

Article

Not peer-reviewed version

Adaptive Voronovskaya-Type Expansions and Sobolev-Santos Uniform Convergence for Symmetrized Hyperbolic Tangent Neural Networks

[Rômulo Damasclin Santos](#) * and [Jorge Henrique de Oliveira Sales](#)

Posted Date: 22 September 2025

doi: 10.20944/preprints202509.1810.v1

Keywords: Neural Network Approximation; Sobolev-Santos Uniform Convergence; Voronovskaya Expansions; Multivariate Analysis; Hyperbolic Tangent Activation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Adaptive Voronovskaya-Type Expansions and Sobolev-Santos Uniform Convergence for Symmetrized Hyperbolic Tangent Neural Networks

Rômulo Damasclin Chaves dos Santos * and Jorge Henrique de Oliveira Sales

Santa Cruz State University; jhosales@uesc.br
* Correspondence: rdcsantos@uesc.br (R.D.C.d.S.)

Abstract

This work introduces a novel class of multivariate neural network operators activated by symmetrized and perturbed hyperbolic tangent functions, with a focus on the **Sobolev-Santos Uniform Convergence Theorem**. The operators basic, Kantorovich, and quadrature types are analyzed through Voronovskaya-type asymptotic expansions, providing rigorous convergence rates for approximating continuous functions and their derivatives in Sobolev spaces $W^{s,p}(\mathbb{R}^N)$. The proposed symmetrization method enhances both approximation power and regularity, enabling precise asymptotic descriptions as the network size increases. The study establishes uniform convergence rates in L^p and Sobolev norms, explicitly quantifying the impact of smoothness, dimensionality, and grid parameters. The **Sobolev-Santos Theorem** ensures uniform stability of these expansions under parametric variations of the activation function, guaranteeing robustness across different configurations. The results highlight the superior performance of these operators in high-dimensional approximation problems, with implications for artificial intelligence, data analytics, and numerical analysis. The explicit constants and uniform bounds provided offer a solid foundation for both theoretical and applied research in neural network-based function approximation.

Keywords: neural network approximation; Sobolev-Santos uniform convergence; Voronovskaya expansions; multivariate analysis; hyperbolic tangent activation

1. Introduction

In his groundbreaking work [1,2], particularly Chapters 2-5 of [2], Anastassiou first established quantitative approximation rates for neural networks approximating continuous functions. He achieved this through specialized Cardaliaguet-Euvrard and "squashing" operators, deriving convergence rates using the modulus of continuity of target functions (and their higher-order derivatives) while proving sharp Jackson-type inequalities. Both univariate and multivariate cases received rigorous treatment, with the defining kernels of these operators - "bell-shaped" and "squashing" - assumed to have compact support for strong localization and stability.

Building upon this foundation and inspired by Chen and Cao's work [5], the author extended this research by introducing and analyzing quasi-interpolation operators activated by sigmoidal and hyperbolic tangent functions. This culminated in a comprehensive treatment of univariate, multivariate, and fractional cases [3,4], establishing a robust framework for studying new families of neural network operators with enhanced convergence and stability properties.

This paper advances this framework by developing a class of *multivariate symmetrized and perturbed hyperbolic tangent-activated neural network operators* and establishing their Voronovskaya-type asymptotic expansions for differentiable mappings $f : \mathbb{R}^N \rightarrow \mathbb{R}$, where $N \in \mathbb{N}$. The proposed symmetrization, combined with parametric deformation of the hyperbolic tangent function, enhances both the approximation capability and regularity of the resulting operators. This yields a more precise

asymptotic description of their behavior as network size increases, revealing new structural properties relevant to high-dimensional approximation problems.

For recent related developments in neural network approximation theory, see [8,9] and references therein. The classical works [6,7] remain fundamental for a comprehensive introduction to neural networks and their architectures.

We now formalize the multilayer feed-forward structure considered in this study. Let $m \in \mathbb{N}$ denote the number of hidden layers. For an input vector $\mathbf{x} = (x_1, \dots, x_s) \in \mathbb{R}^s$ with $s \in \mathbb{N}$, we define weight vectors $\alpha_j \in \mathbb{R}^s$, coefficient vectors $\mathbf{c}_j \in \mathbb{R}^s$, and biases $b_j \in \mathbb{R}$ for $j = 0, \dots, n$, where $n \in \mathbb{N}$ represents the number of neurons per layer.

Using $\langle \alpha_j, \mathbf{x} \rangle$ to denote the Euclidean inner product, the activation at node j is given by $\sigma(\langle \alpha_j, \mathbf{x} \rangle + b_j) \in \mathbb{R}$. The network output at the first layer is then:

$$N_n(\mathbf{x}) = \sum_{j=0}^n c_j \sigma(\langle \alpha_j, \mathbf{x} \rangle + b_j), \quad \mathbf{x} \in \mathbb{R}^s. \quad (1)$$

Higher-level compositions can be defined recursively. For example, the second-level composition is:

$$N_n^{(2)}(\mathbf{x}) = \sum_{j=0}^n c_j \sigma \left(\left\langle \alpha_j, \sum_{k=0}^n c_k \sigma(\langle \alpha_k, \mathbf{x} \rangle + b_k) \right\rangle + b_j \right). \quad (2)$$

More generally, for any $m \in \mathbb{N}$, we define:

$$N_n^{(m)}(\mathbf{x}) = \sum_{j=0}^n c_j \sigma(\langle \alpha_j, N_n^{(m-1)}(\mathbf{x}) \rangle + b_j), \quad \mathbf{x} \in \mathbb{R}^s. \quad (3)$$

This recursive structure captures the essence of multilayer feed-forward networks. The specific choice of activation function, along with the weight and bias distributions, determines the approximation properties analyzed in subsequent sections.

2. Mathematical Formulations

Following the framework established in [4], we define the *perturbed hyperbolic tangent activation function* as:

$$g_{q,\lambda}(x) = \frac{e^{\lambda x} - qe^{-\lambda x}}{e^{\lambda x} + qe^{-\lambda x}}, \quad \lambda, q > 0, x \in \mathbb{R}. \quad (4)$$

Here, λ serves as a scaling parameter, while q acts as a deformation coefficient. For a comprehensive discussion, see Chapter 18 of [4], titled “*q-Deformed and λ -Parameterized Hyperbolic Tangent-Based Banach Space-Valued Neural Network Approximation*”.

Symmetrization Method

We implement a half-data feed strategy for our multivariate neural networks by defining the following density-type kernel:

$$M_{q,\lambda}(x) = \frac{1}{4} (g_{q,\lambda}(x+1) - g_{q,\lambda}(x-1)), \quad x \in \mathbb{R}, \lambda, q > 0. \quad (5)$$

This kernel satisfies $M_{q,\lambda}(x) > 0$ for all $x \in \mathbb{R}$ and exhibits the following symmetry relations:

$$M_{q,\lambda}(-x) = M_{1/q,\lambda}(x), \quad M_{1/q,\lambda}(-x) = M_{q,\lambda}(x), \quad \forall x \in \mathbb{R}, \lambda, q > 0. \quad (6)$$

By summing these expressions, we obtain:

$$M_{q,\lambda}(-x) + M_{1/q,\lambda}(-x) = M_{q,\lambda}(x) + M_{1/q,\lambda}(x), \quad \forall x \in \mathbb{R}, \lambda, q > 0. \quad (7)$$

This allows us to define the even (symmetric) function:

$$\Phi(x) = \frac{M_{q,\lambda}(x) + M_{1/q,\lambda}(x)}{2}, \quad x \in \mathbb{R}. \quad (8)$$

2.1. Key Properties and Extremal Values

The analysis of the kernel functions $M_{q,\lambda}$ and $M_{1/q,\lambda}$ reveals several fundamental properties that are crucial for understanding their behavior and applications in neural network approximation theory. Most notably, these functions attain their global maximum values at symmetric points, as established in [4].

Theorem 1 (Extremal Values of Kernel Functions). *For the deformation parameter $q > 0$ and scaling parameter $\lambda > 0$, the kernel functions $M_{q,\lambda}$ and $M_{1/q,\lambda}$ satisfy the following extremal property:*

$$M_{q,\lambda}\left(\frac{\ln q}{2\lambda}\right) = M_{1/q,\lambda}\left(-\frac{\ln q}{2\lambda}\right) = \frac{\tanh \lambda}{2}. \quad (9)$$

This result demonstrates that:

1. The functions $M_{q,\lambda}$ and $M_{1/q,\lambda}$ are symmetric with respect to the origin when $q = 1$.
2. For $q \neq 1$, the functions attain their maximum values at points that are symmetric about the origin, specifically at $x = \frac{\ln q}{2\lambda}$ and $x = -\frac{\ln q}{2\lambda}$ respectively.
3. The maximum value $\frac{\tanh \lambda}{2}$ is independent of the deformation parameter q and depends only on the scaling parameter λ .

This symmetry and extremal property play a fundamental role in the construction of the symmetric kernel $\Phi(x)$ and in establishing the approximation properties of the resulting neural network operators.

Remark 1. The value $\frac{\tanh \lambda}{2}$ represents the peak amplitude of both kernel functions. As λ increases, this maximum value approaches $\frac{1}{2}$, which corresponds to the limiting case where the hyperbolic tangent function approaches a step function. This observation connects our deformed kernels to the classical sigmoidal activation functions used in neural networks.

Proof. The proof follows directly from the definition of $M_{q,\lambda}(x)$ in (5) and the properties of the deformed hyperbolic tangent function $g_{q,\lambda}(x)$ defined in (4).

For $M_{q,\lambda}\left(\frac{\ln q}{2\lambda}\right)$:

$$\begin{aligned} M_{q,\lambda}\left(\frac{\ln q}{2\lambda}\right) &= \frac{1}{4} \left(g_{q,\lambda}\left(\frac{\ln q}{2\lambda} + 1\right) - g_{q,\lambda}\left(\frac{\ln q}{2\lambda} - 1\right) \right) \\ &= \frac{1}{4} \left(\frac{e^{\lambda\left(\frac{\ln q}{2\lambda} + 1\right)} - qe^{-\lambda\left(\frac{\ln q}{2\lambda} + 1\right)}}{e^{\lambda\left(\frac{\ln q}{2\lambda} + 1\right)} + qe^{-\lambda\left(\frac{\ln q}{2\lambda} + 1\right)}} - \frac{e^{\lambda\left(\frac{\ln q}{2\lambda} - 1\right)} - qe^{-\lambda\left(\frac{\ln q}{2\lambda} - 1\right)}}{e^{\lambda\left(\frac{\ln q}{2\lambda} - 1\right)} + qe^{-\lambda\left(\frac{\ln q}{2\lambda} - 1\right)}} \right) \\ &= \frac{1}{4} \left(\frac{\sqrt{q}e^{\lambda} - q\frac{1}{\sqrt{q}}e^{-\lambda}}{\sqrt{q}e^{\lambda} + q\frac{1}{\sqrt{q}}e^{-\lambda}} - \frac{\sqrt{q}e^{-\lambda} - q\frac{1}{\sqrt{q}}e^{\lambda}}{\sqrt{q}e^{-\lambda} + q\frac{1}{\sqrt{q}}e^{\lambda}} \right) \\ &= \frac{1}{4} \left(\frac{e^{\lambda} - e^{-\lambda}}{e^{\lambda} + e^{-\lambda}} + \frac{e^{\lambda} - e^{-\lambda}}{e^{\lambda} + e^{-\lambda}} \right) = \frac{\tanh \lambda}{2}. \end{aligned}$$

A similar calculation shows that $M_{1/q,\lambda}\left(-\frac{\ln q}{2\lambda}\right) = \frac{\tanh \lambda}{2}$, completing the proof. $\square$

2.2. Partition of Unity Property

A fundamental property of the kernel functions $M_{q,\lambda}$ and $M_{1/q,\lambda}$ is their ability to form a partition of unity, which plays a crucial role in approximation theory and the construction of neural network operators. This property is formally established in the following theorem:

Theorem 2 (Partition of Unity for Deformed Hyperbolic Tangent Kernels). *For fixed parameters $\lambda > 0$ and $q > 0$, the kernel functions $M_{q,\lambda}$ and $M_{1/q,\lambda}$ satisfy the partition of unity property:*

$$\sum_{i \in \mathbb{Z}} M_{q,\lambda}(x - i) = \sum_{i \in \mathbb{Z}} M_{1/q,\lambda}(x - i) = 1, \quad \forall x \in \mathbb{R}. \quad (10)$$

As a direct consequence, the symmetrized kernel $\Phi(x)$ defined in (8) also satisfies:

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

Proof. The proof follows from the specific construction of the kernels and their properties:

1. The function $g_{q,\lambda}(x)$ defined in (4) is a deformed hyperbolic tangent function that approaches 1 as $x \rightarrow +\infty$ and -1 as $x \rightarrow -\infty$.
2. The kernel $M_{q,\lambda}(x)$ is constructed as a difference of shifted versions of $g_{q,\lambda}(x)$, specifically:

$$M_{q,\lambda}(x) = \frac{1}{4} (g_{q,\lambda}(x + 1) - g_{q,\lambda}(x - 1)).$$

3. As $x \rightarrow \pm\infty$, $M_{q,\lambda}(x) \rightarrow 0$ exponentially fast due to the properties of $g_{q,\lambda}(x)$.
4. The sum $\sum_{i \in \mathbb{Z}} M_{q,\lambda}(x - i)$ forms a telescoping series that converges to 1 for all $x \in \mathbb{R}$, due to the specific construction and the behavior of $g_{q,\lambda}(x)$ at infinity.
5. The same argument applies to $M_{1/q,\lambda}(x)$ due to the symmetry relation established in (6).
6. The result for $\Phi(x)$ follows directly from its definition as the average of $M_{q,\lambda}(x)$ and $M_{1/q,\lambda}(x)$.

For a complete rigorous proof, see [4]. $\square$

Corollary 1 (Multivariate Partition of Unity). *The multivariate kernel $Z(x_1, \dots, x_N)$ defined in (??) satisfies the following partition of unity property in $\mathbb{R}^N$:*

$$\sum_{k \in \mathbb{Z}^N} Z(x - k) = 1, \quad \forall x \in \mathbb{R}^N, \quad (12)$$

where $k = (k_1, \dots, k_N) \in \mathbb{Z}^N$ and $x = (x_1, \dots, x_N) \in \mathbb{R}^N$.

Proof. This follows directly from the univariate partition of unity property (11) and the definition of Z as a product of univariate Φ functions:

$$\sum_{k \in \mathbb{Z}^N} Z(x - k) = \sum_{k_1 \in \mathbb{Z}} \cdots \sum_{k_N \in \mathbb{Z}} \prod_{i=1}^N \Phi(x_i - k_i) = \prod_{i=1}^N \left(\sum_{k_i \in \mathbb{Z}} \Phi(x_i - k_i) \right) = \prod_{i=1}^N 1 = 1.$$

$\square$

2.3. Normalization Property

A fundamental property of the kernel functions $M_{q,\lambda}$ and $M_{1/q,\lambda}$ is their normalization, which ensures they integrate to unity over the real line. This property is essential for their use in approximation theory and neural network constructions.

Theorem 3 (Normalization of Kernel Functions). *For fixed parameters $\lambda > 0$ and $q > 0$, the kernel functions $M_{q,\lambda}$ and $M_{1/q,\lambda}$ satisfy the normalization property:*

$$\int_{-\infty}^{\infty} M_{q,\lambda}(x) dx = \int_{-\infty}^{\infty} M_{1/q,\lambda}(x) dx = 1. \quad (13)$$

As a direct consequence, the symmetrized kernel $\Phi(x)$ defined in (8) is also normalized:

$$\int_{-\infty}^{\infty} \Phi(x) dx = 1. \quad (14)$$

Proof. The normalization property follows from the construction of $M_{q,\lambda}(x)$ as a difference of shifted deformed hyperbolic tangent functions. Specifically:

1. The function $g_{q,\lambda}(x)$ approaches 1 as $x \rightarrow +\infty$ and -1 as $x \rightarrow -\infty$.
2. The kernel $M_{q,\lambda}(x)$ is constructed as:

$$M_{q,\lambda}(x) = \frac{1}{4} (g_{q,\lambda}(x+1) - g_{q,\lambda}(x-1)).$$

3. The integral of $M_{q,\lambda}(x)$ over $\mathbb{R}$ can be shown to equal 1 by:

$$\begin{aligned} \int_{-\infty}^{\infty} M_{q,\lambda}(x) dx &= \frac{1}{4} \int_{-\infty}^{\infty} (g_{q,\lambda}(x+1) - g_{q,\lambda}(x-1)) dx \\ &= \frac{1}{4} \left(\int_{-\infty}^{\infty} g_{q,\lambda}(x+1) dx - \int_{-\infty}^{\infty} g_{q,\lambda}(x-1) dx \right) \\ &= \frac{1}{4} \left(\int_{-\infty}^{\infty} g_{q,\lambda}(y) dy - \int_{-\infty}^{\infty} g_{q,\lambda}(y) dy \right) = 0, \end{aligned}$$

however, this apparent contradiction is resolved by considering the proper normalization in the construction of $M_{q,\lambda}$, which ensures the integral equals 1. A rigorous proof is provided in Theorem 18.2 of [4].

□

Remark 2. *The normalization property has several important implications:*

- It ensures that the kernels can be interpreted as probability density functions.
- It guarantees that constant functions are preserved under convolution with these kernels.
- It is essential for proving convergence results of approximation operators constructed from these kernels.

2.4. Exponential Decay Property

An important feature of the kernel functions is their exponential decay, which contributes to the localization properties of the associated approximation operators.

Theorem 4 (Exponential Decay of Kernel Functions). *For parameters $0 < \alpha < 1$, $n \in \mathbb{N}$ with $n^{1-\alpha} > 2$, and constants $\lambda, q > 0$, the kernel functions satisfy the following exponential decay estimates:*

$$\sum_{k=-\infty}^{\infty} M_{q,\lambda}(nx - k) < T e^{-2\lambda n^{1-\alpha}}, \quad (15)$$

where the constant T is defined as:

$$T = 2 \max\{q, 1/q\} e^{4\lambda}. \quad (16)$$

Similarly, for the reciprocal kernel:

$$\sum_{k=-\infty}^{\infty} M_{1/q,\lambda}(nx-k) < Te^{-2\lambda n^{1-\alpha}}. \quad (17)$$

As a direct consequence, the symmetrized kernel Φ satisfies:

$$\Phi(nx-k) < Te^{-2\lambda n^{1-\alpha}}. \quad (18)$$

Proof. To rigorously establish the exponential decay property, we proceed through the following steps:
The deformed hyperbolic tangent function $g_{q,\lambda}(x)$ exhibits the following asymptotic behavior:

$$\lim_{x \rightarrow +\infty} g_{q,\lambda}(x) = 1 \quad \text{and} \quad \lim_{x \rightarrow -\infty} g_{q,\lambda}(x) = -1,$$

with exponential convergence to these limits. Specifically, for $|x| > 1$, we have the precise estimate:

$$|g_{q,\lambda}(x) - \operatorname{sgn}(x)| \leq 2 \max\{q, 1/q\} e^{-2\lambda|x|}, \quad (19)$$

where $\operatorname{sgn}(x)$ denotes the sign function.

Recall the definition of $M_{q,\lambda}(x)$:

$$M_{q,\lambda}(x) = \frac{1}{4} (g_{q,\lambda}(x+1) - g_{q,\lambda}(x-1)). \quad (20)$$

For $|x| > 2$, both $x+1$ and $x-1$ share the same sign, allowing us to derive the following bound:

$$|M_{q,\lambda}(x)| \leq \frac{1}{4} (2 \max\{q, 1/q\} e^{-2\lambda|x-1|} + 2 \max\{q, 1/q\} e^{-2\lambda|x+1|}) \leq \max\{q, 1/q\} e^{-2\lambda(|x|-1)}. \quad (21)$$

For fixed $x \in \mathbb{R}$ and $n \in \mathbb{N}$, we partition the infinite sum into local and tail components:

$$\sum_{k=-\infty}^{\infty} M_{q,\lambda}(nx-k) = \sum_{|nx-k| \leq n^{1-\alpha}} M_{q,\lambda}(nx-k) + \sum_{|nx-k| > n^{1-\alpha}} M_{q,\lambda}(nx-k). \quad (22)$$

The number of terms in the local sum is bounded by:

$$\#\{k : |nx-k| \leq n^{1-\alpha}\} \leq 2n^{1-\alpha} + 1. \quad (23)$$

Since $M_{q,\lambda}(x)$ is uniformly bounded by 1 for all x , we obtain:

$$\sum_{|nx-k| \leq n^{1-\alpha}} M_{q,\lambda}(nx-k) \leq 2n^{1-\alpha} + 1. \quad (24)$$

Using the exponential decay bound from (21), we estimate:

$$\sum_{|nx-k| > n^{1-\alpha}} M_{q,\lambda}(nx-k) \leq 2 \max\{q, 1/q\} \sum_{m=\lfloor n^{1-\alpha} \rfloor}^{\infty} e^{-2\lambda(m-1)}. \quad (25)$$

This geometric series can be bounded as follows:

$$2 \max\{q, 1/q\} \frac{e^{-2\lambda(n^{1-\alpha}-2)}}{1 - e^{-2\lambda}} \leq 2 \max\{q, 1/q\} e^{4\lambda} e^{-2\lambda n^{1-\alpha}}. \quad (26)$$

For $n^{1-\alpha} > 2$, we observe that $2n^{1-\alpha} + 1 < e^{2\lambda n^{1-\alpha}}$. Therefore:

$$\sum_{k=-\infty}^{\infty} M_{q,\lambda}(nx - k) \leq e^{2\lambda n^{1-\alpha}} + 2 \max\{q, 1/q\} e^{4\lambda} e^{-2\lambda n^{1-\alpha}}. \quad (27)$$

Simplifying further, we obtain the desired bound:

$$\sum_{k=-\infty}^{\infty} M_{q,\lambda}(nx - k) < 2 \max\{q, 1/q\} e^{4\lambda} e^{-2\lambda n^{1-\alpha}} = T e^{-2\lambda n^{1-\alpha}}, \quad (28)$$

where T is defined in (16).

The same argument applies to $M_{1/q,\lambda}$ due to the symmetry relation:

$$M_{1/q,\lambda}(x) = M_{q,\lambda}(-x). \quad (29)$$

For the symmetrized kernel Φ , we have:

$$\Phi(nx - k) = \frac{M_{q,\lambda}(nx - k) + M_{1/q,\lambda}(nx - k)}{2} < \frac{T e^{-2\lambda n^{1-\alpha}} + T e^{-2\lambda n^{1-\alpha}}}{2} = T e^{-2\lambda n^{1-\alpha}}, \quad (30)$$

which completes the proof. $\square$

Remark 3 (Implications of Exponential Decay). *The exponential decay property established in Theorem 4 has several significant implications:*

1. *Localization Property:* The kernels $M_{q,\lambda}$ and Φ are effectively localized. For large n , the sum $\sum_k M_{q,\lambda}(nx - k)$ is dominated by terms where $|nx - k|$ is small, enabling efficient numerical approximations.
2. *Numerical Stability:* The exponential decay ensures that truncating the infinite sums in practical computations introduces only exponentially small errors, which is crucial for numerical stability.
3. *Convergence Analysis:* This property is fundamental for establishing convergence rates of the associated approximation operators. It allows for precise control of the tail behavior in error estimates, leading to sharp convergence results.
4. *Sparse Representations:* In numerical implementations, the exponential decay enables efficient sparse representations of the kernel sums, significantly reducing computational complexity from $O(n)$ to $O(\log n)$ in many cases.
5. *Error Bounds:* The explicit form of the decay bound (with constant T defined in (16)) provides concrete error bounds that can be used in the analysis of approximation algorithms.

2.5. Multivariate Extension via Tensor Products

To generalize the univariate kernel construction to higher dimensions, we employ the tensor product approach. This preserves the essential properties of the univariate kernels while enabling rigorous multivariate analysis.

Definition 1 (Multivariate Kernel Construction). *Let $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ be the univariate kernel defined in (8). The corresponding multivariate kernel $Z : \mathbb{R}^N \rightarrow \mathbb{R}$ is defined via the tensor product:*

$$Z(x_1, \dots, x_N) := \prod_{i=1}^N \Phi(x_i), \quad x = (x_1, \dots, x_N) \in \mathbb{R}^N, \quad N \in \mathbb{N}. \quad (31)$$

This construction naturally extends the properties of Φ to the multivariate setting.

Theorem 5 (Fundamental Properties of the Multivariate Kernel). *Let Z be the multivariate kernel defined in (31). Then Z satisfies the following fundamental properties. It is positive, that is,*

$$Z(x) > 0, \quad \forall x \in \mathbb{R}^N. \quad (32)$$

It satisfies a discrete partition of unity:

$$\begin{aligned} \sum_{k \in \mathbb{Z}^N} Z(x - k) &= \sum_{k_1 \in \mathbb{Z}} \cdots \sum_{k_N \in \mathbb{Z}} \prod_{i=1}^N \Phi(x_i - k_i) \\ &= \prod_{i=1}^N \left(\sum_{k_i \in \mathbb{Z}} \Phi(x_i - k_i) \right) = 1, \end{aligned} \quad (33)$$

and it is dilation-invariant in the sense that, for any $n \in \mathbb{N}$,

$$\sum_{k \in \mathbb{Z}^N} Z(nx - k) = \prod_{i=1}^N \left(\sum_{k_i \in \mathbb{Z}} \Phi(nx_i - k_i) \right) = 1. \quad (34)$$

Finally, Z is normalized:

$$\int_{\mathbb{R}^N} Z(x) dx = \int_{\mathbb{R}^N} \prod_{i=1}^N \Phi(x_i) dx = \prod_{i=1}^N \left(\int_{\mathbb{R}} \Phi(x_i) dx_i \right) = 1. \quad (35)$$

Proof. (P1) Positivity: Each $\Phi(x_i) > 0$ implies

$$Z(x) = \prod_{i=1}^N \Phi(x_i) > 0.$$

(P2) Discrete Partition of Unity: Follows directly from (33) and the univariate property (11).

(P3) Dilation Invariance: Follows from (34) and the univariate dilation invariance.

(P4) Normalization: Follows from (35) and the univariate normalization (14), using Fubini's theorem. $\square$

Corollary 2 (Exponential Decay of the Multivariate Kernel). *Let Z be the multivariate kernel defined in (31), and define the sup-norm $\|x\|_\infty := \max_{1 \leq i \leq N} |x_i|$. For any $0 < \beta < 1$ and $n \in \mathbb{N}$ such that $n^{1-\beta} > 2$, the kernel Z exhibits exponential decay outside a neighborhood of size $1/n^\beta$:*

$$\sum_{\substack{k \in \mathbb{Z}^N \\ \|k/n - x\|_\infty > 1/n^\beta}} Z(nx - k) \leq T e^{-2\lambda n^{1-\beta}}, \quad (36)$$

where the constant $T := 2 \max\{q, 1/q\} e^{4\lambda}$ is as defined in (16).

This result shows that the contributions of $Z(nx - k)$ are exponentially small for lattice points k lying outside the sup-norm neighborhood of x , reflecting the strong localization of the multivariate kernel.

Proof. Let $S = \{k \in \mathbb{Z}^N : \|k/n - x\|_\infty > 1/n^\beta\}$. For each $k \in S$, there exists at least one index i_0 with $|k_{i_0}/n - x_{i_0}| > 1/n^\beta$. Then:

$$\begin{aligned} \sum_{k \in S} Z(nx - k) &\leq \sum_{k \in S} \Phi(nx_{i_0} - k_{i_0}) \prod_{i \neq i_0} 1 \\ &\leq N \cdot (2n^{1-\beta} + 1)^{N-1} \cdot T e^{-2\lambda n^{1-\beta}} \\ &\leq T e^{-2\lambda n^{1-\beta}}, \end{aligned}$$

absorbing constants into T for sufficiently large n . $\square$

Theorem 6 (Exponential Decay of the Multivariate Kernel). *Let Z be the multivariate kernel defined in (31), and define the sup-norm $\|x\|_\infty := \max_{1 \leq i \leq N} |x_i|$. Then, for any $0 < \beta < 1$ and $n \in \mathbb{N}$ such that $n^{1-\beta} > 2$, the kernel satisfies the exponential decay estimate*

$$\sum_{\substack{k \in \mathbb{Z}^N \\ \|k/n - x\|_\infty > 1/n^\beta}} Z(nx - k) \leq T e^{-2\lambda n^{1-\beta}}, \quad (37)$$

where $T := 2 \max\{q, 1/q\} e^{4\lambda}$ is the constant appearing in the univariate decay bound.

Proof. Let

$$S := \{k \in \mathbb{Z}^N : \|k/n - x\|_\infty > 1/n^\beta\}. \quad (38)$$

For each $k \in S$, there exists at least one index $i_0 \in \{1, \dots, N\}$ such that

$$|k_{i_0}/n - x_{i_0}| > 1/n^\beta. \quad (39)$$

Using the tensor-product structure of Z , we have

$$\sum_{k \in S} Z(nx - k) = \sum_{k \in S} \prod_{i=1}^N \Phi(nx_i - k_i) \quad (40)$$

$$\leq \sum_{k \in S} \Phi(nx_{i_0} - k_{i_0}) \prod_{i \neq i_0} \Phi(nx_i - k_i) \quad (41)$$

$$\leq \sum_{k \in S} T e^{-2\lambda n^{1-\beta}} \prod_{i \neq i_0} 1 \quad (42)$$

$$= \sum_{k \in S} T e^{-2\lambda n^{1-\beta}}.$$

Next, we bound the number of terms in the sum. For each fixed i_0 , the remaining $N - 1$ coordinates can take at most

$$(2n^{1-\beta} + 1)^{N-1} \quad (43)$$

integer values. Accounting for all N choices of i_0 , we obtain

$$\sum_{k \in S} Z(nx - k) \leq N (2n^{1-\beta} + 1)^{N-1} T e^{-2\lambda n^{1-\beta}}. \quad (44)$$

Since $n^{1-\beta} > 2$, the polynomial factor $(2n^{1-\beta} + 1)^{N-1}$ can be absorbed into the exponential by adjusting T if necessary. Therefore, we conclude

$$\sum_{\substack{k \in \mathbb{Z}^N \\ \|k/n - x\|_\infty > 1/n^\beta}} Z(nx - k) \leq T e^{-2\lambda n^{1-\beta}}, \quad (45)$$

proving the claimed exponential decay. $\square$

2.6. Neural Network Operators

Using the multivariate kernel Z , we define several types of neural network operators that serve as primary tools for function approximation in high-dimensional settings.

Let $f \in C^m(\mathbb{R}^N)$, where $m, N \in \mathbb{N}$. For a multi-index $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{Z}_+^N$ with $|\alpha| = \sum_{i=1}^N \alpha_i$, denote the partial derivative of order $|\alpha|$ by

$$f_\alpha(x) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_N^{\alpha_N}}(x), \quad x \in \mathbb{R}^N. \quad (46)$$

The maximum norm of derivatives of order m is defined as

$$\|f_\alpha\|_{\infty,m}^{\max} = \max_{|\alpha|=m} \|f_\alpha\|_\infty. \quad (47)$$

Let $C_B(\mathbb{R}^N)$ denote the space of continuous and bounded functions on $\mathbb{R}^N$.

Definition 2 (Neural Network Operators). Let $f \in C_B(\mathbb{R}^N)$ be a continuous and bounded function, $n \in \mathbb{N}$, and $x \in \mathbb{R}^N$. Using the multivariate kernel Z , we define the following classes of neural network operators, which generalize classical approximation operators to the multivariate setting:

1. Quasi-interpolation operator:

The quasi-interpolation operator A_n approximates f by evaluating it at the scaled integer lattice points k/n and weighting these evaluations by the kernel:

$$A_n(f, x) = \sum_{k \in \mathbb{Z}^N} f\left(\frac{k}{n}\right) Z(nx - k). \quad (48)$$

This operator is simple and computationally efficient, relying only on pointwise values of f . Its accuracy is controlled by the smoothness of f and the decay properties of Z .

2. Kantorovich-type operator:

The Kantorovich operator K_n generalizes A_n by replacing pointwise evaluations with local integrals over small hypercubes $[k/n, (k+1)/n]$:

$$K_n(f, x) = \sum_{k \in \mathbb{Z}^N} \left(n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t) dt \right) Z(nx - k). \quad (49)$$

This construction improves approximation properties for functions that are less regular or only integrable, and it preserves linear functionals, making it suitable for analysis in L^p spaces.

3. Quadrature-type operator:

To further enhance flexibility and numerical implementation, one can define local quadrature approximations of the integrals in K_n . Let

$$\delta_{n,k}(f) = \sum_{r=0}^{\theta} w_r f\left(\frac{k}{n} + \frac{r}{n\theta}\right), \quad (50)$$

where $\theta = (\theta_1, \dots, \theta_N) \in \mathbb{N}^N$ specifies the number of quadrature points in each dimension, $r = (r_1, \dots, r_N) \in \mathbb{Z}_+^N$ indexes the points, and $w_r = w_{r_1 \dots r_N} \geq 0$ are the corresponding quadrature weights satisfying $\sum_{r=0}^{\theta} w_r = 1$. The quadrature-type operator is then defined as

$$Q_n(f, x) = \sum_{k \in \mathbb{Z}^N} \delta_{n,k}(f) Z(nx - k). \quad (51)$$

This operator interpolates between pointwise and integral-based approximations, providing a practical scheme for high-dimensional problems and allowing for flexible choice of quadrature rules.

Remarks:

1. All three operators rely on the tensor-product kernel Z , which satisfies positivity, partition of unity, and exponential decay properties.
2. A_n is computationally simplest, K_n is more robust for rough functions, and Q_n allows numerical integration with controlled accuracy.
3. These operators provide the foundation for convergence analysis in pointwise, uniform, and L^p senses, as shown in subsequent theorems.

Theorem 7 (Convergence of Multivariate Neural Network Operators). *Let $f \in C^m(\mathbb{R}^N)$ with bounded derivatives up to order m , and let A_n, K_n, Q_n be the operators defined in (48)–(51). Then, for each $x \in \mathbb{R}^N$, we have the following convergence results:*

$$\lim_{n \rightarrow \infty} A_n(f, x) = f(x), \quad (52)$$

$$\lim_{n \rightarrow \infty} K_n(f, x) = f(x), \quad (53)$$

$$\lim_{n \rightarrow \infty} Q_n(f, x) = f(x), \quad (54)$$

with the quantitative error estimates

$$|A_n(f, x) - f(x)| \leq C n^{-m} \|f_\alpha\|_{\infty, m}^{\max}, \quad (55)$$

$$|K_n(f, x) - f(x)| \leq C n^{-m} \|f_\alpha\|_{\infty, m}^{\max}, \quad (56)$$

$$|Q_n(f, x) - f(x)| \leq C n^{-m} \|f_\alpha\|_{\infty, m}^{\max}, \quad (57)$$

where $C > 0$ is a constant depending only on the kernel Z and the dimension N .

Proof. The proof follows standard arguments using Taylor expansions with integral remainders and the tensor-product structure of Z .

For $A_n(f, x)$, write

$$f\left(\frac{k}{n}\right) - f(x) = \sum_{|\alpha| < m} \frac{f_\alpha(x)}{\alpha!} \left(\frac{k}{n} - x\right)^\alpha + R_m\left(f; x, \frac{k}{n}\right), \quad (58)$$

where R_m is the remainder term. Summing against $Z(nx - k)$ and using the moment properties of the kernel yields

$$|A_n(f, x) - f(x)| \leq C n^{-m} \|f_\alpha\|_{\infty, m}^{\max}. \quad (59)$$

For $K_n(f, x)$, the Kantorovich integral averages satisfy

$$n^N \int_{\frac{k}{n}}^{\frac{k+1}{n}} (f(t) - f(x)) dt = \sum_{|\alpha| < m} \frac{f_\alpha(x)}{\alpha!} \frac{1}{n^{|\alpha|}} \sum_{|\beta|=|\alpha|} c_\beta + R_m, \quad (60)$$

where R_m is again bounded by $C n^{-m} \|f_\alpha\|_{\infty, m}^{\max}$.

Similarly, for $Q_n(f, x)$, the local quadrature averages approximate the integral to order n^{-m} under standard assumptions on the weights w_r .

Combining these estimates for all operators concludes the proof. $\square$

Theorem 8 (Fundamental Properties of Neural Network Operators). *Let $f, g \in C_B(\mathbb{R}^N)$ and $a, b \in \mathbb{R}$. For all $n \in \mathbb{N}$ and $x \in \mathbb{R}^N$, each operator $T_n \in \{A_n, K_n, Q_n\}$ defined in (48)–(51) satisfies the following properties:*

The operators are linear, i.e.,

$$T_n(af + bg, x) = a T_n(f, x) + b T_n(g, x). \quad (61)$$

They are positive: if $f(x) \geq 0$ for all $x \in \mathbb{R}^N$, then

$$T_n(f, x) \geq 0, \quad \forall x \in \mathbb{R}^N. \quad (62)$$

Moreover, they reproduce constants: if $f(x) \equiv c \in \mathbb{R}$, then

$$T_n(f, x) = c, \quad \forall x \in \mathbb{R}^N. \quad (63)$$

Proof. Linearity: By definition, each operator is a linear combination of function evaluations or integrals. For A_n :

$$\begin{aligned} A_n(af + bg, x) &= \sum_{k \in \mathbb{Z}^N} (af + bg)\left(\frac{k}{n}\right) Z(nx - k) \\ &= a \sum_{k \in \mathbb{Z}^N} f\left(\frac{k}{n}\right) Z(nx - k) + b \sum_{k \in \mathbb{Z}^N} g\left(\frac{k}{n}\right) Z(nx - k) \\ &= a A_n(f, x) + b A_n(g, x). \end{aligned} \quad (64)$$

Analogous arguments hold for K_n and Q_n using linearity of integrals and sums.

Positivity: Since $Z(nx - k) > 0$ and quadrature weights $w_r \geq 0$, we have

$$T_n(f, x) = \sum_{k \in \mathbb{Z}^N} (\text{non-negative terms}) \geq 0$$

for any $f \geq 0$.

Reproduction of Constants: If $f \equiv c$, then

$$A_n(f, x) = \sum_{k \in \mathbb{Z}^N} c Z(nx - k) = c \sum_{k \in \mathbb{Z}^N} Z(nx - k) = c, \quad (65)$$

$$K_n(f, x) = \sum_{k \in \mathbb{Z}^N} \left(n^N \int_{k/n}^{(k+1)/n} c dt \right) Z(nx - k) = c, \quad (66)$$

$$Q_n(f, x) = \sum_{k \in \mathbb{Z}^N} \left(\sum_r w_r c \right) Z(nx - k) = c, \quad (67)$$

where we used the partition of unity property of Z : $\sum_{k \in \mathbb{Z}^N} Z(nx - k) = 1$. $\square$

Remark 4 (Approximation Order). *The operators defined above exhibit specific approximation properties:*

- If $f \in C^m(\mathbb{R}^N)$ with bounded derivatives, then by a standard Taylor expansion of f about the grid nodes k/n , we obtain:

$$A_n(f, x) = f(x) + O\left(\frac{1}{n}\right), \quad K_n(f, x) = f(x) + O\left(\frac{1}{n}\right),$$

and similarly for $Q_n(f, x)$, depending on the quadrature rule.

- Higher-order moment conditions on Φ or higher-degree quadrature rules can improve this rate, leading to asymptotic Voronovskaya-type expansions as discussed in later sections.

Remark 5 (Interpretation of Kantorovich Operator). *The integral in (49) can also be expressed as:*

$$\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t) dt = \int_0^{\frac{1}{n}} f\left(t + \frac{k}{n}\right) dt, \quad (68)$$

indicating that $K_n(f, x)$ represents a shifted average of f over cubes of side length $1/n$ centered at k/n .

3. Main Results

In this section, we rigorously analyze the approximation properties of the multivariate neural network operators A_n , K_n , and Q_n by deriving Voronovskaya-type asymptotic expansions. Our approach relies on refined multivariate Taylor expansions and precise estimates of the remainder terms.

3.1. Voronovskaya-Type Expansion for Basic Operators

Theorem 9 (Voronovskaya-Type Asymptotic Expansion). *Let $0 < \beta < 1$, $n \in \mathbb{N}$ sufficiently large, $x \in \mathbb{R}^N$, and $f \in C^m(\mathbb{R}^N)$, where $m, N \in \mathbb{N}$. Assume that all partial derivatives of order m are bounded, i.e., $f_\alpha \in C_B(\mathbb{R}^N)$ for all multi-indices $\alpha = (\alpha_1, \dots, \alpha_N)$ with $|\alpha| = \sum_{i=1}^N \alpha_i = m$. Let $0 < \varepsilon \leq m$. Then, for the neural network operator A_n we have*

$$A_n(f, x) - f(x) = \sum_{j=1}^m \sum_{|\alpha|=j} \frac{f_\alpha(x)}{\prod_{i=1}^N \alpha_i!} A_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) + o \left(\frac{1}{n^{\beta(m-\varepsilon)}} \right), \quad (69)$$

which equivalently implies

$$n^{\beta(m-\varepsilon)} \left[A_n(f, x) - f(x) - \sum_{j=1}^m \sum_{|\alpha|=j} \frac{f_\alpha(x)}{\prod_{i=1}^N \alpha_i!} A_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) \right] \rightarrow 0, \quad n \rightarrow \infty. \quad (70)$$

Furthermore, if $f_\alpha(x) = 0$ for all $|\alpha| = 1, \dots, m$, then

$$n^{\beta(m-\varepsilon)} |A_n(f, x) - f(x)| \rightarrow 0, \quad n \rightarrow \infty. \quad (71)$$

Proof. We begin by defining, for $x_0, z \in \mathbb{R}^N$, the univariate path function

$$g_z(t) := f(x_0 + t(z - x_0)), \quad t \in [0, 1].$$

By the chain rule, its j -th derivative satisfies

$$g_z^{(j)}(t) = \left[\sum_{i=1}^N (z_i - x_{0i}) \partial_{x_i} \right]^j f(x_0 + t(z - x_0)), \quad j = 0, 1, \dots, m.$$

Using the standard multivariate Taylor formula with integral remainder:

$$f(z) = \sum_{j=0}^m \frac{g_z^{(j)}(0)}{j!} + \frac{1}{(m-1)!} \int_0^1 (1-\theta)^{m-1} (g_z^{(m)}(\theta) - g_z^{(m)}(0)) d\theta, \quad (72)$$

where

$$g_z^{(j)}(0) = \sum_{|\alpha|=j} \frac{j!}{\prod_{i=1}^N \alpha_i!} \prod_{i=1}^N (z_i - x_{0i})^{\alpha_i} f_\alpha(x_0), \quad j = 0, 1, \dots, m. \quad (73)$$

Applying the operator A_n and setting $z = k/n$, $k \in \mathbb{Z}^N$, we obtain

$$A_n(f, x) - f(x) = \sum_{k \in \mathbb{Z}^N} f\left(\frac{k}{n}\right) Z(nx - k) - f(x) = \sum_{|\alpha| \leq m} \frac{f_\alpha(x)}{\prod_i \alpha_i!} A_n \left(\prod_i \left(\frac{k_i}{n} - x_i\right)^{\alpha_i} \right) + \theta_n,$$

where the remainder term θ_n satisfies the estimate

$$|\theta_n| \leq \frac{2 \|f_\alpha\|_{\infty, m}^{\max} N^m}{m! n^{m\beta}} = O(n^{-m\beta}) = o(1). \quad (74)$$

Consequently, for $0 < \varepsilon \leq m$,

$$|\theta_n| = o(n^{-\beta(m-\varepsilon)}),$$

yielding the desired asymptotic expansion (69)-(70). The case $f_\alpha = 0$ directly implies (90). $\square$

Remark 6. The above theorem rigorously establishes the rate of convergence for $A_n(f, x)$ in terms of the smoothness m of f , the dimensionality N , and the parameter β . The remainder estimate is uniform with respect

to $x \in \mathbb{R}^N$ under the bounded derivative condition, providing a clear path toward the derivation of analogous expansions for Kantorovich-type operators.

3.2. Preparatory Observations

For any multi-index α with $|\alpha| = m$, we adopt the standard combinatorial notation

$$\frac{m!}{\prod_{i=1}^N \alpha_i!}, \quad (75)$$

and define the remainder integrand in the multivariate Taylor expansion as

$$\prod_{i=1}^N (z_i - x_{0i})^{\alpha_i} f_{\alpha}(x_0 + \theta(z - x_0)), \quad 0 \leq \theta \leq 1, \quad (76)$$

which will play a central role in estimating the remainder term.

Under the uniform grid assumption

$$\left| \frac{k_i}{n} - x_i \right| \leq n^{-\beta}, \quad i = 1, \dots, N, \quad (77)$$

we can bound the remainder R as

$$|R| \leq \frac{2 \|f_{\alpha}\|_{\infty, m}^{\max} N^m}{m!} n^{-m\beta}. \quad (78)$$

Moreover, summing over all nodes with the weight function $Z(nx - k)$ yields the global estimate

$$|\theta_n| := \left| \sum_{k \in \mathbb{Z}^N} R Z(nx - k) \right| = O(n^{-m\beta}) = o(n^{-\beta(m-\epsilon)}), \quad n \rightarrow \infty. \quad (79)$$

These estimates complete the proof of Theorem 9 and provide a solid foundation for deriving analogous Voronovskaya-type expansions for Kantorovich-type operators.

4. Voronovskaya-Type Expansions for Kantorovich and Quasi-Interpolation Operators

We now extend the previous results to Kantorovich-type operators K_n and quasi-interpolation neural network operators Q_n . Our goal is to obtain refined multivariate asymptotic expansions, including higher-order remainder estimates.

Theorem 10 (Refined Multivariate Voronovskaya Expansion). *Let $f \in C^{m+2}(\mathbb{R}^N)$ with bounded derivatives up to order $m+2$, and let $0 < \beta < 1$. For $n \in \mathbb{N}$ sufficiently large and $x \in \mathbb{R}^N$, the following expansion holds for the Kantorovich operator K_n :*

$$\begin{aligned} K_n(f, x) - f(x) &= \sum_{j=1}^m \sum_{|\alpha|=j} \frac{f_{\alpha}(x)}{\prod_{i=1}^N \alpha_i!} K_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) \\ &\quad + \sum_{|\alpha|=m+1} \frac{f_{\alpha}(x)}{\prod_{i=1}^N \alpha_i!} K_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) + R_n(x), \end{aligned} \quad (80)$$

where the remainder $R_n(x)$ satisfies the uniform estimate

$$|R_n(x)| \leq C \|f\|_{\infty, m+2}^{\max} \frac{N^{m+2}}{(m+1)!} n^{-(m+1)\beta}, \quad (81)$$

with a constant $C > 0$ independent of x and n . The same expansion holds for Q_n with the corresponding quasi-interpolation weights.

Proof. Define the path function for $x_0, z \in \mathbb{R}^N$:

$$g_z(t) := f(x_0 + t(z - x_0)), \quad t \in [0, 1]. \quad (82)$$

By the chain rule, its derivatives are

$$g_z^{(j)}(t) = \left[\sum_{i=1}^N (z_i - x_{0i}) \partial_{x_i} \right]^j f(x_0 + t(z - x_0)), \quad j = 0, 1, \dots, m+1. \quad (83)$$

Applying the Taylor expansion with integral remainder gives

$$f(z) = \sum_{j=0}^m \frac{g_z^{(j)}(0)}{j!} + \frac{1}{m!} \int_0^1 (1-\theta)^m g_z^{(m+1)}(\theta) d\theta. \quad (84)$$

For a grid point $z = k/n$, we can write

$$f\left(\frac{k}{n}\right) - f(x) = \sum_{j=1}^{m+1} \sum_{|\alpha|=j} \frac{f_\alpha(x)}{\prod_i \alpha_i!} \prod_i \left(\frac{k_i}{n} - x_i\right)^{\alpha_i} + R_{k,n}(x), \quad (85)$$

with the remainder term

$$R_{k,n}(x) = \frac{1}{m!} \int_0^1 (1-\theta)^m \sum_{|\alpha|=m+1} \frac{(m+1)!}{\prod_i \alpha_i!} \prod_i \left(\frac{k_i}{n} - x_i\right)^{\alpha_i} f_\alpha(x + \theta(k/n - x)) d\theta. \quad (86)$$

Using the grid uniformity assumption

$$\left| \frac{k_i}{n} - x_i \right| \leq n^{-\beta}, \quad i = 1, \dots, N, \quad (87)$$

we obtain the bound

$$|R_{k,n}(x)| \leq \frac{\|f\|_{\infty, m+2}^{\max} N^{m+2}}{(m+1)!} n^{-(m+1)\beta}. \quad (88)$$

Finally, applying the operator K_n (or Q_n) and summing over all grid points yields the uniform remainder estimate $R_n(x)$ in (176), completing the proof. $\square$

Remark 7 (Uniform Convergence and Higher-Order Terms). *The above expansion shows that for sufficiently smooth f , the Kantorovich and quasi-interpolation neural network operators achieve the same asymptotic behavior as classical Voronovskaya expansions, with a fully explicit bound for the remainder term. This allows rigorous estimates of the rate of convergence in terms of n , β , N , and the smoothness $m+2$ of f .*

4.1. Implications for Approximation Rates

Let $f \in C^m(\mathbb{R}^N)$ and $0 < \varepsilon \leq m$. From the Voronovskaya-type expansion (85)–(88), we obtain the asymptotic estimate

$$n^{\beta(m-\varepsilon)} |K_n(f, x) - f(x)| = O(n^{-\beta\varepsilon}), \quad n \rightarrow \infty, \quad (89)$$

which provides the precise rate of convergence in terms of the grid parameter β and the smoothness index m .

Furthermore, if all partial derivatives of order $|\alpha| \leq m$ vanish at x , i.e.,

$$f_\alpha(x) = 0, \quad \forall |\alpha| \leq m, \quad (90)$$

then the remainder term dominates the expansion, yielding the higher-order estimate

$$|K_n(f, x) - f(x)| = O(n^{-(m+1)\beta}), \quad n \rightarrow \infty, \quad (91)$$

which explicitly illustrates the gain in the convergence rate due to higher-order smoothness.

These estimates highlight two important aspects:

1. The general rate (89) depends explicitly on the balance between the smoothness m and the chosen parameter ε , giving a controlled decay of the approximation error.
2. The higher-order vanishing condition (90) demonstrates that additional smoothness beyond order m can further accelerate convergence, as shown in (91).

5. Second-Order Multivariate Voronovskaya Expansions and L^p Estimates

We refine the previous results by including second-order multivariate terms and deriving convergence estimates in L^p -norms. This allows a precise comparison among the neural network operators A_n , K_n , and Q_n .

Theorem 11 (Second-Order Multivariate Voronovskaya Expansion). *Let $f \in C^{m+2}(\mathbb{R}^N)$ with bounded derivatives up to order $m+2$, $1 \leq p \leq \infty$, and $0 < \beta < 1$. For sufficiently large $n \in \mathbb{N}$ and $x \in \mathbb{R}^N$, the following expansion holds for the Kantorovich-type operator K_n :*

$$\begin{aligned} K_n(f, x) - f(x) &= \sum_{j=1}^m \sum_{|\alpha|=j} \frac{f_\alpha(x)}{\prod_{i=1}^N \alpha_i!} K_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) \\ &\quad + \sum_{|\alpha|=m+1} \frac{f_\alpha(x)}{\prod_{i=1}^N \alpha_i!} K_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) \\ &\quad + \frac{1}{2} \sum_{|\alpha|=2} \sum_{i,j=1}^N f_{\alpha_i \alpha_j}(x) K_n \left((\cdot - x_i)(\cdot - x_j) \right) (x) + R_n(x), \end{aligned} \quad (92)$$

where the remainder $R_n(x)$ satisfies the uniform and L^p estimate:

$$\|R_n\|_p \leq C \|f\|_{\infty, m+2}^{\max} \frac{N^{m+2}}{(m+1)!} n^{-(m+1)\beta}. \quad (93)$$

The same expansion holds for quasi-interpolation operators Q_n with corresponding weights.

Proof. We extend the path function technique: for $x_0, z \in \mathbb{R}^N$, define

$$g_z(t) := f(x_0 + t(z - x_0)), \quad t \in [0, 1], \quad (94)$$

with derivatives

$$g_z^{(j)}(t) = \left[\sum_{i=1}^N (z_i - x_{0i}) \partial_{x_i} \right]^j f(x_0 + t(z - x_0)), \quad j = 0, \dots, m+2. \quad (95)$$

Applying the multivariate Taylor expansion with integral remainder gives

$$f(z) = \sum_{j=0}^m \frac{g_z^{(j)}(0)}{j!} + \frac{1}{m!} \int_0^1 (1-\theta)^m g_z^{(m+1)}(\theta) d\theta. \quad (96)$$

Decomposing the second-order terms explicitly for cross derivatives yields

$$\frac{1}{2} \sum_{i,j=1}^N f_{x_i x_j}(x) (z_i - x_i)(z_j - x_j). \quad (97)$$

For a uniform grid $z = k/n$ with $|k_i/n - x_i| \leq n^{-\beta}$, the remainder satisfies

$$|R_{k,n}(x)| \leq \frac{\|f\|_{\infty, m+2}^{\max} N^{m+2}}{(m+1)!} n^{-(m+1)\beta}. \quad (98)$$

Finally, summing over the grid points with Kantorovich or quasi-interpolation weights, we obtain the uniform and L^p bound

$$\left\| \sum_k w_k R_{k,n} \right\|_{L^p} = O(n^{-(m+1)\beta}), \quad (99)$$

which completes the proof. $\square$

Corollary 3 (Convergence Rate in L^p Norm). *Under the hypotheses of Theorem 11, for $0 < \varepsilon \leq m$, the following L^p convergence rates hold:*

$$\|K_n(f) - f\|_{L^p} = O(n^{-\beta\varepsilon}), \quad n \rightarrow \infty, \quad (100)$$

$$\|Q_n(f) - f\|_{L^p} = O(n^{-\beta\varepsilon}), \quad n \rightarrow \infty. \quad (101)$$

Moreover, if all derivatives vanish up to order m , i.e., $f_\alpha(x) = 0$ for $|\alpha| \leq m$, the remainder dominates and we obtain the higher-order estimate:

$$\|K_n(f) - f\|_{L^p} = O(n^{-(m+1)\beta}), \quad n \rightarrow \infty, \quad (102)$$

$$\|Q_n(f) - f\|_{L^p} = O(n^{-(m+1)\beta}), \quad n \rightarrow \infty, \quad (103)$$

highlighting the gain due to additional smoothness of f .

Remark 8 (Comparison of Operators). *The asymptotic expansions show that A_n , K_n , and Q_n share the same principal term up to order m , with differences manifesting only in the remainder term. Kantorovich-type operators typically provide better L^p -stability due to integration over the cells, whereas quasi-interpolation operators preserve pointwise accuracy and allow explicit control over cross-derivative contributions.*

Theorem 12 (Voronovskaya-type for Multivariate Kantorovich Operators). *Let $f \in C_b^m(\mathbb{R}^N)$ and K_n be the multivariate Kantorovich operator as in Theorem 6. Then, for $x \in \mathbb{R}^N$, we have*

$$\begin{aligned} K_n(f, x) - f(x) &= \sum_{j=1}^m \sum_{\substack{\alpha \in \mathbb{Z}_+^N \\ |\alpha|=j}} \frac{1}{\alpha!} f_\alpha(x) K_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) \\ &\quad + o \left(\left(\frac{1}{n} + \frac{1}{n^\beta} \right)^{m-\varepsilon} \right), \end{aligned} \quad (104)$$

as $n \rightarrow \infty$, for any $0 < \varepsilon \leq m$, where $\alpha! := \prod_{i=1}^N \alpha_i!$ and f_α denotes the partial derivative of order α .

Proof. Using the multivariate Taylor expansion with integral remainder, for $t \in [0, 1/n]^N$ and $k \in \mathbb{Z}^N$, we write

$$f \left(t + \frac{k}{n} \right) - f(x) - \sum_{j=1}^m \sum_{|\alpha|=j} \frac{1}{\alpha!} f_\alpha(x) \prod_{i=1}^N \left(t_i + \frac{k_i}{n} - x_i \right)^{\alpha_i} = R_{n,k}(t), \quad (105)$$

with the remainder

$$R_{n,k}(t) := m \int_0^1 (1-\theta)^{m-1} \sum_{|\alpha|=m} \frac{1}{\alpha!} \prod_{i=1}^N \left(t_i + \frac{k_i}{n} - x_i\right)^{\alpha_i} \left[f_\alpha(x + \theta(t + \frac{k}{n} - x)) - f_\alpha(x)\right] d\theta. \quad (106)$$

Using the supremum norm of the m -th derivatives, we obtain the estimate

$$|R_{n,k}(t)| \leq \frac{2\|f_\alpha\|_{\infty,m}^{\max} N^m}{m!} \prod_{i=1}^N \left(|t_i| + \left|\frac{k_i}{n} - x_i\right|\right)^{\alpha_i}. \quad (107)$$

Integrating over $[0, 1/n]^N$ and summing over k , we define

$$U_n^*(x) := \sum_{k \in \mathbb{Z}^N} n^N \int_{[0,1/n]^N} R_{n,k}(t) dt Z(nx - k), \quad (108)$$

which represents the total remainder contribution of $K_n(f, x)$. By the above estimate, we have

$$|U_n^*(x)| \leq \frac{2\|f_\alpha\|_{\infty,m}^{\max} N^m}{m!} \left(\frac{1}{n} + \frac{1}{n^\beta}\right)^m, \quad (109)$$

implying

$$|U_n^*(x)| = O\left(\left(\frac{1}{n} + \frac{1}{n^\beta}\right)^m\right) = o(1), \quad (110)$$

and for any $0 < \varepsilon \leq m$,

$$\frac{|U_n^*(x)|}{\left(\frac{1}{n} + \frac{1}{n^\beta}\right)^{m-\varepsilon}} \rightarrow 0, \quad n \rightarrow \infty. \quad (111)$$

The result follows by splitting the operator $K_n(f, x)$ into the main polynomial part and the remainder $U_n^*(x)$. $\square$

Theorem 13 (Voronovskaya-type Expansion for Quadrature Operators). *Let Q_n be the multivariate quadrature-type operator as in Theorem 6, and let $f \in C_b^m(\mathbb{R}^N)$. Then, for each $x \in \mathbb{R}^N$,*

$$Q_n(f, x) - f(x) = \sum_{j=1}^m \sum_{|\alpha|=j} \frac{1}{\alpha!} f_\alpha(x) Q_n\left(\prod_{i=1}^N (\cdot_i - x_i)^{\alpha_i}\right)(x) + o\left(\left(\frac{1}{n} + \frac{1}{n^\beta}\right)^{m-\varepsilon}\right), \quad (112)$$

as $n \rightarrow \infty$, for any $0 < \varepsilon \leq m$.

Proof. We proceed in a series of detailed:

For a fixed quadrature node (k, r) , we expand f around x :

$$f\left(\frac{k}{n} + \frac{r}{n\theta}\right) = f(x) + \sum_{j=1}^m \sum_{|\alpha|=j} \frac{1}{\alpha!} f_\alpha(x) \prod_{i=1}^N \left(\frac{k_i}{n} + \frac{r_i}{n\theta_i} - x_i\right)^{\alpha_i} + R_{n,k,r}, \quad (113)$$

where the remainder $R_{n,k,r}$ admits the integral form

$$R_{n,k,r} = m \int_0^1 (1-\theta)^{m-1} \sum_{|\alpha|=m} \frac{1}{\alpha!} \prod_{i=1}^N \left(\frac{k_i}{n} + \frac{r_i}{n\theta_i} - x_i\right)^{\alpha_i} \left[f_\alpha(x + \theta(\frac{k}{n} + \frac{r}{n\theta} - x)) - f_\alpha(x)\right] d\theta. \quad (114)$$

Using the uniform boundedness of derivatives f_α on $\mathbb{R}^N$, we get

$$|R_{n,k,r}| \leq \frac{2N^m}{m!} \|f_\alpha\|_{\infty,m}^{\max} \prod_{i=1}^N \left|\frac{k_i}{n} + \frac{r_i}{n\theta_i} - x_i\right|^{\alpha_i}. \quad (115)$$

The operator Q_n involves weighted sums over the quadrature points r with weights w_r . Thus, for each k , we have

$$\left| \sum_r w_r R_{n,k,r} \right| \leq \frac{2N^m}{m!} \|f_\alpha\|_{\infty,m}^{\max} \left(\frac{1}{n} + \left\| \frac{k}{n} - x \right\|_\infty \right)^m. \quad (116)$$

Next, summing over $k \in \mathbb{Z}^N$ with the kernel $Z(nx - k)$, which decays sufficiently fast, yields

$$E_n^*(x) := \sum_{k \in \mathbb{Z}^N} \sum_r w_r R_{n,k,r} Z(nx - k) \leq \frac{2N^m}{m!} \|f_\alpha\|_{\infty,m}^{\max} \sum_{k \in \mathbb{Z}^N} Z(nx - k) \left(\frac{1}{n} + \left\| \frac{k}{n} - x \right\|_\infty \right)^m. \quad (117)$$

Using the decay property of Z and standard estimates for lattice sums, we get

$$|E_n^*(x)| = O\left(\left(\frac{1}{n} + \frac{1}{n^\beta}\right)^m\right) = o(1), \quad n \rightarrow \infty. \quad (118)$$

The parameter β reflects the trade-off between kernel truncation error and Taylor remainder. Choosing $\beta = 1/2$ ensures that

$$\frac{|E_n^*(x)|}{\left(\frac{1}{n} + \frac{1}{n^\beta}\right)^{m-\varepsilon}} \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad (119)$$

for any $0 < \varepsilon \leq m$, providing an optimal convergence rate for the remainder.

Combining the main term from the Taylor expansion with the remainder estimate gives

$$Q_n(f, x) = f(x) + \sum_{j=1}^m \sum_{|\alpha|=j} \frac{1}{\alpha!} f_\alpha(x) Q_n\left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i}\right)(x) + E_n^*(x), \quad (120)$$

which establishes the claimed Voronovskaya-type expansion. $\square$

6. Voronovskaya-Type Expansion in Sobolev $W^{s,p}$

Theorem 14 (Voronovskaya Expansion in $W^{s,p}$). *Let $m \in \mathbb{N}$, $0 \leq s \leq m$, and $1 \leq p < \infty$. Assume $f \in W^{m,p}(\mathbb{R}^N)$ and that the moments*

$$M_\alpha := \int_{\mathbb{R}^N} u^\alpha Z(u) du \quad (121)$$

exist for all multi-indices $|\alpha| \leq m$. Then, for the operator A_n ,

$$A_n f(x) - f(x) = \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) + r_{n,s}(x), \quad (122)$$

where $r_{n,s} \in W^{s,p}(\mathbb{R}^N)$ and

$$\|r_{n,s}\|_{W^{s,p}} = o(n^{-(m-s)}), \quad n \rightarrow \infty. \quad (123)$$

Proof. For each lattice point $k \in \mathbb{Z}^N$, the multivariate Taylor expansion with integral remainder gives

$$f\left(\frac{k}{n}\right) = \sum_{|\alpha| \leq m-s} \frac{\partial^\alpha f(x)}{\alpha!} \left(\frac{k}{n} - x\right)^\alpha + R_{n,m}(x, k), \quad (124)$$

where the remainder term admits the integral representation

$$R_{n,m}(x, k) = \sum_{|\alpha|=m-s+1}^m \frac{m!}{\alpha!(m-|\alpha|)!} \int_0^1 (1-t)^{m-|\alpha|} \partial^\alpha f\left(x + t\left(\frac{k}{n} - x\right)\right) \left(\frac{k}{n} - x\right)^\alpha dt. \quad (125)$$

By linearity of A_n and the definition

$$A_n f(x) = \sum_{k \in \mathbb{Z}^N} f\left(\frac{k}{n}\right) Z(nx - k), \quad (126)$$

we have

$$A_n f(x) - f(x) = \sum_{1 \leq |\alpha| \leq m-s} \frac{\partial^\alpha f(x)}{\alpha!} \sum_{k \in \mathbb{Z}^N} \left(\frac{k}{n} - x\right)^\alpha Z(nx - k) + \sum_{k \in \mathbb{Z}^N} R_{n,m}(x, k) Z(nx - k). \quad (127)$$

Using the definition of the moments M_α and the scaling properties of Z , we obtain

$$\sum_{k \in \mathbb{Z}^N} \left(\frac{k}{n} - x\right)^\alpha Z(nx - k) = \frac{M_\alpha}{n^{|\alpha|}} + O(n^{-|\alpha|-1}), \quad (128)$$

recovering the main term of the expansion:

$$\sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha}{\alpha! n^{|\alpha|}} \partial^\alpha f(x). \quad (129)$$

The remainder can be expressed as a discrete convolution:

$$r_{n,s}(x) = \sum_{k \in \mathbb{Z}^N} R_{n,m}(x, k) Z(nx - k). \quad (130)$$

Differentiating under the summation (allowed since Z is smooth and compactly supported) and applying Minkowski's inequality, we get

$$\|\partial^\beta r_{n,s}\|_{L^p} \leq \sum_{|\alpha|=m-s+1}^m \frac{C_\alpha}{\alpha!} \sum_{k \in \mathbb{Z}^N} \int_0^1 \|\partial^\alpha f(x + t(\frac{k}{n} - x))\|_{L^p} \left\| \frac{k}{n} - x \right\|^{|\alpha|} Z(nx - k) dt, \quad (131)$$

for all multi-indices $|\beta| \leq s$.

Passing to the Riemann sum and applying Young's inequality for discrete convolutions yields

$$\|\partial^\beta r_{n,s}\|_{L^p} \leq C n^{-(m-s)} \|\partial^m f\|_{L^p}. \quad (132)$$

Hence, the remainder satisfies (123), and combining (129) with (130) completes the proof. $\square$

7. Quantitative L^p Rate with Explicit Constants

Theorem 15 (Quantitative L^p Approximation Rate). *Let $f \in W^{m,p}(\mathbb{R}^N)$, $1 \leq p \leq \infty$, and assume the kernel moments*

$$M_\alpha := \int_{\mathbb{R}^N} u^\alpha Z(u) du \quad (133)$$

exist and are finite for all multi-indices $|\alpha| \leq m$. Then there exists a constant

$$C := C(N, m, p, Z) := \sum_{|\gamma|=m} \frac{1}{\gamma!} \int_{\mathbb{R}^N} |u|^{|\gamma|} |Z(u)| du \quad (134)$$

such that

$$\left\| A_n f - f - \sum_{1 \leq |\alpha| \leq m-1} \frac{M_\alpha}{\alpha! n^{|\alpha|}} \partial^\alpha f \right\|_{L^p} \leq \frac{C \|f\|_{W^{m,p}}}{n^m}. \quad (135)$$

Proof. For each lattice point $k \in \mathbb{Z}^N$, the multivariate Taylor expansion with integral remainder yields

$$f\left(\frac{k}{n}\right) = f(x) + \sum_{1 \leq |\alpha| \leq m-1} \frac{\partial^\alpha f(x)}{\alpha!} \left(\frac{k}{n} - x\right)^\alpha + R_m(x, k), \quad (136)$$

with the remainder

$$R_m(x, k) = \sum_{|\gamma|=m} \frac{m}{\gamma!} \int_0^1 (1-t)^{m-1} \partial^\gamma f\left(x + t\left(\frac{k}{n} - x\right)\right) \left(\frac{k}{n} - x\right)^\gamma dt. \quad (137)$$

By linearity of A_n ,

$$A_n f(x) - f(x) - \sum_{1 \leq |\alpha| \leq m-1} \frac{M_\alpha}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) = \sum_{k \in \mathbb{Z}^N} R_m(x, k) Z(nx - k) + E_n(x), \quad (138)$$

where

$$E_n(x) = \sum_{1 \leq |\alpha| \leq m-1} \frac{\partial^\alpha f(x)}{\alpha!} \left[\sum_{k \in \mathbb{Z}^N} \left(\frac{k}{n} - x\right)^\alpha Z(nx - k) - \frac{M_\alpha}{n^{|\alpha|}} \right]. \quad (139)$$

Since Z is rapidly decaying, one has $E_n(x) = O(n^{-m})$.

Change variables $u = nx - k$, so $\frac{k}{n} - x = -\frac{u}{n}$. Then

$$\sum_{k \in \mathbb{Z}^N} |R_m(x, k)| |Z(nx - k)| \leq \sum_{|\gamma|=m} \frac{1}{\gamma!} \sum_{u \in \mathbb{Z}^N} \int_0^1 (1-t)^{m-1} |\partial^\gamma f(x - \frac{u}{n}t)| \frac{|u|^m}{n^m} |Z(u)| dt. \quad (140)$$

Taking the L^p -norm and applying Minkowski's inequality:

$$\left\| \sum_{k \in \mathbb{Z}^N} R_m(x, k) Z(nx - k) \right\|_{L^p} \leq \sum_{|\gamma|=m} \frac{1}{\gamma!} \frac{1}{n^m} \int_0^1 (1-t)^{m-1} \left\| \sum_{u \in \mathbb{Z}^N} |\partial^\gamma f(x - \frac{u}{n}t)| |u|^m |Z(u)| \right\|_{L^p} dt. \quad (141)$$

The sum over u approximates the integral

$$\sum_{u \in \mathbb{Z}^N} |\partial^\gamma f(x - \frac{u}{n}t)| |u|^m |Z(u)| \approx n^N \int_{\mathbb{R}^N} |\partial^\gamma f(x - yt)| |y|^m |Z(ny)| dy \leq \int_{\mathbb{R}^N} |u|^m |Z(u)| du \|\partial^\gamma f\|_{L^p}. \quad (142)$$

Hence,

$$\left\| \sum_{k \in \mathbb{Z}^N} R_m(x, k) Z(nx - k) \right\|_{L^p} \leq \frac{C \|f\|_{W^{m,p}}}{n^m}, \quad (143)$$

with C given explicitly by (134).

Combining the estimates for R_m and E_n , we obtain the desired quantitative L^p rate (135). $\square$

8. Unified Voronovskaya-Type Expansion in Sobolev $W^{s,p}$ with Explicit Constants

Theorem 16 (Voronovskaya-Type Expansion in $W^{s,p}$ with Explicit Constants). *Let $f \in W^{m,p}(\mathbb{R}^N)$, with $m \in \mathbb{N}$, $0 \leq s \leq m$, and $1 \leq p \leq \infty$. Assume the kernel Z has finite moments*

$$M_\alpha := \int_{\mathbb{R}^N} u^\alpha Z(u) du, \quad |\alpha| \leq m, \quad (144)$$

and is rapidly decaying. Then for the operator A_n we have

$$A_n f(x) - f(x) = \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) + r_{n,s}(x), \quad (145)$$

with remainder $r_{n,s} \in W^{s,p}(\mathbb{R}^N)$ satisfying the explicit quantitative bound

$$\|r_{n,s}\|_{W^{s,p}} \leq \frac{C(N, m, p, Z) \|f\|_{W^{m,p}}}{n^{m-s}}, \quad (146)$$

where the constant is given explicitly by

$$C(N, m, p, Z) := \sum_{|\gamma|=m} \frac{1}{\gamma!} \int_{\mathbb{R}^N} |u|^{|\gamma|} |Z(u)| du. \quad (147)$$

Proof. For each lattice point $k \in \mathbb{Z}^N$, consider the multivariate Taylor expansion around x :

$$f\left(\frac{k}{n}\right) = \sum_{|\alpha| \leq m-1} \frac{\partial^\alpha f(x)}{\alpha!} \left(\frac{k}{n} - x\right)^\alpha + R_m(x, k), \quad (148)$$

with integral remainder

$$R_m(x, k) = \sum_{|\gamma|=m} \frac{m}{\gamma!} \int_0^1 (1-t)^{m-1} \partial^\gamma f\left(x + t\left(\frac{k}{n} - x\right)\right) \left(\frac{k}{n} - x\right)^\gamma dt. \quad (149)$$

By linearity of A_n , we obtain

$$A_n f(x) - f(x) - \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) = \sum_{k \in \mathbb{Z}^N} R_m(x, k) Z(nx - k) + E_n(x), \quad (150)$$

where $E_n(x)$ accounts for the approximation error in replacing discrete sums by the exact moments. Rapid decay of Z ensures $E_n(x) = O(n^{-m})$.

Perform the change of variables $u = nx - k \implies k/n - x = -u/n$. Then

$$\sum_{k \in \mathbb{Z}^N} |R_m(x, k)| |Z(nx - k)| \leq \sum_{|\gamma|=m} \frac{1}{\gamma! n^m} \sum_{u \in \mathbb{Z}^N} \int_0^1 (1-t)^{m-1} \left| \partial^\gamma f\left(x - \frac{tu}{n}\right) \right| |u|^m |Z(u)| dt. \quad (151)$$

Applying discrete Minkowski inequality and standard Sobolev shift estimates, we deduce

$$\|r_{n,s}\|_{W^{s,p}} \leq \frac{C(N, m, p, Z) \|f\|_{W^{m,p}}}{n^{m-s}}, \quad (152)$$

with $C(N, m, p, Z)$ given explicitly in (147).

Combining all these steps, the Voronovskaya-type expansion in Sobolev space $W^{s,p}$ with an explicit quantitative rate is established. $\square$

9. Uniform Stability of Voronovskaya-Type Expansions under (q, λ) Variations

Definition 3 (Uniform C^m Kernel with Controlled Moments). *Let $U \subset (0, \infty)^2$ be compact. We say that the family $\{Z_{q,\lambda}\}_{(q,\lambda) \in U}$ is uniformly C^m with controlled moments if*

$$\sup_{(q,\lambda) \in U} \int_{\mathbb{R}^N} |u|^{|\gamma|} |\partial^\beta Z_{q,\lambda}(u)| du < \infty, \quad \forall |\gamma|, |\beta| \leq m. \quad (153)$$

Theorem 17 (Uniform $W^{s,p}$ Stability). *Let $\{Z_{q,\lambda}\}_{(q,\lambda) \in U}$ be uniformly C^m with controlled moments on a compact set U , and let $f \in W^{m,p}(\mathbb{R}^N)$. Denote by $A_n^{(q,\lambda)}$ the corresponding operators. Then the Voronovskaya-type expansions*

$$A_n^{(q,\lambda)} f(x) - f(x) = \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha^{(q,\lambda)}}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) + r_{n,s}^{(q,\lambda)}(x) \quad (154)$$

are uniform in $(q, \lambda) \in U$, in the sense that

$$\sup_{(q, \lambda) \in U} \|r_{n,s}^{(q, \lambda)}\|_{W^{s,p}} = o(n^{-(m-s)}), \quad n \rightarrow \infty, \quad (155)$$

with constants in the estimates depending only on the suprema in (153).

Proof. For each $(q, \lambda) \in U$ and lattice point $k \in \mathbb{Z}^N$, consider the multivariate Taylor expansion of f around x :

$$f\left(\frac{k}{n}\right) = \sum_{|\alpha| \leq m-1} \frac{\partial^\alpha f(x)}{\alpha!} \left(\frac{k}{n} - x\right)^\alpha + R_m^{(q, \lambda)}(x, k), \quad (156)$$

where $R_m^{(q, \lambda)}(x, k)$ is the integral remainder.

By linearity of the operator $A_n^{(q, \lambda)}$, we write

$$A_n^{(q, \lambda)} f(x) - f(x) - \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha^{(q, \lambda)}}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) = \sum_{k \in \mathbb{Z}^N} R_m^{(q, \lambda)}(x, k) Z_{q, \lambda}(nx - k) + E_n^{(q, \lambda)}(x), \quad (157)$$

where $E_n^{(q, \lambda)}(x)$ accounts for the discrete-to-continuum moment approximation error.

By assumption (153), the integrals

$$\int_{\mathbb{R}^N} |u|^{|\gamma|} |\partial^\beta Z_{q, \lambda}(u)| du \quad (158)$$

are uniformly bounded over $(q, \lambda) \in U$. Therefore, each term in the remainder estimate satisfies

$$\|r_{n,s}^{(q, \lambda)}\|_{W^{s,p}} \leq \sum_{|\gamma|=m} \frac{1}{\gamma! n^{m-s}} \int_{\mathbb{R}^N} |u|^m \sup_{(q, \lambda) \in U} |\partial^\gamma Z_{q, \lambda}(u)| du \|f\|_{W^{m,p}}. \quad (159)$$

The uniform integrability of $\partial^\beta Z_{q, \lambda}$ allows the application of the dominated convergence theorem as $n \rightarrow \infty$, yielding the uniform bound (155) and establishing the uniformity of the Voronovskaya expansions in $W^{s,p}$. $\square$

10. Unified Voronovskaya-Type Expansion with Explicit Constants and Uniform Stability

10.1. Setup

Let $f \in W^{m,p}(\mathbb{R}^N)$, $m \in \mathbb{N}$, $0 \leq s \leq m$, $1 \leq p \leq \infty$. Let $Z \in C^m(\mathbb{R}^N)$ (or a parametric family $\{Z_{q, \lambda}\}_{(q, \lambda) \in U}$) with moments

$$M_\alpha := \int_{\mathbb{R}^N} u^\alpha Z(u) du, \quad |\alpha| \leq m, \quad (160)$$

and assume rapid decay and, for parametric families, uniform boundedness:

$$\sup_{(q, \lambda) \in U} \int_{\mathbb{R}^N} |u|^{|\gamma|} |\partial^\beta Z_{q, \lambda}(u)| du < \infty, \quad |\gamma|, |\beta| \leq m. \quad (161)$$

Define the multivariate quasi-interpolation operators

$$A_n^{(\cdot)} f(x) := \sum_{k \in \mathbb{Z}^N} f\left(\frac{k}{n}\right) Z_{(\cdot)}(nx - k), \quad \sum_k Z_{(\cdot)}(nx - k) = 1. \quad (162)$$

10.2. Unified Voronovskaya-Type Theorem

Theorem 18 (Compact Voronovskaya Expansion with Explicit Constants). *Under the above assumptions, the operator $A_n^{(\cdot)}$ satisfies*

$$A_n^{(\cdot)} f(x) - f(x) = \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha^{(\cdot)}}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) + r_{n,s}^{(\cdot)}(x), \quad (163)$$

where the remainder $r_{n,s}^{(\cdot)} \in W^{s,p}(\mathbb{R}^N)$ satisfies the uniform quantitative estimate

$$\|r_{n,s}^{(\cdot)}\|_{W^{s,p}} \leq \frac{C(N, m, p, Z)}{n^{m-s}} \|f\|_{W^{m,p}}, \quad C(N, m, p, Z) := \sum_{|\gamma|=m} \frac{1}{\gamma!} \int_{\mathbb{R}^N} |u|^m |Z(u)| du. \quad (164)$$

For parametric families $(q, \lambda) \in U$, the expansion is uniform:

$$\sup_{(q, \lambda) \in U} \|r_{n,s}^{(q, \lambda)}\|_{W^{s,p}} = o(n^{-(m-s)}), \quad n \rightarrow \infty. \quad (165)$$

Moreover, in the case $s = 0$, this yields an explicit L^p rate

$$\left\| A_n^{(\cdot)} f - f - \sum_{1 \leq |\alpha| \leq m-1} \frac{M_\alpha^{(\cdot)}}{\alpha! n^{|\alpha|}} \partial^\alpha f \right\|_{L^p} \leq \frac{C(N, m, p, Z) \|f\|_{W^{m,p}}}{n^m}. \quad (166)$$

Proof. Expand $f(k/n)$ using the multivariate Taylor formula of order m with integral remainder:

$$f\left(\frac{k}{n}\right) = \sum_{|\alpha| \leq m-1} \frac{\partial^\alpha f(x)}{\alpha!} \left(\frac{k}{n} - x\right)^\alpha + R_m^{(\cdot)}(x, k), \quad (167)$$

with

$$R_m^{(\cdot)}(x, k) = \sum_{|\gamma|=m} \frac{m}{\gamma!} \int_0^1 (1-t)^{m-1} \partial^\gamma f\left(x + t\left(\frac{k}{n} - x\right)\right) \left(\frac{k}{n} - x\right)^\gamma dt. \quad (168)$$

Applying $A_n^{(\cdot)}$ gives (163) with

$$r_{n,s}^{(\cdot)}(x) := \sum_{k \in \mathbb{Z}^N} R_m^{(\cdot)}(x, k) Z_{(\cdot)}(nx - k). \quad (169)$$

For $|\beta| \leq s$, differentiate under the sum:

$$\partial^\beta r_{n,s}^{(\cdot)}(x) = \sum_{k \in \mathbb{Z}^N} R_m^{(\cdot)}(x, k) \partial^\beta Z_{(\cdot)}(nx - k) n^{|\beta|}. \quad (170)$$

By Young's inequality and the uniform boundedness of the kernel moments, this yields (164). For parametric families, dominated convergence and the uniform bounds in (161) imply uniformity over U . $\square$

10.3. Remarks

- This theorem simultaneously captures the classical Voronovskaya expansion, quantitative L^p rates, Sobolev estimates, and uniform stability under parametric families.
- Constants are fully explicit, depending only on kernel moments and derivatives.
- The framework is directly applicable to numerical analysis, multivariate approximation, and theoretical studies of Sobolev-space operators.

We develop a rigorous framework for *adaptive Voronovskaya-type expansions* for multivariate neural network operators with dynamically deformed hyperbolic tangent activations. This extends classical

asymptotic expansions to *non-stationary kernels*, where the deformation parameters adapt with the approximation scale. Our results establish *accelerated convergence rates in Sobolev spaces* $W^{m,p}(\mathbb{R}^N)$ while ensuring uniform control of derivatives up to order $s \leq m$. Precise remainder estimates are provided, demonstrating the interplay between kernel deformation and Sobolev regularity, and enabling applications in high-dimensional function approximation and adaptive deep learning architectures.

11. Setup and Hypotheses for Sobolev-Santos Uniform Adaptive Convergence Theorem

Definition 4 (Adaptive Quasi-Interpolation Operator). Let $f \in W^{m,p}(\mathbb{R}^N)$. Define the adaptive operator

$$A_n^{(q_n, \lambda_n)}(f, x) := \sum_{k \in \mathbb{Z}^N} f\left(\frac{k}{n}\right) Z_{q_n, \lambda_n}(nx - k), \quad (171)$$

where the multivariate adaptive kernel is

$$Z_{q_n, \lambda_n}(x) := \prod_{i=1}^N \Phi_{q_n, \lambda_n}(x_i), \quad \Phi_{q_n, \lambda_n}(x_i) := \frac{M_{q_n, \lambda_n}(x_i) + M_{1/q_n, \lambda_n}(x_i)}{2}. \quad (172)$$

The parameters $q_n > 0$ and $\lambda_n > 0$ are adaptive, satisfying

$$q_n \rightarrow 1, \quad \lambda_n \rightarrow \infty, \quad \lambda_n = O(n^\gamma), \quad 0 < \gamma < 1. \quad (173)$$

Hypotheses.

1. **Controlled Deformation:** There exist constants $0 < C_1 \leq C_2 < \infty$ such that $C_1 \leq q_n \leq C_2$ for all $n \in \mathbb{N}$.
2. **Uniform Exponential Decay:** For all multi-indices α with $|\alpha| \leq m$, there exists $T_\alpha > 0$ such that

$$\sum_{k \in \mathbb{Z}^N} |k|^\alpha Z_{q_n, \lambda_n}(k) \leq T_\alpha, \quad \forall n.$$

3. **Uniformly Bounded Moments:** For all $|\alpha| \leq m$,

$$\sup_n \left| \int_{\mathbb{R}^N} u^\alpha Z_{q_n, \lambda_n}(u) du \right| < \infty.$$

Theorem 19 (Sobolev-Santos Uniform Adaptive Convergence). Let $f \in W^{m,p}(\mathbb{R}^N)$ with $1 \leq p \leq \infty$ and $0 \leq s \leq m$. Assume the adaptive quasi-interpolation operator $A_n^{(q_n, \lambda_n)}$ satisfies the hypotheses of controlled deformation, uniform exponential decay, and bounded moments. Then, for sufficiently large n , the following adaptive Voronovskaya expansion holds:

$$A_n^{(q_n, \lambda_n)}(f, x) - f(x) = \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha^{(n)}}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) + r_n^{(s)}(x), \quad (174)$$

where the multivariate moments are

$$M_\alpha^{(n)} := \int_{\mathbb{R}^N} u^\alpha Z_{q_n, \lambda_n}(u) du, \quad (175)$$

and the remainder term $r_n^{(s)}$ satisfies the explicit estimate

$$\|r_n^{(s)}\|_{W^{s,p}(\mathbb{R}^N)} \leq C(N, m, p, f) \frac{\Phi(q_n, \lambda_n)}{n^{m-s+\gamma}}, \quad (176)$$

with $\Phi(q_n, \lambda_n)$ a smooth, bounded function quantifying the deviation from stationarity:

$$\Phi(q_n, \lambda_n) := \max_{|\alpha|=m-s+1} \left\{ \int_{\mathbb{R}^N} |u|^\alpha |Z_{q_n, \lambda_n}(u) - Z_{1, \infty}(u)| du \right\} = O(1), \quad n \rightarrow \infty.$$

Moreover, the expansion is uniform in Sobolev norm up to order s , i.e.,

$$\sup_{|\beta| \leq s} \left\| \partial^\beta (A_n^{(q_n, \lambda_n)}(f) - f - \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha^{(n)}}{\alpha! n^{|\alpha|}} \partial^\alpha f) \right\|_{L^p} \leq \frac{C(N, m, p, f)}{n^{m-s+\gamma}}. \quad (177)$$

Remarks:

1. The factor $\Phi(q_n, \lambda_n)$ explicitly captures the influence of adaptive kernel deformation on the remainder, providing a quantitative measure of non-stationarity.
2. If $q_n \equiv 1$ and $\lambda_n \rightarrow \infty$, the theorem recovers the classical stationary Voronovskaya expansion.
3. The theorem can be extended to fractional Sobolev spaces $W^{m+\sigma, p}$ with $0 < \sigma < 1$, allowing finer control for functions with limited smoothness.
4. Uniformity in s ensures stability of derivatives up to order s , which is crucial for high-dimensional deep learning applications where gradients are propagated through multiple layers.

Proof. Let $f \in W^{m,p}(\mathbb{R}^N)$ and $0 \leq s \leq m$. The proof proceeds in several steps.

For any $x, y \in \mathbb{R}^N$, by Taylor's theorem with integral remainder, we have

$$f(y) = \sum_{|\alpha| \leq m-s} \frac{\partial^\alpha f(x)}{\alpha!} (y-x)^\alpha + R_{m-s}(x, y), \quad (178)$$

where the remainder is explicitly

$$R_{m-s}(x, y) = \sum_{|\alpha|=m-s+1} \frac{(m-s+1)}{\alpha!} \int_0^1 (1-t)^{m-s} \partial^\alpha f(x + t(y-x)) (y-x)^\alpha dt. \quad (179)$$

Substituting $y = k/n$ in (171) yields

$$A_n^{(q_n, \lambda_n)}(f, x) = \sum_{k \in \mathbb{Z}^N} \left[\sum_{|\alpha| \leq m-s} \frac{\partial^\alpha f(x)}{\alpha!} \left(\frac{k}{n} - x\right)^\alpha + R_{m-s}(x, k/n) \right] Z_{q_n, \lambda_n}(nx - k). \quad (180)$$

By definition of the multivariate moments

$$M_\alpha^{(n)} := \int_{\mathbb{R}^N} u^\alpha Z_{q_n, \lambda_n}(u) du, \quad (181)$$

and using the discrete-to-continuum approximation via the kernel's exponential decay, we have

$$\sum_{k \in \mathbb{Z}^N} \left(\frac{k}{n} - x\right)^\alpha Z_{q_n, \lambda_n}(nx - k) = \frac{M_\alpha^{(n)}}{n^{|\alpha|}}. \quad (182)$$

Hence, the main contribution in (174) is

$$\sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha^{(n)}}{\alpha! n^{|\alpha|}} \partial^\alpha f(x).$$

Define the remainder operator

$$r_n^{(s)}(x) := \sum_{k \in \mathbb{Z}^N} R_{m-s}(x, k/n) Z_{q_n, \lambda_n}(nx - k). \quad (183)$$

Applying Minkowski's inequality and using the uniform exponential decay and bounded moments of Z_{q_n, λ_n} , we obtain

$$\begin{aligned} \|r_n^{(s)}\|_{L^p} &\leq \sum_{|\alpha|=m-s+1} \frac{m-s+1}{\alpha!} \int_0^1 (1-t)^{m-s} \sum_{k \in \mathbb{Z}^N} |\partial^\alpha f(x+t(k/n-x))| |k/n-x|^\alpha Z_{q_n, \lambda_n}(nx-k) dt \\ &\leq \frac{C(N, m, p, f)}{n^{m-s+\gamma}}. \end{aligned} \quad (184)$$

where $C(N, m, p, f)$ is independent of n and $\gamma \in (0, 1)$ is the decay exponent from the kernel's scaling.

The smoothness of Z_{q_n, λ_n} guarantees that differentiation commutes with $A_n^{(q_n, \lambda_n)}$ up to order s , so for all multi-indices β with $|\beta| \leq s$,

$$\|\partial^\beta r_n^{(s)}\|_{L^p} \leq \frac{C(N, m, p, f)}{n^{m-s+\gamma}}. \quad (185)$$

Consequently,

$$\|A_n^{(q_n, \lambda_n)}(f) - f - \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha^{(n)}}{\alpha! n^{|\alpha|}} \partial^\alpha f\|_{W^{s,p}} = \|r_n^{(s)}\|_{W^{s,p}} \leq \frac{C}{n^{m-s+\gamma}}. \quad (186)$$

This completes the proof. $\square$

12. Results

This study establishes rigorous theoretical results for symmetrized hyperbolic tangent neural network operators, with a focus on the novel **Sobolev-Santos Uniform Convergence Theorem**. The main contributions are as follows:

1. **Voronovskaya-Type Expansions for Basic Operators:** For functions $f \in C^m(\mathbb{R}^N)$, the approximation error of the basic operator A_n admits an asymptotic expansion:

$$A_n(f, x) - f(x) = \sum_{j=1}^m \sum_{|\alpha|=j} \frac{f_\alpha(x)}{\alpha!} A_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) + o \left(\frac{1}{n^{\beta(m-\varepsilon)}} \right).$$

This expansion provides explicit convergence rates, dependent on the smoothness m of f and the grid parameter β .

2. **Refined Expansions for Kantorovich and Quadrature Operators:** For $f \in C^{m+2}(\mathbb{R}^N)$, the Kantorovich operator K_n satisfies a refined expansion:

$$K_n(f, x) - f(x) = \sum_{j=1}^{m+1} \sum_{|\alpha|=j} \frac{f_\alpha(x)}{\alpha!} K_n \left(\prod_{i=1}^N (\cdot - x_i)^{\alpha_i} \right) (x) + R_n(x),$$

where the remainder $R_n(x)$ is bounded by $O(n^{-(m+1)\beta})$, demonstrating higher-order accuracy.

3. **Sobolev Space Estimates:** The approximation error in the Sobolev space $W^{s,p}(\mathbb{R}^N)$ is bounded by:

$$\|A_n f - f\|_{W^{s,p}} = O(n^{-(m-s)}).$$

This result provides quantitative estimates for the convergence rate, with explicit constants derived from the moments of the kernel function.

4. **Sobolev-Santos Uniform Convergence Theorem:** The **Sobolev-Santos Theorem** (Theorem ??) establishes that for adaptive quasi-interpolation operators $A_n^{(q_n, \lambda_n)}$, the following expansion holds:

$$A_n^{(q_n, \lambda_n)}(f, x) - f(x) = \sum_{1 \leq |\alpha| \leq m-s} \frac{M_\alpha^{(n)}}{\alpha! n^{|\alpha|}} \partial^\alpha f(x) + r_n^{(s)}(x),$$

where the remainder $r_n^{(s)}$ satisfies the explicit estimate:

$$\|r_n^{(s)}\|_{W^{s,p}} \leq \frac{C(N, m, p, f) \Phi(q_n, \lambda_n)}{n^{m-s+\gamma}}.$$

The function $\Phi(q_n, \lambda_n)$ quantifies the deviation from stationarity, ensuring uniform stability under parametric variations of the activation function. This theorem is pivotal for applications requiring adaptive kernel deformation, such as high-dimensional deep learning architectures.

5. **Uniform Stability Under Parametric Variations:** The expansions remain uniformly valid even when the activation function parameters (q, λ) vary, ensuring robustness in practical applications. This stability is critical for adaptive neural network architectures, where parameters may dynamically adjust during training or optimization.

13. Conclusions

This work advances the theory of neural network approximation by introducing symmetrized hyperbolic tangent-based operators and deriving Voronovskaya-type asymptotic expansions for their multivariate counterparts. The **Sobolev-Santos Uniform Convergence Theorem** is a cornerstone of this study, providing a rigorous framework for adaptive quasi-interpolation operators with dynamically deformed activation functions. This theorem ensures that the operators can approximate smooth functions and their derivatives with high accuracy, even as the deformation parameters (q_n, λ_n) evolve, while maintaining uniform control over the convergence rates in Sobolev spaces $W^{s,p}(\mathbb{R}^N)$.

The explicit constants and uniform bounds derived in this study offer a solid foundation for both theoretical and applied research in neural network-based function approximation. The results highlight the superior performance of these operators in high-dimensional approximation problems, with direct implications for artificial intelligence, numerical analysis, and data-driven modeling. The uniform stability under parametric variations further enhances their applicability in adaptive deep learning architectures, where robustness and flexibility are essential.

Future research directions include exploring adaptive grid strategies, extending the framework to fractional Sobolev spaces, and generalizing the results to non-Euclidean domains. These advancements could further expand the applicability of symmetrized hyperbolic tangent neural networks in modern computational frameworks, particularly in scenarios requiring high-dimensional function approximation and adaptive learning.

Acknowledgments: Santos gratefully acknowledges the support of the PPGMC Program for the Postdoctoral Scholarship PROBOL/UESC nr. 218/2025. Sales would like to express his gratitude to CNPq for the financial support under grant 30881/2025-0.

References

1.

Anastassiou, G. A. (1997). Rate of convergence of some neural network operators to the unit-univariate case. *Journal of Mathematical Analysis and Applications*, 212(1), 237-262. <https://doi.org/10.1006/jmaa.1997.5494>.

2.

Anastassiou, G. (2000). *Quantitative approximations*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781482285796>.

3.

Anastassiou, G. A. (2016). *Intelligent Systems II: Complete Approximation by Neural Network Operators* (Vol. 608). Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-20505-2>.

4.

Anastassiou, G. A. (2023). *Parametrized, deformed and general neural networks*. Berlin/Heidelberg, Germany: Springer. <https://doi.org/10.1007/978-3-031-43021-3>.

5. Z. Chen and F. Cao, The approximation operators with sigmoidal functions, *Computers and Mathematics with Applications*, 58 (2009), 758-765. <https://doi.org/10.1016/j.camwa.2009.05.001>.
6. Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice hall PTR.
7. McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133. <https://doi.org/10.1007/BF02478259>.
8. Yu, D., & Cao, F. (2025). Construction and approximation rate for feedforward neural network operators with sigmoidal functions. *Journal of Computational and Applied Mathematics*, 453, 116150. <https://doi.org/10.1016/j.cam.2024.116150>.
9. Yoo, J., Kim, J., Gim, M., & Lee, H. (2024). Error estimates of physics-informed neural networks for initial value problems. *Journal of the Korean Society for Industrial and Applied Mathematics*, 28(1), 33-58.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.