J. Appl. & Pure Math. Vol. **5***(2023), No. 1 - 2, pp. 69 - 93*
https://doi.org/10.23091/japm.2023.069

PARAMETRIZED GUDERMANNIAN FUNCTION RELIED BANACH SPACE VALUED NEURAL NETWORK MULTIVARIATE APPROXIMATIONS

GEORGE A. ANASTASSIOU

Abstract. Here we give 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 treat also the case of approximation by iterated operators of the last four types. These approximations are derived 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 Gudermannian sigmoid function. The approximations are pointwise and uniform. The related feed-forward neural network is with one hidden layer.

AMS Mathematics Subject Classification : 41A17, 41A25, 41A30, 41A36.
Key words and phrases : Parametrized Gudermannian sigmoid function, multivariate neural network approximation, quasi-interpolation operator, Kantorovich type operator, quadrature type operator, multivariate modulus of continuity, iterated 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, 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

Received November 7, 2022. Revised February 8, 2023. Accepted February 15, 2023.

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 [15] of Z. Chen and F. Cao, also by [4]-[12], [16], [17].

The author here performs multivariate parametrized Gudermannian 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 iterated 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 Gudermannian 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 Gudermannian sigmoid function. About neural networks read [18], [19], [20].

2. Background

Here we consider the Gudermannian function ([22]) $gd(x)$ which is defined as follows

$$gd(x) := \int_0^x \frac{dt}{\cosh t} = 2\arctan\left(\tanh\left(\frac{x}{2}\right)\right), \ \forall\, x \in \mathbb{R}. \tag{1}$$

Let $\lambda > 0$, then

$$gd(\lambda x) = \int_0^{\lambda x} \frac{dt}{\cosh t} = 2\arctan\left(\tanh\left(\frac{\lambda x}{2}\right)\right). \tag{2}$$

We will use the following normalized and parametrized function

$$f_\lambda(x) := \frac{2}{\pi} gd(\lambda x) = \frac{4}{\pi}\arctan\left(\tanh\left(\frac{\lambda x}{2}\right)\right) = \tag{3}$$

$$\frac{2}{\pi}\int_0^{\lambda x} \frac{dt}{\cosh t} = \frac{4}{\pi}\int_0^{\lambda x} \frac{dt}{e^t + e^{-t}}, \ x \in \mathbb{R}$$

We will prove that f_λ is a generator sigmoid function with the general properties as in [13]. When $0<\lambda<1$, f_λ is expected to outperform ReLu and Leaky ReLu activation functions.

We notice that

$$\left(\frac{2}{\pi}gd\left(x\right)\right)'=\frac{2}{\pi\cosh x}>0,$$

and

$$f_\lambda'\left(x\right)=\left(\frac{2}{\pi}gd\left(\lambda x\right)\right)'=\frac{2\lambda}{\pi\cosh\lambda x}>0,\quad\forall\, x\in\mathbb{R}. \tag{4}$$

Hence f_λ is strictly increasing on $\mathbb{R}$.

Furthermore we have

$$f_\lambda''\left(x\right)=-\frac{2\lambda^2}{\pi}\frac{\sinh\lambda x}{\left(\cosh\lambda x\right)^2},\quad\forall\, x\in\mathbb{R}. \tag{5}$$

Notice that

$$f_\lambda''\left(x\right)>0 \;\text{ for } x<0, \text{ and}$$
$$f_\lambda''\left(x\right)<0 \;\text{ for } x>0, \text{ and}$$
$$f_\lambda''\left(0\right)=0.$$

Therefore f_λ is stritly concave up for $x<0$, and f_λ is striclty concave down for $x>0$, and $f_\lambda\left(0\right)=0$, with $(0,0)$ the inflection point.

Let $x\to+\infty$, then $\tanh\left(\frac{\lambda x}{2}\right)\to1$ and $\arctan\left(\tanh\left(\frac{\lambda x}{2}\right)\right)\to\frac{\pi}{4}$. Let $x\to-\infty$, then $\tanh\left(\frac{\lambda x}{2}\right)\to-1$ and $\arctan\left(\tanh\left(\frac{\lambda x}{2}\right)\right)\to-\frac{\pi}{4}$.

Clearly, then $f_\lambda\left(+\infty\right)=1$ and $f_\lambda\left(-\infty\right)=-1$, so that $y=\pm1$ are horizontal asymptotes for f_λ.

Also it is $f_\lambda\left(x\right)\geq0$ for $x\geq0$, and $f_\lambda\left(x\right)<0$ for $x<0$. Obviously then $f_\lambda:\mathbb{R}\to[-1,1]$, with $f_\lambda''\in C\left(\mathbb{R}\right)$.

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

We have that

$$f_\lambda\left(-x\right)=\frac{4}{\pi}\arctan\left(\tanh\left(-\frac{\lambda x}{2}\right)\right)=\frac{4}{\pi}\arctan\left(-\tanh\left(\frac{\lambda x}{2}\right)\right)=$$
$$-\frac{4}{\pi}\arctan\left(\tanh\left(\frac{\lambda x}{2}\right)\right)=-f_\lambda\left(x\right),$$

proving

$$f_\lambda\left(-x\right)=-f_\lambda\left(x\right),\quad\forall\, x\in\mathbb{R}. \tag{6}$$

So, indeed, f_λ is a sigmoid function as in [13].

So, all the theory of [13] applies here for f_λ, etc.

We consider the activation function

$$\psi\left(x\right):=\frac{1}{4}\left(f_\lambda\left(x+1\right)-f_\lambda\left(x-1\right)\right),\quad x\in\mathbb{R}, \tag{7}$$

As in [11], p. 285, and [13], we get that $\psi\left(-x\right)=\psi\left(x\right)$, thus ψ is an even function. Since $x+1>x-1$, then $f_\lambda\left(x+1\right)>f_\lambda\left(x-1\right)$, and $\psi\left(x\right)>0$, all $x\in\mathbb{R}$.

We see that

$$\psi(0) = \frac{f_\lambda(1)}{2} = \frac{gd(\lambda)}{\pi}. \quad (8)$$

Let $x > 1$, we have that

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

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

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

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

See that

$$\lim_{x \to +\infty} \psi(x) = \frac{1}{4}\left(f_\lambda(+\infty) - f_\lambda(+\infty)\right) = 0, \quad (9)$$

and

$$\lim_{x \to -\infty} \psi(x) = \frac{1}{4}\left(f_\lambda(-\infty) - f_\lambda(-\infty)\right) = 0. \quad (10)$$

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

Conclusion, ψ is a bell symmetric function with maximum

$$\psi(0) = \frac{gd(\lambda)}{\pi}.$$

We need

Theorem 2.1. *(by [13]) We have that*

$$\sum_{i=-\infty}^{\infty} \psi(x - i) = 1, \quad \forall\, x \in \mathbb{R}. \quad (11)$$

Theorem 2.2. *(by [13]) It holds*

$$\int_{-\infty}^{\infty} \psi(x)\, dx = 1. \quad (12)$$

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

We give

Theorem 2.3. *(by [13]) Let $0 < \alpha < 1$, and $n \in \mathbb{N}$ with $n^{1-\alpha} > 2$. It holds*

$$\sum_{\begin{cases} k = -\infty \\ : |nx - k| \geq n^{1-\alpha} \end{cases}}^{\infty} \psi(nx - k) < \frac{\left(1 - f_\lambda\left(n^{1-\alpha} - 2\right)\right)}{2} = \frac{\left(\pi - 2gd\left(\lambda\left(n^{1-\alpha} - 2\right)\right)\right)}{2\pi}. \quad (13)$$

Notice that

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

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 2.4. *(by [13]) Let $x \in [a,b] \subset \mathbb{R}$ and $n \in \mathbb{N}$ so that $\lceil na \rceil \leq \lfloor nb \rfloor$. It holds*

$$\frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi\left(nx-k\right)} < \frac{1}{\psi\left(1\right)} = \frac{4}{f_{\lambda}\left(2\right)} = \frac{2\pi}{gd\left(2\lambda\right)}, \quad \forall\ x \in \left[a,b\right]. \tag{14}$$

Remark 2.1. (by [13]) We have that

$$\lim_{n\to\infty} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi\left(nx-k\right) \neq 1, \tag{15}$$

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

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

Note 2.1. 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 (11))

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi\left(nx-k\right) \leq 1. \tag{16}$$

We introduce

$$Z\left(x_1,...,x_N\right) := Z\left(x\right) := \prod_{i=1}^{N} \psi\left(x_i\right), \quad x = \left(x_1,...,x_N\right) \in \mathbb{R}^N,\ N \in \mathbb{N}. \tag{17}$$

It has the properties:

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

(ii)

$$\sum_{k=-\infty}^{\infty} Z\left(x-k\right) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z\left(x_1-k_1,...,x_N-k_N\right) = 1, \tag{18}$$

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

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z\left(nx-k\right) = 1, \tag{19}$$

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

and

(iv)

$$\int_{\mathbb{R}^N} Z\left(x\right) dx = 1, \tag{20}$$

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

$$\lceil na \rceil := (\lceil na_1 \rceil, ..., \lceil na_N \rceil), \qquad \lfloor nb \rfloor := (\lfloor nb_1 \rfloor, ..., \lfloor nb_N \rfloor), \tag{21}$$

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

We obviously see that

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

$$\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \cdots \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} \left(\prod_{i=1}^{N} \psi(nx_i - k_i) \right) = \prod_{i=1}^{N} \left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \psi(nx_i - k_i) \right). \tag{22}$$

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

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

$$\sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty \le \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} Z(nx-k) + \sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} Z(nx-k). \tag{23}$$

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 [11], pp. 288-289, we derive that

$$\sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} Z(nx-k) \overset{(13)}{<} \frac{1 - f_\lambda\left(n^{1-\beta} - 2\right)}{2}, \quad 0 < \beta < 1, \tag{24}$$

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

(vi) By Theorem 2.4 we get that

$$0 < \frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)} < \frac{1}{(\psi(1))^N} = \left(\frac{2\pi}{gd(2\lambda)} \right)^N, \tag{25}$$

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

It is also clear that

(vii)

$$\sum_{\begin{cases} k=-\infty \\ \left\| \frac{k}{n}-x \right\|_{\infty} > \frac{1}{n^{\beta}} \end{cases}}^{\infty} Z(nx-k) < \frac{1-f_{\lambda}\left(n^{1-\beta}-2\right)}{2}, \tag{26}$$

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

Furthermore it holds

$$\lim_{n\to\infty} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k) \neq 1, \tag{27}$$

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

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

Let $f \in C\left(\prod_{i=1}^{N} [a_i, b_i], X\right)$, $x = (x_1, ..., x_N) \in \prod_{i=1}^{N} [a_i, b_i]$, $n \in \mathbb{N}$ 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 ($x := (x_1, ..., x_N) \in \left(\prod_{i=1}^{N} [a_i, b_i]\right)$):

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

$$\frac{\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \sum_{k_2=\lceil na_2 \rceil}^{\lfloor nb_2 \rfloor} ... \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} f\left(\frac{k_1}{n}, ..., \frac{k_N}{n}\right) \left(\prod_{i=1}^{N} \psi(nx_i - k_i)\right)}{\prod_{i=1}^{N} \left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \psi(nx_i - k_i)\right)}. \tag{28}$$

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

$$\widetilde{A}_n(g, x) := \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} g\left(\frac{k}{n}\right) Z(nx-k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)}. \tag{29}$$

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

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

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

Furthermore it holds

$$\|A_n(f, x)\|_{\gamma} \leq \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\|f\left(\frac{k}{n}\right)\right\|_{\gamma} Z(nx-k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)} = \widetilde{A}_n\left(\|f\|_{\gamma}, x\right), \tag{30}$$

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

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

So, we have that

$$\|A_n(f,x)\|_{\gamma} \leq \widetilde{A}_n\left(\|f\|_{\gamma}, x\right), \tag{31}$$

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

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

$$A_n(cg, x) = c\widetilde{A}_n(g, x), \ \ \forall\ x \in \prod_{i=1}^{N}[a_i, b_i]. \tag{32}$$

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

$$A_n(c) = c, \ \ \forall\ c \in X. \tag{33}$$

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

For convinience we call

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

$$\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \sum_{k_2=\lceil na_2 \rceil}^{\lfloor nb_2 \rfloor} \cdots \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} f\left(\frac{k_1}{n}, ..., \frac{k_N}{n}\right)\left(\prod_{i=1}^{N} \psi(nx_i - k_i)\right), \tag{34}$$

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

That is

$$A_n(f,x) := \frac{A_n^*(f,x)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)}, \tag{35}$$

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

Hence

$$A_n(f,x) - f(x) = \frac{A_n^*(f,x) - f(x)\left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)\right)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)}. \tag{36}$$

Consequently we derive

$$\|A_n(f,x) - f(x)\|_{\gamma} \overset{(25)}{\leq} \left(\frac{2\pi}{gd(2\lambda)}\right)^N \left\| A_n^*(f,x) - f(x) \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k) \right\|_{\gamma}, \tag{37}$$

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

We will estimate the right hand side of (37).

For the last and others we need

Definition 2.5. ([11], p. 274) Let M be a convex and compact subset of $\left(\mathbb{R}^N, \|\cdot\|_p\right)$, $p \in [1, \infty]$, and $\left(X, \|\cdot\|_\gamma\right)$ be a Banach space. Let $f \in C(M, X)$. We define the first modulus of continuity of f as

$$\omega_1(f, \delta) := \sup_{\substack{x, y \in M: \\ \|x-y\|_p \le \delta}} \|f(x) - f(y)\|_\gamma, \quad 0 < \delta \le diam(M). \tag{38}$$

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

$$\omega_1(f, \delta) = \omega_1(f, diam(M)). \tag{39}$$

Notice $\omega_1(f, \delta)$ is increasing in $\delta > 0$. For $f \in C_B(M, X)$ (continuous and bounded functions) $\omega_1(f, \delta)$ is defined similarly.

Lemma 2.6. *([11], p. 274) We have $\omega_1(f, \delta) \to 0$ as $\delta \downarrow 0$, iff $f \in C(M, X)$, where M is a convex compact subset of $\left(\mathbb{R}^N, \|\cdot\|_p\right)$, $p \in [1, \infty]$.*

Clearly we have also: $f \in C_U\left(\mathbb{R}^N, X\right)$ (uniformly continuous functions), iff $\omega_1(f, \delta) \to 0$ as $\delta \downarrow 0$, where ω_1 is defined similarly to (38). 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,

$$B_n(f, x) := B_n(f, x_1, ..., x_N) := \sum_{k=-\infty}^{\infty} f\left(\frac{k}{n}\right) Z(nx - k) :=$$

$$\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} f\left(\frac{k_1}{n}, \frac{k_2}{n}, ..., \frac{k_N}{n}\right) \left(\prod_{i=1}^{N} \psi(nx_i - k_i)\right), \tag{40}$$

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

$$C_n(f, x) := C_n(f, x_1, ..., x_N) := \sum_{k=-\infty}^{\infty} \left(n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\, dt\right) Z(nx - k) =$$

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

$$\cdot \left(\prod_{i=1}^{N} \psi(nx_i - k_i)\right), \tag{41}$$

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

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 $D_n(f, x)$, $n \in \mathbb{N}$, as follows.

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

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

$$\sum_{r_1=0}^{\theta_1} \sum_{r_2=0}^{\theta_2} ... \sum_{r_N=0}^{\theta_N} w_{r_1, r_2, ... r_N} f\left(\frac{k_1}{n} + \frac{r_1}{n\theta_1}, \frac{k_2}{n} + \frac{r_2}{n\theta_2}, ..., \frac{k_N}{n} + \frac{r_N}{n\theta_N}\right), \quad (42)$$

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

We set

$$D_n(f, x) := D_n(f, x_1, ..., x_N) := \sum_{k=-\infty}^{\infty} \delta_{nk}(f) Z(nx - k) = \quad (43)$$

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

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

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 general sigmoid Neural Network Approximations

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

We give

Theorem 3.1. *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)

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

$$\left(\frac{2\pi}{gd(2\lambda)}\right)^N \left[\omega_1\left(f, \frac{1}{n^\beta}\right) + \left(1 - f_\lambda\left(n^{1-\beta} - 2\right)\right) \left\| \|f\|_\gamma \right\|_\infty\right] =: \lambda_1(n), \quad (44)$$

and

2)

$$\left\| \|A_n(f) - f\|_\gamma \right\|_\infty \leq \lambda_1(n). \quad (45)$$

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* $\max\left(\frac{1}{n^\beta}, \left(1 - f_\lambda\left(n^{1-\beta} - 2\right)\right)\right)$.

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

We make

Remark 3.1. ([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

$$\|g\| := \|g\|_{L_j} := \sup_{\left(\|x\|_{(\mathbb{R}^N)^j}=1\right)} \|g(x)\|_\gamma = \sup \frac{\|g(x)\|_\gamma}{\|x_1\|_p ... \|x_j\|_p}. \tag{46}$$

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 [21]) $f^{(j)} : O \to L_j = L_j\left(\left(\mathbb{R}^N\right)^j; X\right)$ 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 ([14]), ([21], p. 124), we get

$$f(x) = \sum_{j=0}^{m} \frac{f^{(j)}(x_0)(x - x_0)^j}{j!} + R_m(x, x_0), \quad \text{all } x \in M, \tag{47}$$

where the remainder is the Riemann integral

$$R_m(x, x_0) := \int_0^1 \frac{(1-u)^{m-1}}{(m-1)!} \left(f^{(m)}(x_0 + u(x - x_0)) - f^{(m)}(x_0)\right)(x - x_0)^m du, \tag{48}$$

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

We consider

$$w := \omega_1\left(f^{(m)}, h\right) := \sup_{\substack{x,y \in M: \\ \|x-y\|_p \leq h}} \left\|f^{(m)}(x) - f^{(m)}(y)\right\|, \tag{49}$$

$h > 0$.

We obtain

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

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

$$w \|x - x_0\|_p^m \left\lceil \frac{u \|x - x_0\|_p}{h} \right\rceil, \tag{50}$$

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):

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

$$= w\Phi_m\left(\|x-x_0\|_p\right) \tag{51}$$

by a change of variable, where

$$\Phi_m(t) := \int_0^{|t|} \left\lceil \frac{s}{h} \right\rceil \frac{(|t|-s)^{m-1}}{(m-1)!} ds = \frac{1}{m!}\left(\sum_{j=0}^{\infty} (|t|-jh)_+^m\right), \ \forall\, t \in \mathbb{R}, \tag{52}$$

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

Also from there we get

$$\Phi_m(t) \le \left(\frac{|t|^{m+1}}{(m+1)!h} + \frac{|t|^m}{2m!} + \frac{h|t|^{m-1}}{8(m-1)!}\right), \ \ \forall\, t \in \mathbb{R}, \tag{53}$$

with equality true only at $t=0$.

Therefore it holds

$$\|R_m(x,x_0)\|_\gamma \le w\left(\frac{\|x-x_0\|_p^{m+1}}{(m+1)!h} + \frac{\|x-x_0\|_p^m}{2m!} + \frac{h\|x-x_0\|_p^{m-1}}{8(m-1)!}\right), \ \ \forall\, x \in M. \tag{54}$$

We have found that

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

$$\omega_1\left(f^{(m)},h\right)\left(\frac{\|x-x_0\|_p^{m+1}}{(m+1)!h} + \frac{\|x-x_0\|_p^m}{2m!} + \frac{h\|x-x_0\|_p^{m-1}}{8(m-1)!}\right) < \infty, \tag{55}$$

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

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

One can rewrite (55) as follows:

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

$$\omega_1\left(f^{(m)},h\right)\left(\frac{\|\cdot-x_0\|_p^{m+1}}{(m+1)!h} + \frac{\|\cdot-x_0\|_p^m}{2m!} + \frac{h\|\cdot-x_0\|_p^{m-1}}{8(m-1)!}\right), \ \forall\, x_0 \in M, \tag{56}$$

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

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

$$\omega_1\left(f^{(m)}, h\right) \left[\frac{\left(\widetilde{L}_N\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)}{(m+1)!h} + \frac{\left(\widetilde{L}_N\left(\|\cdot - x_0\|_p^{m}\right)\right)(x_0)}{2m!} + \frac{h\left(\widetilde{L}_N\left(\|\cdot - x_0\|_p^{m-1}\right)\right)(x_0)}{8(m-1)!} \right], \tag{57}$$

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

Clearly (57) is valid when $M = \prod_{i=1}^{N} [a_i, b_i]$ and $\widetilde{L}_n = \widetilde{A}_n$, see (29).

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 (28), (29), (31), (32) and (33).

We present the following high order approximation results.

Theorem 3.2. *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)

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

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

$$\left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8}\right], \tag{58}$$

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

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

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

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

3)

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

$$\frac{\omega_1\left(f^{(m)}, r\left(\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right)^{\frac{1}{m+1}}\right)}{rm!}\left(\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right)^{\left(\frac{m}{m+1}\right)} \tag{60}$$

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

and

4)

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

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

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

$$\left\|\left(\widetilde{A}_n\left(\|\cdot - x_0\|_p^{m+1}\right)\right)(x_0)\right\|_{\infty, x_0 \in \prod_{i=1}^{N}[a_i, b_i]}^{\left(\frac{m}{m+1}\right)} \tag{61}$$

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

We give

Corollary 3.3. *(to Theorem 3.2, case of $m = 1$) Then*
1)

$$\left\| \left(A_n (f) \right) (x_0) - f (x_0) \right\|_\gamma \leq \left\| \left(A_n \left(f^{(1)} (x_0) (\cdot - x_0) \right) \right) (x_0) \right\|_\gamma +$$
$$\frac{1}{2r} \omega_1 \left(f^{(1)}, r \left(\left(\widetilde{A}_n \left(\| \cdot - x_0 \|_p^2 \right) \right) (x_0) \right)^{\frac{1}{2}} \right) \left(\left(\widetilde{A}_n \left(\| \cdot - x_0 \|_p^2 \right) \right) (x_0) \right)^{\frac{1}{2}} \tag{62}$$
$$\left[1 + r + \frac{r^2}{4} \right],$$

and
2)

$$\left\| \left\| \left(A_n (f) \right) - f \right\|_\gamma \right\|_{\infty, \prod_{i=1}^N [a_i, b_i]} \leq$$
$$\left\| \left\| \left(A_n \left(f^{(1)} (x_0) (\cdot - x_0) \right) \right) (x_0) \right\|_\gamma \right\|_{\infty, x_0 \in \prod_{i=1}^N [a_i, b_i]} +$$
$$\frac{1}{2r} \omega_1 \left(f^{(1)}, r \left\| \left(\widetilde{A}_n \left(\| \cdot - x_0 \|_p^2 \right) \right) (x_0) \right\|_{\infty, x_0 \in \prod_{i=1}^N [a_i, b_i]}^{\frac{1}{2}} \right)$$
$$\left\| \left(\widetilde{A}_n \left(\| \cdot - x_0 \|_p^2 \right) \right) (x_0) \right\|_{\infty, x_0 \in \prod_{i=1}^N [a_i, b_i]}^{\frac{1}{2}} \left[1 + r + \frac{r^2}{4} \right], \tag{63}$$

$r > 0$.

We make

Remark 3.2. We estimate ($0 < \alpha < 1$, $m, n \in \mathbb{N} : n^{1-\alpha} > 2$),

$$\widetilde{A}_n \left(\| \cdot - x_0 \|_\infty^{m+1} \right) (x_0) = \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z (nx_0 - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z (nx_0 - k)} \overset{(25)}{<}$$
$$\left(\frac{2\pi}{gd (2\lambda)} \right)^N \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z (nx_0 - k) = \tag{64}$$
$$\left(\frac{2\pi}{gd (2\lambda)} \right)^N \left\{ \sum_{\substack{k = \lceil na \rceil \\ \left\{ : \left\| \frac{k}{n} - x_0 \right\|_\infty \leq \frac{1}{n^\alpha} \right.}}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z (nx_0 - k) + \right.$$

$$\left. \sum_{\left\{ \begin{array}{c} k = \lceil na \rceil \\ : \left\| \frac{k}{n} - x_0 \right\|_\infty > \frac{1}{n^\alpha} \end{array} \right.}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z(nx_0 - k) \right\} \overset{(26)}{\leq}$$

$$\left(\frac{2\pi}{gd(2\lambda)} \right)^N \left\{ \frac{1}{n^{\alpha(m+1)}} + \left(\frac{1 - f_\lambda(n^{1-\alpha} - 2)}{2} \right) \|b - a\|_\infty^{m+1} \right\}, \quad (65)$$

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

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

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

$$\left(\frac{2\pi}{gd(2\lambda)} \right)^N \left\{ \frac{1}{n^{\alpha(m+1)}} + \left(\frac{1 - f_\lambda(n^{1-\alpha} - 2)}{2} \right) \|b - a\|_\infty^{m+1} \right\} =: \varphi_1(n) \quad (66)$$

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

And, consequently it holds

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

$$\left(\frac{2\pi}{gd(2\lambda)} \right)^N \left\{ \frac{1}{n^{\alpha(m+1)}} + \left(\frac{1 - f_\lambda(n^{1-\alpha} - 2)}{2} \right) \|b - a\|_\infty^{m+1} \right\} = \varphi_1(n) \to 0, \quad \text{as } n \to +\infty. \quad (67)$$

So, we have that $\varphi_1(n) \to 0$, as $n \to +\infty$. Thus, when $p \in [1, \infty]$, from Theorem 3.2 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

$$\left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) = \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} f^{(j)}(x_0) \left(\frac{k}{n} - x_0 \right)^j Z(nx_0 - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx_0 - k)}. \quad (68)$$

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

$$\left\| f^{(j)}(x_0) \left(\frac{k}{n} - x_0 \right)^j \right\|_\gamma \leq \left\| f^{(j)}(x_0) \right\| \left\| \frac{k}{n} - x_0 \right\|_\infty^j. \quad (69)$$

We further have that

$$\left\| \left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_\gamma \overset{(25)}{<}$$

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

$$\left(\frac{2\pi}{gd(2\lambda)}\right)^N\left(\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\|f^{(j)}(x_0)\right\|\left\|\frac{k}{n}-x_0\right\|_\infty^j Z(nx_0-k)\right)= \tag{70}$$

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

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

$$\left.+\sum_{\substack{k=\lceil na\rceil \\ \left\{:\left\|\frac{k}{n}-x_0\right\|_\infty>\frac{1}{n^\alpha}\right.}}^{\lfloor nb\rfloor}\left\|\frac{k}{n}-x_0\right\|_\infty^j Z(nx_0-k)\right\}\overset{(26)}{\leq} \tag{71}$$

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

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:

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

$$\left(\frac{2\pi}{gd(2\lambda)}\right)^N\left\|f^{(j)}(x_0)\right\|\left\{\frac{1}{n^{\alpha j}}+\left(\frac{1-f_\lambda\left(n^{1-\alpha}-2\right)}{2}\right)\|b-a\|_\infty^j\right\}\leq \tag{72}$$

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

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

We conclude:

In Theorem 3.2, the right hand sides of (60) and (61) converge to zero as $n\to\infty$, for any $p\in[1,\infty]$.

Also in Corollary 3.3, the right hand sides of (62) and (63) converge to zero as $n\to\infty$, for any $p\in[1,\infty]$.

Conclusion 3.1. *We have proved that the left hand sides of (58), (59), (60), (61) and (62), (63) 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 (59). Higher speed of convergence happens also to the left hand side of (58).*

We give

Corollary 3.4. *(to Theorem 3.2) 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 (66) and $\varphi_{2j}(n)$ as in (72), where $n \in \mathbb{N} : n^{1-\alpha} > 2$, $0 < \alpha < 1$, $j = 1, ..., m$. Then*

1)

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

$$\frac{\omega_1\left(f^{(m)}, r(\varphi_1(n))^{\frac{1}{m+1}}\right)}{rm!} (\varphi_1(n))^{\left(\frac{m}{m+1}\right)} \left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8} \right], \tag{73}$$

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

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

$$\frac{\omega_1\left(f^{(m)}, r(\varphi_1(n))^{\frac{1}{m+1}}\right)}{rm!} (\varphi_1(n))^{\left(\frac{m}{m+1}\right)} \left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8} \right], \tag{74}$$

3)

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

$$\frac{\omega_1\left(f^{(m)}, r(\varphi_1(n))^{\frac{1}{m+1}}\right)}{rm!} (\varphi_1(n))^{\left(\frac{m}{m+1}\right)} \tag{75}$$

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

We continue with

Theorem 3.5. *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)

$$\left\| B_n(f,x) - f(x) \right\|_\gamma \leq \omega_1\left(f, \frac{1}{n^\beta}\right) + \left(1 - f_\lambda\left(n^{1-\beta} - 2\right)\right) \left\| \left\| f \right\|_\gamma \right\|_\infty =: \lambda_2(n), \tag{76}$$

2)

$$\left\| \left\| B_n(f) - f \right\|_\gamma \right\|_\infty \leq \lambda_2(n). \tag{77}$$

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* $\lim_{n\to\infty} B_n(f) = f$, *uniformly. The speed of convergence above is* $\max\left(\frac{1}{n^\beta}, \left(1 - f_\lambda\left(n^{1-\beta} - 2\right)\right)\right)$.

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

We give

Theorem 3.6. *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)

$$\left\| C_n(f,x) - f(x) \right\|_\gamma \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + \left(1 - f_\lambda\left(n^{1-\beta} - 2\right)\right) \left\| \left\| f \right\|_\gamma \right\|_\infty =: \lambda_3(n), \tag{78}$$

2)

$$\left\| \left\| C_n(f) - f \right\|_\gamma \right\|_\infty \leq \lambda_3(n). \tag{79}$$

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* $\lim_{n\to\infty} C_n(f) = f$, *uniformly.*

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

We also present

Theorem 3.7. *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)

$$\left\| D_n(f,x) - f(x) \right\|_\gamma \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + \left(1 - f_\lambda\left(n^{1-\beta} - 2\right)\right) \left\| \left\| f \right\|_\gamma \right\|_\infty = \lambda_4(n), \tag{80}$$

2)

$$\left\| \left\| D_n(f) - f \right\|_\gamma \right\|_\infty \leq \lambda_4(n). \tag{81}$$

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. ☐

We make

Definition 3.8. 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

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

$$\begin{cases} B_n(f,x), & \text{if } l_{nk}(f) = f\left(\frac{k}{n}\right), \\ C_n(f,x), & \text{if } l_{nk}(f) = n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\, dt, \\ D_n(f,x), & \text{if } l_{nk}(f) = \delta_{nk}(f). \end{cases} \tag{82}$$

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 3.9. *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 3.3. By (28) 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

$$\left\| \left\|L_n^2(f)\right\|_\gamma \right\|_\infty = \left\| \|L_n(L_n(f))\|_\gamma \right\|_\infty \leq \left\| \|L_n(f)\|_\gamma \right\|_\infty \leq \left\| \|f\|_\gamma \right\|_\infty, \tag{83}$$

etc.

Therefore we get

$$\left\| \left\|L_n^k(f)\right\|_\gamma \right\|_\infty \leq \left\| \|f\|_\gamma \right\|_\infty, \ \forall\, k \in \mathbb{N}, \tag{84}$$

the contraction property.

Also we see that

$$\left\| \left\|L_n^k(f)\right\|_\gamma \right\|_\infty \leq \left\| \left\|L_n^{k-1}(f)\right\|_\gamma \right\|_\infty \leq \ldots \leq \left\| \|L_n(f)\|_\gamma \right\|_\infty \leq \left\| \|f\|_\gamma \right\|_\infty. \tag{85}$$

Here L_n^k are bounded linear operators.

Notation 3.1. *Here $N \in \mathbb{N}$, $0 < \beta < 1$. Denote by*

$$c_N := \begin{cases} \left(\frac{2\pi}{gd(2\lambda)}\right)^N, & \text{if } L_n = A_n, \\ 1, & \text{if } L_n = B_n, C_n, D_n, \end{cases} \tag{86}$$

$$\varphi(n) := \begin{cases} \frac{1}{n^\beta}, & \text{if } L_n = A_n,\ B_n, \\ \frac{1}{n} + \frac{1}{n^\beta}, & \text{if } L_n = C_n, D_n, \end{cases} \tag{87}$$

$$\Omega := \begin{cases} C\left(\prod_{i=1}^{N}[a_i,b_i],X\right), & \text{if } L_n = A_n, \\ C_B\left(\mathbb{R}^N,X\right), & \text{if } L_n = B_n, C_n, D_n, \end{cases} \tag{88}$$

and

$$Y := \begin{cases} \prod_{i=1}^{N}[a_i,b_i], & \text{if } L_n = A_n, \\ \mathbb{R}^N, & \text{if } L_n = B_n, C_n, D_n. \end{cases} \tag{89}$$

We give the condensed

Theorem 3.10. *Let $f \in \Omega$, $0<\beta<1$, $x \in Y$; n, $N \in \mathbb{N}$ with $n^{1-\beta} > 2$. Then*
(i)

$$\|L_n(f,x) - f(x)\|_\gamma \le c_N\left[\omega_1(f,\varphi(n)) + \left(1 - f_\lambda\left(n^{1-\beta}-2\right)\right)\left\|\|f\|_\gamma\right\|_\infty\right] =: \tau(n), \tag{90}$$

where ω_1 is for $p=\infty$,
and
(ii)

$$\left\|\|L_n(f) - f\|_\gamma\right\|_\infty \le \tau(n) \to 0, \text{ as } n \to \infty. \tag{91}$$

For f uniformly continuous and in Ω we obtain

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

pointwise and uniformly.

Proof. By Theorems 3.1, 3.5, 3.6, 3.7. □

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

Remark 3.4. Let $r \in \mathbb{N}$ and L_n as above. We observe that

$$L_n^r f - f = \left(L_n^r f - L_n^{r-1} f\right) + \left(L_n^{r-1} f - L_n^{r-2} f\right) + \left(L_n^{r-2} f - L_n^{r-3} f\right) + \dots + \left(L_n^2 f - L_n f\right) + \left(L_n f - f\right).$$

Then

$$\left\|\|L_n^r f - f\|_\gamma\right\|_\infty \le \left\|\|L_n^r f - L_n^{r-1} f\|_\gamma\right\|_\infty + \left\|\|L_n^{r-1} f - L_n^{r-2} f\|_\gamma\right\|_\infty + \left\|\|L_n^{r-2} f - L_n^{r-3} f\|_\gamma\right\|_\infty + \dots + \left\|\|L_n^2 f - L_n f\|_\gamma\right\|_\infty + \left\|\|L_n f - f\|_\gamma\right\|_\infty =$$
$$\left\|\|L_n^{r-1}(L_n f - f)\|_\gamma\right\|_\infty + \left\|\|L_n^{r-2}(L_n f - f)\|_\gamma\right\|_\infty + \left\|\|L_n^{r-3}(L_n f - f)\|_\gamma\right\|_\infty + \dots + \left\|\|L_n(L_n f - f)\|_\gamma\right\|_\infty + \left\|\|L_n f - f\|_\gamma\right\|_\infty \le r\left\|\|L_n f - f\|_\gamma\right\|_\infty. \tag{92}$$

That is

$$\left\|\|L_n^r f - f\|_\gamma\right\|_\infty \le r\left\|\|L_n f - f\|_\gamma\right\|_\infty. \tag{93}$$

We give

Theorem 3.11. *All here as in Theorem 3.10 and $r \in \mathbb{N}$, $\tau(n)$ as in (90). Then*

$$\left\| \| L_n^r f - f \|_\gamma \right\|_\infty \leq r\tau(n). \tag{94}$$

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

Proof. By (93) and (91). □

We make

Remark 3.5. 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 (87).

Therefore

$$\omega_1(f, \varphi(m_1)) \geq \omega_1(f, \varphi(m_2)) \geq ... \geq \omega_1(f, \varphi(m_r)). \tag{95}$$

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

$$\frac{1 - f_\lambda\left(m_1^{1-\beta} - 2\right)}{2} \geq \frac{1 - f_\lambda\left(m_2^{1-\beta} - 2\right)}{2} \geq ... \geq \frac{1 - f_\lambda\left(m_r^{1-\beta} - 2\right)}{2}. \tag{96}$$

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

We write

$$\begin{gathered} L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1} f\right)\right)\right) - f = \\ L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1} f\right)\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2} f\right)\right) + \\ L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2} f\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_3} f\right)\right) + \\ L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_3} f\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_4} f\right)\right) + ... + \\ L_{m_r}\left(L_{m_{r-1}} f\right) - L_{m_r} f + L_{m_r} f - f = \\ L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\right)\right)\left(L_{m_1} f - f\right) + L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_3}\right)\right)\left(L_{m_2} f - f\right) + \\ L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_4}\right)\right)\left(L_{m_3} f - f\right) + ... + L_{m_r}\left(L_{m_{r-1}} f - f\right) + L_{m_r} f - f. \end{gathered} \tag{97}$$

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

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

(repeatedly applying (83))

$$\leq \left\| \| L_{m_1} f - f \|_\gamma \right\|_\infty + \left\| \| L_{m_2} f - f \|_\gamma \right\|_\infty + \left\| \| L_{m_3} f - f \|_\gamma \right\|_\infty + ... +$$

$$\left\|\left\|L_{m_{r-1}}f-f\right\|_{\gamma}\right\|_{\infty}+\left\|\left\|L_{m_r}f-f\right\|_{\gamma}\right\|_{\infty}=\sum_{i=1}^{r}\left\|\left\|L_{m_i}f-f\right\|_{\gamma}\right\|_{\infty}. \tag{98}$$

That is, we proved

$$\left\|\left\|L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right)-f\right\|_{\gamma}\right\|_{\infty}\leq\sum_{i=1}^{r}\left\|\left\|L_{m_i}f-f\right\|_{\gamma}\right\|_{\infty}. \tag{99}$$

We give

Theorem 3.12. *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*

$$\begin{gathered}\left\|L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right)(x)-f(x)\right\|_{\gamma}\leq\\ \left\|\left\|L_{m_r}\left(L_{m_{r-1}}\left(...L_{m_2}\left(L_{m_1}f\right)\right)\right)-f\right\|_{\gamma}\right\|_{\infty}\leq\\ \sum_{i=1}^{r}\left\|\left\|L_{m_i}f-f\right\|_{\gamma}\right\|_{\infty}\leq\\ c_N\sum_{i=1}^{r}\left[\omega_1\left(f,\varphi\left(m_i\right)\right)+\left(1-f_{\lambda}\left(m_i^{1-\beta}-2\right)\right)\left\|\left\|f\right\|_{\gamma}\right\|_{\infty}\right]\leq\\ rc_N\left[\omega_1\left(f,\varphi\left(m_1\right)\right)+\left(1-f_{\lambda}\left(m_1^{1-\beta}-2\right)\right)\left\|\left\|f\right\|_{\gamma}\right\|_{\infty}\right].\end{gathered} \tag{100}$$

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

Proof. Using (99), (95), (96) and (90), (91). ☐

We continue with

Theorem 3.13. *Let all as in Corollary 3.4, and $r\in\mathbb{N}$. Here $\varphi_3(n)$ is as in (75). Then*

$$\left\|\left\|A_n^r f-f\right\|_{\gamma}\right\|_{\infty}\leq r\left\|\left\|A_n f-f\right\|_{\gamma}\right\|_{\infty}\leq r\varphi_3(n). \tag{101}$$

Proof. By (93) and (75). ☐

Application 3.1. 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.

Conflicts of interest : The author declares no conflict of interest.

Data availability : Not applicable

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 **14** (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.
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 approximation*s, J. Numer. Anal. Approx Theory **51** (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. H. Cartan, *Differential Calculus*, Hermann, Paris, 1971.
15. Z. Chen and F. Cao, *The approximation operators with sigmoidal functions*, Computers and Mathematics with Applications **58** (2009), 758-765.
16. D. Costarelli, R. Spigler, *Approximation results for neural network operators activated by sigmoidal functions*, Neural Networks **44** (2013), 101-106.
17. D. Costarelli, R. Spigler, *Multivariate neural network operators with sigmoidal activation functions*, Neural Networks **48** (2013), 72-77.
18. S. Haykin, *Neural Networks: A Comprehensive Foundation*,Prentice Hall, (2 ed.), New York, 1998.
19. W. McCulloch and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, Bulletin of Mathematical Biophysics **7** (1943), 115-133.
20. T.M. Mitchell, *Machine Learning*, WCB-McGraw-Hill, New York, 1997.
21. L.B. Rall, *Computational Solution of Nonlinear Operator Equations*, John Wiley & Sons, New York, 1969.
22. E.W. Weisstein, *Gudermannian*, MathWorld.

George A. Anastassiou received his B.Sc. degree in Mathematics from Athens University, Greece in 1975. He received his Diploma in Operations Research from Southampton University, UK in 1976. He also received his MA in Mathematics from University of Rochester, USA in 1981. He was awarded his Ph.D. in Mathematics from University of Rochester, USA in 1984. During 1984-86 he served as a visiting assistant professor at the University of Rhode Island, USA. Since 1986 till now 2020, he is a faculty member at the University of Memphis, USA. He is currently a full Professor of Mathematics since 1994. His research area

is "Computational Analysis" in the very broad sense. He has published over 550 research articles in international mathematical journals and over 43 monographs, proceedings and textbooks in well-known publishing houses. Several awards have been awarded to George Anastassiou. In 2007 he received the Honorary Doctoral Degree from University of Oradea, Romania. He is associate editor in over 70 international mathematical journals and editor in-chief in 3 journals, most notably in the well-known "Journal of Computational Analysis and Applications".

Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, U.S.A.
e-mail: ganastss@memphis.edu