

THE COVERING NUMBER FOR SOME MERCER KERNEL HILBERT SPACES ON THE UNIT SPHERE

Sheng Baohuai, Wang Jianli and Chen Zhixiang

Abstract. In the present paper, we estimate the covering number for some reproducing kernel Hilbert spaces on the unit sphere. Both the upper bounds and the lower bounds are provided.

1. INTRODUCTION

It is known that the covering number which is often used to measure the (probability) error or the number of the samples required for the given confidence and the error bounds has been playing an important role in learning theory(see e.g. [1-4]). Some of the investigations on it can be found from F. Cucker and S. Smale[1], R. C. Williamson, A. J. Smola and B. Schölkopf(see [5]), Y. Guo, P. L. Bartlett and J. Shawe-Taylor[6], D. X. Zhou [7]-[8], H.W.Sun and D.X.Zhou[9], etc.

Let $(X, d(\cdot, \cdot))$ be a compact metrics space of the n -dimensional Euclidean space \mathcal{R}^n , and ν be a Borel measure on X . Denote the space of real square integrable functions on X by $L^2_\nu(X)$, and the space of continuous functions by $C(X)$.

Let $K : X \times X \rightarrow \mathcal{R}$ be continuous, symmetric and positive definite, i.e., for any finite set $\{x_1, \dots, x_m\} \subset X$, the matrix $(K(x_i, x_j))_{i,j=1}^m$ is positive definite. We call K a Mercer kernel on $X \times X$. The reproducing kernel Hilbert space \mathcal{H}_K associated with the kernel K is defined to be the closure of the linear span of the set of functions $\{K(t, x) : t \in X\}$ with the inner product satisfying reproducing relation

$$\langle K(\cdot, x), f(\cdot) \rangle_{\mathcal{H}_K} = f(x), \quad \forall x \in X, f \in \mathcal{H}_K.$$

Define a Hilbert-Schmidt integral operator by means of this kernel as

$$L_K(f, x) = \int_X K(x, t)f(t) d\nu(t), \quad x \in X, f \in L^2_\nu(X).$$

Received July 23, 2008, accepted January 14, 2010.

Communicated by J. C. Yao.

2000 *Mathematics Subject Classification*: 41A46, 42C10.

Key words and phrases: Reproducing kernel Hilbert spaces, Covering number, Spherical harmonics. This work was supported (in part) by national NSF (No: 10871226) of China and the NSF (No. Y6100096) of Zhejiang Province.

Then $L_K(f, x)$ is a positive, compact operator and its range lies in $C(X)$.

Let $(\lambda_j)_{j=1}^{+\infty}$ denote the nonincreasing sequence of eigenvalues of L_K and $\{\phi_j\}_{j=1}^{+\infty}$ be corresponding eigenfunctions. Then, the Mercer kernel theorem (see e.g. [1, 10-13]) makes

$$K(x, t) = \sum_{j=1}^{+\infty} \lambda_j \phi_j(x) \phi_j(t), \quad x, t \in X,$$

where the series converges uniformly and absolutely.

\mathcal{H}_K can be imbedded into $C(X)$, and we denote the inclusion as $I_K : \mathcal{H}_K \rightarrow C(X)$. For this facts, see [1] and [13, Chapter 4].

Let $R > 0$ and B_R be the ball of \mathcal{H}_K with radius R :

$$B_R := \{f \in \mathcal{H}_K : \|f\|_K \leq R\}.$$

Then $I_K(B_R) \subset C(X)$. Denote its closure in $C(X)$ as $\overline{I_K(B_R)}$ which is a compact subset in $C(X)$.

Let \mathcal{N} be the set of natural number. S is a compact set in a metric space and $\eta > 0$, the covering number $\mathcal{N}(S, \eta)$ of S is defined to be the minimal integer m such that there exist m disks with radius η covering S . In learning theory, one only need the estimate for the case of $S = B_R$.

Let E be an N -dimensional Banach space. Then, F. Cucker and S. Smale proved in [1] that

$$(1) \quad \ln \mathcal{N}(B_R, \eta) \leq N \ln \left(\frac{4R}{\eta} \right).$$

In [7-8], D. X. Zhou sum up the estimates of $\mathcal{N}(\overline{I_K(B_R)}, \eta)$ to the estimate of the l^2 -norm of the inverse of the Mercer kernel matrix and gave the general lower and upper bound estimates for $\mathcal{N}(\overline{I_K(B_R)}, \eta)$. In particular, some estimates concerning the translation invariant kernels are provided (see also [9] and [14]).

In the present paper, we want to investigate the estimates of $\mathcal{N}(\overline{I_K(B_R)}, \eta)$ for Mercer kernels $K(x, y)$ defined on the unit sphere. For these purposes we need to restate some notions and results with respect to the spherical harmonics.

Let $q \geq 2$ be a given integer which will be fixed throughout this paper and $S^q := \{(x_1, x_2, \dots, x_{q+1})\}$ be the unit sphere in the Euclidean space R^{q+1} ($q \geq 2$) with $d\mu_q$ being the usual volume element on S^q . The volume of S^q is

$$\omega_q = \int_{S^q} d\mu_q = \frac{2\pi^{\frac{q+1}{2}}}{\Gamma(\frac{q+1}{2})}.$$

For a given integer $l \geq 0$, the restriction to S^q of a homogeneous harmonics of order l is called a spherical harmonics of order l . The class of all spherical

harmonics of degree l will be denoted by H_l^q , and the class of all spherical harmonics of degree $l \leq n$ will be denoted by Π_n^q . Of course, $\Pi_n^q = \bigoplus_{l=0}^n H_l^q$.

The dimension of H_l^q is given by (see e.g. [15])

$$d_l^q = \dim H_l^q = \begin{cases} \frac{(2l+q-1)\Gamma(l+q)}{(l+q-1)\Gamma(l+1)\Gamma(q)}, & l \geq 1, \\ 1, & l = 0, \end{cases}$$

and that of Π_n^q is $\sum_{l=0}^n d_l^q = d_n^{q+1} \sim n^q$. Hence, if we choose an orthonormal basis $\{Y_{l,k} : k = 1, 2, \dots, d_l^q\}$ for each H_l^q , then the set $\{Y_{l,k} : l = 0, 1, 2, \dots, n; k = 1, 2, \dots, d_l^q\}$ forms an orthonormal basis for Π_n^q . One has the well known additional formula (see e.g. [16, 17])

$$(2) \quad \sum_{k=1}^{d_l^q} Y_{l,k}(x) \overline{Y_{l,k}(y)} = \frac{d_l^q}{\omega_q} P_l^{q+1}(x \cdot y), \quad l = 0, 1, \dots,$$

where $P_l^{q+1}(x)$ is the degree- l (generalized) Legendre polynomial which are normalized so that $P_l^{q+1}(1) = 1$, and satisfy the orthogonality relations

$$\int_{-1}^1 P_l^{q+1}(x) P_k^{q+1}(x) W_q(x) dx = \frac{\omega_q}{\omega_{q-1} d_l^q} \delta_{l,k},$$

where $W_q(x) = (1 - x^2)^{\frac{q}{2}-1}$. By (2) and the fact $\Pi_n^q = \bigoplus_{l=0}^n H_l^q$ we know for any $p \in \Pi_n^q$ and $x \in S^q$

$$(3) \quad p(x) = \sum_{l=0}^n \frac{d_l^q}{\omega_q} \int_{S^q} p(y) P_l^{q+1}(x \cdot y) d\mu_q(y), \quad x \in S^q.$$

The orthogonal projection $Y_k(f, x)$ of a function $f \in L^1(S^q)$ on H_k^q is defined by

$$Y_k(f, x) = \frac{d_k^q}{\omega_q} \int_{S^q} P_k^{q+1}(x \cdot y) f(y) d\mu_q(y), \quad k = 0, 1, 2, \dots.$$

Correspondingly, we have the following Fourier-Laplace expansion of $f(x)$ (see e.g. [16])

$$f(x) \sim \sum_{k=0}^{\infty} Y_k(f, x), \quad x \in S^q.$$

Let $K(x, y) = \sum_{l=0}^{+\infty} \frac{d_l^q \lambda_l}{\omega_q} P_l^{q+1}(x \cdot y)$, $x, y \in S^q$, $\lambda_l > 0$. Then, by [18, Theorem 17.8] we know for given pairwise distinct points $X_N := \{\xi_1, \xi_2, \dots, \xi_N\} \subset S^q$, the matrix $A_N := (K(\xi_i, \xi_j))_{N \times N}$ is positive definite. Therefore, $K(x, y)$ are Mercer kernels on $S^q \times S^q$ with λ_l being their eigenvalues.

Let $0 < \gamma < 1$. Taking $K_0(x, y) = \frac{1-\gamma^2}{(1-2\gamma x \cdot y + \gamma^2)^{\frac{q+1}{2}}}$, then by [19] or [16, Theorem 1.2.6] we know

$$K_0(x, y) = \sum_{l=0}^{+\infty} \frac{d_l^q \gamma^l}{\omega_q} P_l^{q+1}(x \cdot y), \quad x, y \in S^q.$$

$K_0(x, y)$ shows the existence of the specific Mercer kernel on S^q .

In the present paper, we want to estimate the upper and lower bounds of $\mathcal{N}(I_K(B_R), \eta)$ for the Mercer kernels $K(x, y)$ on $S^q \times S^q$ along the line of [7] and [8].

The paper is organized as follows. The lower bound estimates for $\mathcal{N}(I_K(B_R), \eta)$ are investigated in Section 2. We shall choose the knot point set X_N according to that of [20] and give an upper estimate of $\|A_N^{-1}\|_{l^2(X_N)}$. Some lower estimates for the covering number are thus obtained; In Section 3, we shall construct a kind of local spherical harmonics reproducing basis function, and with which give the upper bound measures for the regularity of the Mercer kernels. The upper bounds for $\mathcal{N}(I_K(B_R), \eta)$ are estimated in two cases. One case is that the eigenvalues λ_l have the decays $\lambda_l d_l^q \sim \frac{1}{(1+l)^\alpha}$, $\alpha > 1$, and the another one is that $\lambda_l d_l^q \sim \frac{1}{e^{(1+l)\alpha}}$, $\alpha > 0$.

Throughout this paper, we shall denote by \mathcal{N} the set of natural numbers. By \mathcal{R}^N we denote the N -dimensional Euclidean space, and by Π_N^q we denote the set of all spherical harmonics of order $\leq N$. The biggest integer which $\leq a$ will be denoted by $[a]$. δ_{ij} is the δ function whose value is 1 if $i = j$ and whose value is 0 if $i \neq j$. We say $A \sim B$ if there are two positive constants $C_1 > 0, C_2 > 0$ such that $C_1 \leq \frac{A}{B} \leq C_2$.

2. THE LOWER BOUND ESTIMATES

In this section, we shall use the general framework given by [8] for estimating the lower bounds of the covering numbers. We first give a lower estimate for $\mathcal{N}(\overline{I_{K_0}(B_R)}, \eta)$.

Theorem 2.1. *Let $K_0(x, y)$ be the Mercer kernel given in Section 1. Then, there are positive constants B_0, A_0 , such that for $0 < \eta \leq \sqrt{\frac{\gamma}{B_0}} R$ there holds*

$$(4) \quad \ln \mathcal{N}(\overline{I_{K_0}(B_R)}, \eta) \geq \left\{ A_0 \left[\frac{1}{2} \log_{\frac{1}{\gamma}} \left(\frac{1}{B_0} \left(\frac{R}{2\eta} \right)^2 \right) - \frac{1}{2} \right]^q - 1 \right\} \ln 2.$$

If the eigenvalues λ_l have suitable decays, then, the lower bounds of $\mathcal{N}(\overline{I_K(B_R)}, \eta)$ may be estimated.

Theorem 2.2 *Let the spherical Mercer $K(x, y)$ defined on $S^q \times S^q$ satisfy $\lambda_l = \frac{1}{(1+l)^\alpha}$, $\alpha > 1$. Then, there are positive constants A_0, B_0 , such that for $0 < \eta \leq \frac{R}{B_0 3^{\frac{\alpha}{2}}}$ there holds*

$$(5) \quad \ln \mathcal{N}(\overline{I_K(B_R)}, \eta) \geq \left\{ A_0 \left[\frac{1}{2} \left(\frac{1}{B_0} \left(\frac{R}{2\eta} \right)^2 \right)^{\frac{1}{\alpha}} - 1 \right]^q - 1 \right\} \ln 2.$$

(4) is similar to the results of example 2 of [8]. (5) agrees with the results for the Sobolev translation-invariant kernels (see [8]).

To show (4) and (5) we give some lemmas. The first one is the general lower estimate for $\mathcal{N}(I_K(B_R), \eta)$ given by D. X. Zhou.

Lemma 2.1. (see [8], [13, Theorem 5.21]). *Let $X \subset \mathcal{R}^n$ be a compact set, K be a Mercer kernel on $X \times X$, $l \in \mathcal{N}$, and $X_l := \{x_1, x_2, \dots, x_l\} \in X$ yield an invertible Gramian matrix $A_{X_l} := (K(x_i, x_j))_{i,j=1}^l$. Then,*

$$(6) \quad \mathcal{N}(\overline{I_K(B_R)}, \frac{\eta}{2}) \geq 2^l - 1,$$

provided that $\|A_{X_l}^{-1}\|_{l^2} \leq \frac{1}{l} \left(\frac{R}{\eta} \right)^2$.

It is known that the Gauss integral and the Marcinkiewicz-Zygmund inequality on spherical harmonic are good tools in constructing zonal and translation networks on the unit sphere (see, e.g., [21-26]) and therefore have been studied by many mathematicians (see e.g., [27-30]). To make the statement more convenient, we cite the form given by Dai Feng in [20].

Given $\varepsilon > 0$, we say a finite subset $\Lambda \subset S^q$ is ε -separable if

$$\min_{\omega, \omega' \in \Lambda, \omega \neq \omega'} d(\omega, \omega') \geq \varepsilon,$$

while we say it is maximal ε -separable if it is ε -separable and satisfies

$$\max_{x \in S^q} \min_{\omega \in \Lambda} d(x, \omega) < \varepsilon.$$

Lemma 2.2. (see [20]). *Let $B(x, r) = \{y \in S^q : d(x, y) \leq r\}$ denote the spherical cap with center $x \in S^q$ and radius $r \in (0, \pi)$. Then, there exists a positive constant ε depending only on q such that for any $\delta \in (0, \varepsilon)$ and any maximal $\frac{\delta}{n}$ -separable subset $\Lambda \subset S^q$ there exists a sequence of positive numbers $\lambda_\omega \sim \text{mes}(B(\omega, \frac{\delta}{n})) \sim \frac{\delta^q}{n^q}$, $(\omega \in \Lambda)$ for which the following holds for all $f \in \Pi_n^q$*

$$(7) \quad \int_{S^q} f(x) d\mu_q(x) = \sum_{\omega \in \Lambda} \lambda_\omega f(\omega),$$

where the cardinality of $|\Lambda|$ satisfies $|\Lambda| \sim n^q \sim \dim \Pi_n^{q+1}$ with the constant of equivalence depending only on q .

By [31, Corollary 3.1] we know there is a constant $C := C(\delta, q)$ such that for every $f \in C(S^q)$ there are $\Lambda_n^*(x) \in \Pi_{cn}^q$ such that

$$\Lambda_n^*(\omega) = f(\omega), \quad \omega \in \Lambda.$$

Since $\delta \in (0, \varepsilon)$ are arbitrary, we may find suitable positive number γ and δ' depending only on q such that $(C + 1) \leq \gamma$ and a maximal $\frac{\delta'}{\gamma n}$ -separable subset $\Lambda \subset S^q$ such that

$$(8) \quad \int_{S^q} f(x) d\mu_q(x) = \sum_{\omega \in \Lambda} \lambda_\omega f(\omega), \quad f \in \Pi_{\gamma n}^q$$

We then have the following Lemma 2.3.

Lemma 2.3. *There is a constant $C > 0$ and a $\Lambda_n^*(x) \in \Pi_{cn}^q$ such that $(C + 1) \leq \gamma$ and for any given real numbers $\{y_\omega\}_{\omega \in \Lambda}$ there holds*

$$(9) \quad \Lambda_n^*(\omega) = y_\omega, \quad \omega \in \Lambda.$$

We now give a method to estimate the norm for the inverse of the Mercer kernel matrices.

Lemma 2.4. *Let the spherical Mercer $K(x, y)$ satisfy $\lambda_l > 0$. Λ is the knot set in Lemma 2.2, and $A_n := (K(\omega, \omega'))_{\omega, \omega' \in \Lambda}$ is the corresponding Mercer kernel matrix. Then, there is a positive B_0 such that*

$$(10) \quad \|A_n^{-1}\|_{l^2} \leq \frac{B_0}{|\Lambda| \min_{0 \leq l \leq n} \lambda_l}.$$

Proof. Let $v = (v_\omega)_{\omega \in \Lambda} \in \mathcal{R}^{|\Lambda|}$. Then, by (2) and the definition of $K(x, y)$ one has

$$\begin{aligned} v^\top A_n v &= \sum_{\omega, \omega' \in \Lambda} v_\omega K(\omega, \omega') v_{\omega'} \\ &= \sum_{\omega, \omega' \in \Lambda} v_\omega \sum_{l=0}^{+\infty} \frac{\lambda_l d_l^q}{\omega_q} P_l^{q+1}(\omega \cdot \omega') v_{\omega'} \\ &= \sum_{l=0}^{+\infty} \frac{\lambda_l d_l^q}{\omega_q} \sum_{\omega, \omega' \in \Lambda} v_\omega P_l^{q+1}(\omega \cdot \omega') v_{\omega'} \end{aligned}$$

$$\begin{aligned}
 &= \sum_{l=0}^{+\infty} \lambda_l \sum_{\omega, \omega' \in \wedge} v_\omega \sum_{i=1}^{d_l^q} Y_{l,i}(\omega) \overline{Y_{l,i}(\omega')} v_{\omega'} \\
 &= \sum_{l=0}^{+\infty} \lambda_l \sum_{i=1}^{d_l^q} \left| \sum_{\omega \in \wedge} v_\omega Y_{l,i}(\omega) \right|^2,
 \end{aligned}$$

where $\{Y_{l,k}(x), l = 0, 1, 2, \dots, d_l^q\}$ is the orthonormal basis given in (2). Then, the Parseval equality for an orthonormal basis in a Hilbert space and (2) makes

$$\begin{aligned}
 v^\top A_n v &\geq \left(\min_{0 \leq l \leq n} \lambda_l \right) \sum_{l=0}^n \sum_{i=1}^{d_l^q} \left| \sum_{\omega \in \wedge} v_\omega Y_{l,i}(\omega) \right|^2 \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \int_{S^q} \left| \sum_{\omega \in \wedge} v_\omega \times \sum_{l=0}^n \sum_{i=1}^{d_l^q} Y_{l,i}(\omega) \overline{Y_{l,i}(x)} \right|^2 d\mu_q(x) \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \int_{S^q} \left| \sum_{\omega \in \wedge} v_\omega \times \sum_{l=0}^n \frac{d_l^q}{\omega_q} P_l^{q+1}(\omega \cdot x) \right|^2 d\mu_q(x) \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \int_{S^q} \left| \sum_{\omega \in \wedge} \frac{v_\omega}{\lambda_\omega} \times \lambda_\omega \times \sum_{l=0}^n \frac{d_l^q}{\omega_q} P_l^{q+1}(\omega \cdot x) \right|^2 d\mu_q(x) \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \int_{S^q} \left| \sum_{\omega \in \wedge} \wedge_n^*(\omega) \lambda_\omega \times \sum_{l=0}^n \frac{d_l^q}{\omega_q} P_l^{q+1}(\omega \cdot x) \right|^2 d\mu_q(x) \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \int_{S^q} \left| \sum_{l=0}^n \frac{d_l^q}{\omega_q} \sum_{\omega} \wedge_n^*(\omega) \times \lambda_\omega \times P_l^{q+1}(\omega \cdot x) \right|^2 d\mu_q(x) \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \int_{S^q} \left| \sum_{l=0}^n \frac{d_l^q}{\omega_q} \int_{S^q} \wedge_n^*(y) P_l^{q+1}(x \cdot y) d\mu_q(y) \right|^2 d\mu_q(x) \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \int_{S^q} \left| \sum_{l=0}^n Y_l(\wedge_n^*, x) \right|^2 d\mu_q(x) \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \int_{S^q} \left| \wedge_n^*(x) \right|^2 d\mu_q(x) \\
 &= \left(\min_{0 \leq l \leq n} \lambda_l \right) \sum_{\omega \in \wedge} \frac{v_\omega^2}{\lambda_\omega} \\
 &\geq \min_{\omega \in \wedge} \frac{1}{\lambda_\omega} \times \left(\min_{0 \leq l \leq n} \lambda_l \right) v^\top v,
 \end{aligned}$$

where we have used (8) and (9). On the other hand, since $\lambda_\omega \sim n^{-q}$ and $n^q \sim |\wedge|$, we have a positive B_0 such that $\min_{\omega \in \wedge} \frac{1}{\lambda_\omega} \geq \frac{|\wedge|}{B_0}$. (10) thus holds. (10) gives an upper bound estimate, itself has independent meaning.

Proof of Theorem 2.1. For the Mercer kernel $K_0(x, y)$ we have $\lambda_l = \gamma^l$. Take

$l = |\wedge|$ in Lemma 2.1, then, it follows that for $0 < \eta \leq R$ (6) holds if n satisfies $\frac{B_0}{\gamma^{2n-1}} \leq \left(\frac{R}{\eta}\right)^2$. Therefore, for $0 < \eta \leq \sqrt{\frac{\gamma}{B_0}}R$ we may choose $n \in \mathcal{N}$ such that $\frac{B_0}{\gamma^{2n-1}} \leq \left(\frac{R}{\eta}\right)^2 \leq \frac{B_0}{\gamma^{2n+1}}$, then, $2n + 1 \geq \log_{\frac{1}{\gamma}} \left(\frac{1}{B_0} \left(\frac{R}{\eta}\right)^2\right)$ and

$$\mathcal{N}(\overline{I_{K_0}(B_R)}, \frac{\eta}{2}) \geq 2^{|\wedge|} - 1 \geq 2^{|\wedge|-1}.$$

Since $|\wedge| \sim n^q$, there is a constant $A_0 > 0$ such that $|\wedge| \geq A_0 n^q$. We then have

$$\mathcal{N}(\overline{I_{K_0}(B_R)}, \frac{\eta}{2}) \geq 2^{A_0 n^q - 1} \geq 2^{A_0 \left[\frac{1}{2} \log_{\frac{1}{\gamma}} \left(\frac{1}{B_0} \left(\frac{R}{\eta}\right)^2 \right) - \frac{1}{2} \right]^q - 1}.$$

(4) thus holds.

Proof of Theorem 2.2. By Lemma 2.4 we know $|\wedge| \|A_n^{-1}\|_{l^2} \leq B_0(2n + 1)^\alpha$. Take $l = |\wedge|$ in Lemma 2.1, then, for $0 < \eta \leq R$ (6) holds if n satisfies $B_0(2n + 1)^\alpha \leq \left(\frac{R}{\eta}\right)^2$. Therefore, for $0 < \eta \leq \frac{R}{B_0^{3\frac{\alpha}{2}}}$ we can choose $n \in \mathcal{N}$ such that $B_0(2n + 1)^\alpha \leq \left(\frac{R}{\eta}\right)^2 \leq B_0(2n + 2)^\alpha$, i.e., $2n + 2 \geq \left(\frac{1}{B_0}\right)^{\frac{1}{\alpha}} \left(\frac{R}{\eta}\right)^{\frac{2}{\alpha}}$. Hence,

$$\mathcal{N}(\overline{I_K(B_R)}, \frac{\eta}{2}) \geq 2^{|\wedge|} - 1 \geq 2^{|\wedge|-1} \geq 2^{A_0 n^q - 1} \geq 2^{A_0 \left[\frac{1}{2} \left(\frac{1}{B_0}\right)^{\frac{1}{\alpha}} \left(\frac{R}{\eta}\right)^{\frac{2}{\alpha}} - 1 \right]^q}.$$

(5) thus holds.

3. THE UPPER BOUND ESTIMATES

In this section we shall show that if the eigenvalues λ_l have suitable decays, the upper bounds of $\mathcal{N}(\overline{I_K(B_R)}, \eta)$ may be estimated.

Theorem 3.1. *If the eigenvalues λ_l of the Mercer kernel $K(x, y)$ defined on $S^q \times S^q$ satisfy $\lambda_l > 0$, $\lambda_l d_l^q \sim \frac{1}{(1+l)^\alpha}$, $l \rightarrow +\infty$, $\alpha \geq q$, then there is a positive constant B_1 such that for $0 < \eta \leq \min\{\frac{R}{2}, 2B_1 \left(\frac{6\delta'}{5\gamma}\right)^{\frac{\alpha-3}{2}} R\}$ there are constants $A > 0$, $A' > 0$, depending on B_1, q and α such that*

$$(11) \quad \ln \mathcal{N}(\overline{I_K(B_R)}, \eta) \leq A' \left(\frac{R}{\eta}\right)^q \ln \left[A \left(\frac{R}{\eta}\right) \right].$$

(11) has the similar form as the Theorem 2 of [9] which and (5) shows that the kernels with eigenvalues of power decays share the same covering number as that of Sobolev space (see [7]).

Theorem 3.2. *If the eigenvalues λ_l of the Mercer kernel $K(x, y)$ defined on $S^q \times S^q$ satisfy $\lambda_l > 0$, and $\lambda_l d_l^q \sim \frac{1}{e^{(1+l)\alpha}}$, $l \rightarrow +\infty$, $\alpha > 0$, then, there are positive numbers $B_2 > 0$, $c_0 > 0$, such that for $0 < \eta \leq \min\{\frac{R}{2}, 2B_2 e^{-\frac{5\gamma c_0}{6\delta^2}} R\}$ there are positive constants $B > 0$, $B' > 0$, $B'' > 0$, such that*

$$(12) \quad \ln \mathcal{N}(\overline{I_K(B_R)}, \eta) \leq B \left(\frac{R}{\eta}\right)^q \ln \left[B' \left(\frac{R}{\eta}\right)\right] + B'' \left(\frac{R}{\eta}\right)^{q+1}.$$

(12) has the similar form as the Proposition 3 of [7]. Since the right side of (11) is smaller than the right side of (12) we know that the Mercer kernel Hilbert space whose eigenvalues have the decay of powers have a smaller covering number than that of the exponential decays. This fact may also be seen from (4) and (5).

To show Theorem 3.1 and Theorem 3.2, we need some lemmas.

Assume that $\{X_N : N \in \mathcal{N}\}$ is a family of finite subsets of X such that

$$d_N := \max_{x \in X} \min_{y \in X_N} d(x, y) \rightarrow 0, \quad N \rightarrow +\infty.$$

This means that the discrete knot X_N become dense in X as N tends to the infinity. Let the function measuring the regularity of K be defined by

$$\begin{aligned} \varepsilon_K(N) : &= \sup_{x \in X} \left[\inf \left(K(x, x) - 2 \sum_{y \in X_N} w_y K(x, y) \right. \right. \\ &\quad \left. \left. + \sum_{y, t \in X_N} w_y K(y, t) w_t : w_y \in \mathcal{R} \right)^{\frac{1}{2}} \right], \end{aligned}$$

the cardinality of the set X_N be ΞX_N and, A_N be the positive definite matrix $A_N := [K(y, t)]_{y, t \in X_N}$. Then, D. X. Zhou gave the following estimate.

Lemma 3.1. (see [7]). *Let $K(x, y)$ be a Mercer kernel on $X \times X$, I_K be given as in Section 1. Then for $0 < \eta \leq \frac{R}{2}$ there holds*

$$(13) \quad \ln \mathcal{N}(\overline{I_K(B_R)}, \eta) \leq (\Xi X_N) \ln \left[8 \|K\|_{\infty}^{\frac{3}{2}} (\Xi X_N) \|A_N^{-1}\|_{l^2(X_N)} \frac{R}{\eta} \right],$$

where N are any integers satisfying $\varepsilon_K(N) \leq \frac{\eta}{2R}$.

Let V be a finite-dimensional vector space with norm $\|\cdot\|_V$ and let $Z \subset V^*$ be a finite set consisting of N functionals. Here, V^* denotes the dual space of V consisting of all linear and continuous functionals defined on V . If a mapping $T : V \rightarrow T(V) \subset \mathcal{R}^N$ defined by $T(v) = \{z(v)\}_{z \in Z}$ is injective, we call T a sampling operator and Z a norming set for V .

Lemma 3.2. (see [18, Theorem 3.4]). *Suppose V is a finite-dimensional normed linear space and $Z = \{z_1, z_2, \dots, z_N\}$ is a norming set for V , T being*

the corresponding sampling operator. For every $\psi \in V^*$ there exists a vector $u \in \mathcal{R}^N$ depending only on ψ such that, for every $v \in V$, $\psi(v) = \sum_{j=1}^N u_j z_j(v)$, and

$$(14) \quad \|u\|_{(\mathcal{R}^N)^*} \leq \|\psi\|_{V^*} \|T^{-1}\|,$$

where

$$\|T^{-1}\| = \sup_{v \in V \setminus \{0\}} \frac{\|v\|_V}{\|T(v)\|_{\mathcal{R}^N}}.$$

We now give a refined Marcinkiewicz-Zygmund inequality for the harmonics, itself has independent meaning.

Lemma 3.3 For the knot set \wedge in Lemma 2.2 we have

$$(15) \quad \frac{1}{6} \|p\|_{\infty(S^q)} \leq \max_{\omega \in \wedge} |p(\omega)| \leq \|p\|_{\infty(S^q)}, \quad p \in \Pi_{[\frac{5\gamma n}{6\delta'}]}^q.$$

Proof. Let $\|p\|_{\infty(S^q)} = |p(x_0)|$, $\omega(x_0)$ be the nearest point of \wedge to x_0 . The Bernstein inequality for the spherical harmonics (see [32]) makes

$$\left| p(x_0) - p(\omega(x_0)) \right| \leq \frac{5\gamma n}{6\delta'} \times d(x_0, \omega(x_0)) \|p\|_{\infty(S^q)}, \quad p \in \Pi_{[\frac{5\gamma n}{6\delta'}]}^q.$$

By the definition of \wedge we know $d(x_0, \omega(x_0)) \leq \frac{\delta'}{\gamma n}$. Hence,

$$\left| p(x_0) - p(\omega(x_0)) \right| \leq \frac{5\gamma n}{6\delta'} \times \frac{\delta'}{\gamma n} \times \|p\|_{\infty(S^q)} \leq \frac{5}{6} \|p\|_{\infty(S^q)}.$$

(15) then holds.

We now give a way to show the existence of the local reproducing basis functions for the harmonics.

Lemma 3.4 Let \wedge be the knot points on S^q in Lemma 2.2. Then, there exist for $\omega \in \wedge$ functions $u_\omega(x) : S^q \rightarrow \mathcal{R}$ such that

- (i). $\sum_{\omega \in \wedge} u_\omega(x) Y(\omega) = Y(x)$ for all $Y(x) \in \Pi_{[\frac{5\gamma n}{6\delta'}]}^q$ and $x \in S^q$;
- (ii). $\sum_{\omega \in \wedge} |u_\omega(x)| \leq 6$ for all $x \in S^q$.

Proof. Let $(V, \|\cdot\|) = (\Pi_n^q, \|\cdot\|_\infty)$ and $(T(V), \|\cdot\|_{T(V)}) = (\mathcal{R}^N, \|\cdot\|_\infty)$. Then, the sampling operator is $T(p) = \{p(\omega)\}_{\omega \in \wedge} \in \mathcal{R}^{|\wedge|}$ and in this case $(\mathcal{R}^{|\wedge|}, \|\cdot\|_\infty)^* = (\mathcal{R}^{|\wedge|}, \|\cdot\|_{l^1})$. For $x \in S^q$ the evaluation functional δ_x , i.e., $\delta_x(p) = p(x)$, belongs to V^* and $|\delta_x(p)| = |p(x)| \leq \|p\|_\infty$.

It follows by (15) that $\|T^{-1}\| \leq 6$. According to Lemma 3.2 there exists functions $u_\omega(x)$, $\omega \in \Lambda$, such that (i) holds and by (14) there holds

$$\sum_{\omega \in \Lambda} |u_\omega(x)| \leq \|\delta_x\| \|T^{-1}\| \leq 6.$$

With the help of Lemma 3.4 we now give a kind of estimate for $\varepsilon_K(N)$.

Lemma 3.5 *Let the Mercer kernel $K(x, y)$ defined on $S^q \times S^q$ satisfies $\lambda_l > 0$, $l = 0, 1, 2, \dots$, and $\sum_{l=0}^{+\infty} \lambda_l d_l^q < +\infty$. Λ is the knot set in Lemma 2.2 and*

$$\begin{aligned} \varepsilon_K(\Lambda) = \sup_{x \in S^q} & \left[\inf \left(K(x, x) - 2 \sum_{\omega \in \Lambda} w_\omega K(x, \omega) \right. \right. \\ & \left. \left. + \sum_{\omega, \omega' \in \Lambda} w_\omega K(\omega, \omega') w_{\omega'} : w_\omega, w_{\omega'} \in \mathcal{R} \right)^{\frac{1}{2}} \right]. \end{aligned}$$

Then,

$$(16) \quad \varepsilon_K(\Lambda) \leq \left(24 \sum_{l=\lceil \frac{5\gamma n}{6\delta} \rceil}^{+\infty} \lambda_l d_l^q \right)^{\frac{1}{2}}.$$

In particular, if $\lambda_l d_l^q \sim \frac{1}{(1+l)^\alpha}$, $l \rightarrow +\infty$, $\alpha > 1$, then there is a positive constant $B_1 > 0$ such that

$$(17) \quad \varepsilon_K(\Lambda) \leq B_1 n^{-\frac{\alpha-1}{2}}.$$

If $\lambda_l d_l^q \sim \frac{1}{e^{\alpha(1+l)}}$, $l \rightarrow +\infty$, $\alpha > 0$, then there is a positive constant $B_2 > 0, c_0 > 0$, such that

$$(18) \quad \varepsilon_K(\Lambda) \leq B_2 e^{-c_0 n}.$$

Proof. By the definition of Λ we know $d_n = \max_{x \in S^q} \min_{\omega \in \Lambda} d(x, \omega) \rightarrow 0$, $n \rightarrow +\infty$. Take $w_\omega = u_\omega(x)$, then

$$\begin{aligned} & \left| K(x, x) - 2 \sum_{\omega \in \Lambda} w_\omega K(x, \omega) + \sum_{\omega, \omega' \in \Lambda} w_\omega K(\omega, \omega') w_{\omega'} \right| \\ = & \left| K(x, x) - 2 \sum_{\omega \in \Lambda} u_\omega(x) K(x, \omega) + \sum_{\omega, \omega' \in \Lambda} u_\omega(x) K(\omega, \omega') u_{\omega'}(x) \right| \\ = & \left| \sum_{l=\lceil \frac{5\gamma n}{6\delta} \rceil+1}^{+\infty} \lambda_l d_l^q \sum_{\omega, \omega' \in \Lambda} u_\omega(x) \left(p_l^{q+1}(1) - 2p_l^{q+1}(x \cdot \omega') + p_l^{q+1}(\omega \cdot \omega') \right) \right| \\ \leq & 24 \sum_{l=\lceil \frac{5\gamma n}{6\delta} \rceil+1}^{+\infty} \lambda_l d_l^q. \end{aligned}$$

When $\lambda_l d_l^q \sim \frac{1}{(1+l)^\alpha}$, $\alpha > 1$, one has a positive constant $C_1 > 0$ such that

$$\begin{aligned} \sum_{l=\lceil \frac{5\gamma n}{6\delta'} \rceil + 1}^{+\infty} \lambda_l d_l^q &\leq C_1 \sum_{l=\lceil \frac{5\gamma n}{6\delta'} \rceil + 1}^{+\infty} \frac{1}{(1+l)^\alpha} \\ &\leq C_1 \int_{\frac{5\gamma n}{6\delta'}}^{+\infty} \frac{dl}{(1+l)^\alpha} \leq (\alpha - 1) \left(1 + \frac{5\gamma n}{6\delta'}\right)^{1-\alpha}. \end{aligned}$$

(17) then holds.

When $\lambda_l d_l^q \sim e^{-\alpha(1+l)}$, $\alpha > 0$, one has a positive constant $C_2 > 0$ such that

$$\begin{aligned} \sum_{l=\lceil \frac{5\gamma n}{6\delta'} \rceil + 1}^{+\infty} \lambda_l d_l^q &\leq C_2 \sum_{l=\lceil \frac{5\gamma n}{6\delta'} \rceil + 1}^{+\infty} e^{-\alpha(1+l)} \\ &\leq C_2 \int_{\frac{5\gamma n}{6\delta'}}^{+\infty} \frac{dl}{e^{\alpha(1+l)}} \leq \frac{C_2}{\alpha e^\alpha} \times e^{-\frac{5\gamma \alpha n}{6\delta'}}. \end{aligned}$$

Hence, (18) holds.

Proof of Theorem 3.1. For $0 < \eta \leq \frac{R}{2}$ we have by $B_1 n^{-\frac{\alpha-1}{2}} \leq \frac{\eta}{2R}$ that $n \geq \left(2B_1\right)^{\frac{1}{\alpha-1}} \left(\frac{R}{\eta}\right)^{\frac{2}{\alpha-1}}$. Therefore, for $0 < \eta \leq \left(2B_1\right)^{\frac{1}{2}} \left(\frac{6\delta'}{5\gamma}\right)^{\frac{\alpha-1}{2}} R$ we have $n \geq \frac{5\gamma}{6\delta'}$, and we can thus choose $n \in \mathcal{N}$ such that

$$\frac{5\gamma}{6\delta'} \leq n \leq \frac{5\gamma}{6\delta'} \times \left(2B_1\right)^{\frac{1}{2}} \left(\frac{6\delta'}{5\gamma}\right)^{\frac{\alpha-1}{2}} \frac{R}{\eta}.$$

By (10) we know $|\wedge| \|A_n^{-1}\|_{l^2} \leq \frac{1}{\min_{0 \leq l \leq n} \lambda_l} \sim n^\alpha d_n^q$. Since $d_n^q \sim n^q, n \rightarrow +\infty$, we have a constant $C_3 > 0$ such that $|\wedge| \|A_n^{-1}\|_{l^2} \leq C_3 n^{\alpha+q}$. Then, (13) makes

$$\ln \mathcal{N}(\overline{I_K(B_R)}, \eta) \leq |\wedge| \ln \left(8 \left\| \sum_{l=0}^{+\infty} \lambda_l d_l^q P_l^{q+1}(x \cdot y) \right\|_\infty^{\frac{3}{2}} \times C_3 n^{\alpha+q} \times \frac{R}{\eta} \right).$$

Also, $\|P_l^{q+1}\|_\infty \leq 1$ makes

$$\left\| \sum_{l=0}^{+\infty} \lambda_l d_l^q P_l^{q+1}(x \cdot y) \right\|_\infty \leq C_1 \sum_{l=0}^{+\infty} \frac{1}{(1+l)^\alpha} = \frac{C_1}{(\alpha - 1)}.$$

We then have

$$\begin{aligned}
 & \ln \mathcal{N}(\overline{I_K(B_R)}, \eta) \\
 & \leq |\wedge| \ln \left(\frac{8C_1^{\frac{3}{2}}}{(\alpha-1)^{\frac{3}{2}}} \times C_3 n^{\alpha+q} \frac{R}{\eta} \right) \\
 & \sim n^q \ln \left(\frac{8\sqrt{C_1^3} C_3 n^{\alpha+q} R}{\sqrt{(\alpha-1)^3} \eta} \right) \\
 & \leq \left(\frac{5\gamma}{6\delta'} \right)^q \times (2B_1)^{\frac{q}{2}} \left(\frac{6\delta'}{5\gamma} \right)^{\frac{(\alpha-1)q}{2}} \left(\frac{R}{\eta} \right)^q \\
 & \quad \times \ln \left(\frac{8\sqrt{C_1^3} C_3 \left(\frac{5\gamma}{6\delta'} \right)^{q+\alpha} \times \left(2B_1 \right)^{\frac{\alpha+q}{2}} \left(\frac{6\delta'}{5\gamma} \right)^{\frac{(\alpha-1)(\alpha+q)}{2}}}{\sqrt{(\alpha-1)^3}} \times \left(\frac{R}{\eta} \right)^{\alpha+q+1} \right)
 \end{aligned}$$

(11) thus holds.

Proof of Theorem 3.2. For $0 < \eta \leq \frac{R}{2}$ we have by $B_2 e^{-c_0 n} \leq \frac{\eta}{2R}$ that

$$n \geq \frac{1}{c_0} \ln \frac{2B_2 R}{\eta}.$$

Thus, for $0 < \eta \leq 2B_2 e^{-\frac{5\gamma c_0}{6\delta'}} R$ we have $n \geq \frac{5\gamma}{6\delta'}$. We can choose $n \in \mathcal{N}$ such that

$$\frac{5\gamma}{6\delta'} \leq n \leq \frac{5\gamma}{6\delta'} \times 2B_2 e^{-\frac{5\gamma c_0}{6\delta'}} \frac{R}{\eta}.$$

Since $|\wedge| \|A_n^{-1}\|_{l^2} \leq \frac{1}{\min_{0 \leq l \leq n} \lambda_l} \sim d_l^q e^{n\alpha} \sim n^q e^{n\alpha}$, $n \rightarrow +\infty$, by the fact $\sum_{l=0}^{+\infty} \frac{1}{e^{\alpha(1+l)}} = \frac{1}{e^\alpha - 1}$ and (13) we have a constant $C_4 > 0$ such that

$$\begin{aligned}
 & \ln \mathcal{N}(\overline{I_K(B_R)}, \eta) \\
 & \leq |\wedge| \ln \left(8 \left\| \sum_{l=0}^{+\infty} \lambda_l d_l^q P_l^{q+1}(x \cdot y) \right\|_{\infty}^{\frac{3}{2}} \times C_4 n^q e^{n\alpha} \frac{R}{\eta} \right) \\
 & \leq |\wedge| \ln \left(\left[\frac{C_2}{e^\alpha - 1} \right]^{\frac{3}{2}} \times C_4 n^q e^{n\alpha} \frac{R}{\eta} \right) \\
 & \sim n^q \ln \left(\left[\frac{C_2}{e^\alpha - 1} \right]^{\frac{3}{2}} \times C_4 n^q \times \exp \left\{ \alpha n \right\} \frac{R}{\eta} \right) \\
 & \leq \left(\frac{10\gamma B_2}{6\delta'} \times e^{-\frac{5\gamma c_0}{6\delta'}} \frac{R}{\eta} \right)^q \times \ln \left(\left[\frac{C_2}{e^\alpha - 1} \right]^{\frac{3}{2}} \times C_4 \times \left(\frac{10\gamma B_2}{6\delta'} \times e^{-\frac{5\gamma c_0}{6\delta'}} \frac{R}{\eta} \right)^q \right. \\
 & \quad \left. \times \exp \left\{ \alpha \times \frac{10\gamma B_2}{6\delta'} \times e^{-\frac{5\gamma c_0}{6\delta'}} \times \frac{R}{\eta} \right\} \frac{R}{\eta} \right).
 \end{aligned}$$

(12) thus holds.

Corollary 3.1 Let $C_5 = 2^4(q+1)(q+2)\left(\frac{\gamma}{1-\gamma}\right)^{q+2}$. Then for $0 < \eta \leq \{\frac{R}{2}, \sqrt{96C_5}R\}$ there are two constants $D > 0, D' > 0$, such that

$$(19) \quad \ln \mathcal{N}(\overline{I_{K_0}(B_R)}, \eta) \leq D' \left(\frac{R}{\eta}\right)^q \ln \left[D \left(\frac{R}{\eta}\right) \right].$$

Proof. In this case, $\lambda_l = \gamma^l = \frac{1}{\left(1 + \left(\frac{1}{\gamma} - 1\right)\right)^l}$, $0 < \gamma < 1$, $l = 0, 1, 2, \dots$.

When $n > q + 1$, $l \geq n + 1 > q + 2$. Therefore,

$$\left(1 + \left(\frac{1}{\gamma} - 1\right)\right)^{l+q} > \binom{l+q}{q+2} \left(\frac{1}{\gamma} - 1\right)^{q+2} = \left(\frac{1}{\gamma} - 1\right)^{q+2} \frac{l(l-1)(l+q-1) d_l^q}{(2l+q)(q+2)(q+1)},$$

and, $\lambda_l < \frac{\gamma^{q+2} (2l+q)(q+2)(q+1)}{(1-\gamma)^{q+2} l(l-1)(l+q-1) d_l^q}$. It follows

$$\begin{aligned} \lambda_l d_l^q &< \frac{\gamma^{q+2} (2l+q)(q+2)(q+1)}{(1-\gamma)^{q+2} l(l-1)(l+q-1)} \\ &\leq \frac{2^4 (q+1)(q+2) \left(\frac{\gamma}{1-\gamma}\right)^{q+2}}{(1+l)^2} = \frac{C_5}{(1+l)^2}, \end{aligned}$$

where $C_5 = 2^4(q+1)(q+2)\left(\frac{\gamma}{1-\gamma}\right)^{q+2}$. By (11) we have for $0 < \eta \leq \{\frac{R}{2}, \sqrt{96C_5}R\}$ there are two constants $D > 0, D' > 0$, such that (19) holds.

ACKNOWLEDGMENT

The authors thank the referees for valuable comments and suggests.

REFERENCES

1. F. Cucker and S. Smale, On the mathematical foundations of learning, *Bulletin (New Series) of the American Mathematical Society*, **39(1)** (2001), 1-49.
2. T. Evgeniou, M. Pontil and T. Poggio, Regularization networks and support vector machines, *Advances in Computational Mathematics*, **13(1)** (2000), 1-50.
3. V. N. Vapnik, *Statistical Learning Theory*, New York, John Wiley and Sons, Inc., 1998.
4. F. Cucker and S. Smale, Best choices for regularization parameters in learning theory: on the bias-variance problem, *Found. Comput. Math.*, **2** (2002), 413-428.
5. R. C. Williamson, A. J. Smola and B. Schölkopf, Generalization performance of regularization networks and support vector machine via entropy numbers of compact operators, *IEEE Trans. Inform. Theory*, **47(6)** (2001), 2516-2532.

6. Y. Guo, P. L. Bartlett, J. Shawe-Taylor, and R. C. Williamson, Covering numbers for support vector machines, *IEEE Trans. Inform. Theory*, **48** (2002), 239-250.
7. D. X. Zhou, The covering number in learning theory, *J. Complexity*, **18** (2002), 739-767.
8. D. X. Zhou, Capacity of reproducing kernel spaces in learning theory, *IEEE Trans. Inform. Theory*, **49(7)** (2003), 1743-1752.
9. H. W. Sun and D. X. Zhou, Reproducing kernel Hilbert spaces associated with analytic translation-invariant Mercer kernels, *J. Fourier Anal. Appl.*, **14** (2008), 89-101.
10. N. Aronszajn, Theory of reproducing kernels, *Trans. Amer. Math. Soc.*, **68** (1950), 337-404.
11. C. Carmeli, E. De Vito and A. Toigo, Vector valued reproducing kernel Hilbert spaces of integrable functions and Mercer theorem, *Analysis and Applications*, **4(4)** (2006), 377-408.
12. H. Sun, Mercer theorem for RKHS on noncompact sets, *J. Complexity*, **21** (2005), 337-349.
13. F. Cucker and D. X. Zhou, *Learning Theory: An Approximation Theory Viewpoint*, Cambridge University, New York, 2007.
14. B. H. Sheng, J. L. Wang and P. Li, The covering number for some Mercer kernel Hilbert spaces, *J. Complexity*, **24** (2008), 241-258.
15. H. Groemer, *Geometric Applications of Fourier Series and Spherical Harmonics*, Cambridge University Press, New York, 1996.
16. K. Y. Wang and L. Q. Li, *Harmonic Analysis and Approximation on the Unit Sphere*, Science Press, Beijing/New York, 2000.
17. C. Muller, *Spherical Harmonic*, Springer-Verlag, Berlin, 1966.
18. H. Wendland, *Scattered Aata Approximation*, Cambridge University Press, Cambridge, 2005.
19. R. S. Womersly and I. H. Sloan, How good can polynomial interpolation on the sphere be? *Advances in Computational Mathematics*, **14(3)** (2001), 195-226.
20. F. Dai, Multivariate polynomial inequalities with respect to doubling weights and A_∞ weights, *J. Functional Analysis*, **235** (2006), 137-170.
21. B. H. Sheng and G. Z. Zhou, A way of constructing spherical approximation operators, *Journal of Systems Science and Mathematical Sciences*, **28(4)** (2008), 456-467, (in Chinese).
22. B. H. Sheng, J. L. Wang and S. P. Zhou, A way of constructing zonal translation network operators with linear bounded operators, *Taiwanese J. Mathematics*, **12(1)** (2008), 77-92.

23. B. H. Sheng, On the degree of approximation by spherical translations, *Acta Mathematicae Applicatae Sinica, English Series*, **22(4)** (2006), 671-680.
24. H. N. Mhaskar, F. J. Narcowich and J. D. Ward, Approximation properties of zonal function networks using scattered data on the sphere, *Advances in Computational Mathematics*, **11** (1999), 121-127.
25. H. N. Mhaskar, F. J. Narcowich and D. Ward, Zonal function network frames on the sphere, *Neural Networks*, **16** (2003), 183-203.
26. B. H. Sheng, The Jackson theorem of approximation by zonal translation networks on the unit sphere, *Advances in Mathematics (Beijing)*, **35(3)** (2006), 325-335, (in Chinese).
27. H. N. Mhaskar, Weighted quadrature formulas and approximation by zonal function networks on the sphere, *J. Complexity*, **22** (2006), 348-370.
28. G. Brown and F. Dai, Approximation of smooth functions on compact two-point homogeneous spaces, *J. Functional Analysis*, **220** (2005), 401-423.
29. G. Brown, F. Dai and Y. S. Sun, Kolmogorov width of classes of smooth functions on the sphere S^{d-1} , *Journal of Complexity*, **18** (2002), 1001-1023.
30. G. Brown, F. Dai and Y. S. Sun, Kolmogorov width of classes of smooth functions on the sphere S^{d-1} , *Advances in Mathematics (Beijing)*, **31(2)** (2002), 181-184.
31. H. N. Mhaskar, F. J. Narcowich, N. Sivakumar and J. D. Ward, Approximation with interpolatory constraints, *Proc. Amer. Math. Soc.*, **130(5)** (2001), 1355-1361.
32. K. Jetter, J. Stöckler and J. D. Ward, Error estimates for scattered data interpolation on spheres, *Mathematics of Computation*, **68(226)** (1999), 733-747.

Sheng Baohuai, Wang Jianli and Chen Zhixiang
Department of Mathematics,
Shaoxing University,
Shaoxing, Zhejiang 312000,
P. R. China
E-mail: shengbaohuai@163.com
bhsheng@usx.edu.cn