Neural Process Lett
DOI 10.1007/s11063-014-9361-x

Error Estimate for Spherical Neural Networks Interpolation

Shaobo Lin · Jinshan Zeng · Zongben Xu

Abstract We present a type of spherical neural network (SNN) with bounded sigmoidal activation function and study its interpolation capability. We find that the provided SNN can exactly interpolate the training samples. Furthermore, based on the special structure of the presented SNN, we can bound the interpolation error by the modulus of smoothness of the target function, which is different from the previous results on the spherical scattered data interpolation problem.

Keywords Spherical neural network · Exact interpolation · Error estimate · Sphere

Mathematics Subject Classification (2000) 41A25 · 41A05 · 41A63

1 Introduction

The problem of fitting scattered data arises mainly from sampling an unknown function defined on the sphere, and comes up frequently in many applications. These applications emerge in many research fields such as geodesy, meteorology, astrophysics and geophysics. More information on this topic can be referred to [10,11], and references therein. A routine way to tackle such a problem is to interpolate the scattered data by a class of functions. The success of this approach is based on many criteria including the cost of producing

The research was supported by the National 973 Programming (2013CB329404), and the Key Program of National Natural Science Foundation of China (Grant No. 11131006).

S. Lin (✉)
School of Mathematics and Information Science, Wenzhou University, Wenzhou 325035, Zhejiang, People's Republic of China
e-mail: sblin1983@gmail.com

S. Lin · J. Zeng · Z. Xu
School of Mathematics and Statistics, Institute for Information and System Sciences, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, People's Republic of China
e-mail: jsh.zeng@gmail.com

Published online: 17 May 2014

the interpolant, the robustness of the interpolation process, and how well the interpolant approximating the target function.

Up till now, there have been several approaches to interpolate spherical scattered data. A classical method is the spherical harmonics (SHs) interpolation. In particular, Womersley and Sloan [27] constructed a spherical harmonic Lagrange interpolant and studied the approximation properties of it; Narcowich and Ward [22] proved that there exist SHs which can both exactly interpolate the samples and near-best approximate the target function. For more details of SHs interpolation, we refer the readers to [12,23,24,26] and references therein.SHs have been extensively used in the context of spherical scattered data fitting, however, there exists a main problem called the curse of dimensionality for this approach, which makes the parameters increase rapidly as the dimensionality of the problem increases.

Several strategies have been suggested to circumvent this problem. The simplest approach is to model a multivariate function defined on the sphere as a sum of univariate functions. Thus, spherical radial basis function networks (SRBFNs) and spherical neural networks (SNNs) come into our sights. For SRBFNs, by using a strictly positive definite univariate function as the activation function, Jetter and Stökler [14] got an error estimate for interpolation by SRBFNs in a spherical Sobolev space. Narcowich and Ward [22] improved the results of [14] by conquering the well-known "barrier of native space" problem. Five years later, Narcowich et al. [23] improved the results of [22] again by deducing a Sobolev error estimate for interpolation by SRBFNs. Some related topics about SRBFNs interpolation can be found in [10,11,17,25,26]. Recently, Lin et al. [19] proved that the essential approximation rate can be improved for a large number of functions by introducing SNNs. In [20], a Jackson-type error estimate for the SNNs approximant was also established. Thus, there naturally arises the following question: what about the interpolation capability of SNNs? In this paper, we focus on giving a detailed study for this problem.

The SNN interpolants in this paper can be mathematically expressed as

$$N_n(x) := \sum_{i=1}^{n} c_i \sigma(g_i(x)), \quad x \in \mathbb{S}^d, \tag{1}$$

where $c_i \in \mathbb{R}$ is the output weight, $g_i : \mathbb{S}^d \to \mathbb{R}$ is the inner processing functions, a typical example of which takes the form as $g_i(x) = w_i \cdot x + b_i$ with $w_i \in \mathbb{R}^d$ and $b_i \in \mathbb{R}$, and $\sigma : \mathbb{R} \to \mathbb{R}$ is the activation function of SNNs. We denote by $\Phi_{\sigma,n}$ the family of SNNs, i.e.

$$\Phi_{\sigma,N} := \left\{ N_n(x) = \sum_{i=1}^{n} c_i \sigma(g_i(x)), \quad x \in \mathbb{S}^d \right\}.$$

Taking account into the fact that the scattered points are located on the sphere, the SNN method can be used to solve the following problem.

Problem. Let $X := \{x_i\}_{i=1}^n$ be a set of distinct points located on the unit sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ and let $f(x_i)$ denote the corresponding values of an unknown target function $f : \mathbb{S}^d \to \mathbb{R}$. Find an SNN formed as (1) such that the following interpolation condition

$$N_n(x_i) = f_i := f(x_i), \quad i = 1, \ldots, n \tag{2}$$

holds.

In [20], we have presented a type of SNN with

$$g_i(x) = A_i g(x) + b_i,$$

and study the approximation properties of this type of SNN, where $g(x)$ is a general distance function between x and the north pole of $\mathbb{S}^d$. In this paper, we will prove that such type of SNN also possesses prominent interpolation capability. In fact, we will prove that there exists an exact interpolation SNN for arbitrary samples. Furthermore, we will deduce an upper bound estimate for the SNN interpolation via taking advantage of the special structure of the SNN.Compared with the previous interpolation methods studied in [14,22,23], the target functions in these methods are assumed to belong to some smooth classes, while the restriction on the target function f in our method can be very mild. Indeed, we will deduce an interpolation error estimate for arbitrary continuous function, and a modulus of smoothness of it will be introduced to describe this error.

The rest of paper is organized as follows. In the next section, we will introduce some preliminaries. In Sect. 3, we will give an existence proof of the exact interpolation SNN. In Sect. 4, we will deduce the interpolation error estimate for the SNN method.

2 Preliminaries

In this section, we give some preliminaries for this paper. We first introduce three quantities associated with the scattered data set X. The mesh norm of X is defined by

$$h_X := \max_{x\in\mathbf{S}^d} \min_j d(x, x_j),$$

where $d(x, y)$ is the geodesic (great circle) distance between the points x and y on $\mathbf{S}^d$. The mesh norm measures the maximum distance any point on $\mathbf{S}^d$ can be from X. The separation radius is defined via

$$q_X := \frac{1}{2} \min_{j\neq k} d(x_j, x_k).$$

This is half of the smallest geodesic distance between any two distinct points in X. It is easy to see that $h_X \geq q_X$, and the equality can hold only for a uniform distribution of point on $\mathbf{S}^1$, the circle. The mesh ratio

$$\tau_X := \frac{h_X}{q_X} \geq 1$$

provides a measure of how uniformly points in X are distributed on $\mathbf{S}^d$.

Now we introduce a general distance corresponding to X on the sphere which can be found in [20]. At first we rearrange the points in X to obey the following three rules:

(A1) x_1 can be chosen arbitrary.
(A2) $d(x_k, x_{k+1}) \leq 4h_X, k = 1, 2, \ldots, n-1$.
(A3) For $j \neq k$, $\overline{x_k x_{k+1}} \cap \overline{x_j x_{j+1}} = \begin{cases} x_{k+1}, & j = k+1, \\ x_k, & j = k-1, \\ \varnothing, & \text{otherwise}, \end{cases}$
where $\overline{x_k x_{k+1}}$ is the segment of minor arc of the great circle from x_k to x_{k+1}.

It was proved in [20] that the arrangement satisfying (A1), (A2) and (A3) exists and can be easily implemented.

From the definition of h_X, it follows that $\mathbf{S}^d \subset \bigcup_{i=1}^n D(x_i, h_X)$, where $D(x, h)$ is the spherical cap with center x and radius h. Thus, for arbitrary $x \in \mathbf{S}^d$ there exists at least one point such that $x \in D(x_k, h_X)$. If we set

$$k := \min\{j : x \in D(x_j, h_X)\}, \tag{3}$$

then for arbitrary $x \in \mathbf{S}^d$, there exists a unique k satisfying (3) such that $x \in D(x_k, h_X)$. For arbitrary points $x, y \in \mathbf{S}^d$, we define a general distance between $x \in D(x_{k_0}, h_X)$ and $y \in D(x_{j_0}, h_X)$ as

$$\overline{d}(x, y) := \begin{cases} d(x, y), & k_0 = j_0, \\ \sum_{i=j_0}^{k_0} d(x_i, x_{i+1}) + d(x_{k_0}, x) + d(x_{j_0}, y), & j_0 < k_0, \\ \sum_{i=k_0}^{j_0} d(x_i, x_{i+1}) + d(x_{k_0}, x) + d(x_{j_0}, y), & k_0 < j_0. \end{cases} \tag{4}$$

It can be found in [20, Section 2] that $\overline{d}(x, y)$ defined in (4) is a distance between x and y. Let $g(x) := \overline{d}(x, x_1)$, then the SNN studied in the paper can be represented as

$$N_n(x) := \sum_{i=1}^{n} c_i \sigma(a_i g(x) + b_i), \tag{5}$$

where $a_i, b_i, c_i \in \mathbb{R}$.

3 Existence of Exact Interpolation SNN

A function $\sigma(\cdot)$ defined on $\mathbb{R}$ is called a sigmoidal function if it satisfies

$$\sigma(t) \to 1 \quad \text{as} \quad t \to +\infty;$$
$$\sigma(t) \to 0 \quad \text{as} \quad t \to -\infty.$$

Neural networks activated by such functions has been proved to possess good approximation and interpolation capability in Euclidean space ([1–8,13,16,18,21]). Let $A > 0$, and denote

$$\delta_\sigma(A) := \sup_{t \geq A} \max\{|1 - \sigma(t)|, |\sigma(-t)|\}.$$

By the definition of the sigmoidal function, it is easy to deduce that $\delta_\sigma(A)$ is non-increasing, and satisfies

$$\lim_{A \to +\infty} \delta_\sigma(A) = 0.$$

This implies that for arbitrary $\varepsilon > 0$, there exists an $A_0 > 0$ such that

$$\delta_\sigma(A) < \varepsilon, \quad \text{for } A \geq A_0. \tag{6}$$

The aim of this section is to prove that there exists an SNN, N_n, formed as (5) such that

$$N_n(x_i) = f_i, \ i = 1, \ldots, n. \tag{7}$$

In order to prove the existence of exact interpolation SNN, we should choose the parameters a_j, b_j, c_j such that (7) holds. But it is not easy to fix all of these parameters. Taking the method used in [21] into account, we can construct a_j and b_j at first, then use (7) to solve the coefficients c_j.

For arbitrary $A > 0$, set

$$a_j = \frac{-2A}{g(x_{j+1}) - g(x_j)}, \quad j = 1, \ldots, n-1,$$
$$a_n = \frac{-2A}{g(x_n) - g(x_{n-1})},$$
$$b_j = \frac{2Ag(x_j)}{g(x_{j+1}) - g(x_j)} + A, \quad j = 1, \ldots, n-1,$$

and

$$b_n = \frac{2Ag(x_n)}{g(x_n) - g(x_{n-1})} + A.$$

Then, we have

$$N_n(x) := \sum_{j=1}^{n-1} c_j \sigma \left(-2A \frac{g(x) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A \right) + c_n \sigma \left(-2A \frac{g(x) - g(x_n)}{g(x_n) - g(x_{n-1})} + A \right).$$

On the other hand, it follows from the definition of $g(\cdot)$ that

$$-2A \frac{g(x_i) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A \geq A, \quad \text{if } i < j,$$
$$-2A \frac{g(x_i) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A = A, \quad \text{if } i = j,$$
$$-2A \frac{g(x_i) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A = -A \quad \text{if } i = j+1,$$

and

$$-2A \frac{g(x_i) - g(x_j)}{g(x_{j+1}) - g(y_j)} + A \leq -A, \quad \text{if } i > j+1.$$

Therefore, for arbitrary $1 \leq i \leq j$, there holds

$$\left| 1 - \sigma \left(-2A \frac{g(x_i) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A \right) \right| \leq \delta_\sigma(A), \tag{8}$$

and for every $j+1 \leq i \leq n$, there holds

$$\left| \sigma \left(-2A \frac{g(x_i) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A \right) \right| \leq \delta_\sigma(A). \tag{9}$$

The following Theorem guarantees the existence of the exact interpolation SNN.

Theorem 1 *Let σ be a sigmoidal function. If we choose A large enough such that*

$$\delta_\sigma(A) < \frac{1}{4n}$$

holds, then there exists a set of real numbers $\{c_j\}_{j=1}^n$ such that $N_n(x)$ satisfying (7).

Proof Denote

$$e_{i,j}(A) := \sigma \left(-2A \frac{g(x_i) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A \right), \; i, j = 1, \ldots, n-1,$$
$$e_{i,n}(A) := \sigma \left(-2A \frac{g(x_i) - g(x_j)}{g(x_n) - g(x_{n-1})} + A \right), \; i = 1, \ldots, n-1,$$
$$e_{n,j}(A) := \sigma \left(-2A \frac{g(x_n) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A \right), \; j = 1, \ldots, n-1,$$
$$e_{n,n}(A) := \sigma(A).$$

Let $M_n(A) := (e_{i,j}(A))_{i,j=1}^n$ be the coefficient matrix of the system of Eq. (7), and

$$D_n(A) := \begin{vmatrix} e_{1,1}(A) & e_{1,2}(A) & \dots & e_{1,n}(A) \\ e_{2,1}(A) & e_{2,2}(A) & \dots & e_{2,n}(A) \\ \dots & \dots & \dots & \dots \\ e_{n,1}(A) & e_{n,2}(A) & \dots & e_{n,n}(A) \end{vmatrix}$$

be its determinant. If we write

$$\begin{aligned} d_{i,j}(A) &= e_{i,j}(A) - e_{i+1,j}(A) \ \ (i, j = 1, \dots, n-1), \\ d_{i,n}(A) &= e_{i,n}(A) - e_{i+1,n}(A) \ \ (i = 1, \dots, n-1), \\ d_{n,j}(A) &= e_{n,j}(A) \ \ (j = 1, \dots, n-1), \ \ d_{n,n}(A) = e_{n,n}(A), \end{aligned}$$

then

$$D_n(A) = \begin{vmatrix} d_{1,1}(A) & d_{1,2}(A) & \dots & d_{1,n}(A) \\ d_{2,1}(A) & d_{2,2}(A) & \dots & d_{2,n}(A) \\ \dots & \dots & \dots & \dots \\ d_{n,1}(A) & d_{n,2}(A) & \dots & d_{n,n}(A) \end{vmatrix}.$$

From the assumption of A, it follows that if $t \geq A$, then

$$|\sigma(-t)| < \frac{1}{4n}, \ \ |1 - \sigma(t)| < \frac{1}{4n}.$$

Thus, by the definition of $d_{i,j}(A)$, we have

$$\begin{aligned} d_{j,j}(A) &= \sigma(A) - \sigma(-A) = 1 - (1 - \sigma(A)) - \sigma(-A) \\ &\geq 1 - \frac{1}{4n} - \frac{1}{4n} = 1 - \frac{1}{2n} \geq \frac{1}{2} \ \ (1 \leq j \leq n-1), \end{aligned}$$

and

$$\begin{aligned} d_{n,n}(A) &= \sigma(A) = 1 - (1 - \sigma(A)) \\ &\geq 1 - \frac{1}{4n} \geq 1 - \frac{1}{2n} \geq \frac{1}{2}. \end{aligned}$$

On the other hand, it follows from (8) and (9) that for arbitrary $1 \leq j \leq n$, there holds

$$\begin{aligned} \sum_{i=1, i \neq j}^{n-1} |d_{i,j}(A)| &= \sum_{i=1, i \neq j}^{n-1} |e_{i,j}(A) - e_{i+1,j}(A)| \\ &= \sum_{i=1}^{j-1} |e_{i,j}(A) - e_{i+1,j}(A)| + \sum_{i=j+1}^{n} |e_{i,j}(A) - e_{i+1,j}(A)| \\ &= \sum_{i=1}^{j-1} |1 - e_{i,j}(A) - (1 - e_{i+1,j}(A))| + \sum_{i=j+1}^{n-1} |e_{i,j}(A) - e_{i+1,j}(A)| \\ &\leq \sum_{i=1}^{j-1} (|1 - e_{i,j}(A)| + |1 - e_{i+1,j}(A)|) + \sum_{i=j+1}^{n-1} (|e_{i,j}(A)| + |e_{i+1,j}(A)|) \\ &< (n-1) \cdot \frac{1}{2n}. \end{aligned}$$

Furthermore,

$$|d_{n,j}| = |e_{n,j}(A)| \le \frac{1}{4n} < \frac{1}{2n}.$$

Therefore,

$$d_{j,j}(A) \ge \frac{1}{2} > \sum_{i=1, i\neq j}^{n} |d_{i,j}(A)|, \quad j = 1, \ldots, n.$$

Then by the strictly diagonally dominant matrices are invertible principle (see [15]), we have

$$D_n(A) \neq 0,$$

which means that the Eq. (7) is solvable. This completes the proof of Theorem 1. □

It follows from Theorem 1 that if A is large enough, then there exists exact interpolation SNN with sigmoidal activation function. Thus, the SNN method is feasible for spherical scattered data interpolation.

4 Interpolation Error Estimate

In this section, we will deduce the interpolation error for the SNN. To this end, we first construct an approximate SNN interpolant and deduce the error estimate for the approximate interpolation. Then, we study the error between the approximate interpolation SNN and the exact interpolation SNN. Under this circumstance, we can deduce an upper bound error estimate for the exact interpolation SNN.

We first need introduce a modulus of smoothness on the sphere. Let $SO(d+1)$ be the (compact) group of rotations on $\mathbf{S}^d$. For $\rho \in SO(d+1)$, the modulus of smoothness on $\mathbf{S}^d$ is defined as

$$\omega(f,t) := \sup_{\rho \in O_t} \max_{x \in \mathbf{S}^d} |f(\rho x) - f(x)|,$$

where

$$O_t := \left\{ \rho \in SO(d+1) : \max_{x \in \mathbf{S}^{d-1}} \arccos(x) \cdot \rho x \le t \right\}.$$

For more details of the modulus of smoothness $\omega(f,t)$, we refer the readers to [9].

The approximate interpolation SNN can be constructed as follows:

$$N_n^a(x) := \sum_{j=1}^{n-1} (f_j - f_{j+1}) \sigma \left(-2A \frac{g(x) - g(x_j)}{g(x_{j+1}) - g(x_j)} + A \right) + f_n \sigma \left(-2A \frac{g(x) - g(x_n)}{g(x_n) - g(x_{n-1})} + A \right).$$

It is obvious that the only difference between $N_n(x)$ and $N_n^a(x)$ is the coefficients, which are given explicitly in the latter one.

The following Lemma 1 [20, Theorem 2] will play an important role in our proof.

Lemma 1 *Let $N_n^a(x)$ be defined above, and σ be a bounded sigmoidal function. If $f \in C(\mathbb{S}^d)$ satisfying $f(x_i) = f_i$, $i = 1, \ldots, n$, then there holds*

$$|f(x) - N_\sigma^a(x)| \le \delta_\sigma(A) \left(\sum_{j=1}^{n-1} |f_j - f_{j+1}| + |f_n| \right) + (9 + 8\|\sigma\|)\pi^{d-1}\tau_X^d \omega(f, h_X),$$

where $\|\sigma\| := \max_{t \in \mathbb{R}} |\sigma(t)|$.

By the help of Lemma 1, we can deduce the following error estimate.

Theorem 2 *Let $N_n(x)$ be the exact interpolation SNN and σ be a bounded sigmoidal function. If $f \in C(\mathbb{S}^d)$ satisfying $f(x_i) = f_i$, $i = 1, \ldots, n$ and A is large enough such that*

$$\delta_\sigma(A) \le \min\left\{ \frac{1}{4n}, \frac{\tau_X^d \omega(f, h_X)}{n\left(\sum_{j=1}^{n-1} |f_j - f_{j+1}| + |f_n|\right)} \right\}, \tag{10}$$

then there exists a constant C depending only on σ and d such that

$$|f(x) - N_n(x)| \le C\tau_X^d \omega(f, h_X), \quad \forall x \in \mathbf{S}^d. \tag{11}$$

Proof of Theorem 2 Since

$$\delta_\sigma(A) \le \frac{\tau_X^d \omega(f, h_X)}{n\left(\sum_{j=1}^{n-1} |f_j - f_{j+1}| + |f_n|\right)},$$

it follows from Lemma 1 that

$$|f(x) - N_n^a(x)| \le C\tau_X^d \omega(f, h_X).$$

Thus, it is sufficient to bound the error between N_n^a and N_n. Since (10) holds, then the system of equations (7) is solvable. We denote its solution as $V_c := (c_1, \ldots, c_n)$, and its coefficient matrix as M. Let $V_f = (f_1, \ldots, f_n)$, then the system of Eq. (7) can be written as

$$MV_c^T = V_f^T, \tag{12}$$

where V^T denotes the transpose of the vector V. Let

$$U := \begin{bmatrix} 1 & 1 & \ldots & 1 & 1 \\ 0 & 1 & \ldots & 1 & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & 1 \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix},$$

then by a simple calculation, we can obtain the inverse matrix of U as

$$U^{-1} = \begin{bmatrix} 1 & -1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & -1 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix}.$$

Write

$$M_n(A) - U = (\alpha_{ij})_{i,j=1}^{n,n},$$

then it follows from (8) and (9) that

$$|\alpha_{ij}| \leq \delta_\sigma(A).$$

Therefore, if we write

$$U^{-1}(M_n(A) - U) = (\beta_{ij})_{i,j=1}^{n,n},$$

then

$$|\beta_{ij}| \leq 2\delta_\sigma(A), \quad |\beta_{nj}| \leq \delta_\sigma(A), \quad i = 1, \ldots, n-1,\, j = 1, \ldots, n. \tag{13}$$

Denoting

$$V_F := (f_1 - f_2, \ldots, f_{n-1} - f_n, f_n), \ \Delta V_c := V_c - V_F,$$

we have

$$U V_F^T = V_f^T.$$

So it follows from (12) that

$$(U + (M_n(A) - U))(V_F^T + \Delta V_c^T) = V_f^T.$$

That is

$$U \Delta V_c^T = -(M_n(A) - U)\Delta V_c^T - (M_n(A) - U)V_F^T.$$

Hence

$$\Delta V_c^T = -U^{-1}(M_n(A) - U)\Delta V_c^T - U^{-1}(M_n(A) - U)V_F^T.$$

The above identity together with (13) yields

$$\sum_{i=1}^{n} |\Delta V_{C_i}| \leq (2n+1)\delta_\sigma(A) \sum_{i=1}^{n} |\Delta V_{C_i}| + (2n+1)\delta_\sigma(A) \left(\sum_{i=1}^{n_1} |f_i - f_{i+1}| + |f_n| \right),$$

which implies

$$\sum_{i=1}^{n} |\Delta V_{C_i}| \leq \frac{(2n+1)\delta_\sigma(A)}{1 - (2n+1)\delta_\sigma(A)} \left(\sum_{j=1}^{n} |f_j - f_{j+1}| + |f_n| \right).$$

Because

$$|N_n(x) - N_n^a(x)| \leq \sum_{i=1}^{n} |\Delta V_{C_i}| \|\sigma\|,$$

then

$$|N_n(x) - N_n^a(x)| \leq \frac{(2n+1)\delta_\sigma(A)\|\sigma\|}{1 - (2n+1)\delta_\sigma(A)} \left(\sum_{j=1}^{n-1} |f_j - f_{j+1}| + |f_n| \right). \tag{14}$$

Furthermore, since

$$\delta_\sigma(A) \leq \min \left\{ \frac{1}{4n}, \frac{\tau^d \omega(f, h_X)}{n \left(\sum_{j=1}^{n-1} |f_j - f_{j+1}| + |f_n| \right)} \right\},$$

there exists a constant depending only on d, σ and f such that

$$|f(x) - N_n(x)| \leq C\tau_X^d \omega(f, h_X),$$

which finishes the proof of Theorem 2. □

In Theorem 2, there are not any smooth assumptions on the target function, which is different from SRBFN method [14,22,23]. Furthermore, it follows from (6) that there always exists a sufficiently large A such that (10) holds.

References

1. Anastassiou GA (2012) Univariate sigmoidal neural network approximation. J Comput Anal Appl 14:659–690
2. Barron AR (1993) Universal approximation bounds for superpositions of a sigmoidal function. IEEE Trans Inf Theory 39:930–945
3. Chen DB (1993) Degree of approximation by superpsitions of a sigmoidal function. Approx Theory Its Appl 9:17–28
4. Chen ZX, Cao FL (2009) The approximation operators with sigmoidal functions. Comput Math Appl 58:758–765
5. Costarelli D, Spigler R (2013) Constructive approximation by superposition of sigmidal functions. Anal Theory Appl 29:169–196
6. Costarelli D, Spigler R (2013) Approximation results for neural network opertors activated by sigmoidal functions. Neural Netw 44:101–106
7. Costarelli D, Spigler R (2013) Multivariate neural network operators with sigmoidal activation functions. Neural Netw 48:72–77
8. Cybenko G (1989) Approximation by superpositions of sigmoidal function. Math Control Signals Syst 2:303–314
9. Ditzian Z (1999) A modulus of smoothness on the unit sphere. J d'Anal Math 79:189–200
10. Fasshauer GE, Schumaker LL (1998) Scattered data fitting on the sphere. In: Dælen M, Lyche T, Schumaker LL (eds) Mathematical methods for curves and surfaces II. Vanderbilt University Press, Nashville, TN, pp 117–166
11. Freeden W, Gervens T, Schreiner M (1998) Constructive approximation on the sphere. Oxford University Press, New York
12. Golitschek M, Light W (2001) Interpolation by polynomials and radial basis functions on spheres. Constr Approx 17:1–18
13. Hahm N, Hong B (2002) Approximation order to a function in $\overline{C}(\mathbf{R})$ by superposition of a sigmoidal function. Appl Math Lett 15:591–597
14. Jetter K, Stökler J, Ward J (1999) Error estimates for scattered data interpolation on spheres. Math Comput 68:733–749
15. Lascaux P, Theodor T (1986) Analyse Numerique Matricielle Apliquée a l'Art de l'Ingenieur. Masson, Paris
16. Lenze B (1992) Constructive multivariate approximation with sigmoidal functions and applications to neural networks, Numerical Methods of Approximation Theory. Birkhauser Verlag, Basel
17. Levesley J, Sun X (2005) Approximation in rough native spaces by shifts of smooth kernels on spheres. J Approx Theory 133:269–283
18. Lewicki G, Marino G (2003) Approximation by superpositions of a sigmoidal function. J Anal Its Appl 22:463–470
19. Lin SB, Cao FL, Xu ZB (2011) Essential rate for approximation by spherical neural networks. Neural Netw 24:752–758
20. Lin SB, Cao FL, Chang XY, Xu ZB (2012) A general radial quasi-interpolation operator on the sphere. J Approx Theory 164:1402–1414
21. Llanas B, Sainz FJ (2006) Constructive approximate interpolation by nerual networks. J Comput Appl Math 188:283–308
22. Narcowich FJ, Ward JD (2002) Scattered data interpolation on spheres: error estimates and locally supported basis functions. SIAM J Math Anal 33:1393–1410
23. Narcowich FJ, Sun XP, Ward JD, Wendland H (2007) Direct and inverse sobolev error estimates forscattered data interpolation via spherical basis functions. Found Comput Math 369–370

24. Sloan I, Womersley R (2000) Constructive polynomial approximation on the sphere. J Approx Theory 103:91–118
25. Sun XP, Cheney EW (1997) Fundamental sets of continuous functions on spheres. Constr Approx 13:245–250
26. Wendland H (2005) Scattered data approximation. In: Cambridge monographs on applied and computational mathematics. Cambridge University Press, Cambridge
27. Womersley R, Sloan I (2001) How good can polynomial interpolation on the sphere be? Adv Comput Math 14:195–226