MATHEMATICAL COMMUNICATIONS
Math. Commun. **18**(2013), 185–207

The construction and approximation of neural networks operators with Gaussian activation function*

ZHIXIANG CHEN[1] AND FEILONG CAO[2,†]

[1] *Department of Mathematics, Shaoxing University, Shaoxing 312 000, Zhejiang Province, P.R. China*
[2] *Department of Mathematics, China Jiliang University, Hangzhou 310 018, Zhejiang Province, P.R. China*

Received August 1, 2011; accepted January 26, 2013

Abstract. This paper studies the construction and approximation of neural network operators with a centered bell-shaped Gaussian activation function. Using a univariate Gaussian function a class of Cardaliaguet-Euvrard type network operators is constructed to approximate the continuous function, and the Jackson type theorems of the approximation and some discussions about the convergence are given. Furthermore, to approximate the multivariate function, a class of bivariate Cardaliaguet-Euvrard type network operators is introduced, and the corresponding estimates of the approximation rate are deduced.
AMS subject classifications: 41A17, 41A25
Key words: neural network, Gaussian function, approximation, modulus of continuity

1. Introduction

Mathematically, feed-forward neural networks (FNNs) with one hidden layer are expressed as

$$N_n(x) = \sum_{j=1}^{n} c_j \sigma(A_j x + b_j), \quad x \in \mathbb{R}^s, \quad s \in \mathbb{N}, \tag{1}$$

where σ is the activation function defined on $\mathbb{R}^d$, and for $1 \leq j \leq n, c_j \in \mathbb{R}$ are the coefficients, A_j are real matrices of $d \times s, b_j \in \mathbb{R}^d$ are the thresholds. If σ is defined on $\mathbb{R}$, then (1) becomes

$$N_n(x) = \sum_{j=1}^{n} c_j \sigma(\langle A_j \cdot x\rangle + b_j), \tag{2}$$

where $\langle A_j \cdot x\rangle$ is the inner product of a_j and $x, b_j \in \mathbb{R}$. If A_j is restricted within diagonal matrices, then $N_n(x)$ is a linear combination of translates and dilates of one or several functions.

*This research was supported by the National Natural Science Foundation of China (Nos. 61272023,61179041).
†Corresponding author. *Email address:* `czx@usx.edu.cn` (Z. Chen), `feilongcao@gmail.com` (F. Cao)

It is well known that FNNs are universal approximators. Theoretically, any continuous function defined on a compact set can be approximated by a FNN to any desired degree of accuracy by increasing the number of hidden neurons. It was proved by Cybenko [17] and Funahashi [19], that any continuous function can be approximated on a compact set with uniform topology by a network of the form given in Equation (1), using any continuous, sigmoidal activation function. Furthermore, various density results on FNN approximations of multivariate functions were later established by many authors using various methods, for more or less general situations: [14] by Chen and Chen, [16] by Chui and Li, [20] by Hartman et al., [21] by Hornik et al., [23] by Leshno et al., [27] by Mhaskar and Micchelli, etc.

Yet a related and important problem is that of complexity: determining the number of neurons required to guarantee that all functions (belonging to a certain class) can be approximated to the prescribed degree of accuracy ϵ. For example, a classical result of Barrom [5] shows that if the function is assumed to satisfy certain conditions expressed in terms of its Fourier transform, and if each of the neurons evaluates a sigmoidal activation function, then at most $\mathcal{O}(\epsilon^{-2})$ neurons are needed to achieve the order of approximation ϵ. Up till now, many authors have published similar results on the complexity of FNN approximations: Cao et al. [11], Chen [13], Ferrari and Stengel [18], Maiorov and Meir [30], Makovoz [31], Mhaskar and Micchelli [28], Suzuki [36], Xu and Cao [38] etc.

A function $b : \mathbb{R} \mapsto \mathbb{R}$ is said to be centered bell-shaped if b belongs to L^1 and its integral is nonzero, if it is nondecreasing on $(-\infty, 0)$ and nonincreasing on $[0, +\infty)$. Let $f : \mathbb{R} \mapsto \mathbb{R}$ be a continuous and bounded function and b be a centered bell-shaped function, Cardaliaguet and Euvrard [12] introduced the neural network operators defined by

$$F_n^1(f,x) = \sum_{k=-n^2}^{n^2} \frac{f\left(\frac{k}{n}\right)}{In^\alpha} b\left(\frac{nx-k}{n^\alpha}\right), \tag{3}$$

where $I = \int_{-\infty}^{\infty} b(t)dt$, which are called Cardaliaguet-Euvrard neural network operators. If $f : \mathbb{R}^s \mapsto \mathbb{R}$ is a continuous and bounded function, then it is natural to define the multivariate Cardaliaguet-Euvrard neural network operators as (see [12])

$$\begin{aligned} &F_n^s(f; x_1, \dots, x_s) \\ &\quad = \sum_{k_1=-n^2}^{n^2} \cdots \sum_{k_s=-n^2}^{n^2} \frac{f\left(\frac{k_1}{n}, \cdots, \frac{k_s}{n}\right)}{In^\alpha} b\left(\frac{nx_1-k_1}{n^\alpha}\right) \times \cdots \times b\left(\frac{nx_s-k_s}{n^\alpha}\right). \end{aligned}$$

Anastassiou [1] considered the centered bell-shaped function with compact support and a nonnegative number, and gave the pointwise estimates for continuous function f defined on $\mathbb{R}$:

$$\begin{aligned} |F_n^1(f,x) - f(x)| \le |f(x)| &\left| \sum_{k=\lceil nx-n^\alpha \rceil}^{[nx+n^\alpha]} \frac{1}{n^\alpha} b\left(\frac{nx-k}{n^\alpha}\right) - 1 \right| \\ &+ b(0)\left(2 + \frac{1}{n^\alpha}\right)\omega\left(f; \frac{1}{n^{1-\alpha}}\right), \end{aligned} \tag{4}$$

where b has support $[-1,1]$ with $\int_{-1}^{1} b(x)dx = 1$, and $\omega(f;t)$ is the first modulus of continuity of f (see [24]). Anastassiou [3] further proved that $F_n^1(f,x) \mapsto f(x)$ as $n \to \infty$. For s-dimensional centered bell-shaped function with support $[-1,1]^s$, Anastassiou [2, 3] obtained the similar pointwise estimate to (4).

For the important sigmoidal activation function defined by

$$s(x) = \frac{1}{1+e^{-x}},$$

it is not difficult to see that the function

$$\phi(x) = \frac{1}{2}(s(x+1) - s(x-1)) \tag{5}$$

is a centered bell-shaped function with support $\mathbb{R}$ (see [4]).

Let $f(x_1,x_2)$ be a continuous function defined on $[-1,1]^2$. Recently, Anastassiou [4] constructed an interesting operator defined by

$$G_n(f;x_1,x_2) = \frac{\sum\limits_{k_1=\lceil -n\rceil}^{[n]} \sum\limits_{k_2=\lceil -n\rceil}^{[n]} f(\frac{k_1}{n},\frac{k_2}{n})\phi(nx_1-k_1)\phi(nx_2-k_2)}{\sum\limits_{k_1=\lceil -n\rceil}^{[n]} \sum\limits_{k_2=\lceil -n\rceil}^{[n]} \phi(nx_1-k_1)\phi(nx_2-k_2)}, \tag{6}$$

and gave the estimate

$$|G_n(f;x_1,x_2)-f(x_1,x_2)| \le (5.250312578)^2 \left(\omega\left(f;\frac{1}{n^\alpha},\frac{1}{n^\alpha}\right) + 6.3984\|f\|_\infty e^{-n(1-\alpha)}\right),$$

where $[\cdot]$ denotes the integral part of a number, $\lceil\cdot\rceil$ the ceiling of a number, $\|f\|_\infty$ the uniform norm of f, and $0<\alpha<1$, and $\omega(f;t_1,t_2)$ is the modulus of continuity of two-variate f defined by (see [34])

$$\omega(f;t_1,t_2) = \sup_{|h_1|\le t_1,|h_2|\le t_2} |f(x_1+h_1,x_2+h_2) - f(x_1,x_2)|.$$

Yet, (6) is an operator of rational type.

It is well known that s-dimensional Gaussian function defined by

$$G(x) = \frac{1}{\pi^{\frac{s}{2}}} e^{-\|x\|^2}, \quad x \in \mathbb{R}^s$$

is a class of important radial basis function (RBF), which is centered bell-shaped function, but has not compact support. Nevertheless, there have been many applications in numerical approximation and neural networks (for example, see [7, 8, 9, 10], [15], [22], [25, 26], [29], [32], [35], [37]).

In this paper, we will construct Cardaliaguet-Euvrard type neural network operators with a Gaussian activation function, and give more refined error estimates of Jackson type. In addition, it is not difficult to see that Cardaliaguet-Eurrard network operators are also a kind of Quasi-interpolation operators without polynomial reproduction (see [6]). Using the method of approximate partition of unity, the quasi-interpolation operators with Gaussian are constructed, and the approximation rates are estimated in [15] and [32]. Therefore, for bell-shaped generating functions, it is natural to study the construction and approximation of quasi-interpolation operators by using the way of the constructive methods of Cardaliaguet-Euvrard operators.

This paper is organized as follows. Section 2 constructs two kinds of Cardaliaguet-Euvrard type neural network operators with univariate Gaussian radial basis activation function, and gives the discussions of rate of approximation on $\mathbb{R}$ and its compact subset, respectively. Section 3 introduces the neural network operators with bivariate Gaussian radial basis activation function and estimates the corresponding rate of approximation. In Section 4 we give numerical results to show the approximation does converge and agree with the rate claimed in the paper.

2. Approximation on $\mathbb{R}$

For a Gaussian function

$$G(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}, \quad x \in \mathbb{R}, \tag{7}$$

and $d > 0$, we set

$$g_d(x) = \frac{1}{\sqrt{\pi} d} e^{-\frac{x^2}{d^2}}, \quad x \in \mathbb{R}.$$

We denote the class of continuous and bounded functions defined on $\mathbb{R}$ by $C(\mathbb{R})$, and construct neural network operators with activation function g_d for $f \in C(\mathbb{R})$ as

$$G^1_{n,d}(f,x) = \sum_{k=\lceil -n^2-n^\alpha \rceil}^{[n^2+n^\alpha]} \frac{f(\frac{k}{n})}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right).$$

To prove our first result, we need the following lemma.

Lemma 1. *If $-n \le x \le n$, we have*

1) *When $\lceil nx \rceil + 1 \le k \le [n^2 + n^\alpha]$, then*

$$\int_{\lceil nx \rceil+1}^{[n^2+n^\alpha]+1} g_d\left(\frac{nx-t}{n^\alpha}\right) dt \le \sum_{k=\lceil nx \rceil+1}^{[n^2+n^\alpha]} g_d\left(\frac{nx-k}{n^\alpha}\right) \le \int_{\lceil nx \rceil}^{[n^2+n^\alpha]} g_d\left(\frac{nx-t}{n^\alpha}\right) dt.$$

2) *When $\lceil -n^2 - n^\alpha \rceil \le k \le [nx] - 1$, then*

$$\int_{\lceil -n^2-n^\alpha \rceil-1}^{[nx]-1} g_d\left(\frac{nx-t}{n^\alpha}\right) dt \le \sum_{k=\lceil -n^2-n^\alpha \rceil}^{[nx]-1} g_d\left(\frac{nx-k}{n^\alpha}\right) \le \int_{\lceil -n^2-n^\alpha \rceil}^{[nx]} g_d\left(\frac{nx-t}{n^\alpha}\right) dt.$$

Proof. 1) Since $-n \le x \le n$, we have

$$\lceil nx \rceil + 1 \le k \le [n^2 + n^\alpha]$$

and

$$x-\frac{k}{n}\le x-\frac{t}{n}\le x-\frac{k-1}{n}\le 0,\quad nx\le k-1\le t\le k.$$

Hence

$$g_d\left(n^{1-\alpha}\left(x-\frac{k}{n}\right)\right)\le g_d\left(n^{1-\alpha}\left(x-\frac{t}{n}\right)\right)\le g_d\left(n^{1-\alpha}\left(x-\frac{(k-1)}{n}\right)\right),$$

that is

$$g_d\left(\frac{nx-k}{n^\alpha}\right)\le g_d\left(\frac{nx-t}{n^\alpha}\right)\le g_d\left(\frac{nx-(k-1)}{n^\alpha}\right).$$

Thus

$$g_d\left(\frac{nx-k}{n^\alpha}\right)\le\int_{k-1}^{k} g_d\left(\frac{nx-t}{n^\alpha}\right)dt.$$

Let $nx\le k\le t\le k+1$, then

$$x-\frac{k+1}{n}\le x-\frac{t}{n}\le x-\frac{k}{n}\le 0.$$

Therefore,

$$g_d\left(\frac{nx-(k+1)}{n^\alpha}\right)\le g_d\left(\frac{nx-t}{n^\alpha}\right)\le g_d\left(\frac{nx-k}{n^\alpha}\right).$$

Hence

$$\int_{k}^{k+1} g_d\left(\frac{nx-t}{n^\alpha}\right)dt\le g_d\left(\frac{nx-k}{n^\alpha}\right).$$

By summation we have

$$\int_{\lceil nx\rceil+1}^{[n^2+n^\alpha]+1} g_d\left(\frac{nx-t}{n^\alpha}\right)dt\le\sum_{k=\lceil nx\rceil+1}^{[n^2+n^\alpha]} g_d\left(\frac{nx-k}{n^\alpha}\right)\le\int_{\lceil nx\rceil}^{[n^2+n^\alpha]} g_d\left(\frac{nx-t}{n^\alpha}\right)dt,$$

for $\lceil nx\rceil+1\le k\le[n^2+n^\alpha]$.

The proof of 2) is similar to 1). We omit the details. □

We now give the first main result.

Theorem 1. *Let $f\in C(\mathbb{R})$, $0<\alpha<1$, and $n^\alpha>2$. Then for $x\in[-n,n]$, we have*

$$\left|G^1_{n,d}(f,x)-f(x)\right|\le\omega\left(f;\frac{1}{n^{1-\alpha}}\right)\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)+6\|f\|_\infty\left(\frac{1}{\sqrt{\pi}dn^\alpha}+de^{-\frac{1}{4d^2}}\right),$$

where $\|f\|_\infty=\sup\limits_{|x|\le n+1}|f(x)|$.

Proof. Since

$$\left|\sum_{k=\lceil -n^2-n^\alpha\rceil}^{[n^2+n^\alpha]}\frac{f(\frac{k}{n})}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)-f(x)\right| \le \sum_{k=\lceil -n^2-n^\alpha\rceil}^{[n^2+n^\alpha]}\frac{|f(\frac{k}{n})-f(x)|}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)$$
$$+\|f\|_\infty\left|\sum_{k=\lceil -n^2-n^\alpha\rceil}^{[n^2+n^\alpha]}\frac{1}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)-1\right|$$
$$=: I_1+\|f\|_\infty I_2,$$

it is easy to see that

$$|I_1| = \sum_{|nx-k|\le n^\alpha}\frac{|f(\frac{k}{n})-f(x)|}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)+\sum_{|nx-k|>n^\alpha}\frac{|f(\frac{k}{n})-f(x)|}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)$$
$$\le \omega\left(f;\frac{1}{n^{1-\alpha}}\right)\sum_{k=\lceil -n^2-n^\alpha\rceil}^{[n^2+n^\alpha]}\frac{1}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)+2\|f\|_\infty\sum_{|nx-k|>n^\alpha}\frac{1}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right).$$

By computation, we can obtain

$$\sum_{k=\lceil -n^2-n^\alpha\rceil}^{[n^2+n^\alpha]}\frac{1}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right) = \frac{1}{\sqrt{\pi}dn^\alpha}\sum_{k=\lceil -n^2-n^\alpha\rceil}^{[n^2+n^\alpha]}e^{-\frac{(nx-k)^2}{d^2n^{2\alpha}}}$$
$$\le \frac{1}{\sqrt{\pi}dn^\alpha}\left(1+\int_{-\infty}^{+\infty}e^{-\frac{x^2}{d^2n^{2\alpha}}}dx\right)$$
$$= 1+\frac{1}{\sqrt{\pi}dn^\alpha},$$

and when $n^\alpha>2$, it follows that

$$\sum_{|nx-k|>n^\alpha}\frac{1}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right) = \frac{1}{\sqrt{\pi}dn^\alpha}\sum_{|nx-k|>n^\alpha}e^{-\frac{(nx-k)^2}{d^2n^{2\alpha}}}$$
$$\le \frac{2}{\sqrt{\pi}dn^\alpha}\int_{n^\alpha-1}^{+\infty}e^{-\frac{x^2}{d^2n^{2\alpha}}}dx$$
$$\le de^{-\frac{1}{4d^2}},$$

here, the inequality (see Section 3.7.3 of [33])

$$\int_a^\infty e^{-x^2}dx < \min\left(\frac{\sqrt{\pi}}{2}e^{-a^2},\frac{1}{2a}e^{-a^2}\right),\quad a>0,$$

is applied. Hence

$$|I_1| \le \omega\left(f;\frac{1}{n^{1-\alpha}}\right)\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)+2\|f\|_\infty de^{-\frac{1}{4d^2}}.$$

On the other hand,

$$\begin{aligned}
I_2 \le & \left| \sum_{k=\lceil nx\rceil+1}^{[n^2+n^\alpha]} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) - \frac{1}{2}\right| + \frac{1}{n^\alpha} g_d\left(\frac{nx-\lceil nx\rceil}{n^\alpha}\right) \\
& + \left| \sum_{k=\lceil -n^2-n^\alpha\rceil}^{[nx]-1} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) - \frac{1}{2}\right| + \frac{1}{n^\alpha} g_d\left(\frac{nx-[nx]}{n^\alpha}\right) \\
=: & \ \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4.
\end{aligned}$$

Obviously,

$$\begin{aligned}
\Delta_2 &= \frac{1}{\sqrt{\pi} d n^\alpha} e^{-\frac{(nx-\lceil nx\rceil)^2}{d^2 n^{2\alpha}}} \le \frac{1}{\sqrt{\pi} d n^\alpha}, \\
\Delta_4 &= \frac{1}{\sqrt{\pi} d n^\alpha} e^{-\frac{(nx-[nx])^2}{d^2 n^{2\alpha}}} \le \frac{1}{\sqrt{\pi} d n^\alpha}.
\end{aligned}$$

Also, the case 1) of Lemma 1 yields that

$$\begin{aligned}
\Lambda_1 := & \frac{1}{\sqrt{\pi} d n^\alpha} \int_{\lceil nx\rceil+1}^{[n^2+n^\alpha]+1} e^{-\frac{(nx-t)^2}{d^2 n^{2\alpha}}} dt - \frac{1}{\sqrt{\pi}} \int_{-\infty}^{0} e^{-t^2} dt \\
\le & \sum_{k=\lceil nx\rceil+1}^{[n^2+n^\alpha]} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) - \frac{1}{2} \\
\le & \frac{1}{\sqrt{\pi} d n^\alpha} \int_{\lceil nx\rceil}^{[n^2+n^\alpha]} e^{-\frac{(nx-t)^2}{d^2 n^{2\alpha}}} dt - \frac{1}{\sqrt{\pi}} \int_{-\infty}^{0} e^{-t^2} dt =: \Lambda_2.
\end{aligned}$$

Therefore,

$$\Delta_1 \le \max\{|\Lambda_1|, |\Lambda_2|\}.$$

From

$$-2 < nx - \lceil nx\rceil - 1 \le 1, \quad nx - [n^2+n^\alpha] - 1 \le -n^\alpha,$$

we get

$$\begin{aligned}
|\Lambda_1| &= \left| \frac{1}{\sqrt{\pi} d n^\alpha} \int_{\lceil nx\rceil+1}^{[n^2+n^\alpha]+1} e^{-\frac{(nx-t)^2}{d^2 n^{2\alpha}}} dt - \frac{1}{\sqrt{\pi} d n^\alpha} \int_{-\infty}^{0} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right| \\
&= \frac{1}{\sqrt{\pi} d n^\alpha} \left| \int_{nx-[n^2+n^\alpha]-1}^{nx-\lceil nx\rceil-1} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt - \int_{-\infty}^{0} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right| \\
&\le \frac{1}{\sqrt{\pi} d n^\alpha} \left(\int_{-2}^{0} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt + \int_{-\infty}^{-n^\alpha} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right) \\
&\le \frac{2}{\sqrt{\pi} d n^\alpha} + d e^{-\frac{1}{4d^2}},
\end{aligned}$$

and

$$\begin{aligned}
|\Lambda_2| &= \left| \frac{1}{\sqrt{\pi} d n^\alpha} \int_{\lceil nx \rceil}^{[n^2+n^\alpha]} e^{-\frac{(nx-t)^2}{d^2 n^{2\alpha}}} dt - \frac{1}{\sqrt{\pi} d n^\alpha} \int_{-\infty}^{0} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right| \\
&= \frac{1}{\sqrt{\pi} d n^\alpha} \left| \int_{nx-[n^2+n^\alpha]}^{nx-\lceil nx \rceil} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt - \int_{-\infty}^{0} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right| \\
&\le \frac{1}{\sqrt{\pi} d n^\alpha} \left(\int_{-1}^{0} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt + \int_{-\infty}^{-n^\alpha+1} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right) \\
&\le \frac{1}{\sqrt{\pi} d n^\alpha} + d e^{-\frac{1}{4d^2}}.
\end{aligned}$$

Combining the estimates of $|\Lambda_1|$ with $|\Lambda_2|$, we have

$$\Delta_1 \le \frac{2}{\sqrt{\pi} d n^\alpha} + d e^{-\frac{1}{4d^2}}.$$

Considering

$$0 \le nx - [nx] < 1, \quad nx - \lceil -n^2 - n^\alpha \rceil \ge n^\alpha - 1,$$

in the similar way we obtain the estimate of Δ_1 that

$$\begin{aligned}
|\Lambda_3| &:= \frac{1}{\sqrt{\pi} d n^\alpha} \left| \int_{nx-[nx]+1}^{nx-\lceil -n^2-n^\alpha \rceil+1} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt - \int_{0}^{+\infty} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right| \\
&\le \frac{1}{\sqrt{\pi} d n^\alpha} \left(\int_{0}^{2} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt + \int_{n^\alpha}^{\infty} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right) \\
&\le \frac{2}{\sqrt{\pi} d n^\alpha} + d e^{-\frac{1}{d^2}},
\end{aligned}$$

and

$$\begin{aligned}
|\Lambda_4| &:= \frac{1}{\sqrt{\pi} d n^\alpha} \left| \int_{nx-[nx]}^{nx-\lceil k=-n^2-n^\alpha \rceil} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt - \int_{0}^{+\infty} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right| \\
&\le \frac{1}{\sqrt{\pi} d n^\alpha} \left(\int_{0}^{1} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt + \int_{n^\alpha-1}^{\infty} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt \right) \\
&\le \frac{1}{\sqrt{\pi} d n^\alpha} + d e^{-\frac{1}{4d^2}}.
\end{aligned}$$

Thus, from the case 2) of Lemma 1 it follows that

$$\begin{aligned}
\Delta_3 &= \left| \sum_{k=\lceil -n^2-n^\alpha \rceil}^{[nx]-1} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha} \right) - \frac{1}{\sqrt{\pi}} \int_{0}^{+\infty} e^{-t^2} dt \right| \\
&\le \max\{|\Lambda_3|, |\Lambda_4|\} \\
&\le \frac{2}{\sqrt{\pi} d n^\alpha} + d e^{-\frac{1}{4d^2}}.
\end{aligned}$$

Hence,

$$I_2 \le \frac{6}{\sqrt{\pi} d n^\alpha} + 2de^{-\frac{1}{4d^2}}.$$

Therefore,

$$\begin{aligned}\left|G^1_{n,d}(f,x) - f(x)\right| &\le \omega\left(f; \frac{1}{n^{1-\alpha}}\right)\left(1 + \frac{1}{\sqrt{\pi} d n^\alpha}\right) \\ &\quad + 6\|f\|_\infty \left(\frac{1}{\sqrt{\pi} d n^\alpha} + de^{-\frac{1}{4d^2}}\right).\end{aligned}$$

□

If we choose d and n such that $n^{\frac{\alpha}{2}} d = 1$, then

$$\begin{aligned}\left|G^1_{n,d}(f,x) - f(x)\right| &\le 2\omega\left(f; \frac{1}{n^{1-\alpha}}\right) + 6\|f\|_\infty (n^{-\frac{\alpha}{2}} + d) \\ &= 2\omega\left(f; \frac{1}{n^{1-\alpha}}\right) + 12d\|f\|_\infty.\end{aligned}$$

Remark 1. *For given n, we choose $f(x) = 1$. Then, when $|x| > n + \frac{n^\alpha+1}{n}$, we have*

$$\sum_{k=\lceil -n^2-n^\alpha \rceil}^{[n^2+n^\alpha]} \frac{f(\frac{k}{n})}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) - f(x) = \sum_{k=\lceil -n^2-n^\alpha \rceil}^{[n^2+n^\alpha]} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) - 1,$$

and

$$\begin{aligned}\sum_{k=\lceil -n^2-n^\alpha \rceil}^{[n^2+n^\alpha]} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) &= \frac{1}{\sqrt{\pi} d n^\alpha} \sum_{k=\lceil -n^2-n^\alpha \rceil}^{[n^2+n^\alpha]} e^{-\frac{(nx-k)^2}{d^2 n^{2\alpha}}} \\ &\le \frac{1}{\sqrt{\pi} d n^\alpha} \int_0^{+\infty} e^{-\frac{t^2}{d^2 n^{2\alpha}}} dt = \frac{1}{2}.\end{aligned}$$

Thus

$$\left|\sum_{k=\lceil -n^2-n^\alpha \rceil}^{[n^2+n^\alpha]} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) - 1\right| \ge \frac{1}{2}.$$

This shows that when $|x| > n + \frac{n^\alpha+1}{n}$, $G^1_{n,d}(f,x) - f(x)$ has no the convergent degree in general.

Remark 2. *If $f \in C(\mathbb{R})$, then for $x \in \mathbb{R}$*

$$G^1_{n,d}(f,x) \to f(x), \quad \text{as } n \to \infty \text{ and } d \to 0 \ (\text{but } n^\alpha d \to \infty),$$

and the above convergence is pointwise.

Remark 3. *From Theorem 1 and Remark 1, we know that an essential change occurs to the error of* $G^1_{n,d}(f,x) - f(x)$ *on* $n < |x| < n + \frac{n^\alpha+1}{n}$. *However, if we construct following operators* $G^2_{n,d}(f,x)$ *defined by*

$$G^2_{n,d}(f,x) = \sum_{k=-\infty}^{\infty} \frac{f(\frac{k}{n})}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right),$$

then the error of $G^2_{n,d}(f,x) - f(x)$ *can be estimated. In fact, since*

$$\begin{aligned}
|G^2_{n,d}(f,x) - f(x)| &= \left|\sum_{k=-\infty}^{\infty} \frac{f(\frac{k}{n})}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) - f(x)\right| \\
&\le \omega\left(f; \frac{1}{n^{1-\alpha}}\right) \sum_{k=-\infty}^{\infty} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) \\
&\quad + 2\|f\|_\infty \sum_{|nx-k|>n^\alpha} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) \\
&\quad + \|f\|_\infty \left|\sum_{k=-\infty}^{\infty} \frac{1}{n^\alpha} g_d\left(\frac{nx-k}{n^\alpha}\right) - 1\right| \\
&\le \omega\left(f; \frac{1}{n^{1-\alpha}}\right)\left(1 + \frac{1}{\sqrt{\pi} d n^\alpha}\right) + 2d\|f\|_\infty e^{-\frac{1}{4d^2}} + \|f\|_\infty I_3,
\end{aligned}$$

and

$$I_3 \le \frac{6}{\sqrt{\pi} d n^\alpha},$$

we have

$$|G^2_{n,d}(f,x) - f(x)| \le 2\omega\left(f; \frac{1}{n^{1-\alpha}}\right) + 8d\|f\|_\infty,$$

provided $n^{\frac{\alpha}{2}} d = 1$.

3. Approximation on compact sets of $\mathbb{R}^2$

Write

$$G_d(t_1,t_2) = g_d(t_1)\, g_d(t_2). \tag{8}$$

For $f(x_1,x_2) \in C([-1,1]^2)$, we construct operators

$$G^3_{n,d}(f;x_1,x_2) = \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]} \sum_{k_2=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]} \frac{f(\frac{k_1}{n+n^\alpha}, \frac{k_2}{n+n^\alpha})}{n^{2\alpha}} G_d\left(\frac{nx_1-k_1}{n^\alpha}, \frac{nx_2-k_2}{n^\alpha}\right).$$

To estimate the error $G^3_{n,d}(f;x_1,x_2) - f(x_1,x_2)$, we need the following lemma.

Lemma 2. *For G_d given by (8), we have*

1)

$$\int_{\lceil -n-n^\alpha\rceil -1}^{[nx_1]-1}\int_{\lceil -n-n^\alpha\rceil -1}^{[nx_2]-1} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2$$
$$\le \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[nx_1]-1}\sum_{k_2=\lceil -n-n^\alpha\rceil}^{[nx_2]-1} G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)$$
$$\le \int_{\lceil -n-n^\alpha\rceil}^{[nx_1]}\int_{\lceil -n-n^\alpha\rceil}^{[nx_2]} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2;$$

2)

$$\int_{\lceil -n-n^\alpha\rceil -1}^{[nx_1]-1}\int_{\lceil nx_2\rceil +1}^{[n+n^\alpha]+1} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2$$
$$\le \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[nx_1]-1}\sum_{k_2=\lceil nx_2\rceil +1}^{[n+n^\alpha]} G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)$$
$$\le \int_{\lceil -n-n^\alpha\rceil}^{[nx_1]}\int_{\lceil nx_2\rceil}^{[n+n^\alpha]} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2;$$

3)

$$\int_{\lceil nx_1\rceil +1}^{[n+n^\alpha]+1}\int_{\lceil -n-n^\alpha\rceil -1}^{[nx_2]-1} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2$$
$$\le \sum_{k_1=\lceil nx_1\rceil +1}^{[n+n^\alpha]}\sum_{k_2=\lceil -n-n^\alpha\rceil}^{[nx_2]-1} G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)$$
$$\le \int_{\lceil nx_1\rceil}^{[n+n^\alpha]}\int_{\lceil -n-n^\alpha\rceil}^{[nx_2]} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2;$$

4)

$$\int_{\lceil nx_1\rceil +1}^{[n+n^\alpha]+1}\int_{\lceil nx_2\rceil +1}^{[n+n^\alpha]+1} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2$$
$$\le \sum_{k_1=\lceil nx_1\rceil +1}^{[n+n^\alpha]}\sum_{k_2=\lceil nx_2\rceil +1}^{[n+n^\alpha]} G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)$$
$$\le \int_{\lceil nx_1\rceil}^{[n+n^\alpha]}\int_{\lceil nx_2\rceil}^{[n+n^\alpha]} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2.$$

Proof. Noting that when $0\le t_i\le t_i'(i=1,2)$, $G_d\left(t_1,t_2\right)\le G_d\left(t_1',t_2'\right)$, and using the way of proving the case 2) of Lemma 1, we can finish the proof of case 1).

Now we prove the case 2). Noting that when $\lceil -n^2-n^\alpha\rceil \le k_1 \le [nx_1]-1, \lceil nx_2\rceil + 1 \le k_2 \le [n^2+n^\alpha]$,

$$\begin{aligned}
&\int_{k_1-1}^{k_1}\int_{k_2}^{k_2+1} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2\\
&\quad\le \int_{k_1-1}^{k_1}\int_{k_2}^{k_2+1} G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2\\
&\quad= \int_{k_2}^{k_2+1} G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_2 \le G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)\\
&\quad\le \int_{k_2-1}^{k_2} G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_2\\
&\quad\le \int_{k_1}^{k_1+1}\int_{k_2-1}^{k_2} G_d\left(\frac{nx_1-t_1}{n^\alpha},\frac{nx_2-t_2}{n^\alpha}\right)dt_1dt_2,
\end{aligned}$$

it is not difficult to complete the proof of case 2).

The proof of case 3) is similar to the proof of case 2).

The proof of case 4) is easy if we note that when $t_i \le t_i' \le 0 (i=1,2), G_d(t_1,t_2) \le G_d(t_1',t_2')$. We omit the details. $\square$

Now we give the main result of this section.

Theorem 2. *Let $f \in C([-1,1]^2)$, and $0<\alpha<1$. Then for $n^\alpha > 2$, we have*

$$\begin{aligned}
|G_{n,d}^3(f;x_1,x_2) - f(x_1,x_2)| &\le \omega\left(f;\frac{2}{n^{1-\alpha}},\frac{2}{n^{1-\alpha}}\right)\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)^2\\
&\quad + 4\|f\|_\infty\left(\left(\frac{3}{2}de^{-\frac{1}{4d^2}}+\frac{1}{\sqrt{\pi}dn^\alpha}\right)\left(3+\frac{1}{\sqrt{\pi}dn^\alpha}\right)^2+\frac{d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}}\right),
\end{aligned}$$

where $\|f\|_\infty = \sup\limits_{(x_1,x_2)\in[-1,1]^2}|f(x_1,x_2)|$.

Proof. It is not difficult to see that

$$\begin{aligned}
&G_{n,d}(f;x_1,x_2) - f(x_1,x_2)\\
&\quad= \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]}\sum_{k_2=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]} \frac{f(\frac{k_1}{n+n^\alpha},\frac{k_2}{n+n^\alpha})-f(x_1,x_2)}{n^{2\alpha}}G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)\\
&\qquad+f(x_1,x_2)\left(\sum_{k_1=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]}\sum_{k_2=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]}\frac{G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}}-1\right)\\
&\quad=: I_4+I_5.
\end{aligned}$$

Write

$$I_4 = \left[\sum_{|nx_1-k_1|\le n^\alpha}\sum_{|nx_2-k_2|\le n^\alpha} + \sum_{|nx_1-k_1|> n^\alpha}\sum_{|nx_2-k_2|\le n^\alpha}\right] + \left[\sum_{|nx_1-k_1|\le n^\alpha}\sum_{|nx_2-k_2|> n^\alpha} + \sum_{|nx_1-k_1|> n^\alpha}\sum_{|nx_2-k_2|> n^\alpha}\right] \frac{f(\frac{k_1}{n+n^\alpha},\frac{k_2}{n+n^\alpha}) - f(x_1,x_2)}{n^{2\alpha}} G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)$$
$$=: I_{41}+I_{42}+I_{43}+I_{44}.$$

From $-1\le x_i\le 1$ and $|nx_i-k_i|\le n^\alpha (i=1,2)$, we get

$$\left|\frac{k_i}{n+n^\alpha}-x_i\right| \le \frac{2n^\alpha}{n+n^\alpha} \le \frac{2}{n^{1-\alpha}},$$

which implies

$$|I_{41}| \le \omega\left(f;\frac{2}{n^{1-\alpha}},\frac{2}{n^{1-\alpha}}\right)\sum_{|nx_1-k_1|\le n^\alpha}\sum_{|nx_2-k_2|\le n^\alpha}\frac{1}{n^{2\alpha}}G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)$$
$$= \omega\left(f;\frac{2}{n^{1-\alpha}},\frac{2}{n^{1-\alpha}}\right)\left(\sum_{|nx-k|\le n^\alpha}\frac{1}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)\right)^2$$
$$\le \omega\left(f;\frac{2}{n^{1-\alpha}},\frac{2}{n^{1-\alpha}}\right)\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)^2,$$
$$|I_{42}| \le 2\|f\|_\infty \sum_{|nx_1-k_1|>n^\alpha}\frac{1}{n^\alpha}g_d\left(\frac{nx_1-k_1}{n^\alpha}\right)\sum_{k=-\infty}^{\infty}\frac{1}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)$$
$$\le 2d\|f\|_\infty e^{-\frac{1}{4d^2}}\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right),$$
$$|I_{43}| \le 2d\|f\|_\infty e^{-\frac{1}{4d^2}}\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right),$$

and

$$|I_{44}| \le 2d\|f\|_\infty e^{-\frac{1}{4d^2}}\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right),$$

where we have used the condition $n^\alpha > 2$ to estimate $\sum_{|nx-k|>n^\alpha}\frac{1}{n^\alpha}g_d\left(\frac{nx-k}{n^\alpha}\right)$. Therefore,

$$|I_4| \le \left(\omega(f;\frac{2}{n^{1-\alpha}},\frac{2}{n^{1-\alpha}}) + 6d\|f\|_\infty e^{-\frac{1}{4d^2}}\right)\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)^2.$$

On the other hand,

$$\begin{aligned}|I_5| \ \le\ & \|f\|_\infty \left| \sum_{k_1=\lceil -n^2-n^\alpha\rceil}^{[n^2+n^\alpha]} \sum_{k_2=\lceil -n^2-n^\alpha\rceil}^{[n^2+n^\alpha]} \frac{G_d\left(\frac{nx_1-k_1}{n^\alpha}, \frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}} - 1 \right| \\ =:& \|f\|_\infty \Delta_5.\end{aligned}$$

Set

$$\begin{aligned}D :=& \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[nx_1]-1} \sum_{k_2=\lceil -n-n^\alpha\rceil}^{[nx_2]-1} + \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[nx_1]-1} \sum_{k_2=\lceil nx_2\rceil+1}^{[n+n^\alpha]} \\ &+ \sum_{k_1=\lceil nx_1\rceil+1}^{[n+n^\alpha]} \sum_{k_2=\lceil -n-n^\alpha\rceil}^{[nx_2]-1} + \sum_{k_1=\lceil nx_1\rceil+1}^{[n+n^\alpha]} \sum_{k_2=\lceil nx_2\rceil+1}^{[n+n^\alpha]}.\end{aligned}$$

Then

$$\begin{aligned}|\Delta_5| \le & \left(D \frac{G_d\left(\frac{nx_1-k_1}{n^\alpha}, \frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}} - 1 \right) + \frac{g_d\left(\frac{nx_1-[nx_1]}{n^\alpha}\right)}{n^{2\alpha}} \sum_{k_2=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]} g_d\left(\frac{nx_2-k_2}{n^\alpha}\right) \\ &+ \frac{g_d\left(\frac{nx_1-\lceil nx_1\rceil}{n^\alpha}\right)}{n^{2\alpha}} \sum_{k_2=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]} g_d\left(\frac{nx_2-k_2}{n^\alpha}\right) \\ &+ \frac{g_d\left(\frac{nx_2-[nx_2]}{n^\alpha}\right)}{n^{2\alpha}} \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]} g_d\left(\frac{nx_1-k_1}{n^\alpha}\right) \\ &+ \frac{g_d\left(\frac{nx_2-\lceil nx_2\rceil}{n^\alpha}\right)}{n^{2\alpha}} \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[n+n^\alpha]} g_d\left(\frac{nx_1-k_1}{n^\alpha}\right) \\ \le & \left(D \frac{G_d\left(\frac{nx_1-k_1}{n^\alpha}, \frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}} - 1 \right) + \frac{4}{\sqrt{\pi} d n^\alpha} \left(1 + \frac{1}{\sqrt{\pi} d n^\alpha}\right).\end{aligned}$$

From the case 1) of Lemma 2 it follows that

$$\begin{aligned}\Omega_L :=& \int_{\lceil -n-n^\alpha\rceil -1}^{[nx_1]-1} \int_{\lceil -n-n^\alpha\rceil -1}^{[nx_2]-1} \frac{1}{n^{2\alpha}} G_d\left(\frac{nx_1-t_1}{n^\alpha}, \frac{nx_2-t_2}{n^\alpha}\right) dt_1 dt_2 \\ \le & \sum_{k_1=\lceil -n-n^\alpha\rceil}^{[nx_1]-1} \sum_{k_2=\lceil -n-n^\alpha\rceil}^{[nx_2]-1} \frac{1}{n^{2\alpha}} G_d\left(\frac{nx_1-k_1}{n^\alpha}, \frac{nx_2-k_2}{n^\alpha}\right) \\ \le & \int_{\lceil -n-n^\alpha\rceil}^{[nx_1]} \int_{\lceil -n-n^\alpha\rceil}^{[nx_2]} \frac{1}{n^{2\alpha}} G_d\left(\frac{nx_1-t_1}{n^\alpha}, \frac{nx_2-t_2}{n^\alpha}\right) dt_1 dt_2 =: \Omega_R.\end{aligned}$$

Since

$$0 \le \frac{nx_i-[nx_i]}{n^\alpha} \le \frac{1}{n^\alpha}, \quad \frac{nx_i-\lceil -n-n^\alpha\rceil}{n^\alpha} > 1 - \frac{1}{n^\alpha}, \quad i=1,2,$$

$$\frac{1}{n^\alpha} \le \frac{nx_i-[nx_i]+1}{n^\alpha} \le \frac{2}{n^\alpha}, \quad \frac{nx_i-\lceil -n-n^\alpha\rceil+1}{n^\alpha} > 1, \quad i=1,2,$$

and

$$\int_0^{+\infty} g_d(t_1)\,dt_1 \int_1^{+\infty} g_d(t_2)\,dt_2 \le \frac{1}{2}\cdot\frac{1}{\sqrt{\pi}}\int_{\frac{1}{d}}^{+\infty} e^{-x^2}dx,$$
$$\int_a^{\infty} e^{-x^2}dx < \min\left(\frac{\sqrt{\pi}}{2}e^{-a^2}, \frac{1}{2a}e^{-a^2}\right), \quad a>0,$$

we obtain

$$\begin{aligned}
\Xi_L &:= \left|\Omega_L - \int_0^{+\infty}\int_0^{+\infty} g_d(t_1)\,g_d(t_2)\,dt_1dt_2\right| \\
&= \left|\left(\int_{\frac{nx_1-[nx_1]}{n^\alpha}}^{\frac{nx_1-\lceil -n-n^\alpha\rceil}{n^\alpha}} \int_{\frac{nx_2-[nx_2]}{n^\alpha}}^{\frac{nx_2-\lceil -n-n^\alpha\rceil}{n^\alpha}} - \int_0^{+\infty}\int_0^{+\infty}\right) g_d(t_1)\,g_d(t_2)\,dt_1dt_2\right| \\
&\le \int_0^{\frac{2}{n^\alpha}} g_d(t_1)\,dt_1 \int_0^{+\infty} g_d(t_2)\,dt_2 + \int_0^{+\infty} g_d(t_1)\,dt_1 \int_0^{\frac{2}{n^\alpha}} g_d(t_2)\,dt_2 \\
&\quad + \int_0^{+\infty} g_d(t_1)\,dt_1 \int_1^{+\infty} g_d(t_2)\,dt_2 + \int_1^{+\infty} g_d(t_1)\,dt_1 \int_0^{+\infty} g_d(t_2)\,dt_2 \\
&\le \frac{2}{\sqrt{\pi}dn^\alpha} + \frac{d}{2\sqrt{\pi}}e^{-\frac{1}{d^2}},
\end{aligned}$$

and

$$\begin{aligned}
\Xi_R &:= \left|\Omega_R - \int_0^{+\infty}\int_0^{+\infty} g_d(t_1)\,g_d(t_2)\,dt_1dt_2\right| \\
&= \left|\left(\int_{\frac{nx_1-[nx_1]+1}{n^\alpha}}^{\frac{nx_1-\lceil -n-n^\alpha\rceil+1}{n^\alpha}} \int_{\frac{nx_2-[nx_2]+1}{n^\alpha}}^{\frac{nx_2-\lceil -n-n^\alpha\rceil+1}{n^\alpha}} - \int_0^{+\infty}\int_0^{+\infty}\right) g_d(t_1)\,g_d(t_2)\,dt_1dt_2\right| \\
&\le \int_0^{\frac{2}{n^\alpha}} g_d(t_1)\,dt_1 \int_0^{+\infty} g_d(t_2)\,dt_2 + \int_0^{+\infty} g_d(t_1)\,dt_1 \int_0^{\frac{1}{n^\alpha}} g_d(t_2)\,dt_2 \\
&\quad + \int_0^{+\infty} g_d(t_1)\,dt_1 \int_{1-\frac{1}{n^\alpha}}^{+\infty} g_d(t_2)\,dt_2 + \int_{1-\frac{1}{n^\alpha}}^{+\infty} g_d(t_1)\,dt_1 \int_0^{+\infty} g_d(t_2)\,dt_2 \\
&\le \frac{1}{\sqrt{\pi}dn^\alpha} + \frac{d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}}.
\end{aligned}$$

Thus

$$\begin{aligned}
&\left|\sum_{k_1=\lceil -n-n^\alpha\rceil}^{[nx_1]-1}\sum_{k_2=\lceil -n-n^\alpha\rceil}^{[nx_2]-1} \frac{G_d\left(\frac{nx_1-k_1}{n^\alpha}, \frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}} - \int_0^{+\infty}\int_0^{+\infty} g_d(t_1)\,g_d(t_2)\,dt_1dt_2\right| \\
&\qquad \le \max\{\Xi_L, \Xi_R\} \le \frac{2}{\sqrt{\pi}dn^\alpha} + \frac{d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}}.
\end{aligned}$$

We use the cases 2), 3) and 4) of Lemma 2 and obtain in a similar way that

$$\left|\sum_{k_1=\lceil -n-n^\alpha\rceil}^{[nx_1]-1}\sum_{k_2=\lceil nx_2\rceil+1}^{[n+n^\alpha]}\frac{G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}}-\int_0^{+\infty}\int_{-\infty}^0 g_d\left(t_1\right)g_d\left(t_2\right)dt_1dt_2\right|$$
$$\le \frac{2}{\sqrt{\pi}dn^\alpha}+\frac{d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}},$$
$$\left|\sum_{k_1=\lceil nx_1\rceil+1}^{[n+n^\alpha]}\sum_{k_2=\lceil -n-n^\alpha\rceil}^{[nx_2]-1}\frac{G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}}-\int_{-\infty}^0\int_0^{+\infty} g_d\left(t_1\right)g_d\left(t_2\right)dt_1dt_2\right|$$
$$\le \frac{2}{\sqrt{\pi}dn^\alpha}+\frac{d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}},$$

and

$$\left|\sum_{k_1=\lceil nx_1\rceil+1}^{[n+n^\alpha]}\sum_{k_2=\lceil nx_2\rceil+1}^{[n+n^\alpha]}\frac{G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}}-\int_{-\infty}^0\int_{-\infty}^0 g_d\left(t_1\right)g_d\left(t_2\right)dt_1dt_2\right|$$
$$\le \frac{2}{\sqrt{\pi}dn^\alpha}+\frac{d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}}.$$

These imply that

$$\left|D\frac{G_d\left(\frac{nx_1-k_1}{n^\alpha},\frac{nx_2-k_2}{n^\alpha}\right)}{n^{2\alpha}}-1\right|\le \frac{8}{\sqrt{\pi}dn^\alpha}+\frac{4d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}}.$$

Therefore, we have

$$\begin{aligned}
&|G^3_{n,d}(f;x_1,x_2)-f(x_1,x_2)|\\
&\quad\le \omega\left(f;\frac{2}{n^{1-\alpha}},\frac{2}{n^{1-\alpha}}\right)\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)^2\\
&\qquad+6d\|f\|_\infty e^{-\frac{1}{4d^2}}\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)^2\\
&\qquad+\|f\|_\infty\left(\frac{4}{\sqrt{\pi}dn^\alpha}\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)+\frac{8}{\sqrt{\pi}dn^\alpha}+\frac{4d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}}\right)\\
&\quad\le \omega\left(f;\frac{2}{n^{1-\alpha}},\frac{2}{n^{1-\alpha}}\right)\left(1+\frac{1}{\sqrt{\pi}dn^\alpha}\right)^2\\
&\qquad+4\|f\|_\infty\left(\left(\frac{3}{2}de^{-\frac{1}{4d^2}}+\frac{1}{\sqrt{\pi}dn^\alpha}\right)\left(3+\frac{1}{\sqrt{\pi}dn^\alpha}\right)^2+\frac{d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}}\right).
\end{aligned}$$

□

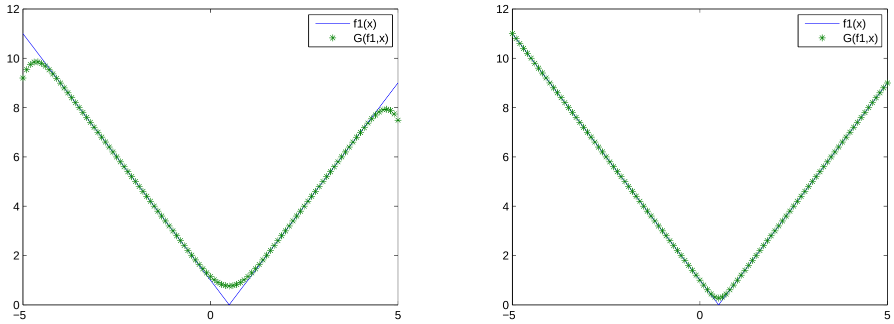
If we choose d and n such that $n^{\frac{\alpha}{2}}d=1$, then

$$|G^3_{n,d}(f;x_1,x_2)-f(x_1,x_2)|\le 4\omega\left(f;\frac{2}{n^{1-\alpha}},\frac{2}{n^{1-\alpha}}\right)+164d\|f\|_\infty.$$

4. Numerical experiments and discussions

In this section, w well give some numerical experiments and discussions.

Example 1. *We take the target function* $f_1(x) = |1 - 2x|$, *and consider the approximation network* $G^1_{n,d}(f_1, x)$ *defined in Section 2. Taking* $n = 5$ *and* $\alpha = 0.5$, *Figure 1 shows the approximation effect with different values of d (*$d = 1.5$ *and* $d = 0.5$, *respectively).*



(a) : $n = 5$, $d = 1.5$, $\alpha = 0.5$ (b) : $n = 5$, $d = 0.5$, $\alpha = 0.5$

Figure 1: *Approximation effect of* $G^1_{n,d}(f_1, x)$ *with different values of d*

We denote the approximation error by $error1 = \max_{x\in[-n,n]} |f_1(x) - G^1_{n,d}(f_1, x)|$. *In Table 1, with* $d = 0.35$ *and* $\alpha = 0.5$, *we give the approximation error for* f_1 *with different values of* n. *If we fix* $n = 5$ *and* $\alpha = 0.5$, *the approximation error with different values of d is shown in Table 2. It can be verified that* $f_1 \in Lip_2 1$ *and* $\|f_1\|_\infty = 11$, *so we have* $\omega(f_1; \frac{1}{n^{1-\alpha}}) \leq \frac{2}{n^{1-\alpha}}$. *We then denote the control error (see the right side of the result in Theorem 1) by*

$$con(error1) = \frac{2}{n^{1-\alpha}}\left(1 + \frac{1}{\sqrt{\pi} d n^\alpha}\right) + 66\left(\frac{1}{\sqrt{\pi} d n^\alpha} + de^{-\frac{1}{4d^2}}\right).$$

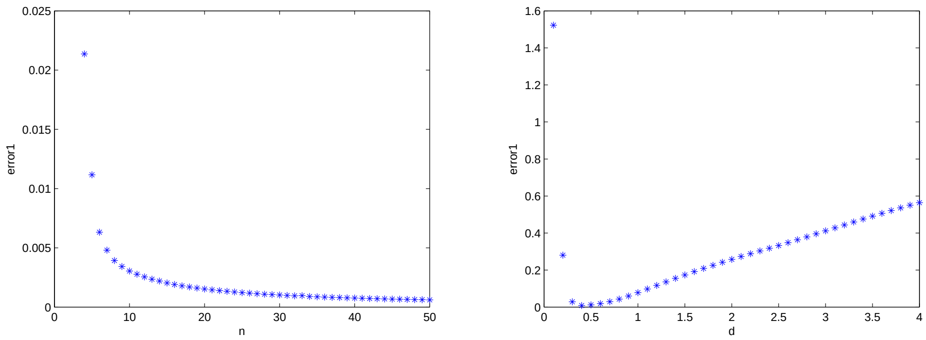
For $d = 0.35$ *and* $\alpha = 0.5$, *Figure 2 (a) shows that the approximation error decreases as n increases. If we fix* $n = 5$ *and* $\alpha = 0.5$, *we can see the changes of the approximation error with different values of d in Figure 2 (b), where the error is smallest when d in close to 0.35 (with n and α fixed). In Figure 3, we show the change of control error about different values of n and d, where the trend is the same as in Figure 2, so that the error estimate in Theorem 1 is reasonable and effective.*

n	5	10	25	50	100	200
error1	0.2137	0.1061	0.0813	0.0543	0.0390	0.0277

Table 1: *Approximation error for* f_1 *with different values of* n

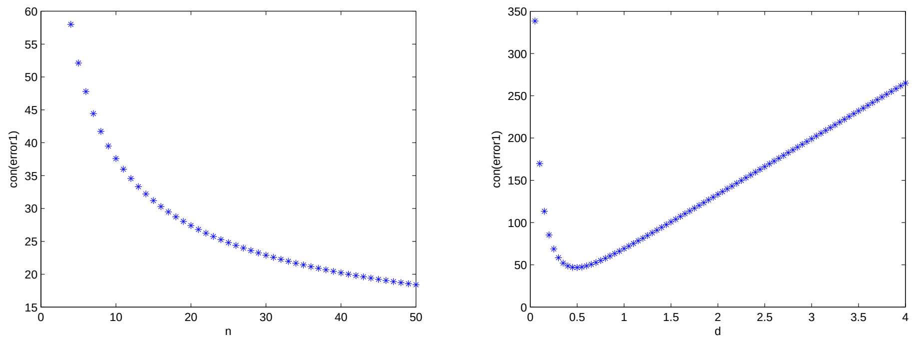
d	0.25	0.35	0.4	0.45	0.5	1
error1	1.0069	0.2137	0.2306	0.2505	0.2722	0.6632

Table 2: *Approximation error for f_1 with different values of d*



(a) : $d = 0.35$, $\alpha = 0.5$ (b) : $n = 5$, $\alpha = 0.5$

Figure 2: *Approximation error of $G^1_{n,d}(f_1, x)$ about different values of n and d*



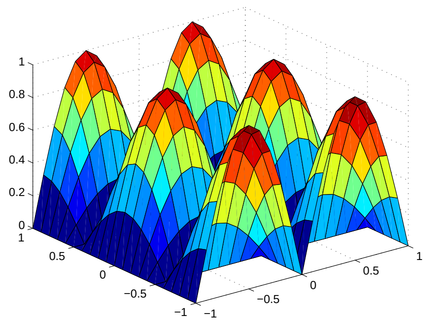
(a) : $d = 0.35, a = 0.5$ (b) : $n = 5, a = 0.5$

Figure 3: *Control error of $G^1_{n,d}(f_1, x)$ about different values of n and d*

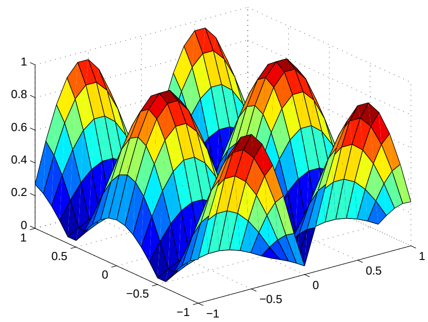
Example 2. *For the case of approximation on compact sets of $\mathbb{R}^2$. We take the target function $f_2(x_1, x_2) = |\sin(\pi x_1)\cos(\pi x_2)|$, $x_1, x_2 \in [-1, 1]$, and consider the approximation network $G^3_{n,d}(f; x_1, x_2)$ defined in Section 3. With $d = 0.35$ and $\alpha = 0.5$, Figure 4 shows the approximation effect with different values of n. We denote the approximation error of f_2 by*

$$\textit{error2} = \max_{x_1, x_2 \in [-1,1]} \left| f_2(x_1, x_2) - G^3_{n,d}(f_2; x_1, x_2) \right|.$$

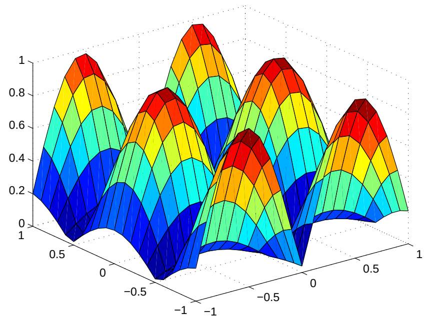
In Table 3, with $d = 0.35$ and $\alpha = 0.5$, we give the approximation error for f_2 with different values of n. If we fix $n = 5$ and $\alpha = 0.5$, the approximation error with different values of d is shown in Table 4.



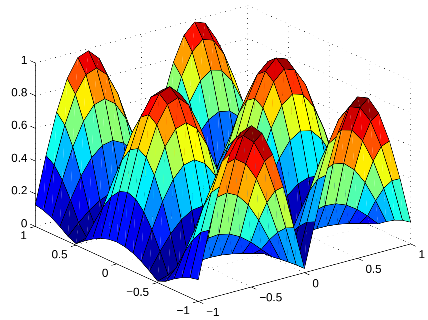
(a) : $f_2(x_1, x_2) = |\sin(\pi x_1)\cos(\pi x_2)|$



(b) : $G^3_{n,d}(f; x_1, x_2), n = 100, d = 0.35, a = 0.5$



(c) : $G^3_{n,d}(f; x_1, x_2), n = 200, d = 0.35, a = 0.5$



(d) : $G^3_{n,d}(f; x_1, x_2), n = 500, d = 0.35, a = 0.5$

Figure 4: *Approximation effect of* $G^3_{n,d}(f; x_1, x_2)$ *with different values of* n

n	5	25	50	100	200	500
error2	0.7892	0.4917	0.3760	0.2803	0.2054	0.1339

Table 3: *Approximation error for* f_2 *with different values of* n, $d = 0.35$, $\alpha = 0.5$

d	0.25	0.35	0.4	0.6	0.8	1.5
error2	0.9696	0.7892	0.7673	0.6975	0.6184	0.6664

Table 4: *Approximation error for* f_2 *with different values of* d, $n = 5$, $\alpha = 0.5$

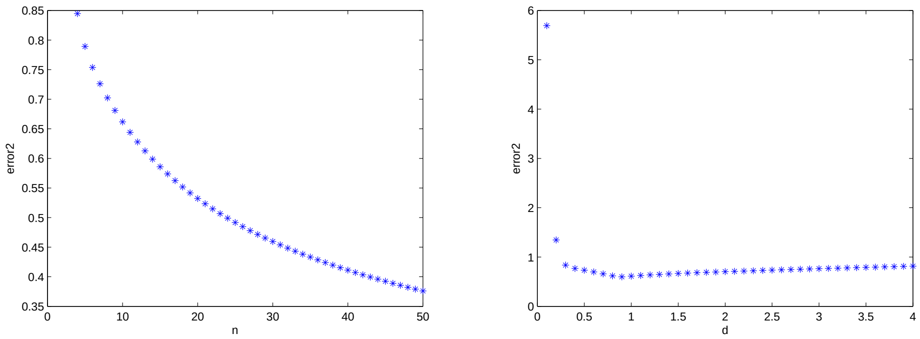
It can be verified that $f_2 \in Lip_{2\pi}1$ *and* $\|f_2\|_\infty = 1$, *so we have*

$$\omega(f_2; \frac{2}{n^{1-\alpha}}, \frac{2}{n^{1-\alpha}}) \leq \frac{8\pi}{n^{1-\alpha}}.$$

We then denote the corresponding control error for f_2 *(see the right-hand side of the* *result in Theorem 2) by*

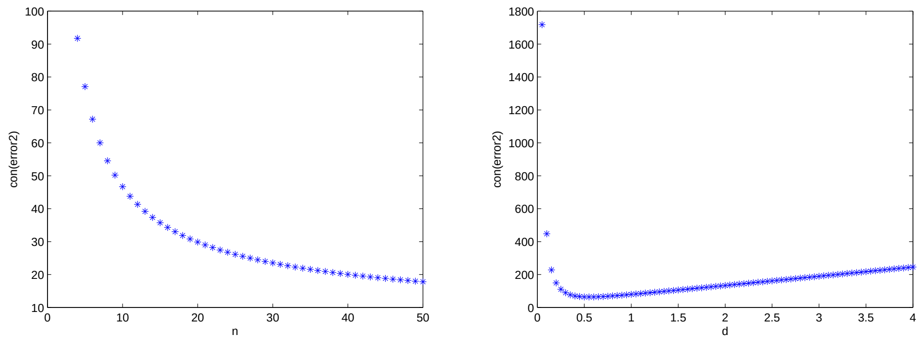
$$
\begin{aligned}
con(error2) = {} & \frac{8\pi}{n^{1-\alpha}}\left(1+\frac{1}{\sqrt{\pi}dn^{\alpha}}\right)^2 \\
& +4\left(\left(\frac{3}{2}de^{-\frac{1}{4d^2}}+\frac{1}{\sqrt{\pi}dn^{\alpha}}\right)\left(3+\frac{1}{\sqrt{\pi}dn^{\alpha}}\right)^2+\frac{d}{\sqrt{\pi}}e^{-\frac{1}{4d^2}}\right).
\end{aligned}
$$

For $d = 0.35$ *and* $\alpha = 0.5$*, Figure 5 (a) shows that the approximation error decreases as* n *increases. If we fix* $n = 5$ *and* $\alpha = 0.5$*, we can see the changes of the approximation error with different values of* d *in Figure 5 (b). In Figure 6, we show the change of control error with different values of* n *and* d*, where the trend is the same as in Figure 5, so that the error estimate in Theorem 2 is reasonable and effective.*



(a) : $d = 0.35$, $\alpha = 0.5$ (b) : $n = 5$, $\alpha = 0.5$

Figure 5: *Approximation error for* f_2 *with different values of* n *and* d



(a) : $d = 0.35$, $\alpha = 0.5$ (b) : $n = 5$, $\alpha = 0.5$

Figure 6: *Control error for* f_2 *with different values of* n *and* d

Remark 4. *It is easy to see that we can obtain corresponding results by replacing* G *given by (7) with* ϕ *defined by (5).*

Remark 5. *For $s > 2$, we can construct network operators similarly to approximate functions $f \in C([-1,1]^s)$ with activation function*

$$G(x) = \frac{1}{\pi^{\frac{s}{2}}} e^{-\|x\|^2} = \frac{1}{\sqrt{\pi}} e^{-x_1^2} \cdots \frac{1}{\sqrt{\pi}} e^{-x_s^2}, \quad x = (x_1, \ldots, x_s) \in \mathbb{R}^s.$$

Remark 6. *For a general centered bell-shaped activation function we can construct network operators and obtain the error estimates, provided that the decay rate of activation function $\sigma(x)$ is given when $|x| \to \infty$.*

Remark 7. *For general centered bell-shaped activation functions we can extend to more than 2 dimensions.*

Remark 8. *Network operators constructed in the paper are the form of feed-forward neural networks with a single hidden layer. Also, they are quasi-interpolation operators. However, they are not quasi-interpolation with polynomial reproduction. Because the coefficients of network operators are related with the target function, the operators include the information of the target function. Therefore, compared to the existing algorithms, such as interpolation operators, we do not need to solve the linear system to obtain the coefficients, which results in lower computational complexity and good convergence. The given numerical results also agree with the rate claimed in theorems.*

Acknowledgement

The authors would like to express their appreciation to the reviewers for their valuable comments and suggestions.

References

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

[2] G. A. Anastassiou, *Rate of convergence of some multivariate neural network operators to the unit*, Comput. Math. Appl. **40**(2000), 1–19.

[3] G. A. Anastassiou, *Quantitative Approximation*, Chapman & Hall/CRC, New York, 2001.

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

[5] A. R. Barrom, *Universial approximation bounds for superpositions of a sigmoidal function*, IEEE Trans. Inform. Theory **39**(1993), 930–945.

[6] R. K. Beatson, W. A. Light, *Quasi-interpolation in the absence of polynomial reproduction*, In: *Numerical Methods of Approximation Theory*, (L. L. Schumaker, Ed.), Birkhäuser, 1992, 21–39.

[7] J. P. Boyd, L. Wang, *Truncated Gaussian RBF differences are always inferior to finite differences of the same stencil width*, Commun. Comput. Phys. **5**(2009), 42–60.

[8] J. P. Boyd, L. Wang, *An analytic approximation to the cardinal functions of Gaussian radial basis functions on a one-dimensional infinite uniform lattice*, Appl. Math. Comput. **215**(2009), 2215–2223.

[9] J. P. BOYD, *Error saturation in Gaussian radial basis function functions on a finite interval*, J. Comput. Appl. Math. **234**(2010), 1435–1441.

[10] J. P. BOYD, L. WANG, *Asymptotic coefficients for Gaussian radial basis function interpolants*, Appl. Math. Comput. **216**(2010), 2394–2407.

[11] F. L. CAO, T. F. XIE, Z. B. XU, *The estimate for approximation error of neural networks:A constructive approach*, Neurocomputing **71**(2008), 626–630.

[12] P. CARDALIAGUET, C. EUVRARD, *Approximation of a function and its derivative with a neural network*, Neural Networks **5**(1992), 207–220.

[13] D. B. CHEN, *Degree of approximation by superpositions of a sigmoidal function*, J. Approx. Theory Appl. **9**(1993), 17–28.

[14] T. P. CHEN, H. CHEN, *Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its qpplication to dynamical system*, IEEE Trans. Neural Networks **6**(1995), 911–917.

[15] Z. X. CHEN, F. L. CAO, *Global errors for approximate approximations with Gaussian kernels on compact intervals*, Appl. Math. Comput. **217**(2010), 725–734.

[16] C. K. CHUI, X. LI, *Approximation by ridge functions and neural networks with one hidden layer*, J. Approx. Theory **70**(1992), 131–141.

[17] G. CYBENKO, *Approximation by superpositions of sigmoidal function*, Math. Control Signals syst. **2**(1989), 303–314.

[18] S. FERRARI, R. F. STENGEL, *Smooth function approximation using neural networks*, IEEE Trans. Neural Networks **16**(2005), 24–38.

[19] K. I. FUNAHASHI, *On the approximate realization of continuous mappings by neural networks*, Neural Networks **2**(1989), 183–192.

[20] E. J. HARTMAN, J. D. KEELER, J. M. KOWALSKI, *Layered neural networks with Gaussian hidden units as universal approximation*, Neural Computation **2**(1990), 210–215.

[21] K. HORNIK, M. STINCHCOMBE, H. WHITE, *Multilayer feedforward networks are universal approximation*, Neural Networks **2**(1989), 359–366.

[22] P. C. KAINEN, V. KURKOVÁ, M. SANGUINETI, *Complexity of Gaussian-radial-basis networks approximating smooth functions*, J. Complexity **25**(2009), 63–74.

[23] M. LESHNO, V. Y. LIN, A. PINKS, S. SCHOCKEN, *Multilayer feedforward networks with a nonpolynomial activation function can approximate any function*, Neural Networks **6**(1993), 861–867.

[24] G. G. LORENTZ, *Approximation of Functions*, Holt, Rinehart and Winston, 1966.

[25] V. MAŹYA, G. SCHMIDT, *On approximate approximations using Gaussian kernels*, IMA J. Num. Anal. **16**(1996), 13–29.

[26] V. MAŹYA, G. SCHMIDT, A*pproximate Approximations*, American Mathematical Society, Providence, 2007.

[27] H. N. MHASKAR, C. A. MICCHELLI, *Approximation by superposition of a sigmoidal function*, Adv. Appl. Math. **13**(1992), 350–373.

[28] H. N. MHASKAR, C. A. MICCHELLI, *Degree of approximation by neural networks with a single hidden layer*, Adv. Appl. Math. **16**(1995), 151–183.

[29] H. N. MHASKAR, *When is approximation by Gaussian networks necessarily a linear process?*, Neural Networks **17**(2004), 989–1001.

[30] V. MAIOROV, R. S. MEIR, *Approximation bounds for smooth functions in $C(R^d)$ by neural and mixture networks*, IEEE Trans. Neural Networks **9**(1998), 969–978.

[31] Y. MAKOVOZ, *Uniform approximation by neural networks*, J. Approx. Theory **95**(1998), 215–228.

[32] F. MÜLLER, W. VARNHORN, *Error estimates for approximate approximation with Gaussian kernels on compact intervals*, J. Approx. Theory **145**(2007), 171–181.

[33] D. S. MITRINOVIC, *Analytic Inequalities*, Springer-Verlag, Berlin, 1970.

[34] S. M. NIKOLSKII, *Approximation of Functions of Several Variables and Imbedding Theorems*, Springer-Verlag, Berlin, 1975.

[35] T. SCHLUMPRECHT, N. SIVAKUMAR, *On the sampling and recovery of bandlimited functions via scattered translates of the Gaussian*, J. Approx. Theory **159**(2009), 128–153.

[36] S. SUZUKI, *Constructive function approximation by three-layer artificial neural networks*, Neural Networks **11**(1998), 1049–1058.

[37] H. WENDLAND, *Scattered Data Approximation*, Cambridge University Press, Cambridge, 2005.

[38] Z. B. XU, F. L. CAO, *The essential order of approximation for neural networks*, Sci. China Ser. F **47**(2004), 97–112.