Acta Math. Univ. Comenianae
Vol. LXXXVIII, 1 (2019), pp. 113–130

QUANTITATIVE APPROXIMATION BY KANTOROVICH-CHOQUET QUASI-INTERPOLATION NEURAL NETWORK OPERATORS

G. A. ANASTASSIOU

Abstract. In this article, we present univariate and multivariate basic approximation by Kantorovich-Choquet type quasi-interpolation neural network operators with respect to supremum norm. This is done with rates using the first univariate and multivariate moduli of continuity. We approximate continuous and bounded functions on $\mathbb{R}^N$, $N \in \mathbb{N}$. When they are also uniformly continuous, we have pointwise and uniform convergences.

1. Introduction

The author in [**2**] and [**3**], see Chapters 2–5, was the first to establish neural network approximations to continuous functions with rates by very specifically defined neural network operators of Cardaliagnet-Euvrard and "Squashing" types, by employing the modulus of continuity of the engaged function or its high order derivative, and producing very tight Jackson type inequalities. He treats there both the univariate and multivariate cases. Defining these operators "bell-shaped" and "squashing" functions are assumed to be 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.

The author inspired by [**12**], continued his studies on neural networks approximation by introducing and using the proper quasi-interpolation operators of sigmoidal and hyperbolic tangent type which resulted into [**4**], [**5**], [**6**], [**7**], [**8**], by treating both the univariate and multivariate cases. He did also the corresponding fractional case [**9**].

The author here performs univariate and multivariate error function based neural network approximations to continuous functions over the whole $\mathbb{R}^N$, $N \in \mathbb{N}$, then he extends his results to complex valued functions. He also finds similar results when the activation function is induced by the sigmoidal and hyperbolic

Received February 27, 2018; revised May 4, 2018.
2010 *Mathematics Subject Classification.* Primary 41A17, 41A25, 41A30, 41A35.
Key words and phrases. error function based; sigmoidal and hyperbolic tangent activation functions; univariate and multivariate quasi-interpolation neural network approximation; Kantorovich-Choquet type operators.

tangent function. All convergences here are with rates expressed via the modulus of continuity of the involved function and given by very tight Jackson type inequalities.

The author comes up with the "right" precisely defined flexible quasi-interpolation, Baskakov-Choquet type integral coefficient neural networks operators associated with the error function, sigmoidal and hyperbolic tangent functions. In preparation to prove our results, we establish important properties of the basic density functions defining our operators. Feed-forward neural networks (FNNs) with one hidden layer, the only type of networks we deal with in this article, are mathematically expressed as

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

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 functions are the error, sigmoidal and hyperbolic tangent functions. About neural networks in general read [**16**], [**17**], [**18**]. We have been greatly inspired by [**15**].

2. Background

Next we present briefly about the Choquet integral.

We make the following definition.

Definition 1. Consider $\Omega \neq \emptyset$ and let $\mathcal{C}$ be a σ-algebra of subsets in Ω.

(i) (see, e.g., [**19**, p. 63]) The set function $\mu\colon \mathcal{C} \to [0, +\infty]$ is called a monotone set function (or capacity) if $\mu(\emptyset) = 0$ and $\mu(A) \leq \mu(B)$ for all $A, B \in \mathcal{C}$, with $A \subset B$. Also, μ is called submodular if

$$\mu(A \cup B) + \mu(A \cap B) \leq \mu(A) + \mu(B) \qquad \text{for all } A, B \in \mathcal{C}.$$

μ is called bounded if $\mu(\Omega) < +\infty$ and normalized if $\mu(\Omega) = 1$.

(ii) (see, e.g., [**19**, p. 233] or [**13**]) If μ is a monotone set function on $\mathcal{C}$ and if $f\colon \Omega \to \mathbb{R}$ is $\mathcal{C}$-measurable (that is, for any Borel subset $B \subset \mathbb{R}$, it follows $f^{-1}(B) \in \mathcal{C}$), then for any $A \in \mathcal{C}$, the Choquet integral is defined by

$$(C) \int_A f \mathrm{d}\mu = \int_0^{+\infty} \mu(F_\beta(f) \cap A) \mathrm{d}\beta + \int_{-\infty}^{0} [\mu(F_\beta(f) \cap A) - \mu(A)] \, \mathrm{d}\beta,$$

where we used the notation $F_\beta(f) = \{\omega \in \Omega : f(\omega) \geq \beta\}$. Notice that if $f \geq 0$ on A, then in the above formula, we get $\int_{-\infty}^{0} = 0$.

The integrals on the right-hand side are the usual Riemann integral.

The function f is called Choquet integrable on A if $(C) \int_A f \mathrm{d}\mu \in \mathbb{R}$.

Next, we list some well known properties of the Choquet integral.

Remark 1. If $\mu\colon \mathcal{C} \to [0, +\infty]$ is a monotone set function, then the following properties hold:

(i) For all $a \geq 0$, we have $(C)\int_A af\mathrm{d}\mu = a \cdot (C)\int_A f\mathrm{d}\mu$ (if $f \geq 0$, then see, e.g., [**19**, Theorem11.2, (5), p. 228], and if f is an arbitrary sign, then see, e.g., [**14**, p. 64, Proposition 5.1, (ii)]).

(ii) For all $c \in \mathbb{R}$ and f of an arbitrary sign, we have (see, e.g., [**19**, pp. 232–233], or [**14**, p. 65]) $(C)\int_A (f + c)\,\mathrm{d}\mu = (C)\int_A f\mathrm{d}\mu + c \cdot \mu\,(A)$.

If μ is submodular too, then for all f, g of an arbitrary sign and lower bounded, we have (see, e.g., [**14**, Theorem 6.3, p. 75])

$$(C)\int_A (f + g)\,\mathrm{d}\mu \leq (C)\int_A f\mathrm{d}\mu + (C)\int_A g\mathrm{d}\mu.$$

(iii) If $f \leq g$ on A then $(C)\int_A f\mathrm{d}\mu \leq (C)\int_A g\mathrm{d}\mu$ (see, e.g., [**19**, Theorem 11.2(3), p. 228] if $f, g \geq 0$ and p. 232 if f, g are of an arbitrary sign).

(iv) Let $f \geq 0$. If $A \subset B$, then $(C)\int_A f\mathrm{d}\mu \leq (C)\int_B f\mathrm{d}\mu$. In addition, if μ is finitely aubadditive, then

$$(C)\int_{A\cup B} f\mathrm{d}\mu \leq (C)\int_A f\mathrm{d}\mu + (C)\int_B f\mathrm{d}\mu.$$

(v) (v) It is immediate that $(C)\int_A 1 \cdot \mathrm{d}\mu(t) = \mu(A)$.

(vi) The formula $\mu(A) = \gamma(M(A))$, where $\gamma\colon [0, 1] \to [0, 1]$ is an increasing and concave function with $\gamma(0) = 0$, $\gamma(1) = 1$ and M is a probability measure (or only finitely additive) on a σ-algebra on Ω (that is, $M(\emptyset) = 0$, $M(\Omega) = 1$ and M is countably additive), gives simple examples of normalized, monotone, and submodular set functions (see, e.g., [**14**, Example 2.1, pp. 16–17]). Such of set functions μ are also called distortsions of countably normalized, additive measures (or distorted measures). For a simple example, we can take $\gamma(t) = \frac{2t}{1+t}$, $\gamma(t) = \sqrt{t}$.

If the above γ function is increasing, concave and satisfies only $\gamma(0) = 0$, then for any bounded Borel measure m, $\mu(A) = \gamma(m(A))$ gives a simple example of bounded, monotone and submodular set function.

(vii) If μ is a countably additive bounded measure, then the Choquet integral $(C)\int_A f\mathrm{d}\mu$ reduces to the usual Lebesgue type integral (see, e.g., [**14**, p. 62] or [**19**, p. 226)].

(viii) If $f \geq 0$, then $(C)\int_A f\mathrm{d}\mu \geq 0$.

(ix) Let $\mu = \sqrt{M}$, where M is the Lebesgue measure on $[0, +\infty)$, then μ is a monotone and submodular set function, furthermore, μ is strictly positive, see [**15**].

(x) If $\Omega = \mathbb{R}^N$, $N \in \mathbb{N}$, we call μ strictly positive if $\mu(A) > 0$, for any open subset $A \subseteq \mathbb{R}^N$.

3. Results-I

Now we consider the (Gauss) error special function ([**1**], [**11**])

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{e}^{-t^2}\,\mathrm{d}t, \qquad x \in \mathbb{R}, \tag{1}$$

which is a sigmoidal type function and a strictly increasing function.

It has the basic properties

$$\operatorname{erf}(0)=0,\quad \operatorname{erf}(-x)=-\operatorname{erf}(x),\quad \operatorname{erf}(+\infty)=1,\quad \operatorname{erf}(-\infty)=-1, \tag{2}$$

and

$$(\operatorname{erf}(x))'=\frac{2}{\sqrt{\pi}}\,\mathrm{e}^{-x^2},\qquad x\in\mathbb{R}, \tag{3}$$

$$\int \operatorname{erf}(x)dx = x\operatorname{erf}(x)+\frac{\mathrm{e}^{-x^2}}{\sqrt{\pi}}+C, \tag{4}$$

where C is a constant.

The error function is related to the cumulative probability distribution function of the standard normal distribution

$$\Phi(x)=\frac{1}{2}+\frac{1}{2}\operatorname{erf}\Big(\frac{x}{\sqrt{2}}\Big).$$

We consider the activation function

$$\chi(x)=\frac{1}{4}\left(\operatorname{erf}(x+1)-\operatorname{erf}(x-1)\right),\qquad x\in\mathbb{R}, \tag{5}$$

and we notice that

$$\begin{aligned}\chi(-x)&=\frac{1}{4}\left(\operatorname{erf}(-x+1)-\operatorname{erf}(-x-1)\right)\\ &=\frac{1}{4}\left(\operatorname{erf}\left(-(x-1)\right)-\operatorname{erf}\left(-(x+1)\right)\right)\\ &=\frac{1}{4}\left(-\operatorname{erf}(x-1)+\operatorname{erf}(x+1)\right)=\chi(x),\end{aligned} \tag{6}$$

thus χ is an even function.

Since $x+1>x-1$, then $\operatorname{erf}(x+1)>\operatorname{erf}(x-1)$, and $\chi(x)>0$, all $x\in\mathbb{R}$.

We see that

$$\chi(0)=\frac{\operatorname{erf}(1)}{2}\simeq\frac{0.843}{2}=0.4215. \tag{7}$$

Let $x>0$, we have

$$\begin{aligned}\chi'(x)&=\frac{1}{4}\Big(\frac{2}{\sqrt{\pi}}\,\mathrm{e}^{-(x+1)^2}-\frac{2}{\sqrt{\pi}}\,\mathrm{e}^{-(x-1)^2}\Big)\\ &=\frac{1}{2\sqrt{\pi}}\Big(\frac{1}{\mathrm{e}^{(x+1)^2}}-\frac{1}{\mathrm{e}^{(x-1)^2}}\Big)=\frac{1}{2\sqrt{\pi}}\Big(\frac{\mathrm{e}^{(x-1)^2}-\mathrm{e}^{(x+1)^2}}{\mathrm{e}^{(x+1)^2}\,\mathrm{e}^{(x-1)^2}}\Big)<0,\end{aligned} \tag{8}$$

proving $\chi'(x)<0$ for $x>0$.

That is, χ is strictly decreasing on $[0,\infty)$, and is strictly increasing on $(-\infty,0]$, and $\chi'(0)=0$.

Clearly the x-axis is the horizontal asymptote on χ.

Conclusion, χ is a bell symmetric function with maximum $\chi(0)\simeq 0.4215$.

We further need the following theorems.

Theorem 2 (**[10]**). *We have that*

$$\sum_{i=-\infty}^{\infty} \chi(x-i) = 1 \qquad \text{for all } x \in \mathbb{R}. \tag{9}$$

Thus

$$\sum_{i=-\infty}^{\infty} \chi(nx-i) = 1 \qquad \text{for all } n \in \mathbb{N}, \quad x \in \mathbb{R}. \tag{10}$$

Furthermore, we get:

Since χ is even, it holds $\sum\limits_{i=-\infty}^{\infty} \chi(i-x) = 1$ for any $x \in \mathbb{R}$.

Hence $\sum_{i=-\infty}^{\infty} \chi(i+x) = 1$ and $\sum_{i=-\infty}^{\infty} \chi(x+i) = 1$ for all $x \in \mathbb{R}$.

Theorem 3 (**[10]**). *It holds*

$$\int_{-\infty}^{\infty} \chi(x)\mathrm{d}x = 1. \tag{11}$$

So $\chi(x)$ is a density function on $\mathbb{R}$.

Theorem 4 (**[10]**). *Let $0 < \alpha < 1$, and $n \in \mathbb{N}$ with $n^{1-\alpha} \geq 3$. It holds*

$$\sum_{\substack{k=-\infty,\\ |nx-k| \geq n^{1-\alpha}}}^{\infty} \chi(nx-k) < \frac{1}{2\sqrt{\pi}(n^{1-\alpha}-2)\,\mathrm{e}^{(n^{1-\alpha}-2)^2}}. \tag{12}$$

We give next definition.

Definition 2. Let $\mathcal{L}$ be the Lebesgue σ-algebra on $\mathbb{R}$, and the set function $\mu\colon \mathcal{L} \to [0,+\infty)$, which is assumed to be monotone, submodular, and strictly positive. Let $f \in C_B^+(\mathbb{R})$ (the set of continuous and bounded functions from $\mathbb{R} \to \mathbb{R}_+$). We define the univariate Kantorovich-Choquet type neural network operator

(13)

$$C_n^\mu(f,x) := \sum_{k=-\infty}^{\infty} \left(\frac{(C)\int_0^{\frac{1}{n}} f\left(t+\frac{k}{n}\right)\mathrm{d}\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]\right)} \right) \chi(nx-k) \quad \text{for all } x \in \mathbb{R},\ n \in \mathbb{N}.$$

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

We notice above that

$$\left\| C_n^\mu(f) \right\|_\infty \leq \|f\|_\infty, \tag{14}$$

so that $C_n^\mu(f,x)$ is well-defined.

Remark 5. Let $t \in \left[0,\frac{1}{n}\right]$ and $x \in \mathbb{R}$, then

$$f\Big(t+\frac{k}{n}\Big) = f\Big(t+\frac{k}{n}\Big) - f(x) + f(x) \leq \Big|f\Big(t+\frac{k}{n}\Big) - f(x)\Big| + f(x),$$

hence

$$(C)\int_0^{\frac{1}{n}} f\Big(t+\frac{k}{n}\Big)\mathrm{d}\mu(t) \le (C)\int_0^{\frac{1}{n}} \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big|\mathrm{d}\mu(t) + (C)\int_0^{\frac{1}{n}} f(x)\mathrm{d}\mu(t)$$

$$= (C)\int_0^{\frac{1}{n}} \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big|\mathrm{d}\mu(t) + f(x)\mu\Big(\Big[0,\frac{1}{n}\Big]\Big).$$

That is,

$$(15)\quad (C)\int_0^{\frac{1}{n}} f\Big(t+\frac{k}{n}\Big)\mathrm{d}\mu(t) - f(x)\mu\Big(\Big[0,\frac{1}{n}\Big]\Big) \le (C)\int_0^{\frac{1}{n}} \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big|\mathrm{d}\mu(t).$$

Similarly, we have

$$f(x) = f(x) - f\Big(t+\frac{k}{n}\Big) + f\Big(t+\frac{k}{n}\Big) \le \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big| + f\Big(t+\frac{k}{n}\Big),$$

hence

$$(C)\int_0^{\frac{1}{n}} f(x)\mu\Big(\mathrm{d}t\Big) \le (C)\int_0^{\frac{1}{n}} \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big|\mathrm{d}\mu(t) + (C)\int_0^{\frac{1}{n}} f\Big(t+\frac{k}{n}\Big)\mu\Big(\mathrm{d}t\Big).$$

That is,

$$f(x)\mu\Big(\Big[0,\frac{1}{n}\Big]\Big) \le (C)\int_0^{\frac{1}{n}} \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big|\mathrm{d}\mu(t) + (C)\int_0^{\frac{1}{n}} f\Big(t+\frac{k}{n}\Big)\mu(\mathrm{d}t)$$

and

$$(16)\quad f(x)\mu\Big(\Big[0,\frac{1}{n}\Big]\Big) - (C)\int_0^{\frac{1}{n}} f\Big(t+\frac{k}{n}\Big)\mu(\mathrm{d}t) \le (C)\int_0^{\frac{1}{n}} \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big|\mathrm{d}\mu(t).$$

By (15) and (16), we derive that

$$(17)\quad \Big|(C)\int_0^{\frac{1}{n}} f\Big(t+\frac{k}{n}\Big)\mu(\mathrm{d}t) - f(x)\mu\Big(\Big[0,\frac{1}{n}\Big]\Big)\Big| \le (C)\int_0^{\frac{1}{n}} \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big|\mathrm{d}\mu(t).$$

In particular, it holds

$$(18)\qquad \left|\left(\frac{(C)\int_0^{\frac{1}{n}} f\big(t+\frac{k}{n}\big)\mu(\mathrm{d}t)}{\mu\big(\big[0,\frac{1}{n}\big]\big)}\right) - f(x)\right| \le \frac{(C)\int_0^{\frac{1}{n}} \Big|f\Big(t+\frac{k}{n}\Big)-f(x)\Big|\mathrm{d}\mu(t)}{\mu\big(\big[0,\frac{1}{n}\big]\big)}.$$

We define

$$\omega_1(f,h) := \sup_{\substack{x,y\in\mathbb{R}\\ |x-y|\le h}} |f(x)-f(y)|, \qquad h>0,$$

the first modulus of continuity of $f \in C_B^+(\mathbb{R})$.

We present the following theorem.

Theorem 6. *Let $f \in C_B^+(\mathbb{R})$, $0<\alpha<1$, $x\in\mathbb{R}$, $n\in\mathbb{N}$ with $n^{1-\alpha}\ge 3$. Then*

i)

$$(19)\quad \sup_{\mu} |C_n^{\mu}(f,x) - f(x)| \le \omega_1\Big(f, \frac{1}{n}+\frac{1}{n^{\alpha}}\Big) + \frac{\|f\|_{\infty}}{\sqrt{\pi}(n^{1-\alpha}-2)\,\mathrm{e}^{(n^{1-\alpha}-2)^2}} =: \rho_{1n},$$

$$\text{(ii) (20)}\qquad \sup_{\mu} \|C_n^{\mu}(f) - f\|_{\infty} \le \rho_{1n}.$$

For $f \in \left(C_B^+(\mathbb{R}) \cap C_u^+(\mathbb{R})\right)$, *we get* $\lim_{n\to\infty} C_n^k(f) = f$ *pointwise and uniformly.*

Here $C_u^+(\mathbb{R})$ *is the set of functions that are uniformly continuous from* $\mathbb{R}$ *into* $\mathbb{R}_+$.

Proof. We observe that

$$
\begin{aligned}
|C_n^\mu(f,x) - f(x)| &= \Big| \sum_{k=-\infty}^{\infty} \Big(\Big(\frac{(C)\int_0^{\frac{1}{n}} f\left(t+\frac{k}{n}\right) \mathrm{d}\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]\right)} \Big) - f(x) \Big) \chi(nx-k) \Big| \\
&\le \sum_{k=-\infty}^{\infty} \Big| \Big(\frac{(C)\int_0^{\frac{1}{n}} f\left(t+\frac{k}{n}\right) \mathrm{d}\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]\right)} \Big) - f(x) \Big| \chi(nx-k) \\
&\overset{(18)}{\le} \sum_{k=-\infty}^{\infty} \Big(\frac{(C)\int_0^{\frac{1}{n}} \left|f\left(t+\frac{k}{n}\right) - f(x)\right| \mathrm{d}\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]\right)} \Big) \chi(nx-k) \qquad (21) \\
&= \sum_{\substack{k=-\infty,\\ x-\frac{k}{n} \le \frac{1}{n^\alpha}}}^{\infty} \Big(\frac{(C)\int_0^{\frac{1}{n}} \left|f\left(t+\frac{k}{n}\right) - f(x)\right| \mathrm{d}\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]\right)} \Big) \chi(nx-k) \\
&\quad + \sum_{\substack{k=-\infty,\\ x-\frac{k}{n} \ge \frac{1}{n^\alpha}}}^{\infty} \Big(\frac{(C)\int_0^{\frac{1}{n}} \left|f\left(t+\frac{k}{n}\right) - f(x)\right| \mathrm{d}\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]\right)} \Big) \chi(nx-k) \qquad (22) \\
&\le \sum_{\substack{k=-\infty,\\ x-\frac{k}{n} \le \frac{1}{n^\alpha}}}^{\infty} \Big(\frac{(C)\int_0^{\frac{1}{n}} \omega_1\left(f, \left|t+\frac{k}{n}-x\right|\right) \mathrm{d}\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]\right)} \Big) \chi(nx-k) \\
&\quad + 2\|f\|_\infty \Bigg(\sum_{\substack{k=-\infty,\\ |nx-k| \ge n^{1-\alpha}}}^{\infty} \chi(|nx-k|) \Bigg) \\
&\overset{(12)}{\le} \sum_{\substack{k=-\infty,\\ |nx-k| \le n^{1-\alpha}}}^{\infty} \Big(\frac{(C)\int_0^{\frac{1}{n}} \omega_1\left(f, |t| + \frac{1}{n^\alpha}\right) \mu(\mathrm{d}t)}{\mu\left(\left[0,\frac{1}{n}\right]\right)} \Big) \chi(nx-k) \\
&\quad + \frac{\|f\|_\infty}{\sqrt{\pi}(n^{1-\alpha}-2)\,\mathrm{e}^{(n^{1-\alpha}-2)^2}} \\
&\le \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\alpha}\right) \Bigg(\sum_{\substack{k=-\infty,\\ |nx-k| \le n^{1-\alpha}}}^{\infty} \chi(nx-k) \Bigg) \qquad (23) \\
&\quad + \frac{\|f\|_\infty}{\sqrt{\pi}(n^{1-\alpha}-2)\,\mathrm{e}^{(n^{1-\alpha}-2)^2}} \\
&\overset{(10)}{\le} \omega_1\Big(f, \frac{1}{n} + \frac{1}{n^\alpha}\Big) + \frac{\|f\|_\infty}{\sqrt{\pi}(n^{1-\alpha}-2)\,\mathrm{e}^{(n^{1-\alpha}-2)^2}},
\end{aligned}
$$

proving the claim. □

Additionally, we give the following definition.

Definition 3. Here $\mathrm{i} = \sqrt{-1}$. Denote $C_B^+(\mathbb{R},\mathbb{C}) = \{f\colon \mathbb{R}\to\mathbb{C} \mid f = f_1 + \mathrm{i}\, f_2,$ where $f_1, f_2 \in C_B^+(\mathbb{R})\}$. For $f \in C_B^+(\mathbb{R},\mathbb{C})$, we set

$$C_n^{\mu}(f,x) := C_n^{\mu}(f_1,x) + \mathrm{i}\, C_n^{\mu}(f_2,x), \qquad \text{for all } n\in\mathbb{N}\ x\in\mathbb{R}. \tag{24}$$

Theorem 7. *Let $f \in C_B^+(\mathbb{R},\mathbb{C})$, $f = f_1 + \mathrm{i}\, f_2$, $0<\alpha<1$, $n\in\mathbb{N}$, $n^{1-\alpha}\ge 3$, $x\in\mathbb{R}$. Then*

(i) (25)
$$\begin{aligned}\sup_{\mu}\left|C_n^{\mu}(f,x) - f(x)\right| \le \Big(\omega_1\Big(f_1, \frac{1}{n} + \frac{1}{n^{\alpha}}\Big) + \omega_1\Big(f_2, \frac{1}{n} + \frac{1}{n^{\alpha}}\Big)\Big)\\ + \frac{(\|f_1\|_{\infty} + \|f_2\|_{\infty})}{\sqrt{\pi}(n^{1-\alpha}-2)\,\mathrm{e}^{(n^{1-\alpha}-2)^2}} =: \rho_{2n}\end{aligned}$$

and

(ii) (26)
$$\sup_{\mu}\left\|C_n^{\mu}(f) - f\right\|_{\infty} \le \rho_{2n}.$$

Proof. We have that

$$\begin{aligned}\left|C_n^{\mu}(f,x) - f(x)\right| &= \left|C_n^{\mu}(f_1,x) + \mathrm{i}\, C_n^{\mu}(f_2,x) - f_1(x) - i f_2(x)\right|\\ &= \left|(C_n^{\mu}(f_1,x) - f_1(x)) + \mathrm{i}\,(C_n^{\mu}(f_2,x) - f_2(x))\right|\\ &\le \left|C_n^{\mu}(f_1,x) - f_1(x)\right| + \left|C_n^{\mu}(f_2,x) - f_2(x)\right|\\ &\overset{(19)}{\le} \Big(\omega_1\Big(f_1, \frac{1}{n} + \frac{1}{n^{\alpha}}\Big) + \frac{\|f_1\|_{\infty}}{\sqrt{\pi}(n^{1-\alpha}-2)\,\mathrm{e}^{(n^{1-\alpha}-2)^2}}\Big)\\ &\quad + \Big(\omega_1\Big(f_2, \frac{1}{n} + \frac{1}{n^{\alpha}}\Big) + \frac{\|f_2\|_{\infty}}{\sqrt{\pi}(n^{1-\alpha}-2)\,\mathrm{e}^{(n^{1-\alpha}-2)^2}}\Big),\end{aligned} \tag{27}$$

proving the claim. □

Remark 8. We introduce

$$Z(x_1,\ldots,x_N) := Z(x) := \prod_{i=1}^{N} \chi(x_i), \qquad x = (x_1,\ldots,x_N)\in\mathbb{R}^N,\ N\in\mathbb{N}. \tag{28}$$

It has the following properties:

(i) (29)
$$Z(x) > 0 \qquad \text{for all } x\in\mathbb{R}^N,$$

(ii) (30)
$$\sum_{k=-\infty}^{\infty} Z(x-k) := \sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\cdots\sum_{k_N=-\infty}^{\infty} Z(x_1-k_1,\ldots,x_N-k_N) = 1,$$

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

(iii) (31)

$$\sum_{k=-\infty}^{\infty} Z(nx-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} Z(nx_1-k_1,\dots,nx_N-k_N) = 1$$

for all $x \in \mathbb{R}^N$, $n \in \mathbb{N}$, and

(iv) (32)
$$\int_{\mathbb{R}^N} Z(x)\mathrm{d}x = 1,$$

that is, Z is a multivariate density function.

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

It is also clear that (see (12))

(v) (33)
$$\sum_{\substack{k=-\infty,\\ \|\frac{k}{n}-x\|_\infty > \frac{1}{n^\beta}}}^{\infty} Z(nx-k) \le \frac{1}{2\sqrt{\pi}(n^{1-\beta}-2)\,\mathrm{e}^{(n^{1-\beta}-2)^2}},$$

$0<\beta<1$, $n \in \mathbb{N} : n^{1-\beta} \ge 3$, $x \in \mathbb{R}^N$.

For $f \in C_B^+\left(\mathbb{R}^N\right)$ (continuous and bounded functions from $\mathbb{R}^N$ into $\mathbb{R}_+$), we define the first modulus of continuity

(34)
$$\omega_1(f,h) := \sup_{\substack{x,y\in\mathbb{R}^N\\ \|x-y\|_\infty \le h}} |f(x)-f(y)|, \qquad h>0.$$

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

(35)
$$\lim_{h\to 0}\omega_1(f,h) = 0.$$

We make next definition.

Definition 4. Let $\mathcal{L}^*$ be the Lebesgue σ-algebra on $\mathbb{R}^N$, $N \in \mathbb{N}$, and the set function $\mu^*\colon \mathcal{L}^* \to [0,+\infty)$, which is assumed to be monotone, submodular and strictly positive. For $f \in C_B^+(\mathbb{R}^N)$, we define the multivariate Kantorovich-Choquet type neural network operator for any $x \in \mathbb{R}^N$

(36)
$$\overline{C}_n^{\mu^*}(f,x) = \overline{C}_n^{\mu^*}(f,x_1,\dots,x_N) := \sum_{k=-\infty}^{\infty} \left(\frac{(C)\int_{[0,\frac{1}{n}]^N} f\left(t+\frac{k}{n}\right)\mathrm{d}\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]^N\right)} \right) Z(nx-k)$$

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

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

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

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

We notice above that

$$\left\|\overline{C}_n^{\mu^*}(f)\right\|_\infty \leq \|f\|_\infty, \tag{37}$$

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

Remark 9. Acting as in the proof of (18), we derive

$$\begin{aligned} &\left|\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right) \mathrm{d}\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]^N\right)}\right) - f(x)\right| \\ &\leq \frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} \left|f\left(t+\frac{k}{n}\right) - f(x)\right| \mathrm{d}\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]^N\right)}. \end{aligned} \tag{38}$$

We present the following theorem.

Theorem 10. *Let $f \in C_B^+\left(\mathbb{R}^N\right)$, $0<\beta<1$, $x \in \mathbb{R}^N$, $N, n \in \mathbb{N}$ with $n^{1-\beta} \geq 3$. Then*

(ii) (39)

$$\sup_{\mu^*}\left|\overline{C}_n^{\mu^*}(f,x) - f(x)\right| \leq \omega_1\Big(f, \frac{1}{n} + \frac{1}{n^\beta}\Big) + \frac{\|f\|_\infty}{\sqrt{\pi}(n^{1-\beta}-2)\,\mathrm{e}^{(n^{1-\beta}-2)^2}} =: \rho_{3n},$$

(ii) (40)

$$\sup_{\mu^*}\left\|\overline{C}_n^{\mu^*}(f) - f\right\|_\infty \leq \rho_{3n}.$$

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

Proof. We observe that

$$\begin{aligned} &\left|\overline{C}_n^{\mu^*}(f,x) - f(x)\right| \\ &= \left|\sum_{k=-\infty}^{\infty}\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right) \mathrm{d}\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]^N\right)}\right) Z(nx-k) - \sum_{k=-\infty}^{\infty} f(x) Z(nx-k)\right| \\ &= \left|\sum_{k=-\infty}^{\infty}\left(\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right) \mathrm{d}\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]^N\right)}\right) - f(x)\right) Z(nx-k)\right| \\ &\leq \sum_{k=-\infty}^{\infty}\left|\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right) \mathrm{d}\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]^N\right)}\right) - f(x)\right| Z(nx-k) \\ &\overset{(38)}{\leq} \sum_{k=-\infty}^{\infty}\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} \left|f\left(t+\frac{k}{n}\right) - f(x)\right| \mathrm{d}\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]^N\right)}\right) Z(nx-k) \end{aligned} \tag{41}$$

$$(42) \quad = \sum_{\substack{k=-\infty,\\ \|\frac{k}{n}-x\|_\infty \le \frac{1}{n^\beta}}}^{\infty} \left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} \left|f\left(t+\frac{k}{n}\right)-f(x)\right| \mathrm{d}\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]^N\right)} \right) Z(nx-k)$$

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

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

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

$$(43) \quad \overset{\text{(v) (33)}}{\le} \omega_1\Big(f, \frac{1}{n}+\frac{1}{n^\beta}\Big) + \frac{\|f\|_\infty}{\sqrt{\pi}\,(n^{1-\beta}-2)\,\mathrm{e}^{(n^{1-\beta}-2)^2}},$$

proving the claim. ☐

Additionally, we give further definition.

Definition 5. Denote by $C_B^+\left(\mathbb{R}^N,\mathbb{C}\right) = \{f\colon \mathbb{R}^N \to \mathbb{C} | f = f_1 + \mathrm{i}\, f_2$, where $f_1, f_2 \in C_B^+\left(\mathbb{R}^N\right)$, $N \in \mathbb{N}\}$. We set for $f \in C_B^+\left(\mathbb{R}^N,\mathbb{C}\right)$ that

$$(44) \qquad \overline{C}_n^{\mu^*}(f,x) := \overline{C}_n^{\mu^*}(f_1,x) + \mathrm{i}\,\overline{C}_n^{\mu^*}(f_2,x) \qquad \text{for all } n\in\mathbb{N},\ x\in\mathbb{R}^N.$$

Theorem 11. *Let* $f \in C_B^+\left(\mathbb{R}^N,\mathbb{C}\right)$, $f = f_1 + \mathrm{i}\, f_2$, $N \in \mathbb{N}$, $0<\beta<1$, $x \in \mathbb{R}^N$, $n \in \mathbb{N}$ *with* $n^{1-\beta} \ge 3$. *Then*

$$\text{(i) (45)} \quad \sup_{\mu^*} \left|\overline{C}_n^{\mu^*}(f,x) - f(x)\right| \le \left(\omega_1\Big(f_1, \frac{1}{n}+\frac{1}{n^\beta}\Big) + \omega_1\Big(f_2, \frac{1}{n}+\frac{1}{n^\beta}\Big)\right) + \frac{\left(\|f_1\|_\infty + \|f_2\|_\infty\right)}{\sqrt{\pi}\,(n^{1-\beta}-2)\,\mathrm{e}^{(n^{1-\beta}-2)^2}} =: \rho_{4n}$$

and

$$\text{(ii) (46)} \qquad \sup_{\mu^*} \left\|\overline{C}_n^{\mu^*}(f) - f\right\|_\infty \le \rho_{4n}.$$

Proof. Similar to Theorem 7, by applying (ii) (39) twice. ☐

4. Results - II

Remark 12. We consider here the sigmoidal function of logarithmic type

$$s_i(x_i) = \frac{1}{1+e^{-x_i}}, \qquad x_i \in \mathbb{R},\ i = 1, \ldots, N,\ x := (x_1, \ldots, x_N) \in \mathbb{R}^N,$$

each has the properties $\lim_{x_i \to +\infty} s_i(x_i) = 1$ and $\lim_{x_i \to -\infty} s_i(x_i) = 0$, $i = 1, \ldots, N$.

These functions play the role of activation functions in the hidden layer of neural networks, also have applications in biology, demography, etc.

As in [**12**], we consider

$$\Phi_i(x_i) := \frac{1}{2}\left(s_i\left(x_i+1\right) - s_i\left(x_i-1\right)\right), \ \ x_i \in \mathbb{R},\ i = 1, \ldots, N. \tag{47}$$

We notice the following properties:

i) $\Phi_i(x_i) > 0$ for all $x_i \in \mathbb{R}$,

ii) $\sum_{k_i=-\infty}^{\infty} \Phi_i\left(x_i - k_i\right) = 1$ for all $x_i \in \mathbb{R}$,

iii) $\sum_{k_i=-\infty}^{\infty} \Phi_i\left(nx_i - k_i\right) = 1$ for all $x_i \in \mathbb{R}$, $n \in \mathbb{N}$,

iv) $\int_{-\infty}^{\infty} \Phi_i(x_i) \mathrm{d}x_i = 1$,

v) Φ_i is a density function,

vi) Φ_i is even: $\Phi_i\left(-x_i\right) = \Phi_i(x_i)$, $x_i \geq 0$, for $i = 1, \ldots, N$.

We see that ([**8**])

$$\Phi_i(x_i) = \left(\frac{e^2-1}{2e^2}\right) \frac{1}{\left(1+e^{x_i-1}\right)\left(1+e^{-x_i-1}\right)}, \qquad i = 1, \ldots, N.$$

vii) Φ_i is decreasing on $\mathbb{R}_+$, and increasing on $\mathbb{R}_-$, $i = 1, \ldots, N$.

Let $0 < \beta < 1$, $n \in \mathbb{N}$. Then as in [**8**], we get

viii)

$$\sum_{\substack{k_i=-\infty, \\ |nx_i-k_i|>n^{1-\alpha}}}^{\infty} \Phi_i\left(nx_i - k_i\right) = \sum_{\substack{k_i=-\infty, \\ |nx_i-k_i|>n^{1-\alpha}}}^{\infty} \Phi_i\left(\left|nx_i - k_i\right|\right) \leq 3.1992\, e^{-n^{(1-\beta)}}, \quad i = 1, \ldots, N.$$

Now, we use the complete multivariate activation function ([**7**])

$$\Phi(x_1, \ldots, x_N) := \Phi(x) := \prod_{i=1}^{N} \Phi_i(x_i), \qquad x \in \mathbb{R}^N. \tag{48}$$

It has the property ([**7**]): $\Phi(x) > 0$ for all $x \in \mathbb{R}^N$. We see that

(i)′ (49)

$$\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \Phi\left(x_1 - k_1, x_2 - k_2, \ldots, x_N - k_N\right)$$
$$= \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \prod_{i=1}^{N} \Phi_i\left(x_i - k_i\right) = \prod_{i=1}^{N} \left(\sum_{k_i=-\infty}^{\infty} \Phi_i\Big(x_i - k_i\Big)\right) = 1.$$

That is,

(ii)′ (50)

$$\sum_{k=-\infty}^{\infty} \Phi(x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \Phi\left(x_1 - k_1, x_2 - k_2, \ldots, x_N - k_N\right) = 1,$$

$k := (k_1, \ldots, k_N)$ for all $x \in \mathbb{R}^N$.

(iii)′ (51)

$$\sum_{k=-\infty}^{\infty} \Phi(nx-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \Phi\left(nx_1 - k_1, nx_2 - k_2, \ldots, nx_N - k_N\right) = 1$$

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

(iv)′ (52)
$$\int_{\mathbb{R}^N} \Phi(x) \mathrm{d}x = 1,$$

that is, Φ is a multivariate density function.

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

For $0 < \beta < 1$ and $n \in \mathbb{N}$, fixed $x \in \mathbb{R}^N$, we have proved ([**7**])

(v)′ (53)
$$\sum_{\substack{k=-\infty, \\ \|\frac{k}{n} - x\|_\infty > \frac{1}{n^\beta}}}^{\infty} \Phi(nx-k) \le 3.1992\, \mathrm{e}^{-n^{(1-\beta)}}.$$

We give the following definition.

Definition 6. Let $\mathcal{L}^*$ be the Lebesgue σ-algebra on $\mathbb{R}^N$, $N \in \mathbb{N}$, and the set function $\mu^*\colon \mathcal{L}^* \to [0, +\infty)$, which is assumed to be monotone, submodular and strictly positive. For $f \in C_B^+(\mathbb{R}^N)$, we define the multivariate Kantorovich-Choquet type neural network operator for any $x \in \mathbb{R}^N$

(54)

$$K_n^{\mu^*}(f,x) = K_n^{\mu^*}(f,x_1,\ldots,x_N)$$

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

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

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

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

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

We notice above that

$$\left\| K_n^{\mu^*}(f) \right\|_\infty \le \|f\|_\infty, \tag{55}$$

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

Remark 13. We also consider the hyperbolic tangent function $\tanh x$, $x \in \mathbb{R}$ (see also [**5**]),

$$\tanh x := \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{56}$$

It has the properties $\tanh 0 = 0$, $-1 < \tanh x < 1$ for all $x \in \mathbb{R}$, and $\tanh(-x) = -\tanh x$. Furthermore, $\tanh x \to 1$ as $x \to \infty$, and $\tanh x \to -1$ as $x \to -\infty$, and it is strictly increasing on $\mathbb{R}$.

This function plays the role of an activation function in the hidden layer of neural networks.

We further consider ([**5**])

$$\Psi(x) := \frac{1}{4}\left(\tanh(x+1) - \tanh(x-1)\right) > 0 \qquad \text{for all } x \in \mathbb{R}. \tag{57}$$

We easily see that $\Psi(-x) = \Psi(x)$, that is, Ψ is even on $\mathbb{R}$. Obviously Ψ is differentiable, thus continuous.

It follows

Proposition 7 ([**5**]). *$\Psi(x)$ for $x \ge 0$, is strictly decreasing.*

Obviously $\Psi(x)$ is strictly increasing for $x \le 0$. Also it holds $\lim_{x\to-\infty} \Psi(x) = 0 = \lim_{x\to\infty} \Psi(x)$.

In fact Ψ has the bell shape with horizontal asymptote the x-axis. So the maximum of Ψ is zero, $\Psi(0) = 0.3809297$.

Theorem 14 ([**5**]). *We have that* $\sum_{i=-\infty}^{\infty} \Psi(x-i) = 1$ *for all* $x \in \mathbb{R}$.

Thus

$$\sum_{i=-\infty}^{\infty} \Psi(nx-i) = 1 \qquad \text{for all } n \in \mathbb{N}, \quad x \in \mathbb{R}.$$

Also it holds

$$\sum_{i=-\infty}^{\infty} \Psi(x+i) = 1 \qquad \text{for all } x \in \mathbb{R}.$$

Theorem 15 ([**5**]). *It holds* $\int_{-\infty}^{\infty} \Psi(x)\mathrm{d}x = 1$.

So $\Psi(x)$ is a density function on $\mathbb{R}$.

Theorem 16 ([**5**]). *Let* $0 < \alpha < 1$ *and* $n \in \mathbb{N}$. *It holds*

$$\sum_{\substack{k=-\infty,\\ |nx-k| \geq n^{1-\alpha}}}^{\infty} \Psi(nx-k) \leq \mathrm{e}^4 \cdot \mathrm{e}^{-2n^{(1-\alpha)}}.$$

Remark 17. In this article, we also use the complete multivariate activation function

(58) $$\Theta(x_1, \ldots, x_N) := \Theta(x) := \prod_{i=1}^{N} \Psi(x_i), \qquad x = (x_1, \ldots, x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}.$$

It has the properties (see [**6**]):
(i) $\Theta(x) > 0$ for all $x \in \mathbb{R}^N$,
(ii) (59)

$$\sum_{k=-\infty}^{\infty} \Theta(x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \ldots \sum_{k_N=-\infty}^{\infty} \Theta(x_1-k_1, x_2-k_2, \ldots, x_N-k_N) = 1,$$

where $k := (k_1, \ldots, k_N)$ for all $x \in \mathbb{R}^N$.

(iii) (60)

$$\begin{aligned} &\sum_{k=-\infty}^{\infty} \Theta(nx-k) \\ &:= \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \Theta(nx_1-k_1, nx_2-k_2, \ldots, nx_N-k_N) = 1 \end{aligned}$$

for all $x \in \mathbb{R}^N$; $n \in \mathbb{N}$.

(iv) (61)

$$\int_{\mathbb{R}^N} \Theta(x)\mathrm{d}x = 1,$$

that is, Θ is a multivariate density function.

By [**6**], we get

(v) (62)

$$\sum_{\substack{k=-\infty,\\ \|\frac{k}{n}-x\|_\infty > \frac{1}{n^\beta}}}^{\infty} \Theta(nx-k) \leq \mathrm{e}^4 \cdot \mathrm{e}^{-2n^{(1-\beta)}},$$

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

Definition 8. Let $\mathcal{L}^*$ be the Lebesgue σ-algebra on $\mathbb{R}^N$, $N \in \mathbb{N}$, and the set function $\mu^* \colon \mathcal{L}^* \to [0, +\infty)$, which is assumed to be monotone, submodular and strictly positive. For $f \in C_B^+(\mathbb{R}^N)$, we define the multivariate Kantorovich-Choquet type neural network operator for any $x \in \mathbb{R}^N$

(63)

$$\begin{aligned}
&L_n^{\mu^*}(f, x) = L_n^{\mu^*}(f, x_1, \ldots, x_N) \\
&:= \sum_{k=-\infty}^{\infty} \left(\frac{(C) \int_{[0, \frac{1}{n}]^N} f(t + \frac{k}{n}) \mathrm{d}\mu^*(t)}{\mu^*\left(\left[0, \frac{1}{n}\right]^N\right)} \right) \Theta(nx - k) \\
&= \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} \left(\frac{(C) \int_0^{\frac{1}{n}} \cdots \int_0^{\frac{1}{n}} f\left(t_1 + \frac{k_1}{n}, t_2 + \frac{k_2}{n}, \ldots, t_N + \frac{k_N}{n}\right) \mathrm{d}\mu^*(t_1, \ldots, t_N)}{\mu^*\left(\left[0, \frac{1}{n}\right]^N\right)} \right) \\
&\cdot \left(\prod_{i=1}^{N} \Psi(nx_i - k_i) \right),
\end{aligned}$$

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

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

We notice above that

$$\left\| L_n^{\mu^*}(f) \right\|_\infty \leq \|f\|_\infty, \tag{64}$$

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

We present the following theorems.

Theorem 18. *Let* $f \in C_B^+(\mathbb{R}^N)$, $0 < \beta < 1$, $x \in \mathbb{R}^N$; $n, N \in \mathbb{N}$. *Then*

(i) (65) $$\sup_{\mu^*} \left| K_n^{\mu^*}(f, x) - f(x) \right| \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + (6.3984) \|f\|_\infty \mathrm{e}^{-n^{(1-\beta)}} =: \lambda_{1n},$$

(ii) (66) $$\sup_{\mu^*} \left\| K_n^{\mu^*}(f) - f \right\|_\infty \leq \lambda_{1n}.$$

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

Proof. As similar to Theorem 10, is omitted. □

Theorem 19. *Let* $f \in C_B^+(\mathbb{R}^N)$, $0 < \beta < 1$, $x \in \mathbb{R}^N$; $n, N \in \mathbb{N}$. *Then*

(i) (67) $$\sup_{\mu^*} \left| L_n^{\mu^*}(f, x) - f(x) \right| \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^\beta}\right) + \|f\|_\infty 2 \mathrm{e}^4 \mathrm{e}^{-2n^{(1-\beta)}} =: \lambda_{2n},$$

(ii) (68) $$\sup_{\mu^*} \left\| L_n^{\mu^*}(f) - f \right\|_\infty \leq \lambda_{2n}.$$

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

Proof. As similar to Theorem 10, is omitted. □

We need next definition.

Definition 9. For $f \in C_B^+(\mathbb{R}^N, \mathbb{C})$ we set

$$K_n^{\mu^*}(f,x) := K_n^{\mu^*}(f_1,x) + \mathrm{i}\, K_n^{\mu^*}(f_2,x),$$
$$L_n^{\mu^*}(f,x) := L_n^{\mu^*}(f_1,x) + \mathrm{i}\, L_n^{\mu^*}(f_2,x)$$

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

We finish with the theorem.

Theorem 20. *Let* $f \in C_B^+\left(\mathbb{R}^N, \mathbb{C}\right)$*,* $f = f_1 + \mathrm{i}\, f_2$*,* $0 < \beta < 1$*,* $x \in \mathbb{R}^N$*,* $n, N \in \mathbb{N}$*. Then*

(i) (70)
$$\sup_{\mu^*}\left|K_n^{\mu^*}(f,x) - f(x)\right| \le \left(\omega_1\left(f_1, \frac{1}{n} + \frac{1}{n^\beta}\right) + \omega_1\left(f_2, \frac{1}{n} + \frac{1}{n^\beta}\right)\right) + \left(\|f_1\|_\infty + \|f_2\|_\infty\right)(6.3984)\, \mathrm{e}^{-n^{(1-\beta)}} =: \lambda_{3n},$$

(ii) (71)
$$\sup_{\mu^*}\left\|K_n^{\mu^*}(f) - f\right\|_\infty \le \lambda_{3n},$$

(iii) (72)
$$\sup_{\mu^*}\left|L_n^{\mu^*}(f,x) - f(x)\right| \le \left(\omega_1\left(f_1, \frac{1}{n} + \frac{1}{n^\beta}\right) + \omega_1\left(f_2, \frac{1}{n}\frac{1}{n^\beta}\right)\right) + \left(\|f_1\|_\infty + \|f_2\|_\infty\right) 2\, \mathrm{e}^4\, \mathrm{e}^{-2n^{(1-\beta)}} =: \lambda_{4n},$$

(iv) (73)
$$\sup_{\mu^*}\left\|L_n^{\mu^*}(f) - f\right\|_\infty \le \lambda_{4n}.$$

Proof. Similar to Theorem 7. □

References

1. Abramowitz M., Stegun I. A., eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York, Dover Publications, 1972.

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

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

4. Anastassiou G. A., *Intelligent Systems: Approximation by Artificial Neural Networks*, Intell. Syst. Ref. Libr., Vol. 19, Springer, Heidelberg, 2011.

5. Anastassiou G. A., *Univariate hyperbolic tangent neural network approximation*, Math. Comput. Modelling **53** (2011), 1111–1132.

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

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

8. Anastassiou G. A., *Univariate sigmoidal neural network approximation*, J. Comput. Anal. Appl. **14**(4) (2012) 659–690.

9. Anastassiou G. A., *Fractional neural network approximation*, Computers and Mathematics with Applications **64** (2012), 1655–1676.

10. G.A. Anastassiou, *Univariate error function based neural network approximation*, Indian J. of Math., **57**(2) (2015), 243–291.

11. Andrews L. C., *Special Functions of Mathematics for Engineers*, Second edition, Mc Graw-Hill, New York, 1992.
12. Chen Z. and Cao F., *The approximation operators with sigmoidal functions*, Computers and Mathematics with Applications, **58** (2009), 758–765.
13. Choquet G., *Theory of capacities*, Ann. Inst. Fourier (Grenoble), **5** (1954), 131–295.
14. Denneberg D., *Non-additive Measure and Integral*, Kluwer, Dordrecht, 1994.
15. Gal S., *Uniform and Pointwise Quantitative Approximation by Kantorovich-Choquet type integral Operators with respect to monotone and submodular set functions*, Mediterr. J. Math. **14**(5) (2017), Art. 205, 12 pp.
16. Haykin I. S., *Neural Networks: A Comprehensive Foundation* (2 ed.), Prentice Hall, New York, 1998.
17. McCulloch W. and Pitts W., *A logical calculus of the ideas immanent in nervous activity*, Bulletin of Mathematical Biophysics, **7** (1943), 115–133.
18. Mitchell T. M., *Machine Learning*, WCB-McGraw-Hill, New York, 1997.
19. Wang Z. and Klir G. J., *Generalized Measure Theory*, Springer, New York, 2009.

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