CUBO, A Mathematical Journal
Vol. 25, no. 03, pp. 411–439, December 2023
DOI: 10.56754/0719-0646.2503.411

Multiple general sigmoids based Banach space valued neural network multivariate approximation

George A. Anastassiou[1,✉]

[1] *Department of Mathematical Sciences, University of Memphis Memphis, TN 38152, U.S.A.*
ganastss@memphis.edu✉

ABSTRACT

Here we present 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 several different among themselves general sigmoid functions. This is done on the purpose to activate as many as possible neurons. The approximations are pointwise and uniform. The related feed-forward neural network is with one hidden layer. We finish with related L_p approximations.

Published: 22 December, 2023
Accepted: 25 October, 2023
Received: 06 October, 2022

RESUMEN

Presentamos aproximaciones multivariadas cuantitativas de funciones multivariadas continuas con valores en un espacio de Banach definidas en una caja o en $\mathbb{R}^N$, $N \in \mathbb{N}$, a través de operadores de redes neuronales multivariados normalizados, de cuasi-interpolación, de tipo Kantorovich y de tipo cuadratura. También tratamos el caso de aproximación usando operadores iterados de los últimos cuatro tipos. Estas aproximaciones se derivan estableciendo desigualdades multidimensionales de tipo Jackson que involucran el módulo de continuidad multivariado de la función comprometida o sus derivadas de Fréchet de alto orden. Nuestros operadores multivariados son definidos usando una función de densidad multidimensional inducida por varias funciones sigmoidales generales diferentes entre sí. Esto se hace con el propósito de activar la mayor cantidad de neuronas posible. Las aproximaciones son puntutales y uniformes. La red neuronal prealimentada relacionada tiene un nivel oculto. Concluimos con aproximaciones L_p relacionadas.

Keywords and Phrases: General sigmoid functions, multivariate neural network approximation, quasi-interpolation operator, Kantorovich type operator, quadrature type operator, multivariate modulus of continuity, abstract approximation, iterated approximation, L_p approximation.

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

1 Introduction

The author in [2, 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.

For this article the author is motivated by the article [14] of Z. Chen and F. Cao, also by [4–12, 15, 16].

The author here performs multivariate multiple general sigmoid functions based neural network approximations to continuous functions over boxes or over the whole $\mathbb{R}^N$, $N \in \mathbb{N}$. Also he does iterated 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.

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 functions induced by multiple general sigmoid functions 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 a general sigmoid function, but here we use a multiple number of them simultaneously for the first time, so we can activate a maximum number of neurons. About neural networks read [17–19].

2 Basics

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

We consider the activation function

$$\psi_i(x) := \frac{1}{4}\left(h_i(x+1) - h_i(x-1)\right), \quad x \in \mathbb{R}, \quad i = 1, \ldots, N. \tag{2.1}$$

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

We see that

$$\psi_i(0) = \frac{h_i(1)}{2}, \quad i = 1, \ldots, N. \tag{2.2}$$

Let $x > 1$, we have that

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

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

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

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

See that

$$\lim_{x \to +\infty} \psi_i(x) = \frac{1}{4}\left(h_i(+\infty) - h_i(+\infty)\right) = 0, \tag{2.3}$$

and

$$\lim_{x \to -\infty} \psi_i(x) = \frac{1}{4}\left(h_i(-\infty) - h_i(-\infty)\right) = 0. \tag{2.4}$$

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

Conclusion, ψ is a bell symmetric function with maximum

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

We need

Theorem 2.1. *We have that*

$$\sum_{i=-\infty}^{\infty} \psi_i(x - i) = 1, \quad \forall x \in \mathbb{R}, \quad i = 1, \ldots, N. \tag{2.5}$$

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

Theorem 2.2. *It holds*

$$\int_{-\infty}^{\infty} \psi_i(x)\, dx = 1, \quad i = 1, \dots, N. \tag{2.6}$$

Proof. Similar to [11, p. 287]. It is omitted. $\square$

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

We give

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

$$\sum_{\substack{k=-\infty \\ |nx-k| \geq n^{1-\alpha}}}^{\infty} \psi_i(nx - k) < \frac{\left(1 - h_i\left(n^{1-\alpha} - 2\right)\right)}{2}, \quad i = 1, \dots, N. \tag{2.7}$$

Notice that

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

Proof. Let $x \geq 1$. That is $0 \leq x - 1 < x + 1$. Applying the mean value theorem we get

$$\psi_i(x) \overset{(2.1)}{=} \frac{1}{4} \cdot 2 \cdot h_i'(\xi) = \frac{h_i'(\xi)}{2}, \tag{2.8}$$

for some $x - 1 < \xi < x + 1$.

Since h_i' is strictly decreasing we obtain $h_i'(\xi) < h_i'(x-1)$ and

$$\psi_i(x) < \frac{h_i'(x-1)}{2}, \quad \forall\, x \geq 1. \tag{2.9}$$

Therefore we have

$$\begin{aligned}
\sum_{\substack{k=-\infty \\ |nx-k| \geq n^{1-\alpha}}}^{\infty} \psi_i(nx-k) &= \sum_{\substack{k=-\infty \\ |nx-k| \geq n^{1-\alpha}}}^{\infty} \psi_i(|nx-k|) < \frac{1}{2} \sum_{\substack{k=-\infty \\ |nx-k| \geq n^{1-\alpha}}}^{\infty} h_i'(|nx-k| - 1) \\
&\leq \frac{1}{2} \int_{(n^{1-\alpha}-1)}^{+\infty} h_i'(x-1)\, d(x-1) = \frac{1}{2} \left(h_i(x-1) \Big|_{(n^{1-\alpha}-1)}^{+\infty} \right) \\
&= \frac{1}{2} \left[h_i(+\infty) - h_i\left(n^{1-\alpha} - 2\right) \right] = \frac{1}{2} \left(1 - h_i\left(n^{1-\alpha} - 2\right) \right).
\end{aligned} \tag{2.10}$$

The claim is proved. $\square$

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

$$\frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi_i (nx-k)} < \frac{1}{\psi_i (1)}, \quad \forall\ x \in [a,b], \quad i = 1, \ldots, N. \tag{2.11}$$

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

Remark 2.5. *We have that*

$$\lim_{n \to \infty} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi_i (nx-k) \neq 1, \quad i = 1, \ldots, N, \tag{2.12}$$

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

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

Note 2.6. *For large enough n we always obtain $\lceil na \rceil \leq \lfloor nb \rfloor$. Also $a \leq \frac{k}{n} \leq b$, iff $\lceil na \rceil \leq k \leq \lfloor nb \rfloor$. In general it holds (by (2.5))*

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \psi_i (nx-k) \leq 1, \quad i = 1, \ldots, N. \tag{2.13}$$

We make

Remark 2.7. *We define*

$$Z(x_1, \ldots, x_N) := Z(x) := \prod_{i=1}^{N} \psi_i (x_i), \quad x = (x_1, \ldots, x_N) \in \mathbb{R}^N, \quad N \in \mathbb{N}. \tag{2.14}$$

It has the properties:

(i)

$$Z(x) > 0, \quad \forall x \in \mathbb{R}^N, \tag{2.15}$$

(ii)

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

$$\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \prod_{i=1}^{N} \psi_i (x_i - k_i) = \prod_{i=1}^{N} \left(\sum_{k_i=-\infty}^{\infty} \psi_i (x_i - k_i) \right) \overset{(2.5)}{=} 1.$$

Hence

$$\sum_{k=-\infty}^{\infty} Z(x-k) = 1. \tag{2.16}$$

That is

(iii)

$$\sum_{k=-\infty}^{\infty} Z(nx-k) = 1, \quad \forall x \in \mathbb{R}^N, \quad n \in \mathbb{N}. \tag{2.17}$$

and

(iv)

$$\int_{\mathbb{R}^N} Z(x)\,dx = \int_{\mathbb{R}^N} \left(\prod_{i=1}^{N} \psi_i(x_i) \right) dx_1 \cdots dx_N = \prod_{i=1}^{N} \left(\int_{-\infty}^{\infty} \psi_i(x_i)\,dx_i \right) \overset{(2.6)}{=} 1, \tag{2.18}$$

thus

$$\int_{\mathbb{R}^N} Z(x)\,dx = 1, \tag{2.19}$$

that is Z is a multivariate density function.

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

$$\lceil na \rceil := (\lceil na_1 \rceil, \ldots, \lceil na_N \rceil)$$

$$\lfloor nb \rfloor := (\lfloor nb_1 \rfloor, \ldots, \lfloor nb_N \rfloor),$$

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

We obviously see that

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k) = \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left(\prod_{i=1}^{N} \psi_i(nx_i - k_i) \right) = \sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \cdots \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} \left(\prod_{i=1}^{N} \psi_i(nx_i - k_i) \right)$$
$$= \prod_{i=1}^{N} \left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \psi_i(nx_i - k_i) \right). \tag{2.20}$$

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

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k) = \sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty \leq \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} Z(nx-k) + \sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\lfloor nb \rfloor} Z(nx-k). \tag{2.21}$$

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

(v) *We notice that*

$$\sum_{\substack{k=\lceil na\rceil \\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}}}^{\lfloor nb\rfloor} Z(nx-k) = \sum_{\substack{k_1=\lceil na_1\rceil \\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}}}^{\lfloor nb_1\rfloor} \cdots \sum_{k_N=\lceil na_N\rceil}^{\lfloor nb_N\rfloor} \left(\prod_{i=1}^{N} \psi_i(nx_i-k_i)\right)$$

$$= \prod_{i=1}^{N} \left(\sum_{\substack{k=\lceil na\rceil \\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}}}^{\lfloor nb_i\rfloor} \psi_i(nx_i-k_i)\right)$$

$$\leq \left(\prod_{\substack{i=1 \\ i\neq r}}^{N} \left(\sum_{k_i=-\infty}^{\infty} \psi_i(nx_i-k_i)\right)\right) \left(\sum_{\substack{k_r=\lceil na_r\rceil \\ \left|\frac{k_r}{n}-x_r\right|>\frac{1}{n^{\beta}}}}^{\lfloor nb_r\rfloor} \psi_r(nx_r-k_r)\right)$$

$$= \left(\sum_{\substack{k_r=\lceil na_r\rceil \\ \left|\frac{k_r}{n}-x_r\right|>\frac{1}{n^{\beta}}}}^{\lfloor nb_r\rfloor} \psi_r(nx_r-k_r)\right) \tag{2.22}$$

$$\leq \sum_{\substack{k_r=-\infty \\ \left|\frac{k_r}{n}-x_r\right|>\frac{1}{n^{\beta}}}}^{\infty} \psi_r(nx_r-k_r) = \sum_{\substack{k_r=-\infty \\ |nx_r-k_r|>n^{1-\beta}}}^{\infty} \psi_r(nx_r-k_r)$$

$$\overset{(2.7)}{<} \frac{1-h_r\left(n^{1-\beta}-2\right)}{2} \leq \max_{i\in\{1,\ldots,N\}} \left(\frac{1-h_i\left(n^{1-\beta}-2\right)}{2}\right),$$

where $0<\beta<1$.

That is we get:

$$\sum_{\substack{k=\lceil na\rceil \\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}}}^{\lfloor nb\rfloor} Z(nx-k) < \max_{i\in\{1,\ldots,N\}} \left(\frac{1-h_i\left(n^{1-\beta}-2\right)}{2}\right), \tag{2.23}$$

$0<\beta<1$*, with* $n\in\mathbb{N}: n^{1-\beta}>2$, $\forall\ x\in\prod_{i=1}^{N}[a_i,b_i]$.

(vi) *It is clear that*

$$\sum_{\substack{k=-\infty \\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}}}^{\infty} Z(nx-k) < \max_{i\in\{1,\ldots,N\}} \left(\frac{1-h_i\left(n^{1-\beta}-2\right)}{2}\right), \tag{2.24}$$

$0<\beta<1$, $n\in\mathbb{N}: n^{1-\beta}>2$, $\forall\ x\in\prod_{i=1}^{N}[a_i,b_i]$.

(*vii*) *By Theorem 2.4 we get that*

$$0 < \frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z\left(nx-k\right)} = \frac{1}{\prod_{i=1}^{N}\left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \psi_i\left(nx_i-k_i\right)\right)} < \frac{1}{\prod_{i=1}^{N}\psi_i\left(1\right)},$$

thus

$$0 < \frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z\left(nx-k\right)} < \frac{1}{\prod_{i=1}^{N}\psi_i\left(1\right)}, \tag{2.25}$$

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

Furthermore it holds

$$\begin{aligned}\lim_{n\to\infty}\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z\left(nx-k\right) &= \lim_{n\to\infty}\prod_{i=1}^{N}\left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor}\psi_i\left(nx_i-k_i\right)\right) \\ &= \prod_{i=1}^{N}\left(\lim_{n\to\infty}\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor}\psi_i\left(nx_i-k_i\right)\right) \neq 1,\end{aligned} \tag{2.26}$$

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

We state

Definition 2.8. *We denote by*

$$\delta_N\left(\beta,n\right) := \max_{i\in\{1,\ldots,N\}}\left(\frac{1-h_i\left(n^{1-\beta}-2\right)}{2}\right), \tag{2.27}$$

where $0<\beta<1$.

We make

Remark 2.9. *Here* $\left(X,\|\cdot\|_{\gamma}\right)$ *is a Banach space.*

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

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

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

$$= \frac{\sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor}\sum_{k_2=\lceil na_2 \rceil}^{\lfloor nb_2 \rfloor}\cdots\sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} f\left(\frac{k_1}{n},\ldots,\frac{k_N}{n}\right)\left(\prod_{i=1}^{N}\psi_i\left(nx_i-k_i\right)\right)}{\prod_{i=1}^{N}\left(\sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor}\psi_i\left(nx_i-k_i\right)\right)}. \tag{2.28}$$

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

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

$$\widetilde{A}_n(g, x) := \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} g\left(\frac{k}{n}\right) Z(nx - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx - k)}. \tag{2.29}$$

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

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

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

$$\|A_n(f, x)\|_\gamma \leq \frac{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\|f\left(\frac{k}{n}\right)\right\|_\gamma Z(nx - k)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx - k)} = \widetilde{A}_n\left(\|f\|_\gamma, x\right), \tag{2.30}$$

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

So, we have that

$$\|A_n(f, x)\|_\gamma \leq \widetilde{A}_n\left(\|f\|_\gamma, x\right), \tag{2.31}$$

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

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

$$A_n(cg, x) = c\widetilde{A}_n(g, x), \quad \forall x \in \prod_{i=1}^{N} [a_i, b_i]. \tag{2.32}$$

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

$$A_n(c) = c, \quad \forall c \in X. \tag{2.33}$$

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

For convenience we call

$$\begin{aligned} A_n^*(f, x) &:= \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} f\left(\frac{k}{n}\right) Z(nx - k) \\ &= \sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \sum_{k_2=\lceil na_2 \rceil}^{\lfloor nb_2 \rfloor} \cdots \sum_{k_N=\lceil na_N \rceil}^{\lfloor nb_N \rfloor} f\left(\frac{k_1}{n}, \dots, \frac{k_N}{n}\right) \left(\prod_{i=1}^{N} \psi_i(nx_i - k_i)\right), \end{aligned} \tag{2.34}$$

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

That is

$$A_n(f,x) := \frac{A_n^*(f,x)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)}, \tag{2.35}$$

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

Hence

$$A_n(f,x) - f(x) = \frac{A_n^*(f,x) - f(x)\left(\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)\right)}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)}. \tag{2.36}$$

Consequently we derive

$$\|A_n(f,x) - f(x)\|_\gamma \overset{(2.25)}{\leq} \left(\prod_{i=1}^{N} \psi_i(1)\right)^{-1} \left\|A_n^*(f,x) - f(x) \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} Z(nx-k)\right\|_\gamma, \tag{2.37}$$

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

We will estimate the right hand side of (2.37).

For the last and others we need

Definition 2.10 ([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*

$$\omega_1(f,\delta) := \sup_{\substack{x,y \in M \\ \|x-y\|_p \leq \delta}} \|f(x) - f(y)\|_\gamma, \quad 0 < \delta \leq diam(M). \tag{2.38}$$

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

$$\omega_1(f,\delta) = \omega_1(f, diam(M)). \tag{2.39}$$

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 2.11 ([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 (2.38). 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,

$$B_n(f, x) := B_n(f, x_1, \ldots, x_N) := \sum_{k=-\infty}^{\infty} f\left(\frac{k}{n}\right) Z(nx - k)$$
$$:= \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} f\left(\frac{k_1}{n}, \frac{k_2}{n}, \ldots, \frac{k_N}{n}\right) \left(\prod_{i=1}^{N} \psi_i(nx_i - k_i)\right), \tag{2.40}$$

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

$$C_n(f, x) := C_n(f, x_1, \ldots, x_N) := \sum_{k=-\infty}^{\infty} \left(n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\, dt\right) Z(nx - k) = \tag{2.41}$$

$$\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \left(n^N \int_{\frac{k_1}{n}}^{\frac{k_1+1}{n}} \int_{\frac{k_2}{n}}^{\frac{k_2+1}{n}} \cdots \int_{\frac{k_N}{n}}^{\frac{k_N+1}{n}} f(t_1, \ldots, t_N)\, dt_1 \ldots dt_N\right) \cdot \left(\prod_{i=1}^{N} \psi_i(nx_i - k_i)\right),$$

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

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

$$\delta_{nk}(f) := \delta_{n, k_1, k_2, \ldots, k_N}(f) := \sum_{r=0}^{\theta} w_r f\left(\frac{k}{n} + \frac{r}{n\theta}\right)$$
$$= \sum_{r_1=0}^{\theta_1} \sum_{r_2=0}^{\theta_2} \cdots \sum_{r_N=0}^{\theta_N} w_{r_1, r_2, \ldots, r_N} f\left(\frac{k_1}{n} + \frac{r_1}{n\theta_1}, \frac{k_2}{n} + \frac{r_2}{n\theta_2}, \ldots, \frac{k_N}{n} + \frac{r_N}{n\theta_N}\right), \tag{2.42}$$

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

We set

$$D_n(f, x) := D_n(f, x_1, \ldots, x_N) := \sum_{k=-\infty}^{\infty} \delta_{nk}(f)\, Z(nx - k) \tag{2.43}$$
$$= \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \delta_{n, k_1, k_2, \ldots, k_N}(f) \left(\prod_{i=1}^{N} \psi_i(nx_i - k_i)\right), \quad \forall x \in \mathbb{R}^N.$$

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)

$$\|A_n(f,x)-f(x)\|_\gamma \le \left(\prod_{i=1}^{N}\psi_i(1)\right)^{-1}\left[\omega_1\left(f,\frac{1}{n^\beta}\right)+2\delta_N(\beta,n)\left\|\|f\|_\gamma\right\|_\infty\right] =: \lambda_1(n), \tag{3.1}$$

and

2)

$$\left\|\|A_n(f)-f\|_\gamma\right\|_\infty \le \lambda_1(n). \tag{3.2}$$

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* $\max\left(\frac{1}{n^\beta},\delta_N(\beta,n)\right)$.

Proof. As similar to [12] is omitted. Use of (2.37). □

We make

Remark 3.2 ([11, pp. 263–266]). *Let* $\left(\mathbb{R}^N,\|\cdot\|_p\right)$, $N\in\mathbb{N}$; *where* $\|\cdot\|_p$ *is the* L_p*-norm,* $1\le p\le\infty$. $\mathbb{R}^N$ *is a Banach space, and* $\left(\mathbb{R}^N\right)^j$ *denotes the* j*-fold product space* $\mathbb{R}^N\times\cdots\times\mathbb{R}^N$ *endowed with the max-norm* $\|x\|_{(\mathbb{R}^N)^j} := \max_{1\le\lambda\le j}\|x_\lambda\|_p$, *where* $x:=(x_1,\ldots,x_j)\in\left(\mathbb{R}^N\right)^j$.

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* $g:\left(\mathbb{R}^N\right)^j\to X$, $j=1,\ldots,m$, *is a Banach space with norm*

$$\|g\| := \|g\|_{L_j} := \sup_{\left(\|x\|_{(\mathbb{R}^N)^j}=1\right)}\|g(x)\|_\gamma = \sup\frac{\|g(x)\|_\gamma}{\|x_1\|_p\cdots\|x_j\|_p}. \tag{3.3}$$

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 [20])* $f^{(j)}:O\to L_j=L_j\left(\left(\mathbb{R}^N\right)^j;X\right)$ *exist and are continuous for* $1\le j\le m$, $m\in\mathbb{N}$.

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

We will work with $f|_M$.

Then, by Taylor's formula [13], [20, p. 124] , we get

$$f(x) = \sum_{j=0}^{m} \frac{f^{(j)}(x_0)(x-x_0)^j}{j!} + R_m(x,x_0), \quad \text{all } x \in M, \tag{3.4}$$

where the remainder is the Riemann integral

$$R_m(x,x_0) := \int_0^1 \frac{(1-u)^{m-1}}{(m-1)!} \left(f^{(m)}(x_0+u(x-x_0)) - f^{(m)}(x_0)\right)(x-x_0)^m \, du, \tag{3.5}$$

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

We consider

$$w := \omega_1\left(f^{(m)},h\right) := \sup_{\substack{x,y\in M \\ \|x-y\|_p \le h}} \left\|f^{(m)}(x) - f^{(m)}(y)\right\|, \tag{3.6}$$

$h > 0$.

We obtain

$$\begin{gathered}\left\|\left(f^{(m)}(x_0+u(x-x_0)) - f^{(m)}(x_0)\right)(x-x_0)^m\right\|_\gamma \le \\ \left\|f^{(m)}(x_0+u(x-x_0)) - f^{(m)}(x_0)\right\| \cdot \|x-x_0\|_p^m \le w\,\|x-x_0\|_p^m \left\lceil \frac{u\,\|x-x_0\|_p}{h} \right\rceil,\end{gathered} \tag{3.7}$$

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

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

$$\|R_m(x,x_0)\|_\gamma \le w\,\|x-x_0\|_p^m \int_0^1 \left\lceil \frac{u\,\|x-x_0\|_p}{h} \right\rceil \frac{(1-u)^{m-1}}{(m-1)!} du = w\Phi_m\left(\|x-x_0\|_p\right) \tag{3.8}$$

by a change of variable, where

$$\Phi_m(t) := \int_0^{|t|} \left\lceil \frac{s}{h} \right\rceil \frac{(|t|-s)^{m-1}}{(m-1)!} ds = \frac{1}{m!}\left(\sum_{j=0}^{\infty} (|t|-jh)_+^m\right), \quad \forall t \in \mathbb{R}, \tag{3.9}$$

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

Also from there we get

$$\Phi_m(t) \le \left(\frac{|t|^{m+1}}{(m+1)!h} + \frac{|t|^m}{2m!} + \frac{h\,|t|^{m-1}}{8(m-1)!}\right), \quad \forall t \in \mathbb{R}, \tag{3.10}$$

with equality true only at $t=0$.

Therefore it holds

$$\left\|R_m(x,x_0)\right\|_\gamma \leq w\left(\frac{\|x-x_0\|_p^{m+1}}{(m+1)!h}+\frac{\|x-x_0\|_p^m}{2m!}+\frac{h\|x-x_0\|_p^{m-1}}{8(m-1)!}\right),\quad \forall x\in M. \tag{3.11}$$

We have found that

$$\left\|f(x)-\sum_{j=0}^m \frac{f^{(j)}(x_0)(x-x_0)^j}{j!}\right\|_\gamma \leq \omega_1\left(f^{(m)},h\right)\left(\frac{\|x-x_0\|_p^{m+1}}{(m+1)!h}+\frac{\|x-x_0\|_p^m}{2m!}+\frac{h\|x-x_0\|_p^{m-1}}{8(m-1)!}\right) < \infty, \tag{3.12}$$

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

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

One can rewrite (3.12) *as follows:*

$$\left\|f(\cdot)-\sum_{j=0}^m \frac{f^{(j)}(x_0)(\cdot-x_0)^j}{j!}\right\|_\gamma \leq \omega_1\left(f^{(m)},h\right)\left(\frac{\|\cdot-x_0\|_p^{m+1}}{(m+1)!h}+\frac{\|\cdot-x_0\|_p^m}{2m!}+\frac{h\|\cdot-x_0\|_p^{m-1}}{8(m-1)!}\right), \tag{3.13}$$

$\forall x_0\in M$, *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

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

$$\omega_1\left(f^{(m)},h\right)\left[\frac{\left(\widetilde{L}_N\left(\|\cdot-x_0\|_p^{m+1}\right)\right)(x_0)}{(m+1)!h}+\frac{\left(\widetilde{L}_N\left(\|\cdot-x_0\|_p^{m}\right)\right)(x_0)}{2m!}+\frac{h\left(\widetilde{L}_N\left(\|\cdot-x_0\|_p^{m-1}\right)\right)(x_0)}{8(m-1)!}\right], \tag{3.14}$$

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

Clearly (3.14) is valid when $M=\prod_{i=1}^N [a_i,b_i]$ and $\widetilde{L}_n=\widetilde{A}_n$, see (2.29).

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 in [11, pp. 268-270]. The operators A_n, $\widetilde{A}_n$ fulfill its assumptions, see (2.28), (2.29), (2.31), (2.32) and (2.33).

We present the following high order approximation results.

Theorem 3.3. *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)

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

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

$$\left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{m r^2}{8} \right], \tag{3.15}$$

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

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

$$\frac{\omega_1 \left(f^{(m)}, r \left(\left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^{m+1} \right) \right)(x_0) \right)^{\frac{1}{m+1}} \right)}{r m!} \left(\left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^{m+1} \right) \right)(x_0) \right)^{\left(\frac{m}{m+1}\right)} \tag{3.16}$$

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

3)

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

$$\frac{\omega_1 \left(f^{(m)}, r \left(\left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^{m+1} \right) \right)(x_0) \right)^{\frac{1}{m+1}} \right)}{r m!} \left(\left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^{m+1} \right) \right)(x_0) \right)^{\left(\frac{m}{m+1}\right)} \tag{3.17}$$

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

4)

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

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

$$\left\| \left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^{m+1} \right) \right)(x_0) \right\|_{\infty, x_0 \in \prod_{i=1}^{N}[a_i,b_i]}^{\left(\frac{m}{m+1}\right)} \left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8} \right]. \tag{3.18}$$

We need

Lemma 3.4. *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, [11, p. 272]. ☐

We make

Remark 3.5. *By [11, Remark 10.4, p. 273], we get that*

$$\left\| \left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^{k} \right) \right)(x_0) \right\|_{\infty, x_0 \in \prod_{i=1}^{N}[a_i,b_i]} \leq \left\| \left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^{m+1} \right) \right)(x_0) \right\|_{\infty, x_0 \in \prod_{i=1}^{N}[a_i,b_i]}^{\left(\frac{k}{m+1}\right)}, \tag{3.19}$$

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

We give

Corollary 3.6 (to Theorem 3.3, case of $m = 1$)**.** *Then*

1)

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

$$+ \frac{1}{2r} \omega_1 \left(f^{(1)}, r \left(\left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^2 \right) \right)(x_0) \right)^{\frac{1}{2}} \right) \left(\left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^2 \right) \right)(x_0) \right)^{\frac{1}{2}} \left[1 + r + \frac{r^2}{4} \right], \tag{3.20}$$

2)

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

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

$$\left\| \left(\widetilde{A}_n \left(\|\cdot - x_0\|_p^2 \right) \right)(x_0) \right\|_{\infty, x_0 \in \prod_{i=1}^{N}[a_i,b_i]}^{\frac{1}{2}} \left[1 + r + \frac{r^2}{4} \right], \quad r > 0. \tag{3.21}$$

We make

Remark 3.7. *We estimate* ($0 < \alpha < 1$, $m, n \in \mathbb{N} : n^{1-\alpha} > 2$),

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

$$\overset{(2.25)}{<} \left(\prod_{i=1}^{N} \psi_i(1) \right)^{-1} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z(nx_0 - k) \tag{3.22}$$

$$= \left(\prod_{i=1}^{N} \psi_i(1) \right)^{-1} \left\{ \sum_{\substack{k=\lceil na \rceil \\ \left\| \frac{k}{n} - x_0 \right\|_\infty \leq \frac{1}{n^\alpha}}}^{\lfloor nb \rfloor} \left\| \frac{k}{n} - x_0 \right\|_\infty^{m+1} Z(nx_0 - k) \right.$$

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

$$\overset{(2.23)}{\leq} \left(\prod_{i=1}^{N} \psi_i(1) \right)^{-1} \left\{ \frac{1}{n^{\alpha(m+1)}} + \delta_N(\alpha, n) \|b - a\|_\infty^{m+1} \right\}, \tag{3.23}$$

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

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

$$\widetilde{A}_n \left(\|\cdot - x_0\|_\infty^{m+1} \right)(x_0) < \left(\prod_{i=1}^{N} \psi_i(1) \right)^{-1} \left\{ \frac{1}{n^{\alpha(m+1)}} + \delta_N(\alpha, n) \|b - a\|_\infty^{m+1} \right\} =: \varphi_1(n) \tag{3.24}$$

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

And, consequently it holds

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

$$= \varphi_1(n) \to 0, \quad \text{as } n \to +\infty. \tag{3.25}$$

So, we have that $\varphi_1(n) \to 0$, *as* $n \to +\infty$. *Thus, when* $p \in [1, \infty]$, *from Theorem 3.3 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

$$\left(\widetilde{A}_n\left(f^{(j)}(x_0)(\cdot - x_0)^j\right)\right)(x_0) = \frac{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} f^{(j)}(x_0)\left(\frac{k}{n}-x_0\right)^j Z(nx_0-k)}{\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor} Z(nx_0-k)}. \tag{3.26}$$

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

$$\left\| f^{(j)}(x_0)\left(\frac{k}{n}-x_0\right)^j\right\|_\gamma \le \left\| f^{(j)}(x_0)\right\| \left\|\frac{k}{n}-x_0\right\|_\infty^j. \tag{3.27}$$

We further have that

$$\begin{aligned}
&\left\|\left(\widetilde{A}_n\left(f^{(j)}(x_0)(\cdot - x_0)^j\right)\right)(x_0)\right\|_\gamma \overset{(2.25)}{<} \\
&\left(\prod_{i=1}^N \psi_i(1)\right)^{-1}\left(\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\| f^{(j)}(x_0)\left(\frac{k}{n}-x_0\right)^j\right\|_\gamma Z(nx_0-k)\right) \le \\
&\left(\prod_{i=1}^N \psi_i(1)\right)^{-1}\left(\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\| f^{(j)}(x_0)\right\|\left\|\frac{k}{n}-x_0\right\|_\infty^j Z(nx_0-k)\right) = \\
&\left(\prod_{i=1}^N \psi_i(1)\right)^{-1}\left\| f^{(j)}(x_0)\right\|\left(\sum_{k=\lceil na\rceil}^{\lfloor nb\rfloor}\left\|\frac{k}{n}-x_0\right\|_\infty^j Z(nx_0-k)\right) =
\end{aligned} \tag{3.28}$$

$$\begin{aligned}
&\left(\prod_{i=1}^N \psi_i(1)\right)^{-1}\left\| f^{(j)}(x_0)\right\|\left\{\sum_{\substack{k=\lceil na\rceil \\ \left\|\frac{k}{n}-x_0\right\|_\infty \le \frac{1}{n^\alpha}}}^{\lfloor nb\rfloor}\left\|\frac{k}{n}-x_0\right\|_\infty^j Z(nx_0-k)\right. \\
&\left. + \sum_{\substack{k=\lceil na\rceil \\ \left\|\frac{k}{n}-x_0\right\|_\infty > \frac{1}{n^\alpha}}}^{\lfloor nb\rfloor}\left\|\frac{k}{n}-x_0\right\|_\infty^j Z(nx_0-k)\right\} \overset{(2.23)}{\le}
\end{aligned} \tag{3.29}$$

$$\left(\prod_{i=1}^N \psi_i(1)\right)^{-1}\left\| f^{(j)}(x_0)\right\|\left\{\frac{1}{n^{\alpha j}} + \delta_N(\alpha,n)\|b-a\|_\infty^j\right\} \to 0, \quad \text{as } n\to\infty.$$

That is

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

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

$$\begin{aligned}\left\|\left(\widetilde{A}_n\left(f^{(j)}(x_0)(\cdot-x_0)^j\right)\right)(x_0)\right\|_\gamma &< \left(\prod_{i=1}^{N}\psi_i(1)\right)^{-1}\left\|f^{(j)}(x_0)\right\|\left\{\frac{1}{n^{\alpha j}}+\delta_N(\alpha,n)\|b-a\|_\infty^j\right\}\\ &\leq \left(\prod_{i=1}^{N}\psi_i(1)\right)^{-1}\left\|f^{(j)}\right\|_\infty\left\{\frac{1}{n^{\alpha j}}+\delta_N(\alpha,n)\|b-a\|_\infty^j\right\}\\ &=: \varphi_{2j}(n)<\infty, \qquad (3.30)\end{aligned}$$

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

We conclude:

In Theorem 3.3, the right hand sides of (3.26) and (3.18) converge to zero as $n\to\infty$, for any $p\in[1,\infty]$.

Also in Corollary 3.6, the right hand sides of (3.20) and (3.21) converge to zero as $n\to\infty$, for any $p\in[1,\infty]$.

Conclusion 3.8. *We have proved that the left hand sides of* (3.15), (3.16), (3.17), (3.18) *and* (3.20), (3.21) *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* (3.16)*. Higher speed of convergence happens also to the left hand side of* (3.15)*.*

We give

Corollary 3.9 (to Theorem 3.3)**.** *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* (3.24) *and $\varphi_{2j}(n)$ as in* (3.30), *where $n\in\mathbb{N}:n^{1-\alpha}>2$, $0<\alpha<1$, $j=1,\ldots,m$. Then*

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

$$\frac{\omega_1\left(f^{(m)},r(\varphi_1(n))^{\frac{1}{m+1}}\right)}{rm!}(\varphi_1(n))^{\left(\frac{m}{m+1}\right)}\left[\frac{1}{(m+1)}+\frac{r}{2}+\frac{mr^2}{8}\right], \qquad (3.31)$$

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

$$\left\| (A_n(f))(x_0) - f(x_0) \right\|_\gamma \leq \frac{\omega_1\left(f^{(m)}, r(\varphi_1(n))^{\frac{1}{m+1}}\right)}{rm!} (\varphi_1(n))^{\left(\frac{m}{m+1}\right)} \left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8}\right], \tag{3.32}$$

3)

$$\left\| \|A_n(f) - f\|_\gamma \right\|_{\infty, \prod_{i=1}^N [a_i,b_i]} \leq \sum_{j=1}^m \frac{\varphi_{2j}(n)}{j!} + \frac{\omega_1\left(f^{(m)}, r(\varphi_1(n))^{\frac{1}{m+1}}\right)}{rm!} (\varphi_1(n))^{\left(\frac{m}{m+1}\right)}$$
$$\cdot \left[\frac{1}{(m+1)} + \frac{r}{2} + \frac{mr^2}{8}\right] =: \varphi_3(n) \to 0, \quad \text{as } n \to \infty. \tag{3.33}$$

We continue with

Theorem 3.10. *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)
$$\|B_n(f,x) - f(x)\|_\gamma \leq \omega_1\left(f, \frac{1}{n^\beta}\right) + 2\delta_N(\beta,n) \left\| \|f\|_\gamma \right\|_\infty =: \lambda_2(n), \tag{3.34}$$

2)
$$\left\| \|B_n(f) - f\|_\gamma \right\|_\infty \leq \lambda_2(n). \tag{3.35}$$

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* $\max\left(\frac{1}{n^\beta}, \delta_N(\beta,n)\right)$.

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

We give

Theorem 3.11. *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)
$$\|C_n(f,x) - f(x)\|_\gamma \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + 2\delta_N(\beta,n) \left\| \|f\|_\gamma \right\|_\infty =: \lambda_3(n), \tag{3.36}$$

2)
$$\left\| \|C_n(f) - f\|_\gamma \right\|_\infty \leq \lambda_3(n). \tag{3.37}$$

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 3.12. *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)
$$\left\| D_n(f,x) - f(x) \right\|_\gamma \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + 2\delta_N(\beta,n) \left\| \|f\|_\gamma \right\|_\infty = \lambda_4(n), \tag{3.38}$$

2)
$$\left\| \left\| D_n(f) - f \right\|_\gamma \right\|_\infty \leq \lambda_4(n). \tag{3.39}$$

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 3.13. *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*

$$F_n(f,x) := \sum_{k=-\infty}^{\infty} l_{nk}(f)\, Z(nx-k) = \begin{cases} B_n(f,x), & \text{if } l_{nk}(f) = f\left(\frac{k}{n}\right), \\ C_n(f,x), & \text{if } l_{nk}(f) = n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\,dt, \\ D_n(f,x), & \text{if } l_{nk}(f) = \delta_{nk}(f). \end{cases} \tag{3.40}$$

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

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 3.14. *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 3.15. *By* (2.28) *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

$$\left\| \left\| L_n^2(f) \right\|_\gamma \right\|_\infty = \left\| \left\| L_n(L_n(f)) \right\|_\gamma \right\|_\infty \leq \left\| \left\| L_n(f) \right\|_\gamma \right\|_\infty \leq \left\| \|f\|_\gamma \right\|_\infty, \tag{3.41}$$

etc.

Therefore we get

$$\left\| \left\| L_n^k (f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty , \quad \forall\ k \in \mathbb{N}, \tag{3.42}$$

the contraction property.

Also we see that

$$\left\| \left\| L_n^k (f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| L_n^{k-1} (f) \right\|_\gamma \right\|_\infty \leq \cdots \leq \left\| \left\| L_n (f) \right\|_\gamma \right\|_\infty \leq \left\| \left\| f \right\|_\gamma \right\|_\infty . \tag{3.43}$$

Here L_n^k are bounded linear operators.

Notation 3.16. *Here $N \in \mathbb{N}$, $0 < \beta < 1$. Denote by*

$$c_N := \begin{cases} \left(\prod_{i=1}^{N} \psi_i (1) \right)^{-1}, & \text{if} \quad L_n = A_n, \\ 1, \quad \text{if } L_n = B_n, C_n, D_n, \end{cases} \tag{3.44}$$

$$\varphi (n) := \begin{cases} \frac{1}{n^\beta}, \ \text{if } L_n = A_n, \quad B_n, \\ \frac{1}{n} + \frac{1}{n^\beta}, \quad \text{if } L_n = C_n, D_n, \end{cases} \tag{3.45}$$

$$\Omega := \begin{cases} C \left(\prod_{i=1}^{N} [a_i, b_i] , X \right), \ \text{if } L_n = A_n, \\ C_B \left(\mathbb{R}^N, X \right), \quad \text{if} \quad L_n = B_n, C_n, D_n, \end{cases} \tag{3.46}$$

and

$$Y := \begin{cases} \prod_{i=1}^{N} [a_i, b_i] , \ \text{if } L_n = A_n, \\ \mathbb{R}^N, \ \text{if } L_n = B_n, C_n, D_n. \end{cases} \tag{3.47}$$

We give the condensed

Theorem 3.17. *Let $f \in \Omega$, $0 < \beta < 1$, $x \in Y$; n, $N \in \mathbb{N}$ with $n^{1-\beta} > 2$. Then*

(i)

$$\left\| L_n (f, x) - f (x) \right\|_\gamma \leq c_N \left[\omega_1 (f, \varphi (n)) + 2\delta_N (\beta, n) \left\| \left\| f \right\|_\gamma \right\|_\infty \right] =: \tau (n) , \tag{3.48}$$

where ω_1 is for $p = \infty$,

(ii)

$$\left\| \left\| L_n (f) - f \right\|_\gamma \right\|_\infty \leq \tau (n) \to 0, \ \text{as } n \to \infty. \tag{3.49}$$

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, 3.10, 3.11, 3.12. □

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

We make

Remark 3.18. *Let $r \in \mathbb{N}$ and L_n as above. We observe that*

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

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

Then

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

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

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

$$+ \ldots + \left\| \|L_n (L_n f - f)\|_\gamma \right\|_\infty + \left\| \|L_n f - f\|_\gamma \right\|_\infty \leq r \left\| \|L_n f - f\|_\gamma \right\|_\infty . \tag{3.50}$$

That is

$$\left\| \|L_n^r f - f\|_\gamma \right\|_\infty \leq r \left\| \|L_n f - f\|_\gamma \right\|_\infty . \tag{3.51}$$

We give

Theorem 3.19. *All here as in Theorem 3.17 and $r \in \mathbb{N}$, $\tau(n)$ as in (3.48). Then*

$$\left\| \|L_n^r f - f\|_\gamma \right\|_\infty \leq r\tau(n). \tag{3.52}$$

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

Proof. By (3.51) and (3.49). □

We make

Remark 3.20. *Let $m_1, \ldots, m_r \in \mathbb{N} : m_1 \leq m_2 \leq \cdots \leq m_r$, $0 < \beta < 1$, $f \in \Omega$. Then $\varphi(m_1) \geq \varphi(m_2) \geq \cdots \geq \varphi(m_r)$, φ as in (3.45).*

Therefore

$$\omega_1(f, \varphi(m_1)) \geq \omega_1(f, \varphi(m_2)) \geq \cdots \geq \omega_1(f, \varphi(m_r)). \tag{3.53}$$

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

$$\delta_N(\beta, m_1) \geq \delta_N(\beta, m_2) \geq \cdots \geq \delta_N(\beta, m_r). \tag{3.54}$$

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

We write

$$
\begin{aligned}
& L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_2}\left(L_{m_1} f\right)\right)\right) - f = \\
& L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_2}\left(L_{m_1} f\right)\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_2} f\right)\right) + \\
& L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_2} f\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_3} f\right)\right) + \\
& L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_3} f\right)\right) - L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_4} f\right)\right) + \dots + \\
& L_{m_r}\left(L_{m_{r-1}} f\right) - L_{m_r} f + L_{m_r} f - f = \\
& L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_2}\right)\right)\left(L_{m_1} f - f\right) + L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_3}\right)\right)\left(L_{m_2} f - f\right) + \\
& L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_4}\right)\right)\left(L_{m_3} f - f\right) + \dots + L_{m_r}\left(L_{m_{r-1}} f - f\right) + L_{m_r} f - f.
\end{aligned}
\tag{3.55}
$$

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

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

(repeatedly applying (3.41)*)*

$$
\begin{aligned}
& \leq \left\|\left\|L_{m_1} f - f\right\|_\gamma\right\|_\infty + \left\|\left\|L_{m_2} f - f\right\|_\gamma\right\|_\infty + \left\|\left\|L_{m_3} f - f\right\|_\gamma\right\|_\infty \\
& + \dots + \left\|\left\|L_{m_{r-1}} f - f\right\|_\gamma\right\|_\infty + \left\|\left\|L_{m_r} f - f\right\|_\gamma\right\|_\infty = \sum_{i=1}^{r} \left\|\left\|L_{m_i} f - f\right\|_\gamma\right\|_\infty .
\end{aligned}
\tag{3.56}
$$

That is, we proved

$$
\left\|\left\|L_{m_r}\left(L_{m_{r-1}}\left(\dots L_{m_2}\left(L_{m_1} f\right)\right)\right) - f\right\|_\gamma\right\|_\infty \leq \sum_{i=1}^{r} \left\|\left\|L_{m_i} f - f\right\|_\gamma\right\|_\infty . \tag{3.57}
$$

We give

Theorem 3.21. *Let $f \in \Omega$; N, $m_1, m_2, \dots, m_r \in \mathbb{N} : m_1 \leq m_2 \leq \dots \leq m_r$, $0 < \beta < 1$; $m_i^{1-\beta} > 2$, $i = 1, \dots, r$, $x \in Y$, and let $(L_{m_1}, \dots, L_{m_r})$ as $(A_{m_1}, \dots, A_{m_r})$ or $(B_{m_1}, \dots, B_{m_r})$ or $(C_{m_1}, \dots, C_{m_r})$ or $(D_{m_1}, \dots, D_{m_r})$, $p = \infty$. Then*

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

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

$$\leq r c_N \left[\omega_1 \left(f, \varphi \left(m_1 \right) \right) + 2\delta_N \left(\beta, m_1 \right) \left\| \|f\|_{\gamma} \right\|_{\infty} \right]. \tag{3.58}$$

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 (3.57), (3.53), (3.54) and (3.48), (3.49). □

We continue with

Theorem 3.22. *Let all as in Corollary 3.9, and* $r \in \mathbb{N}$. *Here* $\varphi_3(n)$ *is as in* (3.33). *Then*

$$\left\| \|A_n^r f - f\|_{\gamma} \right\|_{\infty} \leq r \left\| \|A_n f - f\|_{\gamma} \right\|_{\infty} \leq r \varphi_3(n). \tag{3.59}$$

Proof. By (3.51) and (3.33). □

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

Theorem 3.23. *Let* $p_1 \geq 1$, $f \in C\left(\prod_{i=1}^{N} [a_i, b_i], X \right)$, $0 < \beta < 1$; $N, n \in \mathbb{N}$ *with* $n^{1-\beta} > 2$, *and* $\lambda_1(n)$ *as in (3.1),* ω_1 *is for* $p = \infty$. *Then*

$$\left\| \|A_n(f) - f\|_{\gamma} \right\|_{p_1, \prod_{i=1}^{N} [a_i, b_i]} \leq \lambda_1(n) \left(\prod_{i=1}^{N} (b_i - a_i) \right)^{\frac{1}{p_1}}. \tag{3.60}$$

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

Proof. Obvious, by integrating (3.1), etc. □

It follows

Theorem 3.24. *Let* $p_1 \geq 1$, $f \in C_B\left(\mathbb{R}^N, X\right)$, $0 < \beta < 1$; $N, n \in \mathbb{N}$ *with* $n^{1-\beta} > 2$, *and* ω_1 *is for* $p = \infty$; $\lambda_2(n)$ *as in* (3.34) *and* K *a compact subset of* $\mathbb{R}^N$. *Then*

$$\left\| \|B_n(f) - f\|_{\gamma} \right\|_{p_1, K} \leq \lambda_2(n) |K|^{\frac{1}{p_1}}, \tag{3.61}$$

where $|K| < \infty$, *is the Lebesgue measure of* K.

We notice that $\lim_{n \to \infty} \left\| \|B_n(f) - f\|_{\gamma} \right\|_{p_1, K} = 0$, *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 (3.34), etc. □

Next come

Theorem 3.25. *All as in Theorem 3.24, but now we use $\lambda_3(n)$ of (3.36). Then*

$$\left\| \|C_n(f) - f\|_\gamma \right\|_{p_1,K} \leq \lambda_3(n) |K|^{\frac{1}{p_1}}. \tag{3.62}$$

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

Proof. By (3.36). □

Theorem 3.26. *All as in Theorem 3.24, but now we use $\lambda_4(n)$ of (3.38). Then*

$$\left\| \|D_n(f) - f\|_\gamma \right\|_{p_1,K} \leq \lambda_4(n) |K|^{\frac{1}{p_1}}. \tag{3.63}$$

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

Proof. By (3.38). □

Application 3.27. *A typical application of all of our results is when* $\left(X, \|\cdot\|_\gamma\right) = (\mathbb{C}, |\cdot|)$, *where* $\mathbb{C}$ *are the complex numbers.*

References

[1] G. A. Anastassiou, *Moments in probability and approximation theory*, ser. Pitman Research Notes in Mathematics Series. New York, USA: Longman Scientific & Technical, Harlow; copublished with John Wiley & Sons, Inc., 1993, vol. 287.

[2] G. A. Anastassiou, "Rate of convergence of some neural network operators to the unit-univariate case," *J. Math. Anal. Appl.*, vol. 212, no. 1, pp. 237–262, 1997, doi: 10.1006/jmaa.1997.5494.

[3] G. A. Anastassiou, *Quantitative approximations.* Boca Raton, FL, USA: Chapman & Hall/CRC, 2001.

[4] G. A. Anastassiou, *Intelligent systems: approximation by artificial neural networks*, ser. Intelligent Systems Reference Library. Berlin, Germany: Springer-Verlag, 2011, vol. 19.

[5] G. A. Anastassiou, "Multivariate hyperbolic tangent neural network approximation," *Comput. Math. Appl.*, vol. 61, no. 4, pp. 809–821, 2011, doi: 10.1016/j.camwa.2010.12.029.

[6] G. A. Anastassiou, "Multivariate sigmoidal neural network approximation," *Neural Networks*, vol. 24, no. 4, pp. 378–386, 2011, doi: 10.1016/j.neunet.2011.01.003.

[7] G. A. Anastassiou, "Univariate hyperbolic tangent neural network approximation," *Math. Comput. Modelling*, vol. 53, no. 5–6, pp. 1111–1132, 2011, doi: 10.1016/j.mcm.2010.11.072.

[8] G. A. Anastassiou, "Univariate sigmoidal neural network approximation," *J. Comput. Anal. Appl.*, vol. 14, no. 4, pp. 659–690, 2012, doi: 10.1007/978-3-642-21431-8_1.

[9] G. A. Anastassiou, "Approximation by neural networks iterates," in *Advances in applied mathematics and approximation theory*, ser. Springer Proc. Math. Stat. New York, USA: Springer, New York, 2013, vol. 41, pp. 1–20.

[10] G. A. Anastassiou, *Intelligent systems II: complete approximation by neural network operators*, ser. Studies in Computational Intelligence. Cham, Switzerland: Springer, 2016, vol. 608, doi: 10.1007/978-3-319-20505-2.

[11] G. A. Anastassiou, *Intelligent computations: abstract fractional calculus, inequalities, approximations*, ser. Studies in Computational Intelligence. Cham, Switzerland: Springer, 2018, vol. 734, doi: 10.1007/978-3-319-66936-6.

[12] G. A. Anastassiou, "General multivariate arctangent function activated neural network approximations," *J. Numer. Anal. Approx. Theory*, vol. 51, no. 1, pp. 37–66, 2022.

[13] H. Cartan, *Differential calculus.* Paris, France; Boston, MA, USA: Hermann; Houghton Mifflin Co., 1971.

[14] Z. Chen and F. Cao, "The approximation operators with sigmoidal functions," *Comput. Math. Appl.*, vol. 58, no. 4, pp. 758–765, 2009, doi: 10.1016/j.camwa.2009.05.001.

[15] D. Costarelli and R. Spigler, "Approximation results for neural network operators activated by sigmoidal functions," *Neural Networks*, vol. 44, pp. 101–106, 2013, doi: 10.1016/j.neunet.2013.03.015.

[16] D. Costarelli and R. Spigler, "Multivariate neural network operators with sigmoidal activation functions," *Neural Networks*, vol. 48, pp. 72–77, 2013, doi: 10.1016/j.neunet.2013.07.009.

[17] S. Haykin, *Neural Networks: A Comprehensive Foundation.* New York, USA: Prentice Hall, 1999.

[18] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, pp. 115–133, 1943, doi: 10.1007/bf02478259.

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

[20] L. B. Rall, *Computational solution of nonlinear operator equations.* New York-London-Sydney, USA-UK-Australia: John Wiley & Sons, Inc., 1969, with an appendix by Ramon E. Moore.