)} = \frac{2\pi}{\arctan (\pi\lambda)}, \quad \forall\ x \in [a,b]. \tag{12}$$

Proof. As similar to [11], p. 289 is omitted. ∎

Remark 5 *We have that*

$$\lim_{n\to\infty} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi_\lambda (nx-k) \neq 1, \tag{13}$$

for at least some $x \in [a,b]$.

See [11], p. 290, same reasoning.

Note 6 *For large enough n we always obtain $\lceil na \rceil \leq \lfloor nb \rfloor$. Also $a \leq \frac{k}{n} \leq b$, iff $\lceil na \rceil \leq k \leq \lfloor nb \rfloor$. In general it holds (by (9))*

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi_\lambda (nx-k) \leq 1. \tag{14}$$

We introduce

$$Z_\lambda (x_1, ..., x_N) := Z_\lambda (x) := \prod_{i=1}^{N} \psi_\lambda (x_i), \quad x = (x_1, ..., x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{15}$$

It has the properties:

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

(ii)

$$\sum_{k=-\infty}^{\infty} Z_\lambda (x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z_\lambda (x_1-k_1, ..., x_N-k_N) = 1, \tag{16}$$

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

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_\lambda (nx-k) = 1, \tag{17}$$

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

and

(iv)

$$\int_{\mathbb{R}^N} Z_\lambda (x)\, dx = 1, \tag{18}$$

that is Z is a multivariate density function.

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

$$\begin{aligned} \lceil na \rceil &:= (\lceil na_1 \rceil, ..., \lceil na_N \rceil), \\ \lfloor nb \rfloor &:= (\lfloor nb_1 \rfloor, ..., \lfloor nb_N \rfloor), \end{aligned} \tag{19}$$

where $a := (a_1, ..., a_N)$, $b := (b_1, ..., b_N)$.

We obviously see that

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z_\lambda (nx - k) = \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left(\prod_{i=1}^{N} \psi_\lambda (nx_i - k_i) \right) =$$

$$\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \cdots \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} \left(\prod_{i=1}^{N} \psi_\lambda (nx_i - k_i) \right) = \prod_{i=1}^{N} \left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \psi_\lambda (nx_i - k_i) \right). \tag{20}$$

For $0 < \beta < 1$ and $n \in \mathbb{N}$, a fixed $x \in \mathbb{R}^N$, we have that

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z_\lambda (nx - k) =$$

$$\sum_{\substack{k = \lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_{\infty} \leq \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} Z_\lambda (nx - k) + \sum_{\substack{k = \lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_{\infty} > \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} Z_\lambda (nx - k). \tag{21}$$

In the last two sums the counting is over disjoint vector sets of k's, because the condition $\left\| \frac{k}{n} - x \right\|_{\infty} > \frac{1}{n^\beta}$ implies that there exists at least one $\left| \frac{k_r}{n} - x_r \right| > \frac{1}{n^\beta}$, where $r \in \{1, ..., N\}$.

(v) As in [10], pp. 379-380, we derive that

$$\sum_{\substack{k = \lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_{\infty} > \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} Z_\lambda (nx - k) \overset{(11)}{<} \frac{1 - h_\lambda \left(n^{1-\beta} - 2 \right)}{2}, \quad 0 < \beta < 1, \tag{22}$$

with $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \prod_{i=1}^{N} [a_i, b_i]$.

(vi) By Theorem 4 we get that

$$0 < \frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z_\lambda (nx - k)} < \frac{1}{(\psi_\lambda (1))^N} = \left(\frac{2\pi}{\arctan (\pi \lambda)} \right)^N, \tag{23}$$

$\forall\ x \in \left(\prod_{i=1}^{N}\left[a_i, b_i\right]\right),\ \ n \in \mathbb{N}.$

It is also clear that

(vii)

$$\sum_{\substack{k=-\infty \\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}}}^{\infty} Z_{\lambda}\left(nx-k\right) < \frac{1-h_{\lambda}\left(n^{1-\beta}-2\right)}{2}, \tag{24}$$

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

Furthermore it holds

$$\lim_{n\to\infty} \sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} Z_{\lambda}\left(nx-k\right) \neq 1, \tag{25}$$

for at least some $x \in \left(\prod_{i=1}^{N}\left[a_i, b_i\right]\right)$.

Here $\left(X, \|\cdot\|_{\gamma}\right)$ is a Banach space.

Let $f \in C\left(\prod_{i=1}^{N}\left[a_i, b_i\right], X\right)$, $x=\left(x_1,...,x_N\right) \in \prod_{i=1}^{N}\left[a_i, b_i\right]$, $n \in \mathbb{N}$ such that $\lceil na_i\rceil \leq \lfloor nb_i\rfloor$, $i=1,...,N$.

We introduce and define the following multivariate linear normalized neural network operator ($x:=\left(x_1,...,x_N\right) \in \left(\prod_{i=1}^{N}\left[a_i, b_i\right]\right)$):

$$A_n\left(f, x_1,...,x_N\right) := A_n\left(f,x\right) := \frac{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} f\left(\frac{k}{n}\right) Z_{\lambda}\left(nx-k\right)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} Z_{\lambda}\left(nx-k\right)} =$$

$$\frac{\sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor}\sum_{k_2=\lceil na_2\rceil}^{\lfloor nb_2\rfloor}...\sum_{k_N=\lceil na_N\rceil}^{\lfloor nb_N\rfloor} f\left(\frac{k_1}{n},...,\frac{k_N}{n}\right)\left(\prod_{i=1}^{N}\psi_{\lambda}\left(nx_i-k_i\right)\right)}{\prod_{i=1}^{N}\left(\sum_{k_i=\lceil na_i\rceil}^{\lfloor nb_i\rfloor}\psi_{\lambda}\left(nx_i-k_i\right)\right)}. \tag{26}$$

For large enough $n \in \mathbb{N}$ we always obtain $\lceil na_i\rceil \leq \lfloor nb_i\rfloor$, $i=1,...,N$. Also $a_i \leq \frac{k_i}{n} \leq b_i$, iff $\lceil na_i\rceil \leq k_i \leq \lfloor nb_i\rfloor$, $i=1,...,N$.

When $g \in C\left(\prod_{i=1}^{N}\left[a_i, b_i\right]\right)$ we define the companion operator

$$\widetilde{A}_n\left(g,x\right) := \frac{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} g\left(\frac{k}{n}\right) Z_{\lambda}\left(nx-k\right)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} Z_{\lambda}\left(nx-k\right)}. \tag{27}$$

Clearly $\widetilde{A}_n$ is a positive linear operator. We have that

$$\widetilde{A}_n\left(1,x\right) = 1,\ \ \forall\ x \in \left(\prod_{i=1}^{N}\left[a_i, b_i\right]\right).$$

Notice that $A_n\left(f\right) \in C\left(\prod_{i=1}^{N}\left[a_i, b_i\right], X\right)$ and $\widetilde{A}_n\left(g\right) \in C\left(\prod_{i=1}^{N}\left[a_i, b_i\right]\right)$.

Furthermore it holds

$$\left\|A_n(f,x)\right\|_\gamma \leq \frac{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\|f\left(\frac{k}{n}\right)\right\|_\gamma Z_\lambda(nx-k)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} Z_\lambda(nx-k)} = \widetilde{A}_n\left(\|f\|_\gamma, x\right), \tag{28}$$

$\forall\ x \in \prod_{i=1}^N [a_i, b_i]$.

Clearly $\|f\|_\gamma \in C\left(\prod_{i=1}^N [a_i, b_i]\right)$.

So, we have that

$$\left\|A_n(f,x)\right\|_\gamma \leq \widetilde{A}_n\left(\|f\|_\gamma, x\right), \tag{29}$$

$\forall\ x \in \prod_{i=1}^N [a_i, b_i]$, $\forall\ n \in \mathbb{N}$, $\forall\ f \in C\left(\prod_{i=1}^N [a_i, b_i], X\right)$.

Let $c \in X$ and $g \in C\left(\prod_{i=1}^N [a_i, b_i]\right)$, then $cg \in C\left(\prod_{i=1}^N [a_i, b_i], X\right)$.

Furthermore it holds

$$A_n(cg, x) = c\widetilde{A}_n(g, x), \ \ \forall\ x \in \prod_{i=1}^N [a_i, b_i]. \tag{30}$$

Since $\widetilde{A}_n(1) = 1$, we get that

$$A_n(c) = c, \ \ \forall\ c \in X. \tag{31}$$

We call $\widetilde{A}_n$ the companion operator of A_n.

For convinience we call

$$A_n^*(f,x) := \sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} f\left(\frac{k}{n}\right) Z_\lambda(nx-k) =$$

$$\sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor}\sum_{k_2=\lceil na_2\rceil}^{\lfloor nb_2\rfloor}\cdots\sum_{k_N=\lceil na_N\rceil}^{\lfloor nb_N\rfloor} f\left(\frac{k_1}{n}, ..., \frac{k_N}{n}\right)\left(\prod_{i=1}^N \psi_\lambda(nx_i-k_i)\right), \tag{32}$$

$\forall\ x \in \left(\prod_{i=1}^N [a_i, b_i]\right)$.

That is

$$A_n(f,x) := \frac{A_n^*(f,x)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} Z_\lambda(nx-k)}, \tag{33}$$

$\forall\ x \in \left(\prod_{i=1}^N [a_i, b_i]\right)$, $n \in \mathbb{N}$.

Hence

$$A_n(f,x) - f(x) = \frac{A_n^*(f,x) - f(x)\left(\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} Z_\lambda(nx-k)\right)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} Z_\lambda(nx-k)}. \tag{34}$$

Consequently we derive

$$\left\|A_{n}\left(f,x\right)-f\left(x\right)\right\|_{\gamma}\overset{(23)}{\leq}\left(\frac{2\pi}{\arctan\left(\pi\lambda\right)}\right)^{N}\left\|A_{n}^{*}\left(f,x\right)-f\left(x\right)\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}Z_{\lambda}\left(nx-k\right)\right\|_{\gamma},\tag{35}$$

$\forall\ x\in\left(\prod_{i=1}^{N}\left[a_{i},b_{i}\right]\right)$.

We will estimate the right hand side of (35).

For the last and others we need

Definition 7 *([11], p. 274) Let M be a convex and compact subset of $\left(\mathbb{R}^{N},\left\|\cdot\right\|_{p}\right)$, $p\in[1,\infty]$, and $\left(X,\left\|\cdot\right\|_{\gamma}\right)$ be a Banach space. Let $f\in C\left(M,X\right)$. We define the first modulus of continuity of f as*

$$\omega_{1}\left(f,\delta\right):=\sup_{\substack{x,y\in M:\\ \left\|x-y\right\|_{p}\leq\delta}}\left\|f\left(x\right)-f\left(y\right)\right\|_{\gamma},\quad 0<\delta\leq diam\left(M\right).\tag{36}$$

If $\delta>diam\left(M\right)$, then

$$\omega_{1}\left(f,\delta\right)=\omega_{1}\left(f,diam\left(M\right)\right).\tag{37}$$

Notice $\omega_{1}\left(f,\delta\right)$ is increasing in $\delta>0$. For $f\in C_{B}\left(M,X\right)$ (continuous and bounded functions) $\omega_{1}\left(f,\delta\right)$ is defined similarly.

Lemma 8 *([11], p. 274) We have $\omega_{1}\left(f,\delta\right)\rightarrow0$ as $\delta\downarrow0$, iff $f\in C\left(M,X\right)$, where M is a convex compact subset of $\left(\mathbb{R}^{N},\left\|\cdot\right\|_{p}\right)$, $p\in[1,\infty]$.*

Clearly we have also: $f\in C_{U}\left(\mathbb{R}^{N},X\right)$ (uniformly continuous functions), iff $\omega_{1}\left(f,\delta\right)\rightarrow0$ as $\delta\downarrow0$, where ω_{1} is defined similarly to (36). The space $C_{B}\left(\mathbb{R}^{N},X\right)$ denotes the continuous and bounded functions on $\mathbb{R}^{N}$.

When $f\in C_{B}\left(\mathbb{R}^{N},X\right)$ we define,

$$B_{n}\left(f,x\right):=B_{n}\left(f,x_{1},...,x_{N}\right):=\sum_{k=-\infty}^{\infty}f\left(\frac{k}{n}\right)Z_{\lambda}\left(nx-k\right):=$$

$$\sum_{k_{1}=-\infty}^{\infty}\sum_{k_{2}=-\infty}^{\infty}...\sum_{k_{N}=-\infty}^{\infty}f\left(\frac{k_{1}}{n},\frac{k_{2}}{n},...,\frac{k_{N}}{n}\right)\left(\prod_{i=1}^{N}\psi_{\lambda}\left(nx_{i}-k_{i}\right)\right),\tag{38}$$

$n\in\mathbb{N}$, $\forall\ x\in\mathbb{R}^{N}$, $N\in\mathbb{N}$, the multivariate quasi-interpolation neural network operator.

Also for $f\in C_{B}\left(\mathbb{R}^{N},X\right)$ we define the multivariate Kantorovich type neural network operator

$$C_{n}\left(f,x\right):=C_{n}\left(f,x_{1},...,x_{N}\right):=\sum_{k=-\infty}^{\infty}\left(n^{N}\int_{\frac{k}{n}}^{\frac{k+1}{n}}f\left(t\right)dt\right)Z_{\lambda}\left(nx-k\right)=$$

$$\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}...\sum_{k_N=-\infty}^{\infty}\left(n^N\int_{\frac{k_1}{n}}^{\frac{k_1+1}{n}}\int_{\frac{k_2}{n}}^{\frac{k_2+1}{n}}...\int_{\frac{k_N}{n}}^{\frac{k_N+1}{n}}f\left(t_1,...,t_N\right)dt_1...dt_N\right)$$

$$\cdot\left(\prod_{i=1}^{N}\psi_{\lambda}\left(nx_i-k_i\right)\right), \tag{39}$$

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

Again for $f\in C_B\left(\mathbb{R}^N,X\right)$, $N\in\mathbb{N}$, we define the multivariate neural network operator of quadrature type $D_n\left(f,x\right)$, $n\in\mathbb{N}$, as follows.

Let $\theta=\left(\theta_1,...,\theta_N\right)\in\mathbb{N}^N$, $r=\left(r_1,...,r_N\right)\in\mathbb{Z}_+^N$, $w_r=w_{r_1,r_2,...r_N}\geq 0$, such that $\sum\limits_{r=0}^{\theta}w_r=\sum\limits_{r_1=0}^{\theta_1}\sum\limits_{r_2=0}^{\theta_2}...\sum\limits_{r_N=0}^{\theta_N}w_{r_1,r_2,...r_N}=1$; $k\in\mathbb{Z}^N$ and

$$\delta_{nk}\left(f\right):=\delta_{n,k_1,k_2,...,k_N}\left(f\right):=\sum_{r=0}^{\theta}w_rf\left(\frac{k}{n}+\frac{r}{n\theta}\right)=$$

$$\sum_{r_1=0}^{\theta_1}\sum_{r_2=0}^{\theta_2}...\sum_{r_N=0}^{\theta_N}w_{r_1,r_2,...r_N}f\left(\frac{k_1}{n}+\frac{r_1}{n\theta_1},\frac{k_2}{n}+\frac{r_2}{n\theta_2},...,\frac{k_N}{n}+\frac{r_N}{n\theta_N}\right), \tag{40}$$

where $\frac{r}{\theta}:=\left(\frac{r_1}{\theta_1},\frac{r_2}{\theta_2},...,\frac{r_N}{\theta_N}\right)$.

We set

$$D_n\left(f,x\right):=D_n\left(f,x_1,...,x_N\right):=\sum_{k=-\infty}^{\infty}\delta_{nk}\left(f\right)Z_{\lambda}\left(nx-k\right)= \tag{41}$$

$$\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}...\sum_{k_N=-\infty}^{\infty}\delta_{n,k_1,k_2,...,k_N}\left(f\right)\left(\prod_{i=1}^{N}\psi_{\lambda}\left(nx_i-k_i\right)\right),$$

$\forall\ x\in\mathbb{R}^N$.

In this article we study the approximation properties of $A_n,B_n,C_n,\ D_n$ neural network operators and as well of their iterates. That is, the quantitative pointwise and uniform convergence of these operators to the unit operator I.

3 Multivariate Parametrized Arctangent Neural Network Approximations

Here we present several vectorial neural network approximations to Banach space valued functions given with rates.

We give

Theorem 9 *Let* $f \in C\left(\prod_{i=1}^{N}[a_i,b_i], X\right)$, $0<\beta<1$, $x \in \left(\prod_{i=1}^{N}[a_i,b_i]\right)$, $N,n \in \mathbb{N}$ *with* $n^{1-\beta}>2$. *Then*

1)

$$\|A_n(f,x)-f(x)\|_{\gamma} \leq$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)}\right)^{N}\left[\omega_1\left(f,\frac{1}{n^{\beta}}\right)+\left(1-h_{\lambda}\left(n^{1-\beta}-2\right)\right)\left\|\|f\|_{\gamma}\right\|_{\infty}\right]=:\lambda_1(n), \tag{42}$$

and

2)

$$\left\|\|A_n(f)-f\|_{\gamma}\right\|_{\infty} \leq \lambda_1(n). \tag{43}$$

We notice that $\lim_{n\to\infty} A_n(f) \overset{\|\cdot\|_{\gamma}}{=} f$, *pointwise and uniformly.*

Above ω_1 *is with respect to* $p=\infty$ *and the speed of convergnece is* $\max\left(\frac{1}{n^{\beta}},\left(1-h_{\lambda}\left(n^{1-\beta}-2\right)\right)\right)$.

Proof. As similar to [12] is omitted. See also [14]. ■

We make

Remark 10 *([11], pp. 263-266) Let* $\left(\mathbb{R}^N,\|\cdot\|_p\right)$, $N\in\mathbb{N}$; *where* $\|\cdot\|_p$ *is the* L_p-*norm,* $1\leq p\leq\infty$. $\mathbb{R}^N$ *is a Banach space, and* $\left(\mathbb{R}^N\right)^j$ *denotes the* j-*fold product space* $\mathbb{R}^N\times...\times\mathbb{R}^N$ *endowed with the max-norm* $\|x\|_{(\mathbb{R}^N)^j}:=\max_{1\leq\lambda\leq j}\|x_\lambda\|_p$, *where* $x:=(x_1,...,x_j)\in\left(\mathbb{R}^N\right)^j$.

Let $\left(X,\|\cdot\|_{\gamma}\right)$ *be a general Banach space. Then the space* $L_j:=L_j\left(\left(\mathbb{R}^N\right)^j;X\right)$ *of all* j-*multilinear continuous maps* $g:\left(\mathbb{R}^N\right)^j\to X$, $j=1,...,m$, *is a Banach space with norm*

$$\|g\|:=\|g\|_{L_j}:=\sup_{\left(\|x\|_{(\mathbb{R}^N)^j}=1\right)}\|g(x)\|_{\gamma}=\sup\frac{\|g(x)\|_{\gamma}}{\|x_1\|_p...\|x_j\|_p}. \tag{44}$$

Let M *be a non-empty convex and compact subset of* $\mathbb{R}^k$ *and* $x_0\in M$ *is fixed.*

Let O *be an open subset of* $\mathbb{R}^N$: $M\subset O$. *Let* $f:O\to X$ *be a continuous function, whose Fréchet derivatives (see [22])* $f^{(j)}:O\to L_j=L_j\left(\left(\mathbb{R}^N\right)^j;X\right)$ *exist and are continuous for* $1\leq j\leq m$, $m\in\mathbb{N}$.

Call $(x-x_0)^j:=(x-x_0,...,x-x_0)\in\left(\mathbb{R}^N\right)^j$, $x\in M$.

We will work with $f|_M$.

Then, by Taylor's formula ([15]), ([22], p. 124), we get

$$f(x)=\sum_{j=0}^{m}\frac{f^{(j)}(x_0)(x-x_0)^j}{j!}+R_m(x,x_0),\quad \textit{all } x\in M, \tag{45}$$

where the remainder is the Riemann integral

$$R_m(x, x_0) := \int_0^1 \frac{(1-u)^{m-1}}{(m-1)!} \left(f^{(m)}(x_0 + u(x - x_0)) - f^{(m)}(x_0) \right) (x - x_0)^m \, du, \tag{46}$$

here we set $f^{(0)}(x_0)(x - x_0)^0 = f(x_0)$.

We consider

$$w := \omega_1 \left(f^{(m)}, h \right) := \sup_{\substack{x, y \in M: \\ \|x - y\|_p \leq h}} \left\| f^{(m)}(x) - f^{(m)}(y) \right\|, \tag{47}$$

$h > 0$.

We obtain

$$\left\| \left(f^{(m)}(x_0 + u(x - x_0)) - f^{(m)}(x_0) \right) (x - x_0)^m \right\|_\gamma \leq$$

$$\left\| f^{(m)}(x_0 + u(x - x_0)) - f^{(m)}(x_0) \right\| \cdot \|x - x_0\|_p^m \leq$$

$$w \|x - x_0\|_p^m \left\lceil \frac{u \|x - x_0\|_p}{h} \right\rceil, \tag{48}$$

by Lemma 7.1.1, [1], p. 208, where $\lceil \cdot \rceil$ *is the ceiling.*

Therefore for all $x \in M$ *(see [1], pp. 121-122):*

$$\|R_m(x, x_0)\|_\gamma \leq w \|x - x_0\|_p^m \int_0^1 \left\lceil \frac{u \|x - x_0\|_p}{h} \right\rceil \frac{(1-u)^{m-1}}{(m-1)!} du$$

$$= w \Phi_m \left(\|x - x_0\|_p \right) \tag{49}$$

by a change of variable, where

$$\Phi_m(t) := \int_0^{|t|} \left\lceil \frac{s}{h} \right\rceil \frac{(|t| - s)^{m-1}}{(m-1)!} ds = \frac{1}{m!} \left(\sum_{j=0}^{\infty} (|t| - jh)_+^m \right), \quad \forall \, t \in \mathbb{R},$$

is a (polynomial) spline function, see [1], p. 210-211.

Also from there we get

$$\Phi_m(t) \leq \left(\frac{|t|^{m+1}}{(m+1)!h} + \frac{|t|^m}{2m!} + \frac{h |t|^{m-1}}{8(m-1)!} \right), \quad \forall \, t \in \mathbb{R}, \tag{50}$$

with equality true only at $t = 0$.

Therefore it holds

$$\|R_m(x, x_0)\|_\gamma \leq w \left(\frac{\|x - x_0\|_p^{m+1}}{(m+1)!h} + \frac{\|x - x_0\|_p^m}{2m!} + \frac{h \|x - x_0\|_p^{m-1}}{8(m-1)!} \right), \quad \forall \, x \in M. \tag{51}$$

We have found that

$$\left\| f(x) - \sum_{j=0}^{m} \frac{f^{(j)}(x_0)(x-x_0)^j}{j!} \right\|_\gamma \leq$$

$$\omega_1\left(f^{(m)}, h\right)\left(\frac{\|x-x_0\|_p^{m+1}}{(m+1)!h} + \frac{\|x-x_0\|_p^m}{2m!} + \frac{h\|x-x_0\|_p^{m-1}}{8(m-1)!}\right) < \infty, \quad (52)$$

$\forall\ x, x_0 \in M$.

Here $0 < \omega_1\left(f^{(m)}, h\right) < \infty$, *by* M *being compact and* $f^{(m)}$ *being continuous on* M.

One can rewrite (52) as follows:

$$\left\| f(\cdot) - \sum_{j=0}^{m} \frac{f^{(j)}(x_0)(\cdot-x_0)^j}{j!} \right\|_\gamma \leq$$

$$\omega_1\left(f^{(m)}, h\right)\left(\frac{\|\cdot-x_0\|_p^{m+1}}{(m+1)!h} + \frac{\|\cdot-x_0\|_p^m}{2m!} + \frac{h\|\cdot-x_0\|_p^{m-1}}{8(m-1)!}\right), \ \forall\ x_0 \in M, \quad (53)$$

a pointwise functional inequality on M.

Here $(\cdot - x_0)^j$ *maps* M *into* $\left(\mathbb{R}^N\right)^j$ *and it is continuous, also* $f^{(j)}(x_0)$ *maps* $\left(\mathbb{R}^N\right)^j$ *into* X *and it is continuous. Hence their composition* $f^{(j)}(x_0)(\cdot - x_0)^j$ *is continuous from* M *into* X.

Clearly $f(\cdot) - \sum_{j=0}^{m} \frac{f^{(j)}(x_0)(\cdot-x_0)^j}{j!} \in C(M, X)$, *hence* $\left\| f(\cdot) - \sum_{j=0}^{m} \frac{f^{(j)}(x_0)(\cdot-x_0)^j}{j!} \right\|_\gamma \in C(M)$.

Let $\left\{\widetilde{L}_N\right\}_{N\in\mathbb{N}}$ *be a sequence of positive linear operators mapping* $C(M)$ *into* $C(M)$.

Therefore we obtain

$$\left(\widetilde{L}_N\left(\left\| f(\cdot) - \sum_{j=0}^{m} \frac{f^{(j)}(x_0)(\cdot-x_0)^j}{j!} \right\|_\gamma\right)\right)(x_0) \leq$$

$$\omega_1\left(f^{(m)}, h\right)\left[\frac{\left(\widetilde{L}_N\left(\|\cdot-x_0\|_p^{m+1}\right)\right)(x_0)}{(m+1)!h} + \frac{\left(\widetilde{L}_N\left(\|\cdot-x_0\|_p^{m}\right)\right)(x_0)}{2m!} + \frac{h\left(\widetilde{L}_N\left(\|\cdot-x_0\|_p^{m-1}\right)\right)(x_0)}{8(m-1)!}\right], \quad (54)$$

$\forall\ N \in \mathbb{N}$, $\forall\ x_0 \in M$.

Clearly (54) is valid when $M = \prod_{i=1}^{N} [a_i, b_i]$ and $\widetilde{L}_n = \widetilde{A}_n$, see (27).

All the above is preparation for the following theorem, where we assume Fréchet differentiability of functions.

This will be a direct application of Theorem 10.2, [11], pp. 268-270. The operators A_n, $\widetilde{A}_n$ fulfill its assumptions, see (26), (27), (29), (30) and (31).

We present the following high order approximation results.

Theorem 11 *Let O open subset of $\left(\mathbb{R}^N, \|\cdot\|_p\right)$, $p \in [1, \infty]$, such that $\prod_{i=1}^{N} [a_i, b_i] \subset O \subseteq \mathbb{R}^N$, and let $\left(X, \|\cdot\|_\gamma\right)$ be a general Banach space. Let $m \in \mathbb{N}$ and $f \in C^m(O, X)$, the space of m-times continuously Fréchet differentiable functions from O into X. We study the approximation of $f|_{\prod_{i=1}^{N} [a_i, b_i]}$. Let $x_0 \in \left(\prod_{i=1}^{N} [a_i, b_i]\right)$ and $r > 0$. Then*

1)

$$\left\| (A_n(f))(x_0) - \sum_{j=0}^{m} \frac{1}{j!} \left(A_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_\gamma \leq$$

$$\frac{\omega_1\left(f^{(m)}, r\left(\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right)^{\frac{1}{m+1}}\right)}{rm!} \left(\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right)^{\left(\frac{m}{m+1}\right)}$$

$$\left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8}\right], \tag{55}$$

2) additionally if $f^{(j)}(x_0) = 0$, $j = 1, ..., m$, we have

$$\|(A_n(f))(x_0) - f(x_0)\|_\gamma \leq$$

$$\frac{\omega_1\left(f^{(m)}, r\left(\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right)^{\frac{1}{m+1}}\right)}{rm!} \left(\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right)^{\left(\frac{m}{m+1}\right)} \tag{56}$$

$$\left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8}\right],$$

3)

$$\|(A_n(f))(x_0) - f(x_0)\|_\gamma \leq \sum_{j=1}^{m} \frac{1}{j!} \left\| \left(A_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_\gamma +$$

$$\frac{\omega_1\left(f^{(m)}, r\left(\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right)^{\frac{1}{m+1}}\right)}{rm!} \left(\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right)^{\left(\frac{m}{m+1}\right)} \tag{57}$$

$$\left[\frac{1}{(m+1)}+\frac{r}{2}+\frac{mr^2}{8}\right],$$

and

4)

$$\left\| \left\| A_n(f)-f \right\|_\gamma \right\|_{\infty,\prod_{i=1}^{N}[a_i,b_i]} \leq$$

$$\sum_{j=1}^{m}\frac{1}{j!}\left\| \left\| \left(A_n\left(f^{(j)}(x_0)(\cdot-x_0)^j\right)\right)(x_0)\right\|_\gamma \right\|_{\infty,x_0\in\prod_{i=1}^{N}[a_i,b_i]}+$$

$$\frac{\omega_1\left(f^{(m)},r\left\|\left(\widetilde{A}_n\left(\left\|\cdot-x_0\right\|_p^{m+1}\right)\right)(x_0)\right\|_{\infty,x_0\in\prod_{i=1}^{N}[a_i,b_i]}^{\frac{1}{m+1}}\right)}{rm!}$$

$$\left\|\left(\widetilde{A}_n\left(\left\|\cdot-x_0\right\|_p^{m+1}\right)\right)(x_0)\right\|_{\infty,x_0\in\prod_{i=1}^{N}[a_i,b_i]}^{\left(\frac{m}{m+1}\right)} \tag{58}$$

$$\left[\frac{1}{(m+1)}+\frac{r}{2}+\frac{mr^2}{8}\right].$$

We give

Corollary 12 *(to Theorem 11, case of $m=1$) Then*

1)

$$\left\|(A_n(f))(x_0)-f(x_0)\right\|_\gamma \leq \left\|\left(A_n\left(f^{(1)}(x_0)(\cdot-x_0)\right)\right)(x_0)\right\|_\gamma+$$

$$\frac{1}{2r}\omega_1\left(f^{(1)},r\left(\left(\widetilde{A}_n\left(\left\|\cdot-x_0\right\|_p^2\right)\right)(x_0)\right)^{\frac{1}{2}}\right)\left(\left(\widetilde{A}_n\left(\left\|\cdot-x_0\right\|_p^2\right)\right)(x_0)\right)^{\frac{1}{2}} \tag{59}$$

$$\left[1+r+\frac{r^2}{4}\right],$$

and

2)

$$\left\| \left\| (A_n(f))-f \right\|_\gamma \right\|_{\infty,\prod_{i=1}^{N}[a_i,b_i]} \leq$$

$$\left\| \left\| \left(A_n\left(f^{(1)}(x_0)(\cdot-x_0)\right)\right)(x_0)\right\|_\gamma \right\|_{\infty,x_0\in\prod_{i=1}^{N}[a_i,b_i]}+$$

$$\frac{1}{2r}\omega_1\left(f^{(1)},r\left\|\left(\widetilde{A}_n\left(\left\|\cdot-x_0\right\|_p^2\right)\right)(x_0)\right\|_{\infty,x_0\in\prod_{i=1}^{N}[a_i,b_i]}^{\frac{1}{2}}\right)$$

$$\left\|\left(\widetilde{A}_n\left(\left\|\cdot-x_0\right\|_p^2\right)\right)(x_0)\right\|_{\infty,x_0\in\prod_{i=1}^{N}[a_i,b_i]}^{\frac{1}{2}}\left[1+r+\frac{r^2}{4}\right], \tag{60}$$

$r>0$.

We make

Remark 13 *We estimate* $(0<\alpha<1,\ m,n\in\mathbb{N}:n^{1-\alpha}>2)$,

$$\widetilde{A}_n\left(\|\cdot-x_0\|_\infty^{m+1}\right)(x_0)=\frac{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\|\frac{k}{n}-x_0\right\|_\infty^{m+1}Z_\lambda(nx_0-k)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}Z_\lambda(nx_0-k)}\overset{(23)}{<}$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)}\right)^N\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\|\frac{k}{n}-x_0\right\|_\infty^{m+1}Z_\lambda(nx_0-k)= \tag{61}$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)}\right)^N\left\{\sum_{\substack{k=\lceil na\rceil\\ :\left\|\frac{k}{n}-x_0\right\|_\infty\le\frac{1}{n^\alpha}}}^{\lfloor nb\rfloor}\left\|\frac{k}{n}-x_0\right\|_\infty^{m+1}Z_\lambda(nx_0-k)+\right.$$

$$\left.\sum_{\substack{k=\lceil na\rceil\\ :\left\|\frac{k}{n}-x_0\right\|_\infty>\frac{1}{n^\alpha}}}^{\lfloor nb\rfloor}\left\|\frac{k}{n}-x_0\right\|_\infty^{m+1}Z_\lambda(nx_0-k)\right\}\overset{(24)}{\le}$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)}\right)^N\left\{\frac{1}{n^{\alpha(m+1)}}+\left(\frac{1-h_\lambda\left(n^{1-\alpha}-2\right)}{2}\right)\|b-a\|_\infty^{m+1}\right\}, \tag{62}$$

(where $b-a=(b_1-a_1,...,b_N-a_N)$*).*

We have proved that $\left(\forall\ x_0\in\prod_{i=1}^N[a_i,b_i]\right)$

$$\widetilde{A}_n\left(\|\cdot-x_0\|_\infty^{m+1}\right)(x_0)<$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)}\right)^N\left\{\frac{1}{n^{\alpha(m+1)}}+\left(\frac{1-h_\lambda\left(n^{1-\alpha}-2\right)}{2}\right)\|b-a\|_\infty^{m+1}\right\}=:\varphi_1(n) \tag{63}$$

$(0<\alpha<1,\ m,n\in\mathbb{N}:n^{1-\alpha}>2)$.

And, consequently it holds

$$\left\|\widetilde{A}_n\left(\|\cdot-x_0\|_\infty^{m+1}\right)(x_0)\right\|_{\infty,x_0\in\prod_{i=1}^N[a_i,b_i]}<$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)}\right)^N\left\{\frac{1}{n^{\alpha(m+1)}}+\left(\frac{1-h_\lambda\left(n^{1-\alpha}-2\right)}{2}\right)\|b-a\|_\infty^{m+1}\right\}=\varphi_1(n)\to0,\quad \text{as } n\to+\infty. \tag{64}$$

So, we have that $\varphi_1(n) \to 0$, as $n \to +\infty$. Thus, when $p \in [1, \infty]$, from Theorem 11 we have the convergence to zero in the right hand sides of parts (1), (2).

Next we estimate $\left\| \left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_\gamma$.

We have that

$$\left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) = \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} f^{(j)}(x_0) \left(\frac{k}{n} - x_0 \right)^j Z_\lambda(nx_0 - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z_\lambda(nx_0 - k)}. \tag{65}$$

When $p = \infty$, $j = 1, ..., m$, we obtain

$$\left\| f^{(j)}(x_0) \left(\frac{k}{n} - x_0 \right)^j \right\|_\gamma \leq \left\| f^{(j)}(x_0) \right\| \left\| \frac{k}{n} - x_0 \right\|_\infty^j . \tag{66}$$

We further have that

$$\left\| \left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_\gamma \overset{(23)}{<}$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)} \right)^N \left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| f^{(j)}(x_0) \left(\frac{k}{n} - x_0 \right)^j \right\|_\gamma Z_\lambda(nx_0 - k) \right) \leq$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)} \right)^N \left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| f^{(j)}(x_0) \right\| \left\| \frac{k}{n} - x_0 \right\|_\infty^j Z_\lambda(nx_0 - k) \right) = \tag{67}$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)} \right)^N \left\| f^{(j)}(x_0) \right\| \left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^j Z_\lambda(nx_0 - k) \right) =$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)} \right)^N \left\| f^{(j)}(x_0) \right\| \left\{ \sum_{\begin{cases} k = \lceil na \rceil \\ : \left\| \frac{k}{n} - x_0 \right\|_\infty \leq \frac{1}{n^\alpha} \end{cases}}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^j Z_\lambda(nx_0 - k) \right.$$

$$\left. + \sum_{\begin{cases} k = \lceil na \rceil \\ : \left\| \frac{k}{n} - x_0 \right\|_\infty > \frac{1}{n^\alpha} \end{cases}}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^j Z_\lambda(nx_0 - k) \right\} \overset{(24)}{\leq} \tag{68}$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)} \right)^N \left\| f^{(j)}(x_0) \right\| \left\{ \frac{1}{n^{\alpha j}} + \left(\frac{1 - h_\lambda(n^{1-\alpha} - 2)}{2} \right) \|b - a\|_\infty^j \right\} \to 0, \text{ as } n \to \infty.$$

That is

$$\left\| \left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_\gamma \to 0, \text{ as } n \to \infty.$$

Therefore when $p = \infty$, for $j = 1, ..., m$, we have proved:

$$\left\| \left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_\gamma <$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)} \right)^N \left\| f^{(j)}(x_0) \right\| \left\{ \frac{1}{n^{\alpha j}} + \left(\frac{1 - h_\lambda \left(n^{1-\alpha} - 2 \right)}{2} \right) \|b - a\|_\infty^j \right\} \le \tag{69}$$

$$\left(\frac{2\pi}{\arctan(\pi\lambda)} \right)^N \left\| f^{(j)} \right\|_\infty \left\{ \frac{1}{n^{\alpha j}} + \left(\frac{1 - h_\lambda \left(n^{1-\alpha} - 2 \right)}{2} \right) \|b - a\|_\infty^j \right\} =: \varphi_{2j}(n) < \infty,$$

and converges to zero, as $n \to \infty$.

We conclude:

In Theorem 11, the right hand sides of (57) and (58) converge to zero as $n \to \infty$, for any $p \in [1, \infty]$.

Also in Corollary 12, the right hand sides of (59) and (60) converge to zero as $n \to \infty$, for any $p \in [1, \infty]$.

Conclusion 14 *We have proved that the left hand sides of (55), (56), (57), (58) and (59), (60) converge to zero as $n \to \infty$, for $p \in [1, \infty]$. Consequently $A_n \to I$ (unit operator) pointwise and uniformly, as $n \to \infty$, where $p \in [1, \infty]$. In the presence of initial conditions we achieve a higher speed of convergence, see (56). Higher speed of convergence happens also to the left hand side of (55).*

We give

Corollary 15 *(to Theorem 11) Let O open subset of $\left(\mathbb{R}^N, \|\cdot\|_\infty\right)$, such that $\prod_{i=1}^N [a_i, b_i] \subset O \subseteq \mathbb{R}^N$, and let $\left(X, \|\cdot\|_\gamma\right)$ be a general Banach space. Let $m \in \mathbb{N}$ and $f \in C^m(O, X)$, the space of m-times continuously Fréchet differentiable functions from O into X. We study the approximation of $f|_{\prod_{i=1}^N [a_i, b_i]}$. Let $x_0 \in \left(\prod_{i=1}^N [a_i, b_i] \right)$ and $r > 0$. Here $\varphi_1(n)$ as in (63) and $\varphi_{2j}(n)$ as in (69), where $n \in \mathbb{N} : n^{1-\alpha} > 2$, $0 < \alpha < 1$, $j = 1, ..., m$. Then*

1)

$$\left\| (A_n(f))(x_0) - \sum_{j=0}^m \frac{1}{j!} \left(A_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_\gamma \le$$

$$\frac{\omega_1\left(f^{(m)}, r(\varphi_1(n))^{\frac{1}{m+1}} \right)}{rm!} (\varphi_1(n))^{\left(\frac{m}{m+1}\right)} \left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8} \right], \tag{70}$$

2) additionally, if $f^{(j)}(x_0) = 0$, $j = 1, ..., m$, *we have*

$$\|(A_n(f))(x_0) - f(x_0)\|_\gamma \leq$$

$$\frac{\omega_1\left(f^{(m)}, r(\varphi_1(n))^{\frac{1}{m+1}}\right)}{rm!}(\varphi_1(n))^{\left(\frac{m}{m+1}\right)}\left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8}\right], \tag{71}$$

3)

$$\left\|\|A_n(f) - f\|_\gamma\right\|_{\infty, \prod_{i=1}^{N}[a_i, b_i]} \leq \sum_{j=1}^{m} \frac{\varphi_{2j}(n)}{j!} +$$

$$\frac{\omega_1\left(f^{(m)}, r(\varphi_1(n))^{\frac{1}{m+1}}\right)}{rm!}(\varphi_1(n))^{\left(\frac{m}{m+1}\right)} \tag{72}$$

$$\left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8}\right] =: \varphi_3(n) \to 0, \text{ as } n \to \infty.$$

We continue with

Theorem 16 *Let* $f \in C_B(\mathbb{R}^N, X)$, $0 < \beta < 1$, $x \in \mathbb{R}^N$, $N, n \in \mathbb{N}$ *with* $n^{1-\beta} > 2$, ω_1 *is for* $p = \infty$. *Then*

1)

$$\|B_n(f, x) - f(x)\|_\gamma \leq \omega_1\left(f, \frac{1}{n^\beta}\right) + \left(1 - h_\lambda\left(n^{1-\beta} - 2\right)\right)\left\|\|f\|_\gamma\right\|_\infty =: \lambda_2(n), \tag{73}$$

2)

$$\left\|\|B_n(f) - f\|_\gamma\right\|_\infty \leq \lambda_2(n). \tag{74}$$

Given that $f \in \left(C_U(\mathbb{R}^N, X) \cap C_B(\mathbb{R}^N, X)\right)$, *we obtain* $\lim_{n\to\infty} B_n(f) = f$, *uniformly. The speed of convergence above is* $\max\left(\frac{1}{n^\beta}, \left(1 - h_\lambda\left(n^{1-\beta} - 2\right)\right)\right)$.

Proof. As similar to [12] is omitted. ■

We give

Theorem 17 *Let* $f \in C_B(\mathbb{R}^N, X)$, $0 < \beta < 1$, $x \in \mathbb{R}^N$, $N, n \in \mathbb{N}$ *with* $n^{1-\beta} > 2$, ω_1 *is for* $p = \infty$. *Then*

1)

$$\|C_n(f, x) - f(x)\|_\gamma \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + \left(1 - h_\lambda\left(n^{1-\beta} - 2\right)\right)\left\|\|f\|_\gamma\right\|_\infty =: \lambda_3(n), \tag{75}$$

2)

$$\left\|\|C_n(f) - f\|_\gamma\right\|_\infty \leq \lambda_3(n). \tag{76}$$

Given that $f \in \left(C_U(\mathbb{R}^N, X) \cap C_B(\mathbb{R}^N, X)\right)$, *we obtain* $\lim_{n\to\infty} C_n(f) = f$, *uniformly.*

Proof. As similar to [12] is omitted. ■

We also present

Theorem 18 *Let $f \in C_B\left(\mathbb{R}^N, X\right)$, $0 < \beta < 1$, $x \in \mathbb{R}^N$, $N, n \in \mathbb{N}$ with $n^{1-\beta} > 2$, ω_1 is for $p = \infty$. Then*

1)

$$\left\| D_n(f,x) - f(x) \right\|_\gamma \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + \left(1 - h_\lambda\left(n^{1-\beta} - 2\right)\right) \left\| \|f\|_\gamma \right\|_\infty = \lambda_4(n), \tag{77}$$

2)

$$\left\| \|D_n(f) - f\|_\gamma \right\|_\infty \leq \lambda_4(n). \tag{78}$$

Given that $f \in \left(C_U\left(\mathbb{R}^N, X\right) \cap C_B\left(\mathbb{R}^N, X\right)\right)$, we obtain $\lim_{n\to\infty} D_n(f) = f$, uniformly.

Proof. As similar to [12] is omitted. ■

Next we perform multi layer neural network approximations.

We make

Definition 19 *Let $f \in C_B\left(\mathbb{R}^N, X\right)$, $N \in \mathbb{N}$, where $\left(X, \|\cdot\|_\gamma\right)$ is a Banach space. We define the general neural network operator*

$$F_n(f,x) := \sum_{k=-\infty}^{\infty} l_{nk}(f) Z_\lambda(nx - k) =$$

$$\begin{cases} B_n(f,x), & \text{if } l_{nk}(f) = f\left(\frac{k}{n}\right), \\ C_n(f,x), & \text{if } l_{nk}(f) = n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\,dt, \\ D_n(f,x), & \text{if } l_{nk}(f) = \delta_{nk}(f). \end{cases} \tag{79}$$

Clearly $l_{nk}(f)$ is an X-valued bounded linear functional such that $\|l_{nk}(f)\|_\gamma \leq \left\| \|f\|_\gamma \right\|_\infty$.

Hence $F_n(f)$ is a bounded linear operator with $\left\| \|F_n(f)\|_\gamma \right\|_\infty \leq \left\| \|f\|_\gamma \right\|_\infty$.

We need

Theorem 20 *Let $f \in C_B\left(\mathbb{R}^N, X\right)$, $N \geq 1$. Then $F_n(f) \in C_B\left(\mathbb{R}^N, X\right)$.*

Proof. Very lengthy and as similar to [12] is omitted. ■

Remark 21 *By (26) it is obvious that $\left\| \|A_n(f)\|_\gamma \right\|_\infty \leq \left\| \|f\|_\gamma \right\|_\infty < \infty$, and $A_n(f) \in C\left(\prod_{i=1}^N [a_i, b_i], X\right)$, given that $f \in C\left(\prod_{i=1}^N [a_i, b_i], X\right)$.*

Call L_n any of the operators A_n, B_n, C_n, D_n.

Clearly then

$$\left\| \left\| L_n^2 (f) \right\|_\gamma \right\|_\infty = \left\| \left\| L_n (L_n (f)) \right\|_\gamma \right\|_\infty \leq \left\| \left\| L_n (f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty , \qquad (80)$$

etc.

Therefore we get

$$\left\| \left\| L_n^k (f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty , \ \ \forall \ k \in \mathbb{N}, \qquad (81)$$

the contraction property.

Also we see that

$$\left\| \left\| L_n^k (f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| L_n^{k-1} (f) \right\|_\gamma \right\|_\infty \leq ... \leq \left\| \left\| L_n (f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty . \qquad (82)$$

Here L_n^k are bounded linear operators.

Notation 22 *Here $N \in \mathbb{N}$, $0 < \beta < 1$. Denote by*

$$c_N := \begin{cases} \left(\frac{2\pi}{\arctan(\pi\lambda)} \right)^N , & \text{if } L_n = A_n, \\ 1, & \text{if } L_n = B_n, C_n, D_n, \end{cases} \qquad (83)$$

$$\varphi (n) := \begin{cases} \frac{1}{n^\beta}, & \text{if } L_n = A_n, \ B_n, \\ \frac{1}{n} + \frac{1}{n^\beta}, & \text{if } L_n = C_n, D_n, \end{cases} \qquad (84)$$

$$\Omega := \begin{cases} C \left(\prod_{i=1}^N [a_i, b_i] , X \right) , & \text{if } L_n = A_n, \\ C_B \left(\mathbb{R}^N, X \right) , & \text{if } L_n = B_n, C_n, D_n, \end{cases} \qquad (85)$$

and

$$Y := \begin{cases} \prod_{i=1}^N [a_i, b_i] , & \text{if } L_n = A_n, \\ \mathbb{R}^N, & \text{if } L_n = B_n, C_n, D_n. \end{cases} \qquad (86)$$

We give the condensed

Theorem 23 *Let $f \in \Omega$, $0 < \beta < 1$, $x \in Y$; n, $N \in \mathbb{N}$ with $n^{1-\beta} > 2$. Then*

(i)

$$\left\| L_n (f, x) - f (x) \right\|_\gamma \leq c_N \left[\omega_1 (f, \varphi (n)) + \left(1 - h_\lambda \left(n^{1-\beta} - 2 \right) \right) \left\| \left\| f \right\|_\gamma \right\|_\infty \right] =: \tau_\lambda (n) , \qquad (87)$$

where ω_1 is for $p = \infty$,

and

(ii)

$$\left\| \left\| L_n (f) - f \right\|_\gamma \right\|_\infty \leq \tau_\lambda (n) \to 0, \text{ as } n \to \infty. \qquad (88)$$

For f uniformly continuous and in Ω we obtain

$$\lim_{n \to \infty} L_n (f) = f,$$

pointwise and uniformly.

Proof. By Theorems 9, 16, 17, 18. ■

Next we do iterated multilayer neural network approximation (see also [9]).

We make

Remark 24 *Let $r \in \mathbb{N}$ and L_n as above. We observe that*

$$L_n^r f - f = \left(L_n^r f - L_n^{r-1} f\right) + \left(L_n^{r-1} f - L_n^{r-2} f\right) +$$

$$\left(L_n^{r-2} f - L_n^{r-3} f\right) + ... + \left(L_n^2 f - L_n f\right) + \left(L_n f - f\right).$$

Then

$$\left\| \left\| L_n^r f - f \right\|_\gamma \right\|_\infty \leq \left\| \left\| L_n^r f - L_n^{r-1} f \right\|_\gamma \right\|_\infty + \left\| \left\| L_n^{r-1} f - L_n^{r-2} f \right\|_\gamma \right\|_\infty +$$

$$\left\| \left\| L_n^{r-2} f - L_n^{r-3} f \right\|_\gamma \right\|_\infty + ... + \left\| \left\| L_n^2 f - L_n f \right\|_\gamma \right\|_\infty + \left\| \left\| L_n f - f \right\|_\gamma \right\|_\infty =$$

$$\left\| \left\| L_n^{r-1} \left(L_n f - f\right) \right\|_\gamma \right\|_\infty + \left\| \left\| L_n^{r-2} \left(L_n f - f\right) \right\|_\gamma \right\|_\infty + \left\| \left\| L_n^{r-3} \left(L_n f - f\right) \right\|_\gamma \right\|_\infty$$

$$+ ... + \left\| \left\| L_n \left(L_n f - f\right) \right\|_\gamma \right\|_\infty + \left\| \left\| L_n f - f \right\|_\gamma \right\|_\infty \leq r \left\| \left\| L_n f - f \right\|_\gamma \right\|_\infty. \quad (89)$$

That is

$$\left\| \left\| L_n^r f - f \right\|_\gamma \right\|_\infty \leq r \left\| \left\| L_n f - f \right\|_\gamma \right\|_\infty. \quad (90)$$

We give the following multilayer neural network approximation.

Theorem 25 *All here as in Theorem 23 and $r \in \mathbb{N}$, $\tau_\lambda(n)$ as in (87). Then*

$$\left\| \left\| L_n^r f - f \right\|_\gamma \right\|_\infty \leq r \tau_\lambda(n). \quad (91)$$

So that the speed of convergence to the unit operator of L_n^r is not worse than of L_n.

Proof. By (90) and (88). ■

We make

Remark 26 *Let $m_1, ..., m_r \in \mathbb{N} : m_1 \leq m_2 \leq ... \leq m_r$, $0 < \beta < 1$, $f \in \Omega$. Then $\varphi(m_1) \geq \varphi(m_2) \geq ... \geq \varphi(m_r)$, φ as in (84).*

Therefore

$$\omega_1(f, \varphi(m_1)) \geq \omega_1(f, \varphi(m_2)) \geq ... \geq \omega_1(f, \varphi(m_r)). \quad (92)$$

Assume further that $m_i^{1-\beta} > 2$, $i = 1, ..., r$. Then

$$\frac{1 - h_\lambda\left(m_1^{1-\beta} - 2\right)}{2} \geq \frac{1 - h_\lambda\left(m_2^{1-\beta} - 2\right)}{2} \geq ... \geq \frac{1 - h_\lambda\left(m_r^{1-\beta} - 2\right)}{2}. \quad (93)$$

Let L_{m_i} as above, $i = 1, ..., r$, all of the same kind.

We write

$$L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right) - f =$$
$$L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}f\right)\right) +$$
$$L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}f\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_3}f\right)\right) +$$
$$L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_3}f\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_4}f\right)\right) + ... + \tag{94}$$
$$L_{m_r}\left(L_{m_{r-1}}f\right) - L_{m_r}f + L_{m_r}f - f =$$
$$L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\right)\right)\left(L_{m_1}f - f\right) + L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_3}\right)\right)\left(L_{m_2}f - f\right) +$$
$$L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_4}\right)\right)\left(L_{m_3}f - f\right) + ... + L_{m_r}\left(L_{m_{r-1}}f - f\right) + L_{m_r}f - f.$$

Hence by the triangle inequality property of $\left\|\|\cdot\|_\gamma\right\|_\infty$ we get

$$\left\|\left\|L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right) - f\right\|_\gamma\right\|_\infty \leq$$
$$\left\|\left\|L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\right)\right)\left(L_{m_1}f - f\right)\right\|_\gamma\right\|_\infty +$$
$$\left\|\left\|L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_3}\right)\right)\left(L_{m_2}f - f\right)\right\|_\gamma\right\|_\infty +$$
$$\left\|\left\|L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_4}\right)\right)\left(L_{m_3}f - f\right)\right\|_\gamma\right\|_\infty + ... +$$
$$\left\|\left\|L_{m_r}\left(L_{m_{r-1}}f - f\right)\right\|_\gamma\right\|_\infty + \left\|\left\|L_{m_r}f - f\right\|_\gamma\right\|_\infty$$

(repeatedly applying (80))

$$\leq \left\|\left\|L_{m_1}f - f\right\|_\gamma\right\|_\infty + \left\|\left\|L_{m_2}f - f\right\|_\gamma\right\|_\infty + \left\|\left\|L_{m_3}f - f\right\|_\gamma\right\|_\infty + ... +$$
$$\left\|\left\|L_{m_{r-1}}f - f\right\|_\gamma\right\|_\infty + \left\|\left\|L_{m_r}f - f\right\|_\gamma\right\|_\infty = \sum_{i=1}^{r}\left\|\left\|L_{m_i}f - f\right\|_\gamma\right\|_\infty. \tag{95}$$

That is, we proved

$$\left\|\left\|L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right) - f\right\|_\gamma\right\|_\infty \leq \sum_{i=1}^{r}\left\|\left\|L_{m_i}f - f\right\|_\gamma\right\|_\infty. \tag{96}$$

We give the following multi layer neural network general approximation result.

Theorem 27 *Let $f \in \Omega$; N, $m_1, m_2, ..., m_r \in \mathbb{N} : m_1 \leq m_2 \leq ... \leq m_r$, $0 < \beta < 1$; $m_i^{1-\beta} > 2$, $i = 1, ..., r$, $x \in Y$, and let $(L_{m_1}, ..., L_{m_r})$ as $(A_{m_1}, ..., A_{m_r})$ or $(B_{m_1}, ..., B_{m_r})$ or $(C_{m_1}, ..., C_{m_r})$ or $(D_{m_1}, ..., D_{m_r})$, $p = \infty$. Then*

$$\left\| L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right)(x) - f(x)\right\|_\gamma \leq$$

$$\left\| \left\| L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right) - f\right\|_\gamma \right\|_\infty \leq$$

$$\sum_{i=1}^{r} \left\| \left\| L_{m_i}f - f\right\|_\gamma \right\|_\infty \leq$$

$$c_N \sum_{i=1}^{r} \left[\omega_1\left(f, \varphi\left(m_i\right)\right) + \left(1 - h_\lambda\left(m_i^{1-\beta} - 2\right)\right) \left\| \left\| f \right\|_\gamma \right\|_\infty\right] \leq$$

$$r c_N \left[\omega_1\left(f, \varphi\left(m_1\right)\right) + \left(1 - h_\lambda\left(m_1^{1-\beta} - 2\right)\right) \left\| \left\| f \right\|_\gamma \right\|_\infty\right]. \tag{97}$$

Clearly, we notice that the speed of convergence to the unit operator of the iterated multilayer neural network operator is not worse than the speed of L_{m_1}.

Proof. Using (96), (92), (93) and (87), (88). ■

We continue with

Theorem 28 *Let all as in Corollary 15, and $r \in \mathbb{N}$. Here $\varphi_3(n)$ is as in (72). Then*

$$\left\| \left\| A_n^r f - f\right\|_\gamma \right\|_\infty \leq r \left\| \left\| A_n f - f\right\|_\gamma \right\|_\infty \leq r\varphi_3(n). \tag{98}$$

Proof. By (90) and (72). ■

Application 29 *A typical application of all of our results is when $\left(X, \|\cdot\|_\gamma\right) = (\mathbb{C}, |\cdot|)$, where $\mathbb{C}$ are the complex numbers.*

References

[1] G.A. Anastassiou, *Moments in Probability and Approximation Theory*, Pitman Research Notes in Math., Vol. 287, Longman Sci. & Tech., Harlow, U.K., 1993.

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

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

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

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

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

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

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

[9] G.A. Anastassiou, *Approximation by neural networks iterates,* Advances in Applied Mathematics and Approximation Theory, pp. 1-20, Springer Proceedings in Math. & Stat., Springer, New York, 2013, Eds. G. Anastassiou, O. Duman.

[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, General Multivariate arctangent function activated neural network approximations, J. Numer. Anal. Approx Theory, 51(1) (2022), 37-66.

[13] G.A. Anastassiou, *General sigmoid based Banach space valued neural network approximation*, J. of Computational Analysis and Applications, accepted, 2022.

[14] G.A. Anastassiou, *General sigmoid based Banach space valued neural network multivariate approximations*, submitted, 2022.

[15] H. Cartan, *Differential Calculus*, Hermann, Paris, 1971.

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

[17] D. Costarelli, R. Spigler, *Approximation results for neural network operators activated by sigmoidal functions*, Neural Networks 44 (2013), 101-106.

[18] D. Costarelli, R. Spigler, *Multivariate neural network operators with sigmoidal activation functions*, Neural Networks 48 (2013), 72-77.

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

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

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

[22] L.B. Rall, *Computational Solution of Nonlinear Operator Equations*, John Wiley & Sons, New York, 1969.