Appl. Comput. Harmon. Anal. 32 (2012) 401–412

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Multiscale approximation for functions in arbitrary Sobolev spaces by scaled radial basis functions on the unit sphere

Q.T. Le Gia [a,*], I.H. Sloan [a], H. Wendland [b]

[a] School of Mathematics and Statistics, University of New South Wales, Sydney, Australia
[b] Mathematical Institute, Oxford University, UK

ARTICLE INFO

Article history:
Received 18 April 2011
Revised 25 July 2011
Accepted 30 July 2011
Available online 5 August 2011
Communicated by Charles K. Chui

Keywords:
Multiscale approximation
Radial basis function
Unit sphere

ABSTRACT

In this paper, we prove convergence results for multiscale approximation using compactly supported radial basis functions restricted to the unit sphere, for target functions outside the reproducing kernel Hilbert space of the employed kernel.

1. Introduction

In the geosciences data are often collected at scattered sites on the unit sphere. Moreover, the geophysical data typically occur at many length scales: for example, the topography of central Australia varies slowly, while that of the Himalayan mountains varies rapidly. To handle multiscale data at scattered locations, we consider an approximation scheme based on radial basis functions of different scales, which are generated from a single underlying radial basis function (RBF) Φ using a sequence of scales $\delta_1, \delta_2, \ldots$ with limit zero, with the scale becoming smaller as the point set becomes denser. The scaled RBF is defined by $\Phi_\delta = c_\delta \Phi(\frac{\cdot}{\delta})$, which is then restricted to the unit sphere. There is significant difficulty in dealing with more than one scale at the same time, namely that the associated reproducing kernel Hilbert space (RKHS) or “native space” has a different inner product for each scale. For this reason a multiresolution analysis within a single Hilbert space, of the kind familiar from wavelet analysis, does not seem possible for scaled RBFs on either a sphere or a Euclidean region.

In a recent paper [1], we resolved that issue, and constructed a multiresolution analysis for sufficiently smooth functions on the unit sphere $\mathbb{S}^n$ based on RBFs of different scales. We have proved a convergence result for a multiscale approximation algorithm for functions from the RKHS of the given compactly supported kernel, if the RKHS is a sufficiently smooth Sobolev space.

In this paper, we shall extend our convergence results to rougher target functions. Specifically, if the employed kernel is the reproducing kernel of, say, $H^\sigma(\mathbb{S}^n)$ then the new results will hold for all target functions from $H^\beta(\mathbb{S}^n)$ with $\sigma > \beta > n/2$.

The paper is organized as follows. In Section 2, after reviewing the necessary background on spherical harmonics and spherical kernels defined from radial basis functions, we review Sobolev splines and prove a new result on the asymptotic behavior of the Fourier–Legendre coefficients of spherical kernels defined by restricting the Sobolev splines to the unit

* Corresponding author.
E-mail addresses: qlegia@unsw.edu.au (Q.T. Le Gia), i.sloan@unsw.edu.au (I.H. Sloan), Holger.Wendland@maths.ox.ac.uk (H. Wendland).

1063-5203/$ – see front matter
doi:10.1016/j.acha.2011.07.007

sphere. In Section 3 we review the method of multiscale approximation using scaled compactly supported radial basis functions which was analyzed in [1] for functions lying in the reproducing kernel Hilbert space defined by the spherical kernel. In Section 4 we give a convergence analysis for "rougher" target functions. The main result of the paper is stated in Theorem 4.5.

2. Preliminaries

2.1. Positive definite bizonal kernels on the unit sphere

The unit sphere $\mathbb{S}^n$ is defined by $\mathbb{S}^n = \{\mathbf{x} \in \mathbb{R}^{n+1}\colon\ |\mathbf{x}| = 1\}$, where $|\mathbf{x}|$ denotes the Euclidean norm of $\mathbf{x}$. Bizonal kernels defined on $\mathbb{S}^n \times \mathbb{S}^n$ are kernels that can be represented as $g(\mathbf{x},\mathbf{y}) = \widetilde{g}(\mathbf{x}\cdot\mathbf{y})$ for all $\mathbf{x},\mathbf{y} \in \mathbb{S}^n$, where $\widetilde{g}(t)$ is a continuous function on $[-1,1]$. We shall be concerned exclusively with bizonal kernels of the type

$$g(\mathbf{x},\mathbf{y}) = \sum_{\ell=0}^{\infty} a_\ell P_\ell(n+1;\mathbf{x}\cdot\mathbf{y}), \quad a_\ell > 0,\ \sum_{\ell=0}^{\infty} a_\ell < \infty, \tag{1}$$

where $\{P_\ell(n+1;t)\}_{\ell=0}^{\infty}$ is the sequence of $(n+1)$-dimensional Legendre polynomials normalized to $P_\ell(n+1;1) = 1$.

Thanks to the seminal work of Schoenberg [2] and the later work of [3], we know that such a g is (strictly) positive definite on $\mathbb{S}^n$, that is, the matrix $\mathbf{A} := [g(\mathbf{x}_i,\mathbf{x}_j)]_{i,j=1}^{M}$ is positive definite for every set of distinct points $\{\mathbf{x}_1,\ldots,\mathbf{x}_M\}$ on $\mathbb{S}^n$ and every positive integer M.

For mathematical analysis, sometimes it is convenient to expand the kernel g into a series of spherical harmonics. A detailed discussion on spherical harmonics can be found in [4]. In brief, spherical harmonics are the restriction to $\mathbb{S}^n$ of homogeneous polynomials Y in $\mathbb{R}^{n+1}$ which satisfy $\Delta Y = 0$, where Δ is the Laplacian operator in $\mathbb{R}^{n+1}$. The space of all spherical harmonics of degree ℓ on $\mathbb{S}^n$, denoted by $\mathcal{H}_\ell$, has an orthonormal basis

$$\{Y_{\ell k}\colon\ k = 1,\ldots,N(n,\ell)\},$$

where

$$N(n,0) = 1 \quad \text{and} \quad N(n,\ell) = \frac{(2\ell+n-1)\Gamma(\ell+n-1)}{\Gamma(\ell+1)\Gamma(n)} \quad \text{for } \ell \geqslant 1.$$

The space of spherical harmonics of degree $\leqslant L$ will be denoted by $\mathcal{P}_L := \bigoplus_{\ell=0}^{L}\mathcal{H}_\ell$; it has dimension $N(n+1,L)$. Every function $f \in L_2(\mathbb{S}^n)$ can be expanded in terms of spherical harmonics,

$$f = \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(n,\ell)} \hat{f}_{\ell k} Y_{\ell k}, \qquad \hat{f}_{\ell k} = \int_{\mathbb{S}^n} f\,\overline{Y_{\ell k}}\,dS,$$

where dS is the surface measure of the unit sphere. Using the addition theorem for spherical harmonics (see, for example, [4, p. 18]), we can write

$$g(\mathbf{x},\mathbf{y}) = \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(n,\ell)} \hat{g}(\ell) Y_{\ell k}(\mathbf{x})\overline{Y_{\ell k}(\mathbf{y})}, \quad \text{with } \hat{g}(\ell) = \frac{\omega_n}{N(n,\ell)} a_\ell, \tag{2}$$

where $\omega_n = \int_{\mathbb{S}^n} dS$ is the surface area of $\mathbb{S}^n$. We shall refer to $\hat{g}(\ell)$ as the Fourier symbol of g.

2.2. Kernels defined from radial basis functions

Let $\Phi = \rho(|\cdot|) : \mathbb{R}^{n+1} \to \mathbb{R}$ be a compactly supported radial basis function (RBF) with associated RKHS $H^{\sigma+1/2}(\mathbb{R}^{n+1})$ with $\sigma > n/2$. Examples of such RBFs are Wendland's functions (see [5]).

By restricting the function Φ to the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$, we have a positive definite, bizonal kernel on the unit sphere

$$\phi(\mathbf{x},\mathbf{y}) = \Phi(\mathbf{x}-\mathbf{y}) = \rho(|\mathbf{x}-\mathbf{y}|), \quad \mathbf{x},\mathbf{y} \in \mathbb{S}^n.$$

It is known (see [6]) that the Fourier symbol $\hat{\phi}(\ell)$ satisfies $\hat{\phi}(\ell) \sim (1+\ell)^{-2\sigma}$, i.e. there exist constants $c_1, c_2 > 0$ such that

$$c_1(1+\ell)^{-2\sigma} \leqslant \hat{\phi}(\ell) \leqslant c_2(1+\ell)^{-2\sigma}, \quad \ell \geqslant 0. \tag{3}$$

In the remainder of the paper, we use $c_1, c_2, \ldots$ to denote specific constants while c, C are generic constants that may take different values at each occurrence.

The reproducing kernel Hilbert space (RKHS) of the kernel ϕ is given by

$$\mathcal{N}_\phi := \left\{ f \in \mathcal{D}'(\mathbb{S}^n):\ \|f\|_\phi^2 := \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{|\hat{f}_{\ell k}|^2}{\widehat{\phi}(\ell)} < \infty \right\},$$

where $\mathcal{D}'(\mathbb{S}^n)$ is the set of all tempered distributions defined on $\mathbb{S}^n$. Alternatively, $\mathcal{N}_\phi$ is the completion of $\operatorname{span}\{\phi(\cdot,\mathbf{x}):\ \mathbf{x} \in \mathbb{S}^n\}$ with respect to the norm $\|\cdot\|_\phi$. The norm obviously comes from the following inner product

$$(f,g)_\phi = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{\hat{f}_{\ell k}\hat{g}_{\ell k}}{\hat{\phi}(\ell)}, \quad f,g \in \mathcal{N}_\phi.$$

The kernel ϕ has the reproducing property with respect to this inner product, that is

$$f(\mathbf{x}) = \big(f, \phi(\cdot,\mathbf{x})\big)_\phi, \quad \mathbf{x} \in \mathbb{S}^n,\ f \in \mathcal{N}_\phi.$$

The Sobolev space $H^\sigma(\mathbb{S}^n)$ with real parameter σ is defined by

$$H^\sigma(\mathbb{S}^n) := \left\{ f \in \mathcal{D}'(\mathbb{S}^n):\ \|f\|_{H^\sigma(\mathbb{S}^n)}^2 := \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} (1+\ell)^{2\sigma} |\hat{f}_{\ell k}|^2 < \infty \right\}.$$

For $\sigma = 0$ the space is just $L_2(\mathbb{S}^n)$. Under the condition (3), the norms $\|\cdot\|_\phi$ and $\|\cdot\|_{H^\sigma(\mathbb{S}^n)}$ are equivalent, since

$$c_1^{1/2}\|f\|_\phi \leqslant \|f\|_{H^\sigma(\mathbb{S}^n)} \leqslant c_2^{1/2}\|f\|_\phi. \tag{4}$$

For a given $\delta > 0$, we define the scaled version ϕ_δ of the kernel ϕ by

$$\phi_\delta(\mathbf{x},\mathbf{y}) = \delta^{-n}\Phi\big((\mathbf{x}-\mathbf{y})/\delta\big) = \delta^{-n}\rho\big(|\mathbf{x}-\mathbf{y}|/\delta\big). \tag{5}$$

We can expand ϕ_δ into a series of spherical harmonics

$$\phi_\delta(\mathbf{x},\mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \widehat{\phi_\delta}(\ell) Y_{\ell k}(\mathbf{x})\overline{Y_{\ell k}(\mathbf{y})}, \tag{6}$$

in which the Fourier coefficients satisfy the following condition (see [1, Theorem 6.2])

$$c_1(1+\delta\ell)^{-2\sigma} \leqslant \widehat{\phi_\delta}(\ell) \leqslant c_2(1+\delta\ell)^{-2\sigma}, \quad \ell \geqslant 0, \tag{7}$$

with the coefficients c_1 and c_2 from (3) possibly relaxed so that (7) holds for all $0 < \delta \leqslant 1$.

For a function $f \in H^\sigma(\mathbb{S}^n)$, we define the norm corresponding to the scaled kernel ϕ_δ by

$$\|f\|_{\phi_\delta} = \left(\sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{|\hat{f}_{\ell k}|^2}{\widehat{\phi_\delta}(\ell)} \right)^{1/2}, \tag{8}$$

and the corresponding inner product is

$$(f,g)_{\phi_\delta} = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{\hat{f}_{\ell k}\hat{g}_{\ell k}}{\hat{\phi}_\delta(\ell)}, \quad f,g \in \mathcal{N}_\phi. \tag{9}$$

Clearly the norms $\|\cdot\|_{\phi_\delta}$ for different δ are all equivalent.

From (7) and (8) it follows (see [1, Lemma 3.1]) that

$$c_1^{1/2}\|g\|_{\phi_\delta} \leqslant \|g\|_{H^\sigma(\mathbb{S}^n)} \leqslant c_2^{1/2}\delta^{-n}\|g\|_{\phi_\delta}. \tag{10}$$

2.3. Sobolev splines and associated spherical kernels

Our target function f will be assumed to belong to the Sobolev space $H^\beta(\mathbb{S}^n)$ for some β satisfying $\sigma > \beta > n/2$.

We need an explicit formula for the reproducing kernel of $H^\beta(\mathbb{S}^n)$. This will be defined via a Sobolev spline in $H^{\beta+1/2}(\mathbb{R}^{n+1})$.

For $\nu > 0$, the Sobolev spline (or Matérn function) is defined by

$$S_\nu(\mathbf{x}) = (2\pi)^{-(n+1)/2} \frac{|\mathbf{x}|^\nu K_\nu(|\mathbf{x}|)}{2^{\nu+(n-1)/2}\Gamma(\nu+(n+1)/2)}, \quad \mathbf{x} \in \mathbb{R}^{n+1},$$

where K_ν is the K-Bessel function of order ν. With the Fourier transform of a function $f \in L^1(\mathbb{R}^{n+1})$ defined by

$$\hat{f}(\boldsymbol{\omega}) = (2\pi)^{-(n+1)/2} \int\limits_{\mathbb{R}^{n+1}} f(\mathbf{x}) e^{-i\mathbf{x}\cdot\boldsymbol{\omega}}\, d\mathbf{x},$$

the Fourier transform of S_ν is given by [5, Theorem 6.13]

$$\widehat{S_\nu}(\boldsymbol{\omega}) = (2\pi)^{-(n+1)/2}\big(1+|\boldsymbol{\omega}|^2\big)^{-\nu-(n+1)/2}.$$

By setting $\nu = \beta - n/2$, the kernel $\Psi(\mathbf{x}-\mathbf{y}) := S_\nu(\mathbf{x}-\mathbf{y})$ is hence the reproducing kernel of the Sobolev space $H^{\beta+1/2}(\mathbb{R}^{n+1})$ with respect to the inner product

$$(f, g)_{H^{\beta+1/2}(\mathbb{R}^{n+1})} = \int\limits_{\mathbb{R}^{n+1}} \hat{f}(\boldsymbol{\omega})\overline{\hat{g}(\boldsymbol{\omega})}\big(1+|\boldsymbol{\omega}|^2\big)^{\beta+1/2} d\boldsymbol{\omega}.$$

For $0 < \delta \leqslant 1$, we define the scaled version of Ψ by

$$\Psi_\delta(\mathbf{x}) = \delta^{-n}\Psi(\mathbf{x}/\delta), \quad \mathbf{x} \in \mathbb{R}^{n+1}. \tag{11}$$

Similarly as before, we define the kernel ψ and its scaled version by restricting Ψ on the sphere,

$$\psi(\mathbf{x},\mathbf{y}) = \Psi(\mathbf{x}-\mathbf{y}), \quad \psi_\delta(\mathbf{x},\mathbf{y}) = \Psi_\delta(\mathbf{x}-\mathbf{y}), \quad \mathbf{x},\mathbf{y} \in \mathbb{S}^n. \tag{12}$$

They can be expanded into a series of spherical harmonics as

$$\psi_\delta(\mathbf{x},\mathbf{y}) = \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(n,\ell)} \hat{\psi}_\delta(\ell) Y_{\ell k}(\mathbf{x})\overline{Y_{\ell k}(\mathbf{y})}. \tag{13}$$

We define the following norm

$$\|f\|_{\psi_\delta} := \left(\sum_{\ell=0}^{\infty}\sum_{k=1}^{N(n,\ell)} \frac{|\hat{f}_{\ell k}|^2}{\hat{\psi}_\delta(\ell)}\right)^{1/2}. \tag{14}$$

Then the RKHS of the kernel ψ_δ is defined by

$$\mathcal{N}_{\psi_\delta} := \big\{ f \in \mathcal{D}'(\mathbb{S}^n) \colon \|f\|_{\psi_\delta} < \infty \big\}.$$

The following two lemmas will give more information about this norm.

Lemma 2.1. *Let Ψ be the reproducing kernel of $H^{\beta+1/2}(\mathbb{R}^{n+1})$, and let ψ_δ be defined by restricting Ψ_δ to the unit sphere as in* (12)*. Then the Fourier symbol $\hat{\psi}_\delta(\ell)$ satisfies*

$$c_3(1+\delta\ell)^{-2\beta} \leqslant \hat{\psi}_\delta(\ell) \leqslant c_4(1+\delta\ell)^{-2\beta}, \quad \ell \geqslant 0, \tag{15}$$

where c_3, c_4 are two positive constants independent of δ and ℓ.

Proof. The proof essentially follows the proof of [1, Theorem 6.2] except for one point, namely the proof that $\hat{\psi}_\delta(\ell) \leqslant c$, for $\ell \geqslant 0$, where $c > 0$ is independent of ℓ and δ. The corresponding proof in [1] was given under the assumption that Ψ is compactly supported with support in the unit ball, which is not the case here. A proof for the present definition of Ψ_δ is as follows.

We have

$$\begin{aligned}\hat{\psi}_\delta(\ell) &= \omega_{n-1}\int_{-1}^{1} \Psi_\delta(\sqrt{2-2t})P_\ell(n+1;t)\big(1-t^2\big)^{(n-2)/2}\,dt \\ &= C_{n,\nu}\delta^{-n}\int_{-1}^{1}\left(\frac{\sqrt{2-2t}}{\delta}\right)^{\nu} K_\nu\left(\frac{\sqrt{2-2t}}{\delta}\right)P_\ell(n+1;t)\big(1-t^2\big)^{(n-2)/2}\,dt,\end{aligned}$$

where $C_{n,\nu} := \omega_{n-1}2^{-\nu-(n-1)/2}(2\pi)^{-(n+1)/2}$. Using the substitution

$$r = \frac{\sqrt{2-2t}}{\delta}$$

and the bound

$$1-t^2=(1+t)(1-t)\leqslant 2(1-t)=\delta^2r^2,$$

together with $|P_\ell(n+1;t)|\leqslant 1$ and $K_\nu(r)>0$ for $r>0$ and $\nu\in\mathbb{R}$, we have an upper bound

$$\begin{aligned}\hat{\psi}_\delta(\ell)&\leqslant C_{n,\nu}\int_0^{2/\delta} r^\nu K_\nu(r)r^{n-1}dr\\&\leqslant C_{n,\nu}\int_0^{\infty} r^\nu K_\nu(r)r^{n-1}dr\\&=C_{n,\nu}2^{\nu+n-2}\Gamma(n/2)\Gamma(\nu+n/2),\end{aligned}$$

where in the last step we have used [7, Formula 11.4.22, p. 486]. For the remainder of the proof we refer to [1, Theorem 6.2]. □

3. Multiscale approximation on $\mathbb{S}^n$

Let $X=\{\mathbf{x}_1,\ldots,\mathbf{x}_N\}\subset\mathbb{S}^n$ be a finite set of distinct points in $\mathbb{S}^n$. We define the mesh norm h_X and the separation radius q_X of this point set by

$$h_X=\sup_{\mathbf{x}\in\mathbb{S}^n}\min_{\mathbf{x}_j\in X}\theta(\mathbf{x},\mathbf{x}_j),\qquad q_X=\frac{1}{2}\min_{i\neq j}\theta(\mathbf{x}_i,\mathbf{x}_j),$$

where $\theta(\mathbf{x},\mathbf{y})=\cos^{-1}(\mathbf{x}\cdot\mathbf{y})$ is the geodesic distance on $\mathbb{S}^n$.

We define the interpolation operator $I_{X,\delta}$ associated with the set X and the kernel ϕ_δ by

$$I_{X,\delta}f(\mathbf{x})=\sum_{m=1}^N b_m\phi_\delta(\mathbf{x},\mathbf{x}_m),\qquad I_{X,\delta}f(\mathbf{x}_m)=f(\mathbf{x}_m)\quad\text{for all } \mathbf{x}_m\in X. \tag{16}$$

We have previously obtained in [1] the following result, which extends results obtained for $\delta=1$ in [8] to general $\delta\in(0,1]$.

Theorem 3.1. *(See [1, Theorem 3.2].) Let h_X be the mesh norm of the finite set $X\subset\mathbb{S}^n$, and let $g\in H^\sigma(\mathbb{S}^n)$ vanish on X. Then there exists $c_5>0$ such that, for sufficiently small h_X and all $0<\delta\leqslant 1$,*

$$\|g\|_{L_2(\mathbb{S}^n)}\leqslant c_5\left(\frac{h_X}{\delta}\right)^\sigma\|g\|_{\phi_\delta}, \tag{17}$$

and, for $f\in H^\sigma(\mathbb{S}^n)$,

$$\|f-I_{X,\delta}f\|_{L_2(\mathbb{S}^n)}\leqslant c_5\left(\frac{h_X}{\delta}\right)^\sigma\|f\|_{\phi_\delta}. \tag{18}$$

Suppose $X_1,X_2,\ldots\subset\mathbb{S}^n$ is a sequence of point sets with mesh norms $h_1,h_2,\ldots$ respectively. The mesh norms are assumed to satisfy $h_{j+1}\approx\mu h_j$ for some fixed $\mu\in(0,1)$.

Let $\delta_1,\delta_2,\ldots$ be a decreasing sequence of positive real numbers defined by $\delta_j=\nu h_j$ for some $\nu>0$. Taking the scale proportional to the mesh norm is desirable for both numerical stability and efficiency, since the sparsity of the interpolation matrix is maintained. For every $j=1,2,\ldots$ we define the scaled SBF $\phi_j:=\phi_{\delta_j}$.

We start with a widely spread set of points and use a basis function with scale δ_1 to recover the global behavior of the function f by computing $f_1=s_1:=I_{X_1,\delta_1}f$. The error, or residual, at the first step is $e_1=f-f_1$. To reduce the error, at the next step we use a finer set of points X_2 and a finer scale δ_2, and compute a correction $s_2=I_{X_2,\delta_2}e_1$ and a new approximation $f_2=f_1+s_2$, so that the new residual is $f-f_2=e_1-I_{X_2,\delta_2}e_1$; and so on.

Stated as an algorithm, we set $f_0=0$ and $e_0=f$ and compute for $j=1,2,\ldots,$

$$s_j=I_{X_j,\delta_j}e_{j-1},\qquad f_j=f_{j-1}+s_j,\qquad e_j=e_{j-1}-s_j.$$

Note that the algorithm makes f_j+e_j independent of j, from which it follows that $f_j+e_j=f_0+e_0=f$, allowing us to identify $e_j=f-f_j$ as the error (or the residual) at stage j, and f_j as the approximation to f at the stage j. Since

$$f_j=\sum_{i=1}^j s_i=\sum_{i=1}^j I_{X_i,\delta_i}e_{i-1},$$

we see that the approximation f_j is a linear combination of spherical basis functions at all scales up to level j. We can think of s_j as adding additional "detail" to the approximation f_{j-1} to produce f_j. Also, since $e_j|_{X_j} = 0$, it follows that $f_j = f - e_j$ interpolates f on X_j, or

$$f_j|_{X_j} = f|_{X_j}.$$

The next theorem is the convergence result of [1] for functions from the associated reproducing kernel Hilbert space.

Theorem 3.2. *(See [1, Theorem 4.1].) Let $X_1, X_2, \ldots$ be a sequence of point sets on $\mathbb{S}^n$ with mesh norms $h_1, h_2, \ldots$ satisfying $c\mu h_j \leqslant h_{j+1} \leqslant \mu h_j$ for $j = 1, 2, \ldots$ with fixed $\mu, c \in (0, 1)$ and h_1 sufficiently small. Let $\phi_j = \phi_{\delta_j}$ be a kernel satisfying (7) with scale factor $\delta_j = \nu h_j$ and ν satisfying $1/h_1 \geqslant \nu \geqslant \beta/\mu$ with a fixed $\beta > 0$. Assume that the target function f belongs to $H^\sigma(\mathbb{S}^n)$. Then there is a constant $C = C(\beta)$ independent of j and f such that*

$$\|e_j\|_{\phi_{j+1}} \leqslant \alpha \|e_{j-1}\|_{\phi_j} \quad \text{for } j = 1, 2, \ldots, \tag{19}$$

with $\alpha = C\mu^\sigma$, and hence there exists $c > 0$ such that

$$\|f - f_k\|_{L_2(\mathbb{S}^n)} \leqslant c\alpha^k \|f\|_{H^\sigma(\mathbb{S}^n)} \quad \text{for } k = 1, 2, \ldots.$$

Thus the multiscale approximation f_k converges linearly to f in the L_2 norm if $\mu < C^{-1/\sigma}$.

4. Approximation of rough target functions

In this section, we analyze the multistage approximation method of the last section for target functions outside the RKHS $\mathcal{N}_\phi$. Hence, we prove convergence of the multiscale approximation scheme under the following conditions. The interpolants are computed using the kernel ϕ, which generates $H^\sigma(\mathbb{S}^n)$ with $\sigma > n/2$. The target function f, however, belongs only to $H^\beta(\mathbb{S}^n)$ for some $\sigma > \beta > n/2$.

4.1. A convergence theorem

We need to collect a few auxiliary results before stating the convergence theorem for the multiscale approximation of rougher target functions.

Lemma 4.1. *For $0 < \delta \leqslant 1$ and all $g \in H^\beta(\mathbb{S}^n)$ we have*

$$c_3^{1/2}\|g\|_{\psi_\delta} \leqslant \|g\|_{H^\beta(\mathbb{S}^n)} \leqslant c_4^{1/2}\delta^{-\beta}\|g\|_{\psi_\delta}.$$

Proof. This follows from Lemma 2.1. The proof is essentially that of [1, Lemma 3.1] with σ replaced by β. □

Lemma 4.2. *(See [8, Theorem 3.3].) Let $\tau > n/2$ and let X be a finite set in $\mathbb{S}^n$ with sufficiently small mesh norm h_X. Then, there is a constant $C > 0$, independent of X, such that for all $f \in H^\tau(\mathbb{S}^n)$ satisfying $f|_X = 0$, there holds*

$$\|f\|_{H^\gamma(\mathbb{S}^n)} \leqslant Ch_X^{\tau-\gamma}\|f\|_{H^\tau(\mathbb{S}^n)}, \quad 0 \leqslant \gamma \leqslant \tau.$$

We also need a version of [9, Theorem 5.1], which asserts that for a sufficiently smooth function f there exists a spherical polynomial of degree at most L that interpolates f on a set $X \subset \mathbb{S}^n$ of distinct, scattered points and, simultaneously, is a "good" approximant to f. We need a weak version of such a result for the Sobolev spaces generated by our scaled kernel ψ_δ.

Lemma 4.3. *Let $f \in H^\beta(\mathbb{S}^n)$, $\beta > n/2$ and let X be a finite subset of $\mathbb{S}^n$ with separation radius q_X. Let $\delta \in (0, 1]$ be given. There exists a constant κ, which depends only on n and β, such that if $L \geqslant \kappa \max\{\delta/q_X, 1/\delta\}$, then there is a spherical polynomial $p \in \mathcal{P}_L$ such that $p|_X = f|_X$ and*

$$\|f - p\|_{\psi_\delta} \leqslant 5\|f\|_{\psi_\delta}.$$

We defer the proof of this lemma until the next subsection.

In the next lemma the notation is as in Theorem 3.2 with the addition that $q_j = q_{X_j}$. This lemma makes essential use of Lemma 4.3 in its proof.

Lemma 4.4. *Let $q_j \leqslant h_j \leqslant c_q q_j$ and $\delta_j \geqslant q_j$. Let $\psi_j := \psi_{\delta_j}$, and let $f \in H^\beta(\mathbb{S}^n)$ with $\sigma > \beta > n/2$. Then there exists a constant $C > 0$, independent of X_j, such that*

$$\|e_j\|_{H^\beta(\mathbb{S}^n)} \leqslant C\delta_j^{-\beta}\|e_{j-1}\|_{\psi_j}.$$

Proof. Choose $L = \lceil \kappa \max\{\delta_j/q_j, 1/\delta_j\}\rceil$. Then by Lemma 4.3, there is a spherical polynomial $p \in \mathcal{P}_L$ that interpolates and approximates e_{j-1}, i.e. $p|_{X_j} = e_{j-1}|_{X_j}$ and $\|p - e_{j-1}\|_{\psi_j} \leqslant 5\|e_{j-1}\|_{\psi_j}$. We can use p to split the error,

$$\|e_j\|_{H^\beta(\mathbb{S}^n)} = \|e_{j-1} - I_{X_j,\delta_j} e_{j-1}\|_{H^\beta(\mathbb{S}^n)} \leqslant \|e_{j-1} - p\|_{H^\beta(\mathbb{S}^n)} + \|p - I_{X_j,\delta_j} e_{j-1}\|_{H^\beta(\mathbb{S}^n)}. \tag{20}$$

The first term of (20) can be bounded using Lemmas 4.1 and 4.3,

$$\begin{aligned}\|e_{j-1} - p\|_{H^\beta(\mathbb{S}^n)} &\leqslant c_4^{1/2} \delta_j^{-\beta} \|e_{j-1} - p\|_{\psi_j} \\ &\leqslant 5c_4^{1/2} \delta_j^{-\beta} \|e_{j-1}\|_{\psi_j}.\end{aligned}$$

Since $p|_{X_j} = e_{j-1}|_{X_j}$ the interpolant $I_{X_j,\delta_j} e_{j-1}$ is identical to $I_{X_j,\delta_j} p$. Therefore the second term of (20) can be estimated via Lemma 4.2 and (10),

$$\begin{aligned}\|p - I_{X_j,\delta_j} e_{j-1}\|_{H^\beta(\mathbb{S}^n)} = \|p - I_{X_j,\delta_j} p\|_{H^\beta(\mathbb{S}^n)} &\leqslant C h_j^{\sigma-\beta} \|p - I_{X_j,\delta_j} p\|_{H^\sigma(\mathbb{S}^n)} \\ &\leqslant C h_j^{\sigma-\beta} \delta_j^{-\sigma} \|p - I_{X_j,\delta_j} p\|_{\phi_j} \leqslant C h_j^{\sigma-\beta} \delta_j^{-\sigma} \|p\|_{\phi_j},\end{aligned}$$

where in the last step we used the fact that $I_{X_j,\delta_j} p$ is the orthogonal projection with respect to the inner product $(\cdot,\cdot)_{\phi_j}$ onto $\operatorname{span}\{\phi_{\delta_j}(\cdot,\mathbf{x}) : \mathbf{x} \in X_j\}$. Since $p \in \mathcal{P}_L$, $\delta_j \leqslant 1$ and $\beta > \sigma$, using the definition of the norm $\|\cdot\|_{\phi_j}$ given by (8) and the condition (7) we have

$$\begin{aligned}\|p\|_{\phi_j}^2 &\leqslant \frac{1}{c_1} \sum_{\ell=0}^{L} \sum_{k=1}^{N(n,\ell)} (1 + \delta_j \ell)^{2\sigma} |\widehat{p}_{\ell,k}|^2 \\ &\leqslant \frac{1}{c_1} (1 + \delta_j L)^{2(\sigma-\beta)} \sum_{\ell=0}^{L} \sum_{k=1}^{N(n,\ell)} (1 + \delta_j \ell)^{2\beta} |\widehat{p}_{\ell,k}|^2 \\ &\leqslant C(\delta_j/q_j)^{2(\sigma-\beta)} \|p\|_{\psi_j}^2,\end{aligned}$$

where in the last step we used the property that $\lceil a\rceil \leqslant 2a$ for $a \geqslant 1$. Thus, combining these above estimates together with the fact that $\|p\|_{\psi_j} \leqslant 6\|e_{j-1}\|_{\psi_j}$ we obtain

$$\|p - I_{X_j,\delta_j} p\|_{H^\beta(\mathbb{S}^n)} \leqslant C\delta_j^{-\beta} (h_j/q_j)^{\sigma-\beta} \|e_{j-1}\|_{\psi_j} \leqslant C\delta_j^{-\beta} \|e_{j-1}\|_{\psi_j}. \qquad \square$$

In the following, we are going to prove that the multilevel algorithm also converges for target functions from $H^\beta(\mathbb{S}^n)$. The proof generally follows the line of Theorem 4.1 in [1] with ϕ_j replaced by ψ_j but a new condition (see (ii) below) is needed, for now we need to use Lemma 4.4, which in turn rests on Lemma 4.3.

Theorem 4.5. *Let $X_1, X_2, \ldots$ be a sequence of point sets in $\mathbb{S}^n$ with mesh norms $h_1, h_2, \ldots$ and separation radii $q_1, q_2, \ldots$ satisfying*

(i) $c\mu h_j \leqslant h_{j+1} \leqslant \mu h_j$ *for* $j = 1, 2, \ldots$ *with* $\mu, c \in (0, 1)$,
(ii) $q_j \leqslant h_j \leqslant c_q q_j$ *for* $j = 1, 2, \ldots$ *with* $c_q > 1$.

Let ϕ be a kernel generating $H^\sigma(\mathbb{S}^n)$, i.e. the Fourier symbol of ϕ satisfies (3), *and let $\phi_j := \phi_{\delta_j}$ be defined by* (5) *with scale factor $\delta_j = \nu h_j$ where $1/h_1 \geqslant \nu \geqslant \gamma/\mu \geqslant 1$ with a fixed $\gamma > 0$. Let ψ be a kernel generating $H^\beta(\mathbb{S}^n)$ with $\sigma > \beta > n/2$ and let ψ_j be the scaled version using the same scale factor δ_j. Assume that the target function f belongs to $H^\beta(\mathbb{S}^n)$. Then, there exists a constant $C = C(\gamma) > 0$ such that, with $\alpha = C\mu^\beta$,*

$$\|e_j\|_{\psi_{j+1}} \leqslant \alpha \|e_{j-1}\|_{\psi_j} \quad \textit{for } j = 1, 2, 3, \ldots$$

and hence there exists a constant $c > 0$ so that

$$\|f - f_k\|_{L_2(\mathbb{S}^n)} \leqslant c\alpha^k \|f\|_{H^\beta(\mathbb{S}^n)} \quad \textit{for } k = 1, 2, 3, \ldots.$$

Proof. We start with, using Lemma 2.1,

$$\|e_j\|_{\psi_{j+1}}^2 \leqslant \frac{1}{c_3} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} |\hat{e}_{j,\ell k}|^2 (1 + \delta_{j+1}\ell)^{2\beta} =: \frac{1}{c_3}(S_1 + S_2),$$

with

$$S_1 := \sum_{\ell \leqslant 1/\delta_{j+1}} \sum_{k=1}^{N(n,\ell)} |\hat{e}_{j,\ell k}|^2 (1+\delta_{j+1}\ell)^{2\beta},$$

$$S_2 := \sum_{\ell > 1/\delta_{j+1}} \sum_{k=1}^{N(n,\ell)} |\hat{e}_{j,\ell k}|^2 (1+\delta_{j+1}\ell)^{2\beta}.$$

For the first sum S_1, since $\delta_{j+1}\ell \leqslant 1$,

$$S_1 \leqslant \sum_{\ell \leqslant 1/\delta_{j+1}} \sum_{k=1}^{N(n,\ell)} |\hat{e}_{j,\ell k}|^2 2^{2\beta} \leqslant 2^{2\beta} \|e_j\|^2_{L_2(\mathbb{S}^n)}.$$

Since e_j vanishes on X_j, by Lemma 4.2 we have $\|e_j\|_{L_2(\mathbb{S}^n)} \leqslant C h_j^\beta \|e_j\|_{H^\beta(\mathbb{S}^n)}$. Combining this with Lemma 4.4, which gives $\|e_j\|_{H^\beta(\mathbb{S}^n)} \leqslant C\delta_j^{-\beta}\|e_{j-1}\|_{\psi_j}$, we have

$$\|e_j\|^2_{L_2(\mathbb{S}^n)} \leqslant C\left(\frac{h_j}{\delta_j}\right)^{2\beta} \|e_{j-1}\|^2_{\psi_j} \leqslant C\mu^{2\beta}\|e_{j-1}\|^2_{\psi_j}.$$

For the sum S_2, since $\delta_{j+1}\ell > 1$ we have,

$$(1+\delta_{j+1}\ell)^{2\beta} < (2\delta_{j+1}\ell)^{2\beta} \leqslant (2\delta_{j+1})^{2\beta}(1+\ell)^{2\beta}.$$

So

$$S_2 \leqslant \sum_{\ell > 1/\delta_{j+1}} \sum_{k=1}^{N(n,\ell)} (2\delta_{j+1})^{2\beta} |\hat{e}_{j,\ell k}|^2 (1+\ell)^{2\beta} \leqslant (2\delta_{j+1})^{2\beta} \|e_j\|^2_{H^\beta(\mathbb{S}^n)}.$$

Using Lemma 4.4 again, which gives $\|e_j\|_{H^\beta(\mathbb{S}^n)} \leqslant C\delta_j^{-\beta}\|e_{j-1}\|_{\psi_j}$, we obtain

$$S_2 \leqslant C\left(\frac{\delta_{j+1}}{\delta_j}\right)^{2\beta} \|e_{j-1}\|^2_{\psi_j} \leqslant C\mu^{2\beta}\|e_{j-1}\|^2_{\psi_j}.$$

Putting all the estimates together, we obtain

$$\|e_j\|_{\psi_{j+1}} \leqslant C\mu^\beta \|e_{j-1}\|_{\psi_j} \leqslant \alpha \|e_{j-1}\|_{\psi_j},$$

which is the first statement of the theorem. Finally, since $e_k = f - f_k$ vanishes on X_k, we have by Lemma 4.2 and Lemma 4.1 that

$$\begin{aligned}
\|f - f_k\|_{L_2} = \|e_k\|_{L_2} &\leqslant C h_k^\beta \|e_k\|_{H^\beta(\mathbb{S}^n)} \\
&\leqslant C h_k^\beta \delta_{k+1}^{-\beta} \|e_k\|_{\psi_{k+1}} \\
&\leqslant C \|e_k\|_{\psi_{k+1}} \\
&\leqslant C\alpha \|e_{k-1}\|_{\psi_k},
\end{aligned}$$

since $h_k/\delta_{k+1} = h_k/(\nu h_{k+1}) \leqslant 1/(c\gamma)$. Now we can apply the argument k times to obtain, with the aid of Lemma 4.1,

$$\|f - f_k\|_{L_2} \leqslant C\alpha^k \|e_0\|_{\psi_{k+1}} = C\alpha^k \|f\|_{\psi_{k+1}} \leqslant C\alpha^k \|f\|_{H^\beta(\mathbb{S}^n)}. \qquad \square$$

4.2. Existence of a polynomial interpolant

We are now going to prove Lemma 4.3. To this end, we have to collect a number of auxiliary results, which are mainly adaptation of results from [9]. The idea there is to "lift" the norms involving spherical basis functions (SBFs) on $\mathbb{S}^n$ to those for RBFs on $\mathbb{R}^{n+1}$.

We will once again work with the bizonal kernel $\psi(\mathbf{x},\mathbf{y}) = \Psi(\mathbf{x}-\mathbf{y})$ for $\mathbf{x},\mathbf{y} \in \mathbb{S}^n$ where Ψ is the reproducing kernel of the Sobolev space $H^{\beta+1/2}(\mathbb{R}^{n+1})$.

For a given $\tau > 0$, we define the function $\Psi(\tau;\cdot)$ via its Fourier transform,

$$\widehat{\Psi}(\tau;\boldsymbol{\omega}) := \begin{cases} \widehat{\Psi}(\boldsymbol{\omega}), & |\boldsymbol{\omega}| \leqslant \tau, \\ 0, & |\boldsymbol{\omega}| > \tau. \end{cases} \tag{21}$$

Hence, $\Psi(\tau;\cdot)$ is given by

$$\Psi(\tau;\mathbf{x}) := (2\pi)^{-(n+1)/2}\int\limits_{\mathbb{R}^{n+1}} \widehat{\Psi}(\tau;\boldsymbol{\omega})e^{i\mathbf{x}\cdot\boldsymbol{\omega}}\,d\boldsymbol{\omega} = (2\pi)^{-(n+1)/2}\int\limits_{|\boldsymbol{\omega}|\leqslant\tau} \widehat{\Psi}(\boldsymbol{\omega})e^{i\mathbf{x}\cdot\boldsymbol{\omega}}\,d\boldsymbol{\omega}.$$

A scaled version of $\Psi(\tau;\mathbf{x})$ is defined by

$$\Psi_\delta(\tau;\mathbf{x}) := \delta^{-n}\Psi(\delta\tau;\mathbf{x}/\delta). \tag{22}$$

Consequently, the Fourier transform of $\Psi_\delta(\tau;\cdot)$ scales as

$$\widehat{\Psi_\delta}(\tau;\boldsymbol{\omega}) = \delta\widehat{\Psi}(\delta\tau;\delta\boldsymbol{\omega}) = \begin{cases} \widehat{\Psi_\delta}(\boldsymbol{\omega}) = \delta\widehat{\Psi}(\delta\boldsymbol{\omega}), & |\boldsymbol{\omega}|\leqslant\tau, \\ 0, & |\boldsymbol{\omega}|>\tau. \end{cases} \tag{23}$$

The bizonal kernel $\psi_{\delta;\tau}(\mathbf{x},\mathbf{y}) := \Psi_\delta(\tau;\mathbf{x}-\mathbf{y})|_{\mathbf{x},\mathbf{y}\in\mathbb{S}^n}$ is an SBF with Fourier-Legendre coefficients given by (see [6, Theorem 4.1 and Corollary 4.3])

$$\hat{\psi}_{\delta;\tau}(\ell) = (2\pi)^{(n+1)/2}\int_0^\tau t\widehat{\Psi_\delta}(t)J^2_{\ell+(n-1)/2}(t)\,dt.$$

Besides this, we will also need the truncation of ψ_δ to $\mathcal{P}_L$ given by

$$\psi_\delta^L(\mathbf{x},\mathbf{y}) := \sum_{\ell=0}^{L}\sum_{k=1}^{N(n,\ell)}\hat{\psi}_\delta(\ell)Y_{\ell,k}(\mathbf{x})Y_{\ell,k}(\mathbf{y}).$$

The following result is [9, Lemma 4.2], with Ψ and ψ being replaced by Ψ_δ and ψ_δ, respectively.

Proposition 4.6. *Let $X=\{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$ be a set of scattered points on $\mathbb{S}^n$ and $\alpha\in\mathbb{R}^N$. Let Ψ be the reproducing kernel for $H^{\beta+1/2}(\mathbb{R}^{n+1})$ and ψ be its associated spherical kernel. Let the scale parameter $\delta\in(0,1]$ be given. Then*

$$\Bigg\|\sum_{j=1}^N\alpha_j\psi_\delta(\cdot,\mathbf{x}_j)\Bigg\|_{\psi_\delta} = \Bigg\|\sum_{j=1}^N\alpha_j\Psi_\delta(\cdot-\mathbf{x}_j)\Bigg\|_{\Psi_\delta}. \tag{24}$$

If there is a constant $c>0$ such that when $\tau\leqslant cL$ we have $\hat{\psi}_{\delta;\tau}(\ell)\leqslant\frac{1}{2}\widehat{\psi}_\delta(\ell)$ for all $\ell>L$, then we have

$$\Bigg\|\sum_{j=1}^N\alpha_j\big(\psi_\delta-\psi_\delta^L\big)(\cdot,\mathbf{x}_j)\Bigg\|^2_{\psi_\delta} \leqslant 2\Bigg\|\sum_{j=1}^N\alpha_j\big(\Psi_\delta(\mathbf{x}-\mathbf{x}_j)-\Psi_\delta(\tau;\cdot-\mathbf{x}_j)\big)\Bigg\|^2_{\Psi_\delta}. \tag{25}$$

We are now going to show that there is indeed such a constant $c>0$ such that when $\tau\leqslant cL$ we have $\hat{\psi}_{\delta;\tau}(\ell)\leqslant\frac{1}{2}\hat{\psi}_\delta(\ell)$ for all $\ell>L$. For this, we need the following lemma.

Lemma 4.7. *Let $\widehat{\Psi}\in L^1(\mathbb{R}^{n+1})$ be radial and nonnegative, and let $\ell\geqslant0$. If $0<\delta\leqslant1,\tau>0$ and $\nu=\ell+(n-1)/2$, then*

$$\hat{\psi}_{\delta;\tau}(\ell) = (2\pi)^{(n+1)/2}\int_0^\tau t\widehat{\Psi_\delta}(t)J_\nu^2(t)\,dt \leqslant \frac{2(2\pi)^{(n+1)/2}e\delta^{-n}\Psi(0)(e\pi)^n}{\omega_n(\nu+1)^n}\left(\frac{e\tau}{2(\nu+1)}\right)^{2\nu-n+1}.$$

Proof. This follows from [9, Lemma 4.3] by replacing Ψ with Ψ_δ and using $\Psi_\delta(0)=\delta^{-n}\Psi(0)$. □

Since ψ is a reproducing kernel for $H^\beta(\mathbb{S}^n)$ with $\beta>n/2$, condition (15) and Lemma 4.7 yield

$$\frac{\hat{\psi}_{\delta;\tau}(\ell)}{\hat{\psi}_\delta(\ell)} \leqslant C_{\beta,n}(1+\delta\ell)^{2\beta}(\delta\ell)^{-n}\left(\frac{e\tau}{2\ell}\right)^{2\ell}.$$

If we now set $\ell_0 = \lceil 1/\delta \rceil$ then we have for all $\ell \geqslant \ell_0$ that

$$\frac{\hat{\psi}_{\delta;\tau}(\ell)}{\hat{\psi}_\delta(\ell)} \leqslant C_{\beta,n} 2^{2\beta} (\delta\ell)^{2\beta-n} \left(\frac{e\tau}{2\ell}\right)^{2\ell}.$$

From this we can conclude that there is an $L_0 \geqslant \ell_0$ such that for all $\ell \geqslant L \geqslant L_0$ and $\tau \leqslant e^{-1}L$ we have

$$\frac{\hat{\psi}_{\delta;\tau}(\ell)}{\hat{\psi}_\delta(\ell)} \leqslant C_{\beta,n} 2^{2\beta} (\delta\ell)^{2\beta-n} 2^{-2\ell} \leqslant \frac{1}{2}. \tag{26}$$

Thus the conclusions of Proposition 4.6 hold.

We now state a proposition, whose proof can be found in [10, Theorem 2.1] or [11, Prop. 3.1]. It provides the theoretical framework for proving the existence of an interpolant that is also a good approximant.

Proposition 4.8. *Let $\mathcal{Y}$ be a (possibly complex) Banach space, $\mathcal{V}$ a subspace of $\mathcal{Y}$, and Z^* a finite-dimensional subspace of $\mathcal{Y}^*$, the dual of $\mathcal{Y}$. If there is a $\gamma > 1$ such that for every $z^* \in Z^*$,*

$$\|z^*\|_{\mathcal{Y}^*} \leqslant \gamma \|z^*|_{\mathcal{V}}\|_{\mathcal{V}^*}, \tag{27}$$

then for $y \in \mathcal{Y}$ there exists $v \in \mathcal{V}$ such that v interpolates y on Z^*; that is, $z^*(y) = z^*(v)$ for all $z^* \in Z^*$. In addition, v approximates y in the sense that $\|y - v\|_{\mathcal{Y}} \leqslant (1 + 2\gamma)\mathrm{dist}(y, \mathcal{V})$.

We will apply this result to the case in which the underlying space is the RKHS of the scaled kernel ψ_δ, that is $\mathcal{Y} = \mathcal{N}_{\psi_\delta}$. We will take $Z^* = \mathrm{span}\{\delta_{\mathbf{x}_j}\}_{j=1}^N$, where the points are distinct and come from a finite set $X \subset \mathbb{S}^n$. Finally, we take $\mathcal{V} = \mathcal{P}_L$, the span of the spherical harmonics of degree $\ell \leqslant L$, with L at the moment fixed.

The quantities in Proposition 4.8 can be put in terms of the reproducing kernel ψ_δ. First, we observe that

$$\left\|\sum_{j=1}^N \alpha_j \delta_{\mathbf{x}_j}\right\|_{\mathcal{N}^*_{\psi_\delta}} = \left\|\sum_{j=1}^N \alpha_j \psi_\delta(\cdot, \mathbf{x}_j)\right\|_{\psi_\delta}.$$

Second, we have

$$\left\|\sum_{j=1}^N \alpha_j \delta_{\mathbf{x}_j}\bigg|_{\mathcal{P}_L}\right\|_{\mathcal{P}_L^*} = \left\|\sum_{j=1}^N \alpha_j \psi_\delta^L(\cdot, \mathbf{x}_j)\right\|_{\psi_\delta},$$

see formula (5.3) in [9]. Moreover, since $\psi_\delta^L(\cdot, \mathbf{x}_j)$ and $\psi_\delta(\cdot, \mathbf{x}_k) - \psi_\delta^L(\cdot, \mathbf{x}_k)$ are orthogonal in $\mathcal{N}_{\psi_\delta}$ for all j, k, we can use the Pythagorean theorem to obtain

$$\left\|\sum_{j=1}^N \alpha_j \psi_\delta^L(\cdot, \mathbf{x}_j)\right\|_{\psi_\delta}^2 = \left\|\sum_{j=1}^N \alpha_j \psi_\delta(\cdot, \mathbf{x}_j)\right\|_{\psi_\delta}^2 - \left\|\sum_{j=1}^N \alpha_j \left[\psi_\delta(\cdot, \mathbf{x}_j) - \psi_\delta^L(\cdot, \mathbf{x}_j)\right]\right\|_{\psi_\delta}^2.$$

From this and the quantities above, it follows that Proposition 4.8 applied to our situation yields the following: if we can find a $\gamma > 1$ such that for every linear functional $z^* = \sum_j \alpha_j \delta_{\mathbf{x}_j}$ we have

$$\frac{\|\sum_{j=1}^N \alpha_j [\psi_\delta(\cdot, \mathbf{x}_j) - \psi_\delta^L(\cdot, \mathbf{x}_j)]\|_{\psi_\delta}^2}{\|\sum_{j=1}^N \alpha_j \psi_\delta(\cdot, \mathbf{x}_j)\|_{\psi_\delta}^2} \leqslant 1 - \frac{1}{\gamma^2}, \tag{28}$$

then for $f \in \mathcal{N}_{\psi_\delta}$ there exists a polynomial p so that $f(\mathbf{x}_j) = p(\mathbf{x}_j)$ for $1 \leqslant j \leqslant N$. In addition, p approximates f in the sense that $\|f - p\|_{\psi_\delta} \leqslant (1 + 2\gamma)\|f\|_{\psi_\delta}$.

An explicit value of γ will be given in the proof of Lemma 4.3, as follows.

Proof of Lemma 4.3. We have, by using Proposition 4.6 and (26) that

$$\frac{\|\sum_{j=1}^N \alpha_j [\psi_\delta(\cdot, \mathbf{x}_j) - \psi_\delta^L(\cdot, \mathbf{x}_j)]\|_{\psi_\delta}^2}{\|\sum_{j=1}^N \alpha_j \psi_\delta(\cdot, \mathbf{x}_j)\|_{\psi_\delta}^2} \leqslant \frac{2\|\sum_{j=1}^N \alpha_j [\Psi_\delta(\cdot - \mathbf{x}_j) - \Psi_\delta(\tau; \cdot - \mathbf{x}_j)]\|_{\Psi_\delta}^2}{\|\sum_{j=1}^N \alpha_j \Psi_\delta(\cdot - \mathbf{x}_j)\|_{\Psi_\delta}^2},$$

where we take $\tau = e^{-1}L$ and $L \geqslant L_0$ as required for (26). Using the reproducing property of Ψ_δ, we have

$$\left\| \sum_{j=1}^{N} \alpha_j \big[\Psi_\delta(\cdot - \mathbf{x}_j) - \Psi_\delta(\tau; \cdot - \mathbf{x}_j)\big] \right\|_{\Psi_\delta}^2 = \sum_{j,k=1}^{N} \alpha_j\alpha_k \big[\Psi_\delta(\mathbf{x}_j - \mathbf{x}_k) - \Psi_\delta(\tau; \mathbf{x}_j - \mathbf{x}_k)\big]$$
$$= \delta^{-n} \sum_{j,k=1}^{N} \alpha_j\alpha_k \left[\Psi\left(\frac{\mathbf{x}_j - \mathbf{x}_k}{\delta}\right) - \Psi\left(\tau; \frac{\mathbf{x}_j - \mathbf{x}_k}{\delta}\right)\right].$$

Similarly, using the reproducing property of Ψ_δ again, we have

$$\left\| \sum_{j=1}^{N} \alpha_j \Psi_\delta(\cdot - \mathbf{x}_j) \right\|_{\Psi_\delta}^2 = \sum_{j,k=1}^{N} \alpha_j\alpha_k \Psi_\delta(\mathbf{x}_j - \mathbf{x}_k) = \delta^{-n} \sum_{j,k=1}^{N} \alpha_j\alpha_k \Psi\left(\frac{\mathbf{x}_j - \mathbf{x}_k}{\delta}\right).$$

By using an estimate from the proof of [12, Lemma 3.3] on the scaled data set $Y = X/\delta$, we obtain the following estimate

$$\frac{\sum_{j,k=1}^{N} \alpha_j\alpha_k \big[\Psi\big(\frac{\mathbf{x}_j-\mathbf{x}_k}{\delta}\big) - \Psi\big(\tau; \frac{\mathbf{x}_j-\mathbf{x}_k}{\delta}\big)\big]}{\sum_{j,k=1}^{N} \alpha_j\alpha_k \Psi\big(\frac{\mathbf{x}_j-\mathbf{x}_k}{\delta}\big)} \leqslant C(\tau Q_X/\delta)^{n+1-2(\beta+1/2)},$$

where $C = C(\beta, n+1)$ and Q_X is the Euclidean separation radius for X as a subset of $\mathbb{R}^{n+1}$. For any discrete set X, we know that Q_X is comparable to q_X, indeed $q_X \geqslant Q_X \geqslant (2/\pi)q_X$. Combining all of the results and using $\tau = e^{-1}L$ yields

$$\frac{\| \sum_{j=1}^{N} \alpha_j [\psi_\delta(\cdot, \mathbf{x}_j) - \psi_\delta^L(\cdot, \mathbf{x}_j)] \|_{\psi_\delta}^2}{\| \sum_{j=1}^{N} \alpha_j \psi_\delta(\cdot, \mathbf{x}_j) \|_{\psi_\delta}^2} \leqslant C'(Lq_X/\delta)^{n-2\beta},$$

where $C' = 2^{n+1-2\beta} C/(e\pi)^{n-2\beta}$. We can choose $\kappa > 0$ such that $C'\kappa^{n-2\beta} \leqslant \frac{3}{4}$. Since $\beta > n/2$ we have $n - 2\beta < 0$, thus, if $Lq_X/\delta \geqslant \kappa$, then $C'(Lq_X/\delta)^{n-2\beta} \leqslant \frac{3}{4}$. Moreover, by enlarging κ if necessary, we can also ensure that the condition $L \geqslant L_0$ required for (26) is satisfied. This means that (28) holds with $\gamma = 2$ provided that $L \geqslant \kappa \max\{q_X/\delta, 1/\delta\}$.

Hence, we have shown that we can apply Proposition 4.8 in this situation, which finishes the proof. $\square$

4.3. Multiresolution analysis

For a given RBF Φ in $\mathbb{R}^{n+1}$ and a sequence of point sets $X_1, X_2, \ldots \subset \mathbb{S}^n$ and a sequence of scales $\delta_1, \delta_2, \ldots$ converging to 0, the multiscale approximation scheme in Section 3 can be put into the following multiresolution framework.

For $j \geqslant 1$, let W_j and V_j be the linear subspaces of $H^\sigma(\mathbb{S}^n)$ defined by

$$W_j := \operatorname{span}\big\{\phi_{\delta_j}(\cdot, \mathbf{x}) \colon \mathbf{x} \in X_j\big\}$$

and

$$V_j := \operatorname{span}\big\{\phi_{\delta_i}(\cdot, \mathbf{x}) \colon \mathbf{x} \in X_i, \quad i \leqslant j\big\},$$

where ϕ_δ is as in (5). Thus

$$V_1 \subset V_2 \subset \cdots \subset V_j \subset V_{j+1} \subset \cdots \subset H^\sigma\big(\mathbb{S}^n\big).$$

In the language of wavelets, W_j is the wavelet space and V_j the scale space. The space V_j can be written as a direct sum of spaces W_i,

$$V_j = \bigoplus_{1 \leqslant i \leqslant j} W_i, \quad j \geqslant 1$$

and hence if $V_0 = \{0\}$,

$$V_j = V_{j-1} \bigoplus W_j, \quad j \geqslant 1.$$

That the sum is direct follows under appropriate conditions from the following lemma [1, Lemma 5.1].

Lemma 4.9. *Let Φ be a compactly supported RBF as in* [5] *and let ϕ_{δ_i} for $i = 1, 2, \ldots$ be scaled SBFs constructed as in* (5) *where $\delta_1, \delta_2, \ldots$ are distinct scales with $\delta_i \leqslant 1$. Let $X_i = \{\mathbf{x}_{i,1}, \ldots, \mathbf{x}_{i,N_i}\} \subset \mathbb{S}^n$ for $i \geqslant 1$ be a set of N_i distinct points. Then for $j \geqslant 1$ and $a_{i,k} \in \mathbb{R}$*

$$\sum_{i=1}^{j} \sum_{k=1}^{N_i} a_{i,k} \phi_{\delta_i}(\cdot, \mathbf{x}_{i,k}) = 0 \textit{ implies } a_{i,k} = 0 \textit{ for } i = 1, \ldots, j;\ k = 1, \ldots, N_i. \tag{29}$$

The sequence of spaces $\{V_j\}$ is ultimately dense in $L_2(\mathbb{S}^n)$, in the sense of the following theorem [1, Theorem 5.2].

Theorem 4.10. *Let $X_1, X_2, \ldots$ be a sequence of point sets on $\mathbb{S}^n$ with mesh norms $h_1, h_2, \ldots$ satisfying $c\mu h_j \leqslant h_{j+1} \leqslant \mu h_j$ for $j = 1, 2, \ldots$ with fixed $\mu, c \in (0, 1)$ and h_1 sufficiently small. Let ϕ_{δ_j} be a kernel satisfying (7) with scale factor $\delta_j = \nu h_j$ and ν satisfying $1/h_1 \geqslant \nu \geqslant \beta/\mu$ with a fixed $\beta > 0$. For all μ sufficiently small*

$$\text{the closure of } \bigcup_{j=1}^{\infty} V_j \text{ with respect to the norm } \|\cdot\|_{L_2} \text{ is } L_2\big(\mathbb{S}^n\big).$$

Acknowledgment

The authors are grateful to the Australian Research Council for its support.

References

[1] Q.T. Le Gia, I.H. Sloan, H. Wendland, Multiscale analysis in Sobolev spaces on the sphere, SIAM J. Numer. Anal. 48 (2010) 2065–2090.
[2] I.J. Schoenberg, Positive definite function on spheres, Duke Math. J. 9 (1942) 96–108.
[3] Y. Xu, E.W. Cheney, Strictly positive definite functions on spheres, Proc. Amer. Math. Soc. 116 (1992) 977–981.
[4] C. Müller, Spherical Harmonics, Lecture Notes in Math., vol. 17, Springer-Verlag, Berlin, 1966.
[5] H. Wendland, Scattered Data Approximation, Cambridge University Press, Cambridge, 2005.
[6] F.J. Narcowich, J.D. Ward, Scattered data interpolation on spheres: error estimates and locally supported basis functions, SIAM J. Math. Anal. 33 (2002) 1393–1410.
[7] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, National Bureau of Standards, Washington, DC, 1970.
[8] Q.T. Le Gia, F.J. Narcowich, J.D. Ward, H. Wendland, Continuous and discrete least-square approximation by radial basis functions on spheres, J. Approx. Theory 143 (2006) 124–133.
[9] F.J. Narcowich, X. Sun, J.D. Ward, H. Wendland, Direct and inverse Sobolev error estimates for scattered data interpolation via spherical basis functions, Found. Comput. Math. 7 (2007) 369–390.
[10] H.N. Mhaskar, F.J. Narcowich, N. Sivakumar, J.D. Ward, Approximation with interpolatory constraints, Proc. Amer. Math. Soc. 130 (2002) 1355–1364.
[11] F.J. Narcowich, J.D. Ward, Scattered-data interpolation on $\mathbb{R}^n$: error estimates for radial basis and band-limited functions, SIAM J. Math. Anal. 36 (2004) 284–300.
[12] F.J. Narcowich, J.D. Ward, H. Wendland, Sobolev error estimates and a Bernstein inequality for scattered-data interpolation via radial basis functions, Constr. Approx. 24 (2006) 175–186.