Parametrized hyperbolic tangent based Banach space valued multivariate multi layer neural network approximations

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

Seda Karateke
Department of Computer Engineering, Faculty of Engineering,
Istanbul Topkapı University, Istanbul, Zeytinburnu 34087, Turkey.
sedakarateke@topkapi.edu.tr

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 hyperbolic tangent sigmoid function. The approximations are pointwise, uniform and L_p. The related feed-forward neural networks are with one or multi hidden layers.

2020 AMS Mathematics Subject Classification: 41A17, 41A25, 41A30, 41A36.

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

RGMIA Res. Rep. Coll. **25** (2022), Art. 122, 30 pp. Received 14/11/22

L_p approximations.

1 Introduction

G.A. Anastassiou 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 Cardaliaguet-Euvrard and "Squashing" types, by employing the modulus of continuity of the engaged function or its high order derivative, and producing very tight Jackson type inequalities. He treats there both the univariate and multivariate cases. The defining these operators "bell-shaped" and "squashing" functions are assumed to be 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.

Motivations for this work are the article [22] of Z. Chen and F. Cao, and [4]-[19], [23], [24].

Here we perform a parametrized hyperbolic tangent sigmoid function based neural network multivariate approximation to continuous functions over boxes or over the whole $\mathbb{R}^N$, $N \in \mathbb{N}$, and also iterated, multi layer and L_p approximations. 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.

We come 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 hyperbolic tangent sigmoid function.

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 neural network models, the activation function is based on the hyperbolic tangent sigmoid function. About neural networks read [25]-[27].

2 Background

We consider here the generalized hyperbolic tangent function $\tanh \lambda x$, $x \in \mathbb{R}$, $\lambda > 0$:

$$\tanh \lambda x = \frac{e^{\lambda x} - e^{-\lambda x}}{e^{\lambda x} + e^{-\lambda x}}. \tag{1}$$

It is $\tanh \lambda 0 = 0, -1 < \tanh \lambda x < 1, \forall\ x \in \mathbb{R}$, and $\tanh \lambda (-x) = -\tanh \lambda x$. Furthermore we have $\tanh \lambda (\infty) = 1$ and $\tanh \lambda (-\infty) = -1$, and $\tanh \lambda x$ is strictly increasing on $\mathbb{R}$, with

$$\frac{d}{dx} \tanh \lambda x = \frac{\lambda}{\cos^2 \lambda x} > 0. \tag{2}$$

The induced activation function will be

$$\theta (x) := \frac{1}{4} (\tanh \lambda (x+1) - \tanh \lambda (x-1)) > 0, \forall x \in \mathbb{R}, \tag{3}$$

with $\theta (x) = \theta (-x)$.

Clearly $\theta (x)$ is differentiable and thus it is continuous.

Proposition 1 *$\theta (x)$ is strictly decrasing on $(0, \infty)$ and strictly increasing on $(-\infty, 0]$. We have that $\theta (-\infty) = \theta (\infty) = 0$. So that θ has the bell shape with horizontal asymptote the x-axis. The maximum of θ is*

$$\theta (0) = \frac{\tanh \lambda}{2}. \tag{4}$$

We mention

Theorem 2 *([20]) It holds*

$$\sum_{i=-\infty}^{\infty} \theta (x - i) = 1, \quad \forall\ x \in \mathbb{R}. \tag{5}$$

Theorem 3 *([20]) We have that*

$$\int_{-\infty}^{\infty} \theta (x)\, dx = 1. \tag{6}$$

So that θ is a density function on $\mathbb{R}$.

Theorem 4 *([20]) Let $0 < \alpha < 1$, $\lambda > 0$ and $n \in \mathbb{N}$. It holds*

$$\sum_{\begin{cases} k = -\infty \\ : |nx - k| \geq n^{1-\alpha} \end{cases}}^{\infty} \theta (nx - k) < e^{4\lambda} e^{-2\lambda n^{(1-\alpha)}}. \tag{7}$$

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

Theorem 5 *([20]) Let $x \in [a,b] \subset \mathbb{R}$ and $n \in \mathbb{N}$, so that $\lceil na \rceil \leq \lfloor nb \rfloor$. Then*

$$\frac{1}{\sum\limits_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \theta(nx-k)} < \frac{4}{\tanh 2\lambda} = \frac{1}{\theta(1)}. \tag{8}$$

We make

Remark 6 *([20])*
(i) We have that

$$\lim_{n\to\infty} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \theta(nx-k) \neq 1, \tag{9}$$

for at least some $x \in [a,b]$.
(ii) Let $[a,b] \subset \mathbb{R}$. For large n we always have $\lceil na \rceil \leq \lfloor nb \rfloor$. Also $a \leq \frac{k}{n} \leq b$, iff $\lceil na \rceil \leq k \leq \lfloor nb \rfloor$.
In general it holds

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \theta(nx-k) \leq 1. \tag{10}$$

We introduce

$$Z(x_1,...,x_N) := Z(x) := \prod_{i=1}^{N} \theta(x_i), \quad x = (x_1,...,x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{11}$$

It has the properties:
(i) $Z(x) > 0, \ \forall \ x \in \mathbb{R}^N$,
(ii)

$$\sum_{k=-\infty}^{\infty} Z(x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z(x_1-k_1,...,x_N-k_N) = 1, \tag{12}$$

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

$$\sum_{k=-\infty}^{\infty} Z(nx-k) = 1, \tag{13}$$

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

(iv)

$$\int_{\mathbb{R}^N} Z(x)\,dx = 1, \tag{14}$$

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

$$\begin{gathered} \lceil na \rceil := (\lceil na_1 \rceil, ..., \lceil na_N \rceil), \\ \lfloor nb \rfloor := (\lfloor nb_1 \rfloor, ..., \lfloor nb_N \rfloor), \end{gathered} \tag{15}$$

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

We obviously see that

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

$$\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} ... \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} \left(\prod_{i=1}^{N} \theta(nx_i - k_i) \right) = \prod_{i=1}^{N} \left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \theta(nx_i - k_i) \right). \tag{16}$$

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

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

$$\sum_{\begin{cases} k = \lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty \le \frac{1}{n^\beta} \end{cases}}^{\lfloor nb \rfloor} Z(nx-k) + \sum_{\begin{cases} k = \lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\lfloor nb \rfloor} Z(nx-k). \tag{17}$$

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, Theorem 4 we derive that

$$\sum_{\begin{cases} k = \lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\lfloor nb \rfloor} Z(nx-k) \overset{(7)}{<} e^{4\lambda} e^{-2\lambda n^{(1-\beta)}}, \;\; 0 < \beta < 1, \;\; \lambda > 0. \tag{18}$$

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

(vi) By Theorem 5 we get that

$$0 < \frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)} < \left(\frac{4}{\tanh 2\lambda} \right)^N, \tag{19}$$

$\lambda > 0$, $\forall\ x \in \left(\prod_{i=1}^{N} [a_i, b_i]\right)$, $n \in \mathbb{N}$.

It is also clear that

(vii)

$$\sum_{\begin{cases} k = -\infty \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\infty} Z(nx - k) < e^{4\lambda} e^{-2\lambda n^{(1-\beta)}}, \tag{20}$$

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

Furthermore it holds

$$\lim_{n \to \infty} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx - k) \neq 1, \tag{21}$$

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

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

$$\frac{\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \sum_{k_2=\lceil na_2 \rceil}^{\lfloor nb_2 \rfloor} ... \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} f\left(\frac{k_1}{n}, ..., \frac{k_N}{n}\right) \left(\prod_{i=1}^{N} \theta(nx_i - k_i)\right)}{\prod_{i=1}^{N} \left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \theta(nx_i - k_i)\right)}. \tag{22}$$

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

$$\widetilde{A}_n(g, x) := \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} g\left(\frac{k}{n}\right) Z(nx - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx - k)}. \tag{23}$$

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

$$\left\| A_n(f,x) \right\|_\gamma \leq \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| f\left(\frac{k}{n}\right) \right\|_\gamma Z(nx-k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)} = \widetilde{A}_n\left(\|f\|_\gamma, x \right), \tag{24}$$

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

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

So, we have that

$$\left\| A_n(f,x) \right\|_\gamma \leq \widetilde{A}_n\left(\|f\|_\gamma, x \right), \tag{25}$$

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

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

$$A_n(cg, x) = c\widetilde{A}_n(g,x), \ \ \forall\ x \in \prod_{i=1}^N [a_i, b_i]. \tag{26}$$

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

$$A_n(c) = c, \ \ \forall\ c \in X. \tag{27}$$

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

For convenience we call

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

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

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

That is

$$A_n(f,x) := \frac{A_n^*(f,x)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)}, \tag{29}$$

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

Hence

$$A_n(f,x) - f(x) = \frac{A_n^*(f,x) - f(x)\left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k) \right)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)}. \tag{30}$$

Consequently we derive

$$\left\| A_n(f,x) - f(x) \right\|_{\gamma} \overset{(19)}{\leq} \left(\frac{4}{\tanh 2\lambda} \right)^{N} \left\| A_n^{*}(f,x) - f(x) \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k) \right\|_{\gamma}, \tag{31}$$

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

We will estimate the right hand side of (31).

For the last and others we need

Definition 7 *([15], 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*

$$\omega_1(f,\delta) := \sup_{\substack{x,y \in M: \\ \|x-y\|_p \leq \delta}} \|f(x) - f(y)\|_{\gamma}, \quad 0 < \delta \leq diam(M). \tag{32}$$

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

$$\omega_1(f,\delta) = \omega_1(f, diam(M)). \tag{33}$$

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 8 *([15], 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 (32). 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,

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

$$\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} f\left(\frac{k_1}{n}, \frac{k_2}{n}, ..., \frac{k_N}{n}\right) \left(\prod_{i=1}^{N} \theta(nx_i - k_i) \right), \tag{34}$$

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

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

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

$$\cdot\left(\prod_{i=1}^{N}\theta\left(nx_i-k_i\right)\right), \tag{35}$$

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

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

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

$$\sum_{r_1=0}^{\theta_1}\sum_{r_2=0}^{\theta_2}...\sum_{r_N=0}^{\theta_N}w_{r_1,r_2,...r_N}f\left(\frac{k_1}{n}+\frac{r_1}{n\theta_1},\frac{k_2}{n}+\frac{r_2}{n\theta_2},...,\frac{k_N}{n}+\frac{r_N}{n\theta_N}\right), \tag{36}$$

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

We set

$$D_n\left(f,x\right):=D_n\left(f,x_1,...,x_N\right):=\sum_{k=-\infty}^{\infty}\delta_{nk}\left(f\right)Z\left(nx-k\right)= \tag{37}$$

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

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

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 Hyperbolic Tangent Induced Banach Space Valued Network Approximations

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

We give

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

1)

$$\left\|A_n(f,x)-f(x)\right\|_\gamma \leq \left(\frac{4}{\tanh 2\lambda}\right)^N \left[\omega_1\left(f,\frac{1}{n^\beta}\right)+\frac{2e^{4\lambda}\left\|\|f\|_\gamma\right\|_\infty}{e^{2\lambda(n^{1-\beta})}}\right] =: \Omega_1(n), \tag{38}$$

and

2)

$$\left\|\|A_n(f)-f\|_\gamma\right\|_\infty \leq \Omega_1(n). \tag{39}$$

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

Proof. We observe that

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

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

$$\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left(f\left(\frac{k}{n}\right)-f(x)\right)Z(nx-k). \tag{40}$$

Thus

$$\|\Delta(x)\|_\gamma \leq \sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\|f\left(\frac{k}{n}\right)-f(x)\right\|_\gamma Z(nx-k) =$$

$$\sum_{\substack{k=\lceil na\rceil \\ \left\|\frac{k}{n}-x\right\|_\infty \leq \frac{1}{n^\beta}}}^{\lfloor nb\rfloor}\left\|f\left(\frac{k}{n}\right)-f(x)\right\|_\gamma Z(nx-k) +$$

$$\sum_{\substack{k=\lceil na\rceil \\ \left\|\frac{k}{n}-x\right\|_\infty > \frac{1}{n^\beta}}}^{\lfloor nb\rfloor}\left\|f\left(\frac{k}{n}\right)-f(x)\right\|_\gamma Z(nx-k) \overset{(13)}{\leq}$$

$$\omega_1\left(f,\frac{1}{n^\beta}\right)+2\left\|\|f\|_\gamma\right\|_\infty \sum_{\begin{cases} k=\lceil na\rceil \\ \left\|\frac{k}{n}-x\right\|_\infty>\frac{1}{n^\beta}\end{cases}}^{\lfloor nb\rfloor} Z(nx-k) \overset{(18)}{\leq}$$

$$\omega_1\left(f,\frac{1}{n^\beta}\right)+2\left\|\|f\|_\gamma\right\|_\infty e^{4\lambda}e^{-2\lambda n^{(1-\beta)}},\ 0<\beta<1,\ \lambda>0. \tag{41}$$

So that

$$\|\Delta(x)\|_\gamma \leq \omega_1\left(f,\frac{1}{n^\beta}\right)+\frac{2e^{4\lambda}\left\|\|f\|_\gamma\right\|_\infty}{e^{2\lambda n^{(1-\beta)}}}. \tag{42}$$

Now using (31) we finish the proof. ■

We make

Remark 10 *([15], 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* $\|x\|_{(\mathbb{R}^N)^j}:=\max_{1\leq\rho\leq j}\|x_\rho\|_p$, *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* $V_j:=V_j\left(\left(\mathbb{R}^N\right)^j;X\right)$ *of all* j*-multilinear continuous maps* $g:\left(\mathbb{R}^N\right)^j\to X$, $j=1,...,m$, *is a Banach space with norm*

$$\|g\|:=\|g\|_{V_j}:=\sup_{\left(\|x\|_{(\mathbb{R}^N)^j}=1\right)}\|g(x)\|_\gamma=\sup\frac{\|g(x)\|_\gamma}{\|x_1\|_p\cdots\|x_j\|_p}. \tag{43}$$

Let M *be a non-empty convex and compact subset of* $\mathbb{R}^N$ *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 [28])* $f^{(j)}:O\to V_j=V_j\left(\left(\mathbb{R}^N\right)^j;X\right)$ *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 ([21]), ([28], p. 124), we get

$$f(x)=\sum_{j=0}^m\frac{f^{(j)}(x_0)(x-x_0)^j}{j!}+R_m(x,x_0),\quad \text{all } x\in M, \tag{44}$$

where the remainder is the Riemann integral

$$R_m(x,x_0):=\int_0^1\frac{(1-u)^{m-1}}{(m-1)!}\left(f^{(m)}(x_0+u(x-x_0))-f^{(m)}(x_0)\right)(x-x_0)^m\,du, \tag{45}$$

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

We consider

$$\omega := \omega_1\left(f^{(m)}, h\right) := \sup_{\substack{x,y\in M:\\ \|x-y\|_p \le h}} \left\| f^{(m)}(x) - f^{(m)}(y) \right\|, \tag{46}$$

$h > 0$.

We obtain

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

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

$$\omega \|x-x_0\|_p^m \left\lceil \frac{u\|x-x_0\|_p}{h} \right\rceil, \tag{47}$$

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):*

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

$$= \omega \Phi_m\left(\|x-x_0\|_p\right) \tag{48}$$

by a change of variable, where

$$\Phi_m(t) := \int_0^{|t|} \left\lceil \frac{s}{h} \right\rceil \frac{(|t|-s)^{m-1}}{(m-1)!} ds = \frac{1}{m!} \left(\sum_{j=0}^{\infty} (|t|-jh)_+^m \right), \quad \forall\, t \in \mathbb{R}, \tag{49}$$

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

Also from there we get

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

with equality true only at $t = 0$.

Therefore it holds

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

We have found that

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

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

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

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:

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

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

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

Therefore we obtain

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

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

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

Clearly (54) is valid when $M=\prod_{i=1}^{N}[a_i,b_i]$ and $\widetilde{S}_n=\widetilde{A}_n$, see (23).

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, [15], pp. 268-270. The operators A_n, $\widetilde{A}_n$ fulfill its assumptions, see (22), (23), (25), and (26).

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)

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

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

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

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

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

We need

Lemma 12 *The function* $\left(\widetilde{A}_n\left(\|\cdot-x_0\|_p^m\right)\right)(x_0)$ *is continuous in* $x_0\in\left(\prod_{i=1}^{N}[a_i,b_i]\right)$, $m\in\mathbb{N}$.

Proof. By Lemma 10.3, [15], p. 272.

Remark 13 *By Remark 10.4 [15], p.273, we get that*

$$\left\|\left(\widetilde{A}_n\left(\|\cdot-x_0\|_p^k\right)\right)(x_0)\right\|_{\infty,x_0\in\prod_{i=1}^{N}[a_i,b_i]}\le\left\|\left(\widetilde{A}_n\left(\|\cdot-x_0\|_p^{m+1}\right)\right)(x_0)\right\|_{\infty,x_0\in\prod_{i=1}^{N}[a_i,b_i]}^{\left(\frac{k}{m+1}\right)}, \tag{59}$$

for all $k=1,...,m$.

■

We give

Corollary 14 *(to Theorem 11, case of* $m=1$*) Then*

1)

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

$$\frac{1}{2r}\omega_1\left(f^{(1)},r\left(\left(\widetilde{A}_n\left(\|\cdot-x_0\|_p^2\right)\right)(x_0)\right)^{\frac{1}{2}}\right)\left(\left(\widetilde{A}_n\left(\|\cdot-x_0\|_p^2\right)\right)(x_0)\right)^{\frac{1}{2}} \tag{60}$$

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

and

2)

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

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

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

$$\left\| \left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^2 \right) \right) (x_0) \right\|_{\infty, x_0 \in \prod_{i=1}^{N} [a_i, b_i]}^{\frac{1}{2}} \left[1 + r + \frac{r^2}{4} \right], \tag{61}$$

$r > 0$.

We make

Remark 15 *We estimate* $0 < \alpha < 1$, $\lambda > 0$, $m, n \in \mathbb{N} : n^{1-\alpha} > 2$,

$$\widetilde{A}_n \left(\|\cdot - x_0\|_\infty^{m+1} \right) (x_0) = \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z(nx_0 - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx_0 - k)} \overset{(19)}{<}$$

$$\left(\frac{4}{\tanh 2\lambda} \right)^N \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z(nx_0 - k) = \tag{62}$$

$$\left(\frac{4}{\tanh 2\lambda} \right)^N \left\{ \sum_{\substack{k = \lceil na \rceil \\ : \left\| \frac{k}{n} - x_0 \right\|_\infty \le \frac{1}{n^\alpha}}}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z(nx_0 - k) + \right.$$

$$\left. \sum_{\substack{k = \lceil na \rceil \\ : \left\| \frac{k}{n} - x_0 \right\|_\infty > \frac{1}{n^\alpha}}}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z(nx_0 - k) \right\} \overset{(20)}{\le}$$

$$\left(\frac{4}{\tanh 2\lambda} \right)^N \left\{ \frac{1}{n^{\alpha(m+1)}} + \frac{e^{4\lambda} \|b - a\|_\infty^{m+1}}{e^{2\lambda(n^{1-\beta})}} \right\}, \tag{63}$$

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

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

$$\widetilde{A}_n \left(\|\cdot - x_0\|_\infty^{m+1} \right) (x_0) < \left(\frac{4}{\tanh 2\lambda} \right)^N \left\{ \frac{1}{n^{\alpha(m+1)}} + \frac{e^{4\lambda} \|b - a\|_\infty^{m+1}}{e^{2\lambda(n^{1-\beta})}} \right\} =: \Lambda_1(n) \tag{64}$$

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

And, consequently it holds

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

$$\left(\frac{4}{\tanh 2\lambda} \right)^N \left\{ \frac{1}{n^{\alpha(m+1)}} + \frac{e^{4\lambda} \left\| b-a \right\|_{\infty}^{m+1}}{e^{2\lambda (n^{1-\beta})}} \right\} = \Lambda_1(n) \to 0, \quad \text{as } n \to +\infty. \tag{65}$$

So, we have that $\Lambda_1(n)\to 0$, 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

$$\left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) = \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} f^{(j)}(x_0) \left(\frac{k}{n} - x_0 \right)^j Z(nx_0 - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx_0 - k)}. \tag{66}$$

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

$$\left\| f^{(j)}(x_0) \left(\frac{k}{n} - x_0 \right)^j \right\|_{\gamma} \leq \left\| f^{(j)}(x_0) \right\| \left\| \frac{k}{n} - x_0 \right\|_{\infty}^j. \tag{67}$$

We further have that

$$\left\| \left(\widetilde{A}_n \left(f^{(j)}(x_0)(\cdot - x_0)^j \right) \right)(x_0) \right\|_{\gamma} \overset{(19)}{<}$$

$$\left(\frac{4}{\tanh 2\lambda} \right)^N \left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| f^{(j)}(x_0) \left(\frac{k}{n} - x_0 \right)^j \right\|_{\gamma} Z(nx_0 - k) \right) \leq$$

$$\left(\frac{4}{\tanh 2\lambda} \right)^N \left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| f^{(j)}(x_0) \right\| \left\| \frac{k}{n} - x_0 \right\|_{\infty}^j Z(nx_0 - k) \right) = \tag{68}$$

$$\left(\frac{4}{\tanh 2\lambda} \right)^N \left\| f^{(j)}(x_0) \right\| \left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_{\infty}^j Z(nx_0 - k) \right) =$$

$$\left(\frac{4}{\tanh 2\lambda} \right)^N \left\| f^{(j)}(x_0) \right\| \left\{ \sum_{\substack{k = \lceil na \rceil \\ \left\{ : \left\| \frac{k}{n} - x_0 \right\|_{\infty} \leq \frac{1}{n^{\alpha}} \right.}}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_{\infty}^j Z(nx_0 - k) \right.$$

$$+ \sum_{\left\{\begin{array}{c} k=\lceil na\rceil \\ :\left\|\frac{k}{n}-x_0\right\|_\infty > \frac{1}{n^\alpha} \end{array}\right.}^{\lfloor nb\rfloor} \left\|\frac{k}{n}-x_0\right\|_\infty^j Z\left(nx_0-k\right)\Bigg\} \overset{(20)}{\leq} \tag{69}$$

$$\left(\frac{4}{\tanh 2\lambda}\right)^N \left\|f^{(j)}\left(x_0\right)\right\| \left\{\frac{1}{n^{\alpha j}}+\frac{e^{4\lambda}\left\|b-a\right\|_\infty^j}{e^{2\lambda\left(n^{1-\beta}\right)}}\right\} \to 0, \text{ as } n\to\infty.$$

That is

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

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

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

$$\left(\frac{4}{\tanh 2\lambda}\right)^N \left\|f^{(j)}\left(x_0\right)\right\| \left\{\frac{1}{n^{\alpha j}}+\frac{e^{4\lambda}\left\|b-a\right\|_\infty^j}{e^{2\lambda\left(n^{1-\beta}\right)}}\right\} \leq$$

$$\left(\frac{4}{\tanh 2\lambda}\right)^N \left\|f^{(j)}\left(x_0\right)\right\|_\infty \left\{\frac{1}{n^{\alpha j}}+\frac{e^{4\lambda}\left\|b-a\right\|_\infty^j}{e^{2\lambda\left(n^{1-\beta}\right)}}\right\} =: \Lambda_{2j}\left(n\right) < \infty, \tag{70}$$

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

Conclusion 16 *We have proved that the left hand sides of (55), (56), (57), (58) and (60), (61) 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 further give

Corollary 17 *(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 $f\in C^m\left(O,X\right)$, 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* $\Lambda_1(n)$ *as in (65) and* $\Lambda_{2j}(n)$ *as in (70), where* $n \in \mathbb{N} : n^{1-\alpha} > 2$, $0 < \alpha < 1$, $\lambda > 0$, $j = 1, ..., m$. *Then*

1)

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

$$\frac{\omega_1\left(f^{(m)}, r(\Lambda_1(n))^{\frac{1}{m+1}}\right)}{rm!} (\Lambda_1(n))^{\left(\frac{m}{m+1}\right)} \left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8} \right], \tag{71}$$

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

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

$$\frac{\omega_1\left(f^{(m)}, r(\Lambda_1(n))^{\frac{1}{m+1}}\right)}{rm!} (\Lambda_1(n))^{\left(\frac{m}{m+1}\right)} \left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8} \right], \tag{72}$$

3)

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

$$\frac{\omega_1\left(f^{(m)}, r(\Lambda_1(n))^{\frac{1}{m+1}}\right)}{rm!} (\Lambda_1(n))^{\left(\frac{m}{m+1}\right)}$$

$$\left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8} \right] =: \Lambda_3(n) \to 0, \text{ as } n \to \infty. \tag{73}$$

We continue with

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

1)

$$\|B_n(f, x) - f(x)\|_\gamma \leq \omega_1\left(f, \frac{1}{n^\beta}\right) + \frac{e^{4\lambda} \left\| \|f\|_\gamma \right\|_\infty}{e^{2\lambda(n^{1-\beta})}} =: \Omega_2(n), \tag{74}$$

2)

$$\left\| \|B_n(f) - f\|_\gamma \right\|_\infty \leq \Omega_2(n). \tag{75}$$

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

Proof. We have that

$$B_n(f,x) - f(x) \overset{(13)}{=} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{n}\right) Z(nx-k) - f(x) \sum_{k=-\infty}^{\infty} Z(nx-k) = \tag{76}$$

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

Hence

$$\left\|B_n(f,x) - f(x)\right\|_\gamma \leq \sum_{k=-\infty}^{\infty} \left\|f\left(\frac{k}{n}\right) - f(x)\right\|_\gamma Z(nx-k) =$$

$$\sum_{\begin{cases} k = -\infty \\ \left\|\frac{k}{n} - x\right\|_\infty \leq \frac{1}{n^\beta} \end{cases}}^{\infty} \left\|f\left(\frac{k}{n}\right) - f(x)\right\|_\gamma Z(nx-k) +$$

$$\sum_{\begin{cases} k = -\infty \\ \left\|\frac{k}{n} - x\right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\infty} \left\|f\left(\frac{k}{n}\right) - f(x)\right\|_\gamma Z(nx-k) \overset{(13)}{\leq}$$

$$\omega_1\left(f, \frac{1}{n^\beta}\right) + 2\left\|\|f\|_\gamma\right\|_\infty \sum_{\begin{cases} k = -\infty \\ \left\|\frac{k}{n} - x\right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\infty} Z(nx-k) \overset{(20)}{\leq}$$

$$\omega_1\left(f, \frac{1}{n^\beta}\right) + \frac{2e^{4\lambda}\left\|\|f\|_\gamma\right\|_\infty}{e^{2\lambda\left(n^{1-\beta}\right)}}, \tag{77}$$

proving the claim. ■

We give

Theorem 19 *Let $f \in C_B\left(\mathbb{R}^N, X\right)$, $0 < \beta < 1$, $x \in \mathbb{R}^N$, $\lambda > 0, N, n \in \mathbb{N}$ with $n^{1-\beta} > 2$, ω_1 is for $p = \infty$. Then*

1)

$$\left\|C_n(f,x) - f(x)\right\|_\gamma \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + \frac{2e^{4\lambda}\left\|\|f\|_\gamma\right\|_\infty}{e^{2\lambda\left(n^{1-\beta}\right)}} =: \Omega_3(n), \tag{78}$$

2)

$$\left\|\left\|C_n(f) - f\right\|_\gamma\right\|_\infty \leq \Omega_3(n). \tag{79}$$

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. We notice that

$$\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\,dt = \int_{\frac{k_1}{n}}^{\frac{k_1+1}{n}} \int_{\frac{k_2}{n}}^{\frac{k_2+1}{n}} \dots \int_{\frac{k_N}{n}}^{\frac{k_N+1}{n}} f(t_1, t_2, \dots, t_N)\, dt_1 dt_2 \dots dt_N =$$

$$\int_0^{\frac{1}{n}} \int_0^{\frac{1}{n}} \dots \int_0^{\frac{1}{n}} f\left(t_1 + \frac{k_1}{n}, t_2 + \frac{k_2}{n}, \dots, t_N + \frac{k_N}{n}\right) dt_1 \dots dt_N = \int_0^{\frac{1}{n}} f\left(t + \frac{k}{n}\right) dt. \tag{80}$$

Thus it holds (by (35))

$$C_n(f, x) = \sum_{k=-\infty}^{\infty} \left(n^N \int_0^{\frac{1}{n}} f\left(t + \frac{k}{n}\right) dt\right) Z(nx - k). \tag{81}$$

We observe that

$$\|C_n(f, x) - f(x)\|_\gamma =$$

$$\left\|\sum_{k=-\infty}^{\infty} \left(n^N \int_0^{\frac{1}{n}} f\left(t + \frac{k}{n}\right) dt\right) Z(nx - k) - \sum_{k=-\infty}^{\infty} f(x) Z(nx - k)\right\|_\gamma =$$

$$\left\|\sum_{k=-\infty}^{\infty} \left(\left(n^N \int_0^{\frac{1}{n}} f\left(t + \frac{k}{n}\right) dt\right) - f(x)\right) Z(nx - k)\right\|_\gamma =$$

$$\left\|\sum_{k=-\infty}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left(f\left(t + \frac{k}{n}\right) - f(x)\right) dt\right) Z(nx - k)\right\|_\gamma \leq \tag{82}$$

$$\sum_{k=-\infty}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left\|f\left(t + \frac{k}{n}\right) - f(x)\right\|_\gamma dt\right) Z(nx - k) =$$

$$\sum_{\begin{cases} k = -\infty \\ \left\|\frac{k}{n} - x\right\|_\infty \leq \frac{1}{n^\beta} \end{cases}}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left\|f\left(t + \frac{k}{n}\right) - f(x)\right\|_\gamma dt\right) Z(nx - k) +$$

$$\sum_{\begin{cases} k = -\infty \\ \left\|\frac{k}{n} - x\right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \left\|f\left(t + \frac{k}{n}\right) - f(x)\right\|_\gamma dt\right) Z(nx - k) \leq$$

$$\sum_{\begin{cases} k = -\infty \\ \left\|\frac{k}{n} - x\right\|_\infty \leq \frac{1}{n^\beta} \end{cases}}^{\infty} \left(n^N \int_0^{\frac{1}{n}} \omega_1\left(f, \|t\|_\infty + \left\|\frac{k}{n} - x\right\|_\infty\right) dt\right) Z(nx - k) +$$

$$2\left\|\|f\|_{\gamma}\right\|_{\infty}\left(\sum_{\begin{cases} k=-\infty \\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}\end{cases}}^{\infty} Z\left(|nx-k|\right)\right)\leq$$

$$\omega_1\left(f,\frac{1}{n}+\frac{1}{n^{\beta}}\right)+\frac{2e^{4\lambda}\left\|\|f\|_{\gamma}\right\|_{\infty}}{e^{2\lambda\left(n^{1-\beta}\right)}}, \tag{83}$$

proving the claim. ■

We also present

Theorem 20 *Let $f\in C_B\left(\mathbb{R}^N,X\right)$, $0<\beta<1$, $x\in\mathbb{R}^N$, $\lambda>0, N,n\in\mathbb{N}$ with $n^{1-\beta}>2$, ω_1 is for $p=\infty$. Then*

1)

$$\|D_n(f,x)-f(x)\|_{\gamma}\leq\omega_1\left(f,\frac{1}{n}+\frac{1}{n^{\beta}}\right)+\frac{2e^{4\lambda}\left\|\|f\|_{\gamma}\right\|_{\infty}}{e^{2\lambda\left(n^{1-\beta}\right)}}=:\Omega_4(n), \tag{84}$$

2)

$$\left\|\|D_n(f)-f\|_{\gamma}\right\|_{\infty}\leq\Omega_4(n). \tag{85}$$

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. We have that (by (37))

$$\|D_n(f,x)-f(x)\|_{\gamma}=\left\|\sum_{k=-\infty}^{\infty}\delta_{nk}(f)Z(nx-k)-\sum_{k=-\infty}^{\infty}f(x)Z(nx-k)\right\|_{\gamma}=$$

$$\left\|\sum_{k=-\infty}^{\infty}\left(\delta_{nk}(f)-f(x)\right)Z(nx-k)\right\|_{\gamma}=\left\|\sum_{k=-\infty}^{\infty}\omega_r\left(f\left(\frac{k}{n}+\frac{r}{n\theta}\right)-f(x)\right)Z(nx-k)\right\|_{\gamma}\leq$$

$$\sum_{k=-\infty}^{\infty}\left(\sum_{r=0}^{\theta}\omega_r\left\|f\left(\frac{k}{n}+\frac{r}{n\theta}\right)-f(x)\right\|_{\gamma}\right)Z(nx-k)=$$

$$\sum_{\begin{cases} k=-\infty \\ \left\|\frac{k}{n}-x\right\|_{\infty}\leq\frac{1}{n^{\beta}}\end{cases}}^{\infty}\left(\sum_{r=0}^{\theta}\omega_r\left\|f\left(\frac{k}{n}+\frac{r}{n\theta}\right)-f(x)\right\|_{\gamma}\right)Z(nx-k)+$$

$$\sum_{\begin{cases} k=-\infty \\ \left\| \frac{k}{n}-x \right\|_{\infty} > \frac{1}{n^{\beta}} \end{cases}}^{\infty} \left(\sum_{r=0}^{\theta} \omega_r \left\| f\left(\frac{k}{n} + \frac{r}{n\theta} \right) - f(x) \right\|_{\gamma} \right) Z(nx-k) \leq$$

$$\sum_{\begin{cases} k=-\infty \\ \left\| \frac{k}{n}-x \right\|_{\infty} \leq \frac{1}{n^{\beta}} \end{cases}}^{\infty} \left(\sum_{r=0}^{\theta} \omega_r \left\| f\left(\frac{k}{n} + \frac{r}{n\theta} \right) - f(x) \right\|_{\gamma} \right) Z(nx-k) +$$

$$2 \left\| \|f\|_{\gamma} \right\|_{\infty} \left(\sum_{\begin{cases} k=-\infty \\ \left\| \frac{k}{n}-x \right\|_{\infty} > \frac{1}{n^{\beta}} \end{cases}}^{\infty} (Z(nx-k)) \right) \leq$$

$$\omega_1 \left(f, \frac{1}{n} + \frac{1}{n^{\beta}} \right) + \frac{2e^{4\lambda} \left\| \|f\|_{\gamma} \right\|_{\infty}}{e^{2\lambda(n^{1-\beta})}} = \Omega_4(n),$$

proving the claim. ■

Next we perform multi layer neural network approximations.

We make

Definition 21 *Let $f \in C_B(\mathbb{R}^N, X)$, $N \in \mathbb{N}$, where $\left(X, \|\cdot\|_{\gamma}\right)$ is a Banach space. We define the general neural network operator*

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

$$\begin{cases} B_n(f,x), & \text{if } l_{nk}(f) = f\left(\frac{k}{n}\right), \\ C_n(f,x), & \text{if } l_{nk}(f) = n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\,dt, \\ D_n(f,x), & \text{if } l_{nk}(f) = \delta_{nk}(f). \end{cases} \tag{86}$$

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

Proof. Lengthy and similar to the proof of Theorem 11 of [18], as such is omitted. ■

Remark 23 *By (22) it is obvious that* $\left\| \left\| A_n(f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty < \infty$, *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 K_n *any of the operators* A_n, B_n, C_n, D_n.

Clearly then

$$\left\| \left\| K_n^2(f) \right\|_\gamma \right\|_\infty = \left\| \left\| K_n(K_n(f)) \right\|_\gamma \right\|_\infty \leq \left\| \left\| K_n(f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty, \tag{87}$$

etc.

Therefore we get

$$\left\| \left\| K_n^k(f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty, \quad \forall\, k \in \mathbb{N}, \tag{88}$$

the contraction property.

Also we see that

$$\left\| \left\| K_n^k(f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| K_n^{k-1}(f) \right\|_\gamma \right\|_\infty \leq \ldots \leq \left\| \left\| K_n(f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty. \tag{89}$$

Here K_n^k *are bounded linear operators.*

Notation 24 *Here* $N \in \mathbb{N}$, $0 < \beta < 1$. *Denote by*

$$c_N := \begin{cases} \left(\frac{4}{\tanh 2\lambda} \right)^N, & \text{if } K_n = A_n, \\ 1, & \text{if } K_n = B_n, C_n, D_n, \end{cases} \tag{90}$$

$$\Lambda(n) := \begin{cases} \frac{1}{n^\beta}, & \text{if } K_n = A_n,\ B_n, \\ \frac{1}{n} + \frac{1}{n^\beta}, & \text{if } K_n = C_n, D_n, \end{cases} \tag{91}$$

$$\Gamma := \begin{cases} C\left(\prod_{i=1}^{N} [a_i, b_i], X \right), & \text{if } K_n = A_n, \\ C_B\left(\mathbb{R}^N, X \right), & \text{if } K_n = B_n, C_n, D_n, \end{cases} \tag{92}$$

and

$$Y := \begin{cases} \prod_{i=1}^{N} [a_i, b_i], & \text{if } K_n = A_n, \\ \mathbb{R}^N, & \text{if } K_n = B_n, C_n, D_n. \end{cases} \tag{93}$$

We give the condensed

Theorem 25 *Let* $f \in \Gamma$, $0 < \beta < 1$, $x \in Y$; n, $\lambda > 0$; $N \in \mathbb{N}$ *with* $n^{1-\beta} > 2$. *Then*

(i)

$$\left\| K_n(f, x) - f(x) \right\|_\gamma \leq c_N \left[\omega_1(f, \Lambda(n)) + \frac{2e^{4\lambda} \left\| \left\| f \right\|_\gamma \right\|_\infty}{e^{2\lambda(n^{1-\beta})}} \right] =: \tau(n), \tag{94}$$

where ω_1 is for $p = \infty$,

and

(ii)

$$\left\| \left\| K_n (f) - f \right\|_\gamma \right\|_\infty \leq \tau (n) \to 0, \text{ as } n \to \infty. \tag{95}$$

For f uniformly continuous and in Γ we obtain

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

pointwise and uniformly.

Proof. By Theorems 9, 18, 19, 20. ■

Next we do iterated, multi layer neural network approximation. (see also [10]).

We make

Remark 26 *Let $r \in \mathbb{N}$ and K_n as above. We observe that*

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

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

Then

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

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

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

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

That is

$$\left\| \left\| K_n^r f - f \right\|_\gamma \right\|_\infty \leq r \left\| \left\| K_n f - f \right\|_\gamma \right\|_\infty. \tag{96}$$

We give

Theorem 27 *All here as in Theorem 25 and $r \in \mathbb{N}$, $\tau (n)$ as in (94). Then*

$$\left\| \left\| K_n^r f - f \right\|_\gamma \right\|_\infty \leq r \tau (n). \tag{97}$$

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

Proof. As similar to [18] is omitted. ■

Remark 28 *Let $m_1, m_2, ..., m_r \in \mathbb{N} : m_1 \leq m_2 \leq ... \leq m_r$, $0 < \beta < 1, \lambda > 0$, $f \in \Gamma$. Then*

$$\Lambda(m_1) \geq \Lambda(m_2) \geq ... \geq \Lambda(m_r), \ \Lambda \text{ as in (91).}$$

Therefore

$$\omega_1(f, \Lambda(m_1)) \geq \omega_1(f, \Lambda(m_2)) \geq ... \geq \omega_1(f, \Lambda(m_r)).$$

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

$$\frac{e^{4\lambda}}{e^{2\lambda m_1^{(1-\beta)}}} \geq \frac{e^{4\lambda}}{e^{\lambda m_2^{(1-\beta)}}} \geq ... \geq \frac{e^{4\lambda}}{e^{\lambda m_r^{(1-\beta)}}}.$$

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

$$K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_2}\left(K_{m_1 f}\right)\right)\right) - f =$$

$$K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_2}\left(K_{m_1 f}\right)\right)\right) - K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_2} f\right)\right) +$$
$$K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_2} f\right)\right) - K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_3} f\right)\right) +$$
$$K_{m_r}\left(K_{m_{r-1}}\left(...K_{m} 3 f\right)\right) - K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_4} f\right)\right) + ...+$$
$$K_{m_r}\left(K_{m_{r-1}} f\right) - K_{m_r} f + K_{m_r} f - f =$$
$$K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_2}\right)\right)\left(K_{m_1} f - f\right) + K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_3}\right)\right)\left(K_{m_2} f - f\right) +$$
$$K_{m_r}\left(K_{m_{r-1}}\left(...K_{m_4}\right)\right)\left(K_{m} 3 f - f\right) + ... + K_{m_r}\left(K_{m_{r-1}} f - f\right) + K_{m_r} f - f.$$

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

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

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

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

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

$$\left\| \left\| K_{m_r}\left(K_{m_{r-1}} f - f\right) \right\|_\gamma \right\|_\infty + \left\| \left\| K_{m_r} f - f \right\|_\gamma \right\|_\infty \leq$$

(repeatedly applying (87))

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

$$\left\| \left\| K_{m_{r-1}} f - f \right\|_\gamma \right\|_\infty + \left\| \left\| K_{m_2} f - f \right\|_\gamma \right\|_\infty + \left\| \left\| K_{m_3} f - f \right\|_\gamma \right\|_\infty + ...+$$

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

That is, we proved

$$\left\| \left\| K_{m_r} \left(K_{m_{r-1}} \left(... K_{m_2} \left(K_{m_1} f \right) \right) \right) - f \right\|_{\gamma} \right\|_{\infty} \leq_{i=1}^{r} \left\| \left\| K_{m_i} f - f \right\|_{\gamma} \right\|_{\infty}. \tag{98}$$

We also present

Theorem 29 *Let $f \in \Gamma$; m, N, $m_1, m_2, ..., m_r \in \mathbb{N} : m_1 \leq m_2 \leq ... \leq m_r$, $0 < \beta < 1, \lambda > 0$; $m_i^{(1-\beta)} > 2$, $i = 1, ..., r$, $x \in Y$, and let $(K_{m_1}, ..., K_{m_r})$ 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*

$$\begin{gathered}
\left\| K_{m_r} \left(K_{m_{r-1}} \left(... K_{m_2} \left(K_{m_1} f \right) \right) \right) (x) - f (x) \right\|_{\gamma} \leq \\
\left\| \left\| K_{m_r} \left(K_{m_{r-1}} \left(... K_{m_2} \left(K_{m_1} f \right) \right) \right) - f \right\|_{\gamma} \right\|_{\infty} \leq \\
\sum_{i=1}^{r} \left\| \left\| K_{m_i} f - f \right\|_{\gamma} \right\|_{\infty} \leq \\
c_N \sum_{i=1}^{r} \left[\omega_1 \left(f, \Lambda \left(m_i \right) \right) + \frac{2e^{4\lambda} \left\| \left\| f \right\|_{\gamma} \right\|_{\infty}}{e^{2\lambda m_i^{(1-\beta)}}} \right] \leq \\
r c_N \left[\omega_1 \left(f, \Lambda \left(m_1 \right) \right) + \frac{2e^{4\lambda} \left\| \left\| f \right\|_{\gamma} \right\|_{\infty}}{e^{2\lambda m_1^{(1-\beta)}}} \right].
\end{gathered} \tag{99}$$

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

Proof. As similar to [18] is omitted. ■

We continue with

Theorem 30 *Let all as in Corollary 17, and $r \in \mathbb{N}$. Here $\Lambda_3(n)$ is as in (73). Then*

$$\left\| \left\| A_n^r f - f \right\|_{\gamma} \right\|_{\infty} \leq r \left\| \left\| A_n f - f \right\|_{\gamma} \right\|_{\infty} \leq r \Lambda_3 (n). \tag{100}$$

Proof. As similar to [18] is omitted. ■

Next we present some L_{p_1}, $p_1 \geq 1$, approximation related results.

Theorem 31 *Let $p_1 \geq 1$, $f \in C\left(\prod_{i=1}^{n} [a_i, b_i], X \right)$, $0 < \beta < 1, \lambda > 0$; $N, n \in \mathbb{N}$ with $n^{1-\beta} > 2$, and $\Omega_1(n)$ as in (38), ω_1 is for $p = \infty$. Then*

$$\left\| \left\| A_n f - f \right\|_{\gamma} \right\|_{p_1, \prod_{i=1}^{n} [a_i, b_i]} \leq \Omega_1 (n) \left(\prod_{i=1}^{n} (b_i - a_i) \right)^{\frac{1}{p_1}}. \tag{101}$$

We notice that $\lim_{n \to \infty} \left\| \left\| A_n f - f \right\|_{\gamma} \right\|_{p_1, \prod_{i=1}^{n} [a_i, b_i]} = 0$.

Proof. Obvious, by integrating (38), etc. ■

It follows

Theorem 32 *Let* $p_1 \geq 1$, $f \in C_B\left(\mathbb{R}^N, X\right)$, $0 < \beta < 1, \lambda > 0$; $N, n \in \mathbb{N}$ *with* $n^{1-\beta} > 2$, *and* ω_1 *is for* $p = \infty$; $\Omega_2(n)$ *as in (74) and* P *a compact set of* $\mathbb{R}^N$. *Then*

$$\left\| \|B_n f - f\|_\gamma \right\|_{p_1,\ P} \leq \Omega_2(n) |P|^{\frac{1}{p_1}}, \tag{102}$$

where $|P| < \infty$, *is the Lebesgue measure of* P. *We notice that* $\lim_{n \to \infty} \left\| \|B_n f - f\|_\gamma \right\|_{p_1, P} = 0$ *for* $f \in \left(C_U\left(\mathbb{R}^N, X\right) \cap C_B\left(\mathbb{R}^N, X\right)\right)$.

Proof. By integrating (74), etc. ■

Next come

Theorem 33 *All as in Theorem 32, but we use* $\Omega_3(n)$ *of (78). Then*

$$\left\| \|C_n f - f\|_\gamma \right\|_{p_1,\ P} \leq \Omega_3(n) |P|^{\frac{1}{p_1}}. \tag{103}$$

We have that $\lim_{n \to \infty} \left\| \|C_n f - f\|_\gamma \right\|_{p_1,\ P} = 0$ *for* $f \in \left(C_U\left(\mathbb{R}^N, X\right) \cap C_B\left(\mathbb{R}^N, X\right)\right)$.

Proof. By (78). ■

Theorem 34 *All as in Theorem 32, but we use* $\Omega_4(n)$ *of (84). Then*

$$\left\| \|D_n f - f\|_\gamma \right\|_{p_1,\ P} \leq \Omega_4(n) |P|^{\frac{1}{p_1}}. \tag{104}$$

We have that $\lim_{n \to \infty} \left\| \|D_n f - f\|_\gamma \right\|_{p_1,\ P} = 0$ *for* $f \in \left(C_U\left(\mathbb{R}^N, X\right) \cap C_B\left(\mathbb{R}^N, X\right)\right)$.

Proof. By (84). ■

Application 35 *A typical application of all of our results is when* $\left(X, \|\cdot\|_\gamma\right) = (\mathbb{C}, |\cdot|)$, *where* $\mathbb{C}$ *is the set of 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, *Intelligent 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, *Fractional neural network approximation*, Computers and Mathematics with Applications, 64 (2012), 1655-1676.

[10] 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.

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

[12] G.A. Anastassiou, *Strong Right Fractional Calculus for Banach space valued functions*, 'Revista Proyecciones, Vol. 36, No. 1 (2017), 149-186.

[13] G.A. Anastassiou, *Vector fractional Korovkin type Approximations*, Dynamic Systems and Applications, 26(2017), 81-104.

[14] G.A. Anastassiou, *A strong Fractional Calculus Theory for Banach space valued functions*, Nonlinear Functional Analysis and Applications (Korea), 22(3) (2017), 495-524.

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

[16] G.A. Anastassiou, *Nonlinearity: Ordinary and Fractional Approximations by Sublinear and Max-Product Operators*, Springer, Heidelberg, New York, 2018.

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

[18] G.A. Anastassiou, *General multimariate arctangent function activated neural network approximations*, Journal of Numerical Analysis and Approximation Theory 51(1) (2022), 37-66.

[19] G.A. Anastassiou, *Abstract multivariate algebraic function activated neural network approximations*, accepted to Journal of Computational Analysis and Applications, 2022.

[20] G.A. Anastassiou and S. Karateke, *Parametrized hyperbolic tangent induced Banach space valued ordinary and fractional neural network approximation, submitted*, 2022.

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

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

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

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

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

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

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

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