Uyan et al. *Journal of Inequalities and Applications* (2024) 2024:125
https://doi.org/10.1186/s13660-024-03199-x

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH — Open Access

Interpolation for neural-network operators activated with a generalized logistic-type function

Hande Uyan[1], Abdullah Ozan Aslan[1], Seda Karateke[2*] and İbrahim Büyükyazıcı[3]

*Correspondence:
[2]Department of Software Engineering, Faculty of Engineering and Natural Sciences, Istanbul Atlas University, 34408, Kagithane, Istanbul, Türkiye
Full list of author information is available at the end of the article

Abstract

This paper defines a family of neural-network interpolation operators. The first derivative of generalized logistic-type functions is considered as a density function. Using the first-order uniform approximation theorem for continuous functions defined on the finite intervals, the interpolation properties of these operators are presented. A Kantorovich-type variant of the operators $F_n^{a,\varepsilon}$ is also introduced. The approximation of Kantorovich-type operators in L_p spaces with $1 \leq p \leq \infty$ is studied. Further, different combinations of the parameters of our generalized logistic-type activation function $\theta_{s,a}$ are examined to see which parameter values might give us a more efficient activation function. By choosing suitable parameters for the operator $F_n^{a,\varepsilon}$ and the Kantorovich variant of the operator $F_n^{a,\varepsilon}$, the approximation of various function examples is studied.

Keywords: Generalized logistic-type function; Neural-Network (NN) operators; Interpolation; Uniform approximation; Order of approximation

1 Introduction

Neural Networks (NNs) are widely used in numerous areas such as visual recognition, healthcare, astronomical physics, geology, cybersecurity, and many more. As the most widely used neural networks, Feedforward Neural Networks (FNNs) have been extensively studied thanks to their universal approximation capabilities. In theoretical terms, we can state that a continuous function within any compact set can be approximated to an arbitrary degree by FNNs, as long as the number of neurons is chosen large enough [11].

The mathematical expression of the NNs with one hidden layer is:

$$N_{\sigma,n}(x) = \sum_{i=1}^{n} c_i \sigma\left(a_i \cdot x + b_i\right), x = (x_1, x_2, \ldots, x_s) \in \mathbb{R}^s, \quad 1 \leq i \leq n, n \in \mathbb{N}, \tag{1}$$

where $a_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,s}) \in \mathbb{R}^s$ are the connection weights, $c_i \in \mathbb{R}$ are the coefficients, $a_i \cdot x$ is the inner product of a_i and x, and also σ is the activation function. See Fig. 1 for the architecture of the neural network $N_{\sigma,n}(x)$.

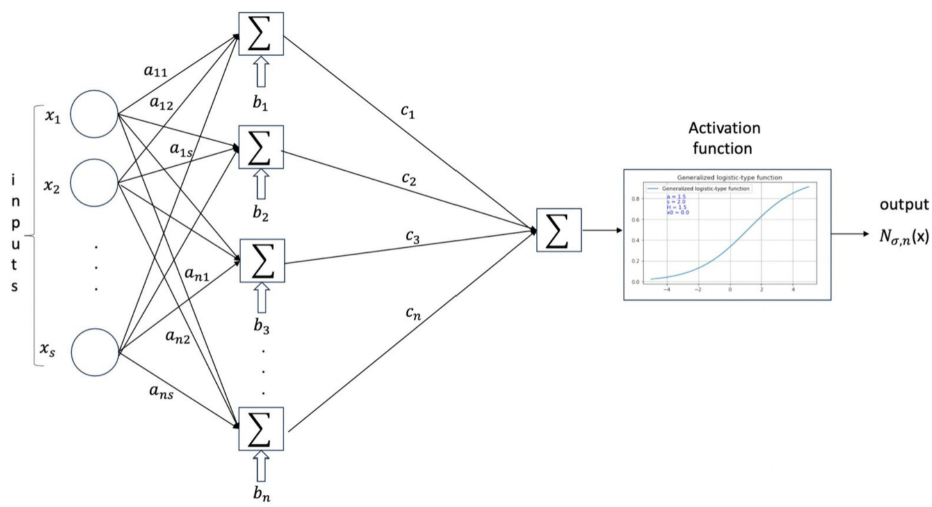

Figure 1 The structural diagram of the neural network $N_{\sigma,n}$ [23]

Thus far, some upper bounds on the approximation error of FNNs in the uniform metric and L_p metric have been studied in [9–11, 14], and so on. Moreover, the theory of neural-network (NN) operators activated by sigmodial functions has been extensively studied in recent years, see e.g. [1, 6, 8, 10, 12, 13, 15]. Concerning applications to mathematics, more precisely to approximation theory, several papers have been published, see, e.g., [7, 17–20, 22, 23]. The main results proved in this field are characterized by a nonconstructive approach, where for any given function f, the various elements that make up a neural network that approximates f in some sense, such as coefficients, weights, and thresholds cannot be determined in practice.

The NN operators studied here interpolate any given measurable and bounded function f on finite sets of uniform spaced notes taken on $[\alpha,\beta]$. In addition, the order of approximation is estimated for continuous functions using the modulus of continuity of the function to be approximated. We introduce some modifications to the classical definition of the NN operators studied in [2–4, 16]. In particular, the definition of the density functions generated by generalized logistic-type functions has been modified, together with the values of the coefficients, the weights, and the thresholds of recommended operators.

The paper includes five sections. In the second section, we introduce the generalized logistic-type function with its first and second derivatives, and determine some limit properties. In the third section, we constract the operator $F_n^{a,\varepsilon}$, which is considered as a density function, and establish the interpolation properties of it. Moreover, we present the uniform first-order approximation theorem for continuous functions defined on the finite intervals. Next, we introduce a Kantorovich-type variant of the $F_n^{a,\varepsilon}$ operators. Later, the approximation results of these operators are given in the spaces L_p with $1 \leq p \leq \infty$ and $C[\alpha,\beta]$, respectively. In the fourth one, varied parameter combinations of the proposed generalized logistic-type activation function are analyzed to determine for which parameter values we are able to obtain a more feasible activation function. Identifying convenient parameters for the operators $F_n^{a,\varepsilon}$ and the Kantorovich variant of $F_n^{a,\varepsilon}$, the numerical approximation results of some functions are studied enriched with graphs. In the fifth and concluding section, we present the key findings of the study and their implications.

2 Generalized logistic-type function

Let $s > 0$ be a real number, and also $a > 0$, $H > 1$ be parameters, the generalized logistic-type function is defined in [5] as follows:

$$\theta_{s,a}(x) = \frac{1}{1 + sH^{-ax}} = \frac{H^{ax}}{s + H^{ax}}, \tag{2}$$

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

The first and second derivatives of the generalized logistic-type function $\theta_{s,a}(x)$ are given below.

Let $a, s > 0$ be the parameters, and $H > 1$.

We have

$$\begin{aligned} \theta'_{s,a}(x) &= \left(\frac{1}{1 + sH^{-ax}}\right)' \\ &= \frac{sa\,(lnH)\,H^{-ax}}{(1 + sH^{-ax})^2} \end{aligned} \tag{3}$$

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

Moreover, if we take the second derivative of (2) for $x \in \mathbb{R}$, then we have

$$\theta''_{s,a}(x) = \frac{sa^2\,(ln\,(H))^2\,(s^2H^{-ax} - H^{ax})}{(H^{ax} + 2s + s^2H^{-ax})^2}. \tag{4}$$

Proposition 1 *Let a and s be the parameters such that $a > 0$, $s > 0$, and also $H > 1$, with $\theta_{s,a}(x)$ given in (2).*

Now, let us find the first derivative of $\theta_{s,a}(x)$:

$$\begin{aligned} \theta'_{s,a}(x) &= sa\,(lnH)\,H^{-ax}(1 + sH^{-ax})^{-2} \\ &= a\,(lnH)\,\theta_{s,a}(x)(1 - \theta_{s,a}(x)). \end{aligned} \tag{5}$$

Then, proceeding to find the second derivative of the function $\theta_{s,a}(x)$

$$\begin{aligned} \theta''_{s,a}(x) &= a\,(lnH)\,(\theta_{s,a}(x) - \theta^2_{s,a}(x))' \\ &= a^2\,(ln\,(h))^2\,\theta_{s,a}(x)(1 - \theta_{s,a}(x))(1 - 2\theta_{s,a}(x)) \end{aligned} \tag{6}$$

is obtained. The generalized logistic-type function $\theta_{s,a}(x)$ has the following properties:

$$\begin{aligned} &\lim_{x\to+\infty} \theta_{s,a}(x) = \lim_{x\to+\infty} \frac{H^{ax}}{s + H^{ax}} = 1, \\ &\lim_{x\to-\infty} \theta_{s,a}(x) = \lim_{x\to-\infty} \frac{H^{ax}}{s + H^{ax}} = 0, \\ &\lim_{x\to 0} \theta_{s,a}(x) = \lim_{x\to 0} \frac{H^{ax}}{s + H^{ax}} = \frac{1}{1 + s}; s > 0, \\ &\lim_{x\to 0} \theta'_{s,a}(x) = \lim_{x\to 0} \frac{sa\,(lnH)}{H^{ax}(1 + sH^{-ax})^2} = \frac{sa\,(lnH)}{(1 + s)^2}, \\ &\lim_{x\to\pm\infty} \theta'_{s,a}(x) = \lim_{x\to\pm\infty} \frac{sa\,(lnH)}{H^{ax} + s^2H^{-ax} + 2s} = 0. \end{aligned}$$

In the following section, we define neural-network (NN) interpolation operators activated by the generalized logistic-type function. The interpolation properties of these operators are obtained along with the uniform approximation theorem with order, for continuous functions defined on bounded intervals. Additionally, we establish a Kantorovich-type variant of these operators and prove some approximation theorems for both $C[\alpha,\beta]$ and L_p spaces, $1 \leq p < \infty$.

3 Interpolation and approximation results

Before defining the NN interpolation operator, let us introduce the related activation function and density function, respectively.

Now, let $\theta_{s,a}\colon \mathrm{R} \rightarrow [0,1]$ be the generalized logistic-type function as follows:

$$\theta_{s,a}(x) = \frac{1}{1 + sH^{-ax}} = \frac{H^{ax}}{s + H^{ax}}. \tag{7}$$

Remark 2 The generalized logistic-type function is an expanded version of a logistic-type function. This function provides an opportunity to obtain a more effective activation function with appropriate parameter selections.

If $s = 1$ is taken in expression (3), this corresponds to a special case of the generalized logistic-type function defined by [5], and the resulting expression can be considered as a density function:

$$\theta'_{1,a}(x) = \frac{a\,(lnH)\,H^{-ax}}{(1 + H^{-ax})^2}. \tag{8}$$

The function $\theta'_{1,a}(x)$ satisfies the essential properties as follows:

($\boldsymbol{\psi}_1$) $\theta'_{1,a}(x)$ is an even function;

($\boldsymbol{\psi}_2$) $\theta'_{1,a}(x)$ is nondecreasing for $x < 0$ and nonincreasing for $x \geq 0$;

($\boldsymbol{\psi}_3$) $supp\left(\theta'_{1,a}\right) \subseteq [-U_{a,\varepsilon}, U_{a,\varepsilon}]$.

Finally, we observe that $\theta'_{1,a}\left(\frac{U_{a,\varepsilon}}{2}\right) > 0$, where

$$U_{a,\varepsilon} = \frac{\ln\left[\left(\frac{a(lnH)}{\varepsilon} - 2\right) + \sqrt{\left(\frac{a(lnH)}{\varepsilon}\right)^2 - 4\left(\frac{a(lnH)}{\varepsilon}\right)}\right] - \ln(2)}{(lnH)\,a}, \tag{9}$$

when the condition $\frac{a(lnH)}{\varepsilon} \geq 4$ is satisfied.

This equation is solved using the equality $\frac{a(lnH)H^{-ax}}{(1+H^{-ax})^2} = \varepsilon$. If we put $H^{ax} =: y$, $\frac{a(lnH)}{\varepsilon} = c$ the solution of the equation $y\left(1 + \frac{1}{y}\right)^2 = c$ gives (9).

Now, we may introduce the NN interpolation operators, which are based on the generalized logistic-type functions, and study their important properties.

Definition 3 Let $f : [\alpha,\beta] \rightarrow \mathbb{R}$ be a bounded and measurable function and $n \in \mathbb{N}^+$. The NN interpolation operators, activated by the generalized logistic-type functions and acting upon f, are given by

$$F_n^{a,\varepsilon}\left(f,x\right) = \frac{\sum_{k=0}^{n} f\left(x_k\right)\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)}{\sum_{k=0}^{n}\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)}, \quad x \in [\alpha,\beta], \tag{10}$$

where the x_ks are the uniform spaced notes given by $x_k = \alpha + kh$, $k = 0, 1, \ldots, n$ with $h = \frac{\beta-\alpha}{n}$.

Remark 4 Different density functions may be employed in NN interpolation operators. For example, Costarelli [16] defines the NN interpolation operator ϕ_R by considering a special linear combination of shifted ramp functions as a density function as follows:

$$G_n\left(g, x\right) = \frac{\sum_{k=0}^{n} g\left(x_k\right) \phi_R\left(\frac{\mathrm{n}(x-x_k)}{b-a}\right)}{\sum_{k=0}^{n} \phi_R\left(\frac{n(x-x_k)}{b-a}\right)}, \quad x \in [a, b],$$

where the x_ks are the uniform spaced notes defined by $x_k = a + kh$, $k = 0, 1, \ldots, n$, with $h = \frac{\beta-\alpha}{n}$. Here, we consider function $\theta'_{1,a}$ as a density function and $\theta_{s,a}$ as the activation function to construct the operator $F_n^{a,\varepsilon}$. Figure 10 shows that for the function f, the operator $F_n^{a,\varepsilon}$ gives a better approximation than the operator G_n. Also, for the function g, *the operator* G_n gives a better approximation than the operator $F_n^{a,\varepsilon}$.

We aim to give the following lemma that is of great importance for the proof of the theorems in the rest of the paper.

Lemma 5 *The operator* $F_n^{a,\varepsilon}$ *satisfies the following essential properties*:

(***i***) *For every* $x \in [\alpha, \beta]$, *the following inequality holds*

$$\sum_{k=0}^{n} \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right) \geq \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n|x-x_t|}{\beta-\alpha}\right).$$

(***ii***) *If we choose* $t \in \{0, 1, \ldots, n\}$ *as a proper index such that* $|x - x_t| \leq \frac{h}{2}$, *then*

$$\sum_{k=0}^{n} \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right) \geq \theta'_{1,a}\left(\frac{U_{a,\varepsilon}}{2}\right)$$

is satisfied.

(***iii***) *For every bounded measurable function* $f : [\alpha, \beta] \to \mathrm{R}$, *we have the inequality*:

$$\left|F_n^{a,\varepsilon}\left(f, x\right)\right| \leq \|f\|_\infty,$$

where $\|f\|_\infty = ess \sup_{x\in[\alpha,\beta]} \left|f(x)\right|$.

Proof (***i***) It is clear from ($\boldsymbol{\psi_1}$).

(***ii***) follows from the observation that $|x - x_t| \leq \frac{h}{2}$ for $t \in \{0, 1, \ldots, n\}$,

$$U_{a,\varepsilon}\frac{n|x-x_t|}{\beta-\alpha} \leq U_{a,\varepsilon}\frac{nh}{2(\beta-\alpha)} = \frac{U_{a,\varepsilon}}{2}$$

and by property ($\boldsymbol{\psi_2}$) we have:

$$\sum_{k=0}^{n} \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right) \geq \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n|x_t-x_k|}{\beta-\alpha}\right) \geq \theta'_{1,a}\left(\frac{U_{a,\varepsilon}}{2}\right) > 0.$$

(***iii***) is obtained as follows: for every $x \in [\alpha, \beta]$, since

$$|F_n^{a,\varepsilon}(f,x)| \leq \|f\|_\infty \frac{\sum_{k=0}^{n} \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right)}{\sum_{k=0}^{n} \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right)}, \tag{11}$$

where $\|f\|_\infty = ess\sup_{x\in\alpha,\beta]} \left|f(x)\right|$, we obtain the desired result, thus the proof is complete. □

Now, we can prove the following theorem, which represents the interpolation properties of the operators $F_n^{a,\varepsilon}$ in $C[\alpha,\beta] := \{f | f : [\alpha,\beta] \to \mathbb{R} \text{ continuous}\}$.

3.1 Approximation results in the space $C[\alpha, \beta]$

Theorem 6 *Let $f : [\alpha,\beta] \to \mathbb{R}$ be a bounded and measurable function and $n \in \mathbb{N}^+$. Then, for every $t = 0, 1, \dots, n$.*

$$F_n^{a,\varepsilon}\left(f, x_t\right) = f\left(x_t\right).$$

Proof Let $t = 0, 1, \dots, n$ be fixed. Initially, we note that when $k = t$, we obtain:

$$\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right) = \theta'_{1,a}(0) = \frac{a\ln(H)}{4}.$$

For $k \neq t$, we have the following:

$$U_{a,\varepsilon}\frac{n|x_t-x_k|}{\beta-\alpha} \geq U_{a,\varepsilon}\frac{nh}{\beta-\alpha} = U_{a,\varepsilon}.$$

Therefore, by using the properties $(\boldsymbol{\psi}_1)$ and $(\boldsymbol{\psi}_2)$ we obtain:

$$0 \leq \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right) = \theta'_{1,a}\left(U_{a,\epsilon}\frac{n|x_t-x_k|}{\beta-\alpha}\right) \leq \theta'_{1,a}(U_{a,\varepsilon}) = 0.$$

Consequently, we reach the conclusion:

$$\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right) = \begin{cases} \frac{aln(H)}{4}, & t = k, \\ 0, & t \neq k, \end{cases} \tag{12}$$

for every $t,k = 0,1,\dots,n$. Thus, (12) leads to the following:

$$F_n^{a,\varepsilon}(f,x_t) = \frac{f(x_t)\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right)}{\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x_t-x_k)}{\beta-\alpha}\right)} = f(x_t),$$

for every $t = 0,1,\dots n$. This proves the desired result.

In what follows, for every continuous function f on the bounded interval $[\alpha,\beta]$, the following uniform approximation theorem with order can also be proven. We consider the first difference with step u,

$$\Delta_u f(x) = f(x+u) - f(x),$$

of the function f, and put

$$\omega(f,\delta) = w(\delta) = \sup_{x,x+u\in[\alpha,\beta],|u|\leq\delta} \left|f(x+u) - f(x)\right|, \tag{13}$$

for any function $f \in C[\alpha,\beta]$. The function $w(\delta)$, which is called the modulus of continuity of f, is defined for $0 \leq \delta \leq l$ such that $l = \beta - \alpha$. □

Theorem 7 *Let $f \in C[\alpha,\beta]$, then*

$$\|F_n^{a,\epsilon}(f,.) - f(.)\|_\infty \leq \frac{2}{\theta'_{1,a}(\frac{U_{a,\varepsilon}}{2})}\omega\left(f, \frac{\beta-\alpha}{n}\right),$$

for every $n \in \mathbb{N}^+$.

Proof For any fixed $x \in [\alpha,\beta]$ there exists $t \in \{0,1,\dots,n-1\}$ such that $x_t \leq x \leq x_{t+1}$ and by property ($\mathbf{\Omega_2}$), we can express the result as follows:

$$\begin{aligned}
\left|F_n^{a,\varepsilon}(f,x) - f(x)\right| &= \frac{\left|\sum_{k=0}^n f(x_k)\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right) - f(x)\sum_{k=0}^n \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)\right|}{\sum_{k=0}^n \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)} \\
&\leq \frac{1}{\theta'_{1,a}\left(\frac{U_{a,\varepsilon}}{2}\right)} \sum_{k=0}^n \left|f(x_k) - f(x)\right| \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right) \\
&= \frac{1}{\theta'_{1,a}\left(\frac{U_{a,\varepsilon}}{2}\right)} \Bigg[\sum_{\substack{k=0 \\ k\neq i,i+1}}^n \left|f(x_k) - f(x)\right| \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right) \\
&\quad + \left|f(x_i) - f(x)\right| \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_t)}{\beta-\alpha}\right) \\
&\quad + \left|f(x_{i+1}) f(x)\right| \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_{t+1})}{\beta-\alpha}\right)\Bigg] \\
&= \frac{I_1 + I_2 + I_3}{\theta'_{1,a}\left(\frac{U_{a,\varepsilon}}{2}\right)}.
\end{aligned}$$

It is easy to see that, for $k \neq t, t+1$ we have $U_{a,\varepsilon}\frac{n|x-x_k|}{\beta-\alpha} \geq U_{a,\varepsilon}\frac{nh}{\beta-\alpha} = U_{a,\varepsilon}$. Then, by ($\boldsymbol{\psi}_1$), ($\boldsymbol{\psi}_2$), and ($\boldsymbol{\psi}_3$) we obtain $\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right) = \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n|x-x_k|}{\beta-\alpha}\right) = 0$. The observation above implies that $I_1 = 0$.

Note that $|x_t - x| \leq h$ and $|x_{t+1} - x| \leq h$, we can state;

$$|f(x_t) - f(x)| \leq \omega(f,h) = \omega\left(f, \frac{\beta-\alpha}{n}\right)$$

and similarly,

$$|f(x_t) - f(x)| \leq \omega(f,h) = \omega\left(f, \frac{\beta-\alpha}{n}\right).$$

Then, we finally have $I_1 + I_2 + I_3 = I_2 + I_3 \leq \frac{1}{\theta'_{1,a}\left(\frac{U_{a,\varepsilon}}{2}\right)} 2\omega\left(f, \frac{\beta-\alpha}{n}\right)$, and that completes the proof. □

Now, we investigate the approximation results of NN operators in the space L_p with $1 \leq p \leq \infty$. We define a Kantorovich-type variant of $F_n^{a,\varepsilon}$. Later, the theorem of approximation by the Kantorovich-type operator in L_p is given.

3.2 Approximation results in the space $L_p[\alpha, \beta]$

Let $1 \leq p < \infty$. The space $L_p[\alpha, \beta]$ consists of all measurable functions f for which the following is finite

$$\begin{cases} \|f\|_p = \left(\int_\alpha^\beta |f(x)|^p \, dx\right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \|f\|_\infty = ess\sup |f(x)|_{x\in[\alpha,\beta]}, & p = \infty. \end{cases}$$

For any $f \in L_p[\alpha, \beta]$, the modulus of continuity of f is defined as follows:

$$\omega(f, \delta)_p = \begin{cases} \sup_{|u|\leq\delta} \left(\int_\alpha^\beta |f(x+u) - f(x)|^p \, dx\right)^{\frac{1}{p}}, & 1 \leq p < \infty \\ ess\sup_{x, x+u\in[\alpha,\beta], |u|\leq\delta} |f(x+u) - f(x)|, & p = \infty. \end{cases}$$

Taking into account the following assumption

$$\sum_{k=0}^{n} \theta'_{1,a}\left(U_{a,\varepsilon} \frac{n(x - x_k)}{\beta - \alpha}\right) = 1, \tag{14}$$

in (10), we can define the Kantorovich variant of the operator $F_n^{a,\varepsilon}$ as follows:

$$K_n^{a,\varepsilon}(f, x) = \frac{n+1}{\beta - \alpha} \sum_{k=0}^{n} \int_{m_k}^{m_{k+1}} f(\eta) d\eta \theta'_{1,a}\left(U_{a,\varepsilon} \frac{n(x - x_k)}{\beta - \alpha}\right),$$

where

$$m_k = \alpha + \frac{(\beta - \alpha)k}{n+1}, \quad k = 0, 1, \ldots, n+1.$$

Remark 8 The Kantorovich-type variant of operators may also vary with respect to activation functions, which lead to different results in function approximations. For example, Qian and Yu [23] define the Kantorovich the variant as

$$T_{n,\sigma}(f, x) = \frac{n+1}{b - a} \sum_{k=0}^{n} \int_{y_k}^{y_{k+1}} f(t) dt \phi\left(\frac{2m}{h}(x - x_k)\right),$$

with φ as the activation function, and where

$$y_k = a + \frac{(b - a)k}{n+1}, k = 0, 1, \ldots, n+1.$$

Based on this, in the fourth section, two operators are compared for a discontinuous function and analyzed to determine which one gives a better approximation.

Now, we show that operators $K_n^{a,\varepsilon}$ are bounded in $L_p[\alpha, \beta]$.

Lemma 9 *Let* $1 \le p \le \infty$. *If* $f \in L_p[\alpha,\beta]$ *then, we have*:

$$\|K_n^{a,\varepsilon}(f)\|_p \le (a\,(\ln H))^{\frac{1}{p}}\,\|f\|_p. \tag{15}$$

Proof When $1 \le p < \infty$, for any $x \in [\alpha,\beta]$, it holds that $F_n^{a,\epsilon}(1,x) = 1$. Applying Hölder's inequality for $p > 1$, we obtain

$$\begin{aligned}
\|K_n^{a,\epsilon}(f)\|_p^p &= \int_\alpha^\beta \left|K_n^{a,\varepsilon}(f)\right|^p dx \\
&\le \int_\alpha^\beta \left(\frac{n+1}{\beta-\alpha}\right)^p \left(\sum_{k=0}^{n} \left|\int_{m_k}^{m_{k+1}} f(\eta)\,d\eta\right| \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)\right) \\
&\quad \left(\sum_{k=0}^{n} \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)\right)^{p-1} dx \\
&= \int_\alpha^\beta \left(\frac{n+1}{\beta-\alpha}\right)^p \sum_{k=0}^{n} \left|\int_{m_k}^{m_{k+1}} f(\eta)\,d\eta\right| \\
&\quad \times \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right) dx \\
&\le \frac{n+1}{\beta-\alpha}\int_\alpha^\beta \sum_{k=0}^{n} \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right) \int_{m_k}^{m_{k+1}} \left|f(\eta)\right|^p d\eta dx. \\
&= \frac{n+1}{\beta-\alpha}\sum_{j=0}^{n-1}\sum_{k=0}^{n} \int_{x_j}^{x_{j+1}} \theta'_{1,a}\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right) dx \\
&\quad \times \int_{m_k}^{m_{k+1}} \left|f(\eta)\right|^p d\eta.
\end{aligned}$$

Using the expression $\theta'_{1,a}\left(U_{a,\varepsilon}\frac{n|x-x_k|}{\beta-\alpha}\right) = 0$ for $k \neq t, t+1$, and the fact that $\|\theta'_{1,a}\| = \frac{a(lnH)}{4}$, we have:

$$\begin{aligned}
\|K_n^{a,\varepsilon}(f)\|_p^p &\le \frac{n+1}{\beta-\alpha}\frac{a\,(lnH)}{4}\sum_{j=0}^{n-1}\sum_{k=j,j+1}(x_{j+1}-x_j)\int_{m_k}^{m_{k+1}}\left|f(\eta)\right|^p d\eta \\
&\le \frac{n+1}{n}\frac{a\,(lnH)}{4}\left(\sum_{j=0}^{n-1}\int_{m_j}^{m_{j+1}}\left|f(\eta)\right|^p d\eta + \sum_{j=0}^{n-2}\int_{m_{j+1}}^{m_{j+2}}\left|f(\eta)\right|^p d\eta\right) \\
&\le a\,(\ln H)\,\|f\|_p^p
\end{aligned}$$

and this proves (15) for $1 \le p < \infty$.

Let us recall the definition of the well-known Steklov function $f_h(x)$:

For an integrable function f defined on the closed and bounded interval $[\alpha,\beta]$, the function f_h is defined as:

$$f_h(x) = \frac{1}{h}\int_x^{x+h} f(u)du = \frac{1}{h}\int_0^h f(x+u)du, \quad \alpha \le x \le \beta - h,$$

where $h = \frac{\beta-\alpha}{n}$. This function, expressed in the given form, is called the Steklov function associated with the interval $[\alpha,\beta]$ [24]. The Steklov function f_h has a derivative:

$$f_h'(x) = \frac{1}{h}(f(x+h) - f(x)) \tag{16}$$

and it holds almost everywhere in $[\alpha,\beta-h]$. Then, f_h is uniformly continuous on the $[\alpha,\beta-h]$, we have:

$$\|f - f_h\|_{L_p[\alpha,\beta-h]} \leq \omega(f,h)_p, \quad 1 \leq p \leq \infty, \tag{17}$$

$$\|f_h'\|_{L_p[\alpha,\beta-h]} \leq \frac{1}{h}\omega(f,h)_p, \quad 1 \leq p \leq \infty,$$

where $\omega(f,h)$ is the modulus of continuity of f_h.

Set

$$\hat{f}_h(x) = \begin{cases} f_h(x), & \alpha \leq x \leq \beta - h \\ f_h(\beta - h), & \beta - h < x \leq \beta. \end{cases} \tag{18}$$

Then, it holds almost everywhere in $[\alpha,\beta]$ that

$$(\hat{f}_h(x))' = \begin{cases} f_h'(x), & \alpha \leq x \leq \beta - h \\ 0, & \beta - h < x \leq \beta. \end{cases} \tag{19}$$

□

Lemma 10 ([23, 25]) *Let* $1 \leq p \leq \infty$. *If* $f \in L_p[\alpha,\beta]$. *Then,*

$$\|f - \hat{f}_h(x)\|_p \leq (1 + 2^{\frac{1}{p}})\omega(f,h)_p. \tag{20}$$

$$h\|\hat{f}_h(x)'\| \leq \omega(f,h)_p. \tag{21}$$

Lemma 11 *Let* $f \in L_p[\alpha,\beta]$, $1 \leq p \leq \infty$. *Then,*

$$\|K_n^{a,\varepsilon}\hat{f}_h(x) - \hat{f}_h\|_p \leq a(\ln H) \times 2^{\frac{1}{p}}\omega(f,h)_p.$$

Proof For $j = 0,1,\dots,n-1$, observing that $m_j \leq x_j \leq m_{j+1} \leq x_{j+1} \leq m_{j+2}$, $t = 0,1,\dots,n-1$, for $1 \leq p \leq \infty$ we have

$$\int_{x_j}^{x_{j+1}} \left| \sum_{k=0}^{n} \int_{m_k}^{m_{k+1}} \left(\hat{f}_h(\eta) - \hat{f}_h(x)\right) d\eta \left(\theta_{1,a}'\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)\right)\right|^p dx$$

$$= \int_{x_j}^{x_{j+1}} \left| \sum_{k=j,j+1} \int_{m_k}^{m_{k+1}} \left(\hat{f}_h(\eta) - \hat{f}_h(x)\right) d\eta \left(\theta_{1,a}'\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)\right)\right|^p dx$$

$$\leq \int_{x_j}^{x_{j+1}} \left(\left| \sum_{k=j,j+1} \int_{m_k}^{m_{k+1}} \left| \int_{\eta}^{x} (\hat{f}_h(\lambda))' d\lambda \right| d\eta \right| \left(\theta_{1,a}'\left(U_{a,\varepsilon}\frac{n(x-x_k)}{\beta-\alpha}\right)\right)^p \right) dx$$

$$\begin{aligned}
&\leq \left(\frac{a\,(lnH)}{2}\right)^p \int_{x_j}^{x_{j+1}} \left(\left| \int_{m_j}^{m_{j+1}} \int_{m_j}^{x_{j+1}} \left| (\hat{f}_h(\lambda))' \right| d\lambda \right|^p d\eta \right. \\
&\quad \times \left. \left| \int_{m_{j+1}}^{m_{j+2}} \int_{x_j}^{m_{j+2}} |(\hat{f}_h(\lambda))'| d\lambda \right|^p d\eta \right) dx \\
&\leq \left(\frac{a\,(lnH)}{2}\right)^p \frac{\beta-\alpha}{n} \left(\frac{\beta-\alpha}{n+1}\right)^p \left(\left| \int_{m_j}^{x_{j+1}} \left| (\hat{f}_h(\lambda))' \right| d\lambda \right|^p + \left| \int_{x_j}^{m_{j+2}} |(\hat{f}_h(\lambda))'| d\lambda \right|^p \right) \\
&\leq (a\,(lnH))^p 2^{1-p} \frac{\beta-\alpha}{n} \left(\frac{\beta-\alpha}{n+1}\right)^p \left(\int_{m_j}^{m_{j+2}} \left| (\hat{f}_h(\lambda))' \right| d\lambda \right)^p \\
&\leq (a\,(lnH))^p \frac{\beta-\alpha}{n} \left(\frac{\beta-\alpha}{n+1}\right)^{2p-1} \int_{m_j}^{m_{j+2}} \left| (\hat{f}_h(\lambda))' \right|^p d\lambda.
\end{aligned}$$

Note that the last inequality uses Hölder's inequality for $p > 1$, consequently,

$$\begin{aligned}
&\|K_n^{a,\varepsilon}\left(\hat{f}_h\right) - \hat{f}_h\|_p^p \\
&= \left(\frac{n+1}{\beta-\alpha}\right)^p \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} \left| \sum_{k=0}^{n} \left(\theta_{1,a}' \left(U_{a,\varepsilon} \frac{n\,(x-x_k)}{\beta-\alpha} \right) \right) \int_{m_j}^{m_{j+1}} ((\hat{f}_h(\eta) - \hat{f}_h(x)) d\eta \right|^p \\
&\leq (ah\,(lnH))^p \sum_{j=0}^{n-1} \int_{m_j}^{m_{j+2}} \left| (\hat{f}_h(\lambda))' \right|^p d\lambda \\
&\leq 2(ah\,(\ln H))^p \|\hat{f}_h(\lambda)'\|_p^p \\
&\leq 2 \times (a\,(\ln H))^p \omega(f,h)_p^p,
\end{aligned}$$

where in the last inequality, (20) is used. When $p = \infty$, we have for any $x \in [x_j, x_{j+1}]$ that

$$\begin{aligned}
\left| K_n^{a,\varepsilon}\left(\hat{f}_h, x\right) - \hat{f}_h\,(x) \right| &= \frac{n+1}{\beta-\alpha} \left| \sum_{k=0}^{n} \left(\theta_{1,a}' \left(U_{a,\varepsilon} \frac{n\,(x-x_k)}{\beta-\alpha} \right) \right) \int_{m_k}^{m_{k+1}} (\hat{f}_h(\eta) - \hat{f}_h(x) d\eta \right| \\
&\leq \frac{a\,(lnH)}{4} \frac{n+1}{\beta-\alpha} \sum_{k=j,j+1} \int_{m_k}^{m_{k+1}} \left| \int_t^x (\hat{f}_h(\lambda))' du \right| d\eta \\
&\leq \frac{a\,(lnH)}{4} \left(\int_{m_j}^{x_{j+1}} \left| (\hat{f}_h(\lambda)' \right| d\lambda + \int_{x_j}^{m_{j+2}} \left| \left(\hat{f}_h(\lambda)\right)' \right| d\lambda \right) \\
&\leq \frac{a\,(\ln H)}{4} \|(\hat{f}_h)'\| \left(\left| x_{j+1} - m_j \right| + \left| m_{j+2} - x_j \right| \right) \\
&\leq ah\,(\ln H)\, \|(\hat{f}_h)'\| \\
&\leq a\,(\ln H)\, \omega(f,h)_\infty. \qquad \square
\end{aligned}$$

Theorem 12 *If $f \in L_p[\alpha,\beta]$, $(1 \leq p \leq \infty)$, then*

$$\|K_n^{a,\epsilon}(f) - f\|_p \leq \left(1 + (a\,(\ln H) + 1) \times 2^{\frac{1}{p}} + (a\,(\ln H))^{\frac{1}{p}} + (2a\,(\ln H))^{\frac{1}{p}}\right) \omega(f,h)_p. \tag{22}$$

Proof By using (15), Lemma 10, and Lemma 11, we have:

$$
\begin{aligned}
\|K_n^{a,\varepsilon}(f)-f\|_p &\le \left\|K_n^{a,\varepsilon}(f-\hat{f}_h)\right\| + \left\|K_n^{a,\varepsilon}(\hat{f}_h)-\hat{f}_h)\right\|_p + \left\|f-\hat{f}_h\right\|_p \\
&\le ((a(\ln H))^{\frac{1}{p}}+1)\|f-\hat{f}_h\|_p + a(\ln H)\times 2^{\frac{1}{p}}\omega(f,h)_p \\
&\le \left(\left((a(\ln H))^{\frac{1}{p}}+1\right)(1+2^{\frac{1}{p}}) + a(\ln H)\times 2^{\frac{1}{p}}\right)\omega(f,h)_p \\
&\le \left(1+(a(\ln H)+1)\times 2^{\frac{1}{p}} + (a(\ln H))^{\frac{1}{p}} + (2a(\ln H))^{\frac{1}{p}}\right)\omega(f,h)_p. \qquad \square
\end{aligned}
$$

4 Numerical examples

In this section, we observe how the generalized logistic-type function evolves for different parameter values. These graphical results have been generated using the computer programming language Python 3.8 (see [6, 21] and references therein). As can be seen in the graphs below, we can make the following inferences for the parameters. While the parameters H and a increase, the support interval decreases. The approximation behavior of the operators $F_n^{a,\varepsilon}$ and $K_n^{a,\varepsilon}$ to the functions with respect to the choice of parameters n, a, and H are investigated. In Fig. 10, the operators $F_n^{a,\varepsilon}$ and "G_n" [16] are compared. For the function f, the operator $F_n^{a,\varepsilon}$ performs a better approximation; while for g, the operator "G_n" demonstrates a better approximation. Then, the approximation performances of the operators $K_n^{a,\varepsilon}$ and $T_{n,\sigma}$ to the function h and z is analyzed and it is seen that the operator $K_n^{a,\varepsilon}$ gives a preferable approximation in Fig. 11.

Let us dive into the above-mentioned analysis. First, in Fig. 2 the support values are calculated for $(H,\ a,\ \varepsilon)$ under the condition $\frac{a(lnH)}{\varepsilon} \ge 4$ for values $(1.2,\ 1, 0.01)$, $(1.2,\ 1,\ 0.001)$, $(1.2,\ 1,\ 0.0001)$, $(2,\ 1,\ 0.01)$, $(2,\ 1,\ 0.001)$, $(2,\ 1,\ 0.0001)$. Again in Fig. 3, the support

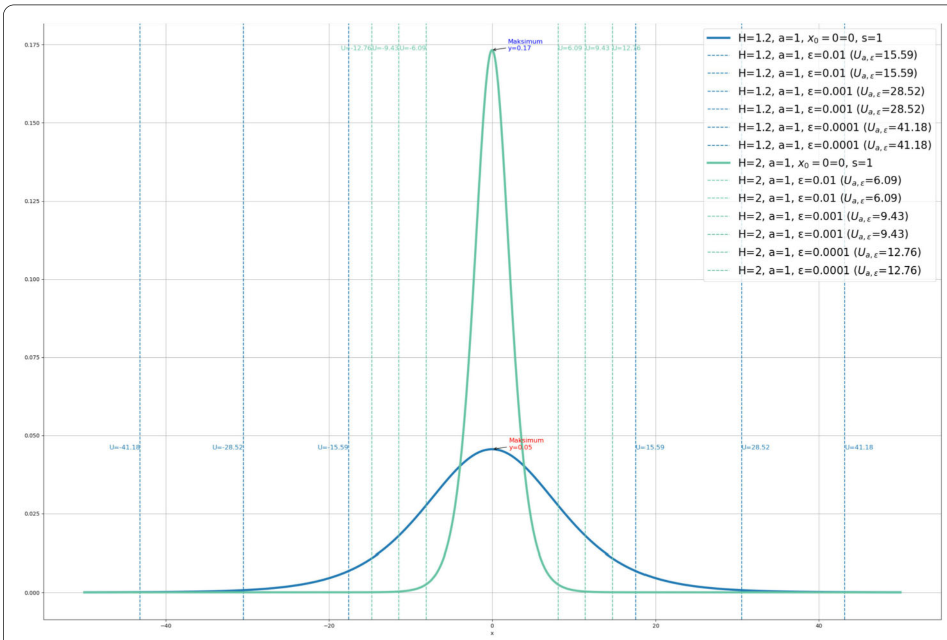

Figure 2 Comparison of support range and maximum point for different values of *H*

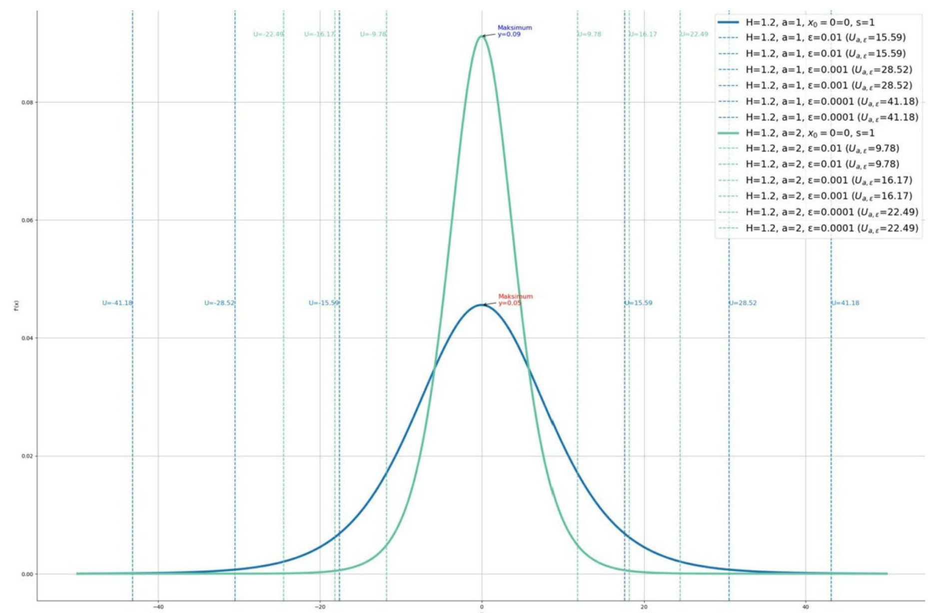

Figure 3 Comparison of support range and maximum point for different values of *a*

Table 1 Maximum error of Operator 1 used in Fig. 4 for different values of *n*

Operator type	*a*	*H*	*n*	Maximum error
Operator 1	0.8	1.01	20	0.6348
	0.8	1.01	50	0.2506
	0.8	1.01	100	0.1416

Table 2 Maximum error of Operator 1 used in Fig. 5 for different values of *H*

Operator type	*a*	*H*	*n*	Maximum error
Operator 1	2.5	1.01	50	0.0841
	2.5	100	50	0.1744

Table 3 Maximum error of Operator 1 used in Fig. 6 for different values of *a*

Operator type	*a*	*H*	*n*	Maximum error
Operator 1	0.1	1.1	50	0.2086
	100	1.1	50	0.1724

values for (H, a, ε) are calculated under the same condition $\frac{a(lnH)}{\varepsilon} \geq 4$ for another six values $(1.2, 1, 0.01)$, $(1.2, 1, 0.001)$, $(1.2, 1, 0.0001)$, $(1.2, 2, 0.01)$, $(1.2, 2, 0.001)$, $(1.2, 2, 0.0001)$, respectively.

Secondly, Tables 1–3 represent maximum error values of Operator 1 used in Figs. 4–6; the subsequent Tables 4–6 demonstrate the maximum error values of Operator 3 inserted in Figs. 7–9 for specific values of n, H, and a. Lastly, Tables 7 and 8 yield other numerical results.

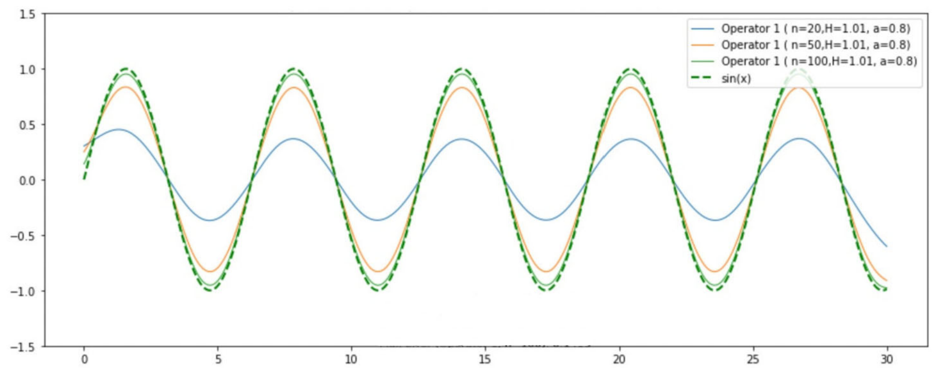

Figure 4 Approximation of Operator 1 to *sin* (x) for different values of n

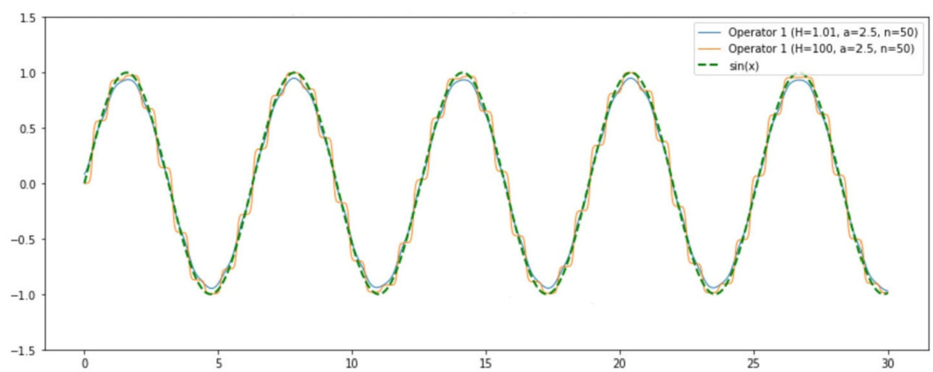

Figure 5 Approximation of Operator 1 to sin (x) for different values of H

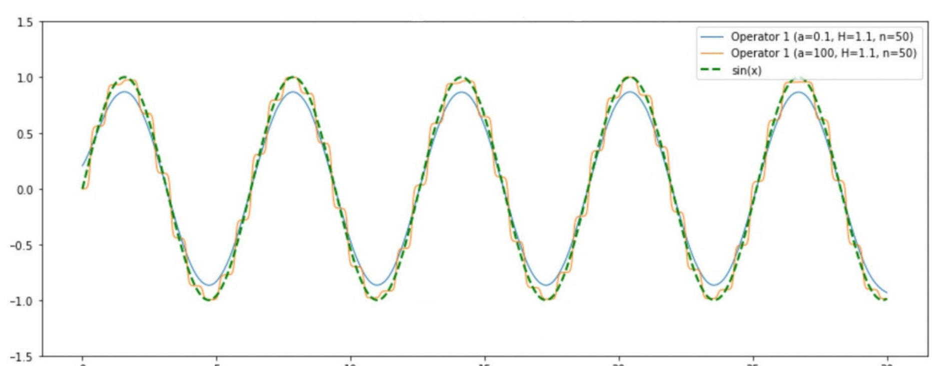

Figure 6 Approximation of Operator 1 to sin (x) for different values of a

5 Conclusion

In this paper, the first derivative of the generalized logistic-type function $\theta_{s,a}$ for $a, s > 0$ is considered and also used as the density function into the NN interpolation operators $F_n^{a,\varepsilon}$. The derivative of the generalized logistic-type function has extremely small values at $+\infty$ and $-\infty$. This leads to the vanishing gradient problem, so the weight and bias values in the cells cannot be updated efficiently. Therefore, no efficient "learning" occurs. In fact,

Table 4 Maximum error of Operator 3 used in Fig. 7 for different values of n

Operator type	a	H	n	Maximum error
Operator 3	0.5	100	20	16.2001
	0.5	100	100	15.4426

Table 5 Maximum error of Operator 3 used in Fig. 8 for different values of H

Operator type	a	H	n	Maximum error
Operator 3	0.1	2	100	17.6894
	0.1	100	100	31.8502

Table 6 Maximum error of Operator 3 used in Fig. 9 for different values of a

Operator type	a	H	n	Maximum error
Operator 3	0.1	100	100	18.3994
	2	100	100	31.8502

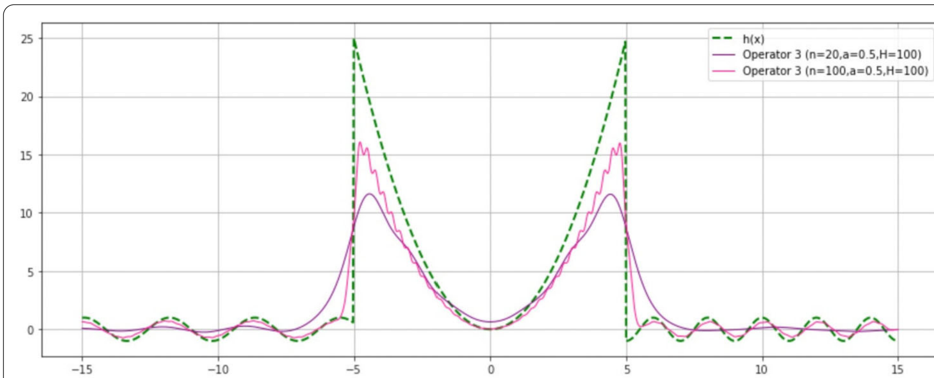

Figure 7 Approximation of Operator 3 to $h(x)$ for different values of n

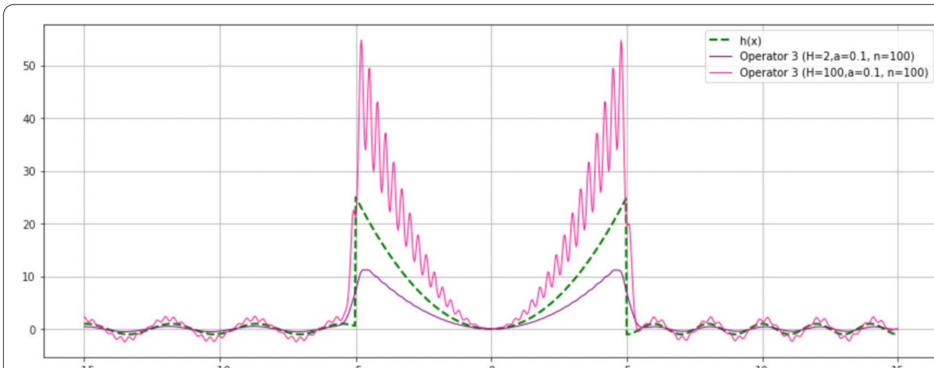

Figure 8 Approximation of Operator 3 to $h(x)$ for different values of H

a wider support interval allows us to achieve better results. Our support values vary according to the H, a, ε. As can be seen from the graphs in the article; as the parameters H and a decrease, our support range increases, so we can create a better logistic function by making appropriate choices. Furthermore, when the parameter values of H, a, ε, and n of

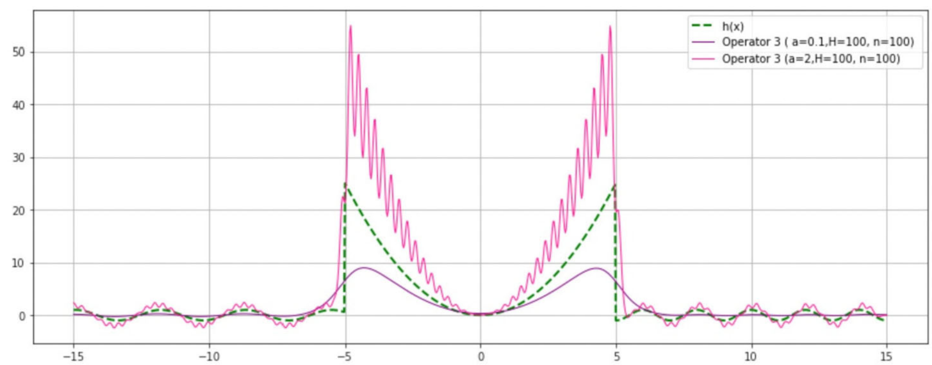

Figure 9 Approximation of Operator 3 to $h(x)$ for different values of a

Table 7 Maximum error of Operator 1 and Operator 2 used in Fig. 10 for specific values of a, H, and n

Operator type	a	H	n	Functions		Maximum error	
Operator 1 ($F_n^{a,\varepsilon}$)	0.1	1.01	20	f	g	**0.7065**	0.1104
Operator 2 (G_n)	0.1	1.01	20	f	g	0.9916	**0.0381**

Table 8 Maximum error of Operator 3 and Operator 4 used in Fig. 11 for some values of a, H, m, and n

Operator type	a	H	m, n	Functions		Maximum error	
Operator 3 ($F_n^{a,\varepsilon}$)	0.5	100	10,100	z	h	**0.5076**	**15.4426**
Operator 4 ($T_{n,\sigma}$)	0.5	100	10,100	z	h	0.5879	15.8871

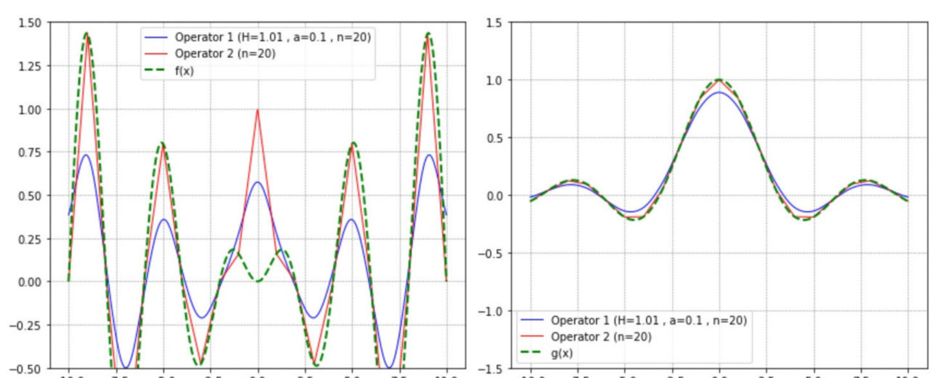

Figure 10 Comparison of approximations of $F_n^{a,\varepsilon}$ (Operator 1) and G_n (Operator 2) to functions f and g, defined as: $f(x) = \begin{cases} \frac{\sin(\frac{\pi}{2}x)}{\frac{\pi}{2}x}, & x \neq 0 \\ 1, & x = 0 \end{cases}$ $g(x) = \begin{cases} \frac{\sin(x)}{x}, & x \neq 0 \\ 1, & x = 0 \end{cases}$

the operator $F_n^{a,\varepsilon}$ are chosen appropriately, the approximation to the function improves. Here, the appropriate parameter selection and maximum error may vary according to the functions selected. The proposed operator $F_n^{a,\varepsilon}$ shows a better approximation performance than the other NN interpolation operators for some functions. Finally, the approximation of the $F_n^{a,\varepsilon}$ to the functions h and z is analyzed and its maximum error is found.

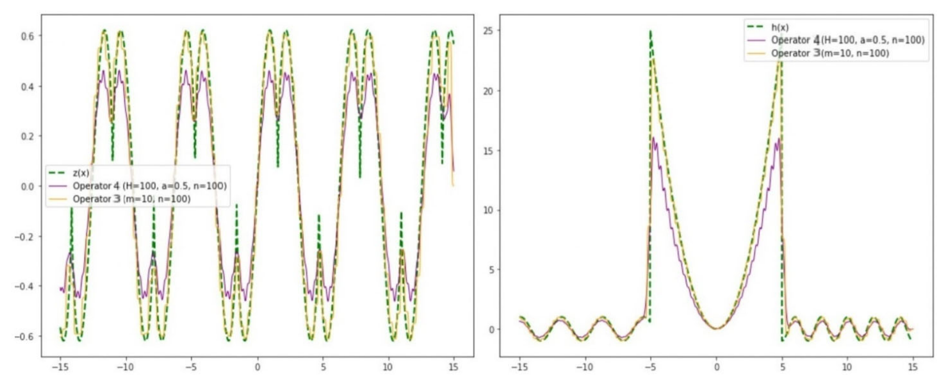

Figure 11 Comparison of approximations of the $F_n^{a,\varepsilon}$ (Operator 3) and $T_{n,\sigma}$ (Operator 4) to functions z and h defined as: $z(x) = \sin(x)\sqrt{|\cos(x)|}$ $\quad h(x) = \begin{cases} \sin(2x), & x < -5 \\ x^2, & -5 \leq x < 5 \\ \cos(\pi x), & x \geq 5. \end{cases}$

Acknowledgements
We acknowledge the editorial team's and reviewers' contributions, which strengthened the clarity and consistency of this research paper.

Author contributions
H.U. , A. O. A., S. K., and İ.B. conceived of the presented idea. H.U. , A. O. A, S. K., and İ.B. developed the theory and performed the computations. H.U. , A. O. A, S. K., and İ.B. verified the methods. All authors discussed the results and contributed to the final manuscript.

Funding
Not applicable.

Data Availability
No datasets were generated or analysed during the current study.

Declarations

Competing interests
The authors declare no competing interests.

Author details
[1]Graduate School of Natural and Applied Sciences, Ankara University, Ankara, Türkiye. [2]Department of Software Engineering, Faculty of Engineering and Natural Sciences, Istanbul Atlas University, 34408, Kagithane, Istanbul, Türkiye. [3]Department of Mathematics, Faculty of Science, Ankara University, Ankara, Türkiye.

Received: 11 April 2024 Accepted: 2 September 2024 Published online: 27 September 2024

References
1. Anastassiou, G.A.: Intelligent Systems: Approximation by Artificial Neural Networks, Intelligent Systems Reference Library vol. 19. Springer, Berlin (2011)
2. Anastassiou, G.A.: Univariate hyperbolic tangent neural network approximation. Math. Comput. Model. **53**(5–6), 1111–1132 (2011)
3. Anastassiou, G.A.: Multivariate hyperbolic tangent neural network approximation. Comput. Math. Appl. **61**(4), 809–821 (2011)
4. Anastassiou, G.A.: Multivariate sigmoidal neural network approximation. Neural Netw. **24**, 378–386 (2011)
5. Anastassiou, G.A.: Banach space valued multivariate multi layer neural network approxi mation based on q-deformed and-parametrized A-generalized logistic function. In: λ-Parametrized, Deformed and General Neural Networks. Studies in Computational Intelligence, vol. 1116, pp. 65–394. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-43021-3_15
6. Anastassiou, G.A., Coroianu, L., Gal, S.G.: Approximation by a nonlinear Cardaliaguet-Euvrard neural network operator of max-product kind. J. Comput. Anal. Appl. **12**(2), 396–406 (2010)
7. Barron, A.R.: Universal approximation bounds for superpositions of a sigmoidal function. IEEE Trans. Inf. Theory **39**, 930–945 (1993)
8. Cao, F., Chen, Z.: The approximation operators with sigmoidal functions. Comput. Math. Appl. **58**(4), 758–765 (2009)
9. Cao, F.L., Xie, T.F., Xu, Z.B.: The estimate for approximation error of neural networks: a constructive approach. Neurocomputing **71**, 626–630 (2008)

10. Cao, F.L., Xu, Z.B., Li, Y.M.: Pointwise approximation for neural networks. Lect. Notes Comput. Sci. **3496**, 39–44 (2005)
11. Cao, F.L., Zhang, R.: The errors of approximation for feedforward neural networks in the L_p metric. Math. Comput. Model. **49**, 1563–1572 (2009)
12. Cardaliaguet, P., Euvrard, G.: Approximation of a function and its derivative with a neural network. Neural Netw. **5**(2), 207–220 (1992)
13. Cheang, G.H.L.: Approximation with neural networks activated by ramp sigmoids. J. Approx. Theory **162**, 1450–1465 (2010)
14. Chui, C.K., Li, X.: Approximation by ridge functions and neural networks with one hidden layer. J. Approx. Theory **70**, 131–141 (1992)
15. Costarelli, D.: Neural network operators: constructive interpolation of multivariate functions. Neural Netw. **67**, 28–36 (2015)
16. Costarelli, D., Spigler, R.: Approximation results for neural network operators activated by sigmoidal functions. Neural Netw. **44**, 101–106 (2013)
17. Costarelli, D., Spigler, R.: Approximation by series of sigmoidal functions with applications to neural networks. Ann. Mat. Pura Appl. **194**(1), 289–306 (2015)
18. Cybenko, G.: Approximation by superpositions of a sigmoidal function. Math. Control Signals Syst. **2**, 303–314 (1989)
19. Gripenberg, G.: Approximation by neural network with a bounded number of nodes at each level. J. Approx. Theory **122**(2), 260–266 (2003)
20. Kainen, P.C., Kurkova, V.: An integral upper bound for neural network approx imation. Neural Comput. **21**, 2970–2989 (2009)
21. Karateke, S., Zontul, M., Mishra, V.N., et al.: On the approximation by Stancu-type bivariate Jakimovski–Leviatan–Durrmeyer operators. La Matematica (2024). https://doi.org/10.1007/s44007-023-00083-w
22. Pinkus, A.: Approximation theory of the MLP model in neural networks. Acta Numer. **8**, 143–195 (1999)
23. Qian, Y., Yu, D.: Rates of approximation by neural network interpolation operators. Appl. Math. Comput. **418** (2022)
24. Steklov, V.A.: On the asymptotic representation of certain functions defined by a linear differential equation of the second order, and their application to the problem of expanding an arbitrary function into a series of these functions. Kharkov (1957). (In Russian)
25. Yu, D.S., Zhou, P.: Rates of approximation by neural networks with four layers. J. Comput. Anal. Appl. **18**, 551–558 (2015)

Publisher's Note