Zooming from global to local: a multiscale RBF approach

Q. T. Le Gia\textsuperscript{1} · I. H. Sloan\textsuperscript{1} · H. Wendland\textsuperscript{2}

Received: 29 May 2014 / Accepted: 9 November 2016 / Published online: 18 November 2016
© Springer Science+Business Media New York 2016

Abstract Because physical phenomena on Earth’s surface occur on many different length scales, it makes sense when seeking an efficient approximation to start with a crude global approximation, and then make a sequence of corrections on finer and finer scales. It also makes sense eventually to seek fine scale features locally, rather than globally. In the present work, we start with a global multiscale radial basis function (RBF) approximation, based on a sequence of point sets with decreasing mesh norm, and a sequence of (spherical) radial basis functions with proportionally decreasing scale centered at the points. We then prove that we can “zoom in” on a region of particular interest, by carrying out further stages of multiscale refinement on a local region. The proof combines multiscale techniques for the sphere from Le Gia, Sloan and Wendland, SIAM J. Numer. Anal. 48 (2010) and Applied Comp. Harm. Anal. 32 (2012), with those for a bounded region in $\mathbb{R}^d$ from Wendland, Numer. Math. 116 (2010). The zooming in process can be continued indefinitely, since the condition numbers of matrices at the different scales remain bounded. A numerical example illustrates the process.
Keywords  Radial basis function · Approximation · Interpolation · Unit sphere

Mathematics Subject Classification (2010)  41A05 · 65D05

1 Introduction

In many modern areas of geosciences, such as geomagnetic or gravitational field modeling, the problem of interpolation from scattered data on the sphere arises naturally. Such problems are often of multiscale nature, so one would like models that can be used to draw conclusions globally as well as locally. For example, in modeling the global gravitational field one should be able to see the general nature of the global field as well as local gravitational anomalies.

Multiscale interpolation and approximation for functions on the unit sphere has been considered by a number of authors using different techniques. Some authors used wavelets defined on spheres [2, 25], but these are not suitable for scattered data. Other authors have proposed kernel methods based on truncations of the expansions of some special kernels into spherical harmonics [5, 16, 19]; these methods require a quadrature scheme on the sphere which can integrate spherical polynomials exactly, but the construction of a good quadrature based on scattered data is itself a non-trivial problem [4, 6, 8, 11, 17].

In recent articles [12, 13], we proposed a multiscale interpolation framework using radial basis functions for functions that lie in Sobolev spaces defined on the unit sphere. The theory underlying our multiscale method will work for scattered data. In this paper we introduce a “zooming-in” framework, which allows the multiscale algorithm to model the data from the global scale and then zoom in to local regions. We do this by combining multiscale techniques for the sphere with those for a bounded region established in [29]. We develop a procedure for generating a sequence of radial basis functions with different scales that can provide enhanced resolution in a local region while maintaining an accurate representation globally.

The paper is organised as follows. In Section 2 we review necessary materials about Sobolev spaces on spheres and positive definite kernels defined via radial basis functions (RBFs). The global and local multiscale algorithm using spherical RBFs is then introduced in Section 3. A convergence results for functions in the native space in given there. The next section, Section 4, deals with convergence results for function in Sobolev spaces with lesser smoothness. In Section 5 we provide an analysis for the computational cost of the multiscale method. Finally we conclude the paper with numerical experiments given in Section 6.

2 Preliminaries

In this section, we will introduce necessary materials for the main results presented in the paper.
2.1 Sobolev spaces on the unit sphere

Let $S^d$ be the unit sphere in $\mathbb{R}^{d+1}$. Denote the inner product in $L_2(S^d)$ by

$$\langle v, w \rangle := \int_{S^d} vw dS,$$

where $dS$ is the surface measure on the unit sphere, and denote the measure of the whole sphere by $\omega_d$ (so, for example, $\omega_2 = 4\pi$). Recall [18] that a spherical harmonic is the restriction to $S^d$ of a homogeneous harmonic polynomial $Y(x)$ in $\mathbb{R}^{d+1}$.

The space of all spherical harmonics of degree $\ell$ on $S^d$, denoted by $H_\ell$, has an $L_2$ orthonormal basis

$$\{Y_{\ell,k} : k = 1, \ldots, N(d, \ell)\},$$

where $N(d, 0) = 1$ and $N(d, \ell) = \frac{(2\ell + d - 1)\Gamma(\ell + d - 1)}{\Gamma(\ell + 1)\Gamma(d)}$ for $\ell \geq 1$.

The space of spherical harmonics of degree $\leq L$ will be denoted by $P_L := \bigoplus_{\ell=0}^L H_\ell$; i.e., a dimension $N(d+1, L)$.

Every function $f \in L_2(S^d)$ can be expanded in terms of spherical harmonics,

$$f = \sum_{\ell=0}^\infty \sum_{k=1}^{N(d, \ell)} \hat{f}_{\ell,k} Y_{\ell,k}, \quad \hat{f}_{\ell,k} = \langle f, Y_{\ell,k} \rangle.$$

For a non-negative parameter $\sigma$, the Sobolev space $H^\sigma(S^d)$ may be defined by

$$H^\sigma(S^d) := \left\{ f \in L_2(S^d) : \|f\|_{H^\sigma(S^d)}^2 := \sum_{\ell=0}^\infty \sum_{k=1}^{N(d, \ell)} (1 + \ell)^{2\sigma} |\hat{f}_{\ell,k}|^2 < \infty \right\}. \quad (2)$$

Note that $H^0(S^d) = L_2(S^d)$.

Sobolev spaces on $S^d$ can also be defined using local charts (see [15]). Here we use a specific atlas of charts, as in [9].

Let $z$ be a given point on $S^d$. The spherical cap centered at $z$ of radius $\theta$ is defined by

$$G(z, \theta) := \{y \in S^d : \cos^{-1}(z \cdot y) < \theta\}, \quad \theta \in (0, \pi),$$

where $z \cdot y$ denotes the Euclidean inner product of $z$ and $y$ in $\mathbb{R}^{d+1}$.

Let $\hat{n}$ and $\hat{s}$ denote the north and south poles of $S^d$, respectively. Then a simple cover for the sphere is provided by

$$U_1 = G(\hat{n}, \theta_0) \quad \text{and} \quad U_2 = G(\hat{s}, \theta_0), \quad \theta_0 \in (\pi/2, 2\pi/3). \quad (3)$$

The stereographic projection $\sigma_{\hat{n}}$ of the punctured sphere $S^d \setminus \{\hat{n}\}$ onto $\mathbb{R}^d$ is defined as a mapping that maps $x \in S^d \setminus \{\hat{n}\}$ to the intersection of the equatorial hyperplane $\{z = 0\}$ and the extended line that passes through $x$ and $\hat{n}$. The stereographic projection $\sigma_{\hat{s}}$ based on $\hat{s}$ can be defined analogously. We set

$$\psi_1 = \frac{1}{\tan(\theta_0/2)} \sigma_{\hat{n}}|_{U_1} \quad \text{and} \quad \psi_2 = \frac{1}{\tan(\theta_0/2)} \sigma_{\hat{n}}|_{U_2}, \quad (4)$$
so that \( \psi_k, k = 1, 2, \) maps \( U_k \) onto \( B(0, 1) \), the unit ball in \( \mathbb{R}^d \). We conclude that \( \mathcal{A} = \{ U_k, \psi_k \}_{k=1,2} \) is a \( C^\infty \) atlas of covering coordinate charts for the sphere. It is known (see [22]) that the stereographic coordinate charts \( \{ \psi_k \}_{k=1,2} \) as defined in Eq. 4 map spherical caps to Euclidean balls, but in general concentric spherical caps are not mapped to concentric Euclidean balls. The projection \( \psi_k, \) for \( k = 1, 2, \) does not distort too much the geodesic distance between two points \( x, y \in \mathbb{S}^d \), as shown in [10].

With the atlas so defined, we define the map \( \pi_k \) which takes a real-valued function \( g \) with compact support in \( U_k \) into a real-valued function on \( \mathbb{R}^d \) by

\[
\pi_k(g)(x) = \begin{cases} 
g \circ \psi_k^{-1}(x), & \text{if } x \in B(0, 1), \\
0, & \text{otherwise}.
\end{cases}
\]

Let \( \{ \chi_k : \mathbb{S}^d \to \mathbb{R} \}_{k=1,2} \) be a partition of unity subordinated to the atlas, i.e., a pair of non-negative infinitely differentiable functions \( \chi_k \) on \( \mathbb{S}^d \) with compact support in \( U_k \), such that \( \sum_k \chi_k = 1 \). For any function \( f : \mathbb{S}^d \to \mathbb{R} \), we can use the partition of unity to write

\[
f = \sum_{k=1}^{2} \chi_k f, \quad \text{where } (\chi_k f)(x) = \chi_k(x) f(x), \quad x \in \mathbb{S}^d.
\]

The Sobolev space \( H^\sigma(\mathbb{S}^d) \) is then the set

\[
\left\{ f \in L_2(\mathbb{S}^d) : \pi_k(\chi_k f) \in H^\sigma(\mathbb{R}^d) \quad \text{for } k = 1, 2 \right\},
\]

which is equipped with the norm

\[
\| f \|_{H^\sigma(\mathbb{S}^d)} := \left( \sum_{k=1}^{2} \| \pi_k(\chi_k f) \|_{H^\sigma(\mathbb{R}^d)}^2 \right)^{1/2}.
\]

This norm is equivalent to the \( H^\sigma(\mathbb{S}^d) \) norm given in Eq. 2 (see [15]). From now on we will use only the \( \| \cdot \| \) notation for the equivalent norms.

We recall that [1] the Sobolev space \( H^\sigma(\mathbb{R}^d) \) is the set

\[
\left\{ f \in L_2(\mathbb{R}^d) : \int_{\mathbb{R}^d} |\mathcal{F}(f)(\omega)|^2 \left( 1 + \| \omega \|_2^2 \right)^\sigma d\omega < \infty \right\},
\]

where \( \mathcal{F}(f) \) is the usual Fourier transform

\[
\mathcal{F}(f)(\omega) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x)e^{-ix^T\omega} dx.
\]

Before introducing local Sobolev spaces on subdomains of the unit sphere, let us recall a few key definitions on Sobolev spaces defined on a given bounded domain \( D \) in \( \mathbb{R}^d \). For a given non-negative integer \( m \), the Sobolev space \( H^m(D) \) consist of all \( f \) with weak derivatives \( D^\alpha f \in L_2(D), |\alpha| \leq m \). The semi-norms and norms are defined by

\[
|f|_{H^m(D)} = \left( \sum_{|\alpha| = m} \| D^\alpha f \|_{L_2(D)}^2 \right)^{1/2} \quad \text{and} \quad \| f \|_{H^m(D)} = \left( \sum_{|\alpha| \leq m} \| D^\alpha f \|_{L_2(D)}^2 \right)^{1/2}.
\]
For $m \in \mathbb{N}_0$, $0 < s < 1$, the fractional Sobolev spaces $H^{m+s}(D)$ is defined to be the set of all $f$ for which the following semi-norm and norm

$$|f|_{H^{m+s}(D)} := \left( \sum_{|\alpha| = m} \int_D \int_D \frac{|D^\alpha f(x) - D^\alpha f(y)|^2}{\|x - y\|^{d+2s}} \right)^{1/2}$$

$$\|f\|_{H^{m+s}(D)} := \left( \|f\|_{H^m(D)}^2 + |f|_{H^{m+s}(D)}^2 \right)^{1/2}$$

are finite.

Let $\Omega \subset \mathbb{S}^d$ be an open connected set with sufficiently smooth boundary. In order to define the spaces on $\Omega$, let $D_k = \psi_k(\Omega \cap U_k)$ for $k = 1, 2$. The local Sobolev space $H^\sigma(\Omega)$ is defined to be the set

$$\{ f \in L^2(\Omega) : \pi_k(\chi_k f)|_{D_k} \in H^\sigma(D_k) \text{ for } k = 1, 2, D_k \neq \emptyset \},$$

which is equipped with the norm

$$\|f\|_{H^\sigma(\Omega)} = \left( \sum_{k=1}^2 \|\pi_k(\chi_k f)|_{D_k}\|_{H^\sigma(D_k)}^2 \right)^{1/2}$$

where, if $\Omega = \emptyset$, then we adopt the convention that $\|\cdot\|_{H^\sigma(D_k)} = 0$.

We observe, following [9], that there exists a positive constant $C_A$, depending on $A$ and the partition of unity $\{\chi_1, \chi_2\}$, such that the geodesic distance of supp $\chi_k$ from the boundary of $U_k$ is strictly greater than $C_A$. A spherical cap $G(z, \theta)$ with $\theta < C_A/3$ will have its closure being a subset of at least one of the open subsets $U_1$ or $U_2$, defined by Eq. 3, and if the cap $G(z, \theta)$ is not a subset of one of these subsets, say $U_2$, then its intersection with supp $\chi_2$ must be empty.

Now we state an extension theorem for a spherical cap on the sphere.

**Theorem 2.1 (Extension operator)** Let $\Omega = G(z, \theta)$ be a spherical cap for some $z \in \mathbb{S}^d$ and $\theta < C_A/3$. There is an extension operator $E : H^\nu(\Omega) \rightarrow H^\nu(\mathbb{S}^d)$ for all $\nu \geq 0$, with $E$ independent of $\nu$, such that

1. $Ef|_{\Omega} = f$ for all $f \in H^\nu(\Omega)$,
2. $\|Ef\|_{H^\nu(\mathbb{S}^d)} \leq C_\nu \|f\|_{H^\nu(\Omega)}$.

**Proof** The case of $\nu$ being an integer was proved in [9, Theorem 4.3]. The framework for the case of fractional order $\nu$ is also available in [9] even if the explicit statement is not given there. For the sake of completeness, we give the proof here.

When $\nu$ is not an integer, let $k$ be the non-negative integer for which $\nu = k + s$, with $s \in (0, 1)$. By [9, Theorem 4.3], there is an extension operator which maps $H^{k+i}(\Omega)$ to $H^{k+i}(\mathbb{S}^d)$ and there are constants $C_{k,i}$ for $i = 0, 1$ so that

$$\|Ef\|_{H^{k+i}(\mathbb{S}^d)} \leq C_{k,i} \|f\|_{H^{k+i}(\Omega)}, \quad i = 0, 1.$$

Using the operator interpolation property (see [27]) we conclude that $E$ is a bounded linear map from $H^\nu(\Omega)$ to $H^\nu(\mathbb{S}^d)$ and

$$\|Ef\|_{H^\nu(\mathbb{S}^d)} \leq C_{k,0}^{1-s} C_{k,1}^s \|f\|_{H^\nu(\Omega)}.$$
Property 1) follows from the fact that $H^v(S^d) \subset H^k(S^d)$ and $H^v(\Omega) \subset H^k(\Omega)$. □

2.2 Positive definite kernels on the unit sphere

A continuous function $\Phi : S^d \times S^d \to \mathbb{R}$ we call a positive semi-definite kernel [24, 30] on $S^d$ if it satisfies the following conditions:

(i) $\Phi$ is continuous,
(ii) $\Phi(x, y) = \Phi(y, x)$ for all $x, y \in S^d$,
(iii) For any set of distinct scattered points $X = \{x_1, \ldots, x_K\} \subset S^d$, the symmetric $K \times K$-matrix $[\Phi(x_p, x_q)]$ is positive semi-definite.

We call $\Phi$ positive definite if the matrix is positive definite.

We will work with a zonal kernel $\Phi$ defined in terms of a univariate function $\phi : [-1, 1] \to \mathbb{R}$ by

$$\Phi(x, y) = \phi(x \cdot y) \quad \text{for all} \quad x, y \in S^d. \quad (7)$$

Following M"uller [18], let $P_\ell(t)$ denote the Legendre polynomial of degree $\ell$ for $\mathbb{R}^{d+1}$, that is $\{P_\ell(t)\}_{\ell=0}^{\infty}$ is a sequence of polynomials that satisfies the following orthogonal property ([18, Lemma 10]):

$$\int_{-1}^{1} P_\ell(t) P_k(t)(1 - t^2)^{(d-2)/2} dt = \frac{\omega_d}{\omega_{d-1}} \frac{1}{N(d, \ell)} \delta_{\ell,k},$$

where $\omega_d$ is the surface area of $S^d$, $N(d, \ell)$ has been introduced in Section 2.1 and $\delta_{\ell,k}$ is the Kronecker’s symbol. We expand $\phi(t)$ in a Fourier–Legendre series

$$\phi(t) = \frac{1}{\omega_d} \sum_{\ell=0}^{\infty} N(d, \ell) \hat{\phi}(\ell) P_\ell(t), \quad (8)$$

where

$$\hat{\phi}(\ell) = \frac{1}{\omega_{d-1}} \int_{-1}^{1} \phi(t) P_\ell(t)(1 - t^2)^{(d-2)/2} dt.$$

Due to the addition formula for spherical harmonics [18, page 10]

$$\sum_{k=1}^{N(n,\ell)} Y_{\ell,k}(x) Y_{\ell,k}(y) = \frac{N(d, \ell)}{\omega_d} P_\ell(x \cdot y), \quad (9)$$

the kernel $\Phi$ can be represented as

$$\Phi(x, y) = \sum_{\ell=0}^{\infty} \sum_{k=0}^{N(d,\ell)} \hat{\phi}(\ell) Y_{\ell,k}(x) Y_{\ell,k}(y), \quad x, y \in S^d.$$ 

and since $P_\ell(1) = 1$ we find that

$$\|\Phi(x, \cdot)\|_{H^\sigma(S^d)}^2 = \frac{1}{\omega_d} \sum_{\ell=0}^{\infty} (1 + \ell)^{2\sigma} (\hat{\phi}(\ell))^2 N(d, \ell), \quad \text{for all} \quad x \in S^d. \quad (10)$$

Chen et al. [3] proved that the kernel $\Phi$ is positive definite if and only if $\hat{\phi}(\ell) \geq 0$ for all $\ell \geq 0$ and $\hat{\phi}(\ell) > 0$ for infinitely many even values of $\ell$ and infinitely many
odd values of $\ell$; see also Schoenberg [24] and Xu and Cheney [30]. Here, we assume there is a $\sigma > d/2$ and there are positive constants $c_1$ and $c_2$ such that
\[
c_1(1 + \ell)^{-2\sigma} \leq \hat{\phi}(\ell) \leq c_2(1 + \ell)^{-2\sigma}, \quad \text{for all } \ell \geq 0, \tag{11}
\]
hence, $\Phi$ is positive definite. Also, since $N(d, \ell) = O(\ell^{d-1})$ as $\ell \to \infty$, the sum (10) is finite for each fixed $x \in S^d$. Thus the function $y \mapsto \Phi(x, y)$ belongs to $H^\sigma(S^d)$. Moreover, this function is continuous by the Sobolev imbedding theorem.

The reproducing kernel Hilbert space (RKHS) (also called the native space) induced by $\Phi$ is defined to be
\[
N_\Phi = \left\{ f \in L_2(S^2) : \|f\|^2_\Phi := \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} |\hat{f}_{\ell,k}|^2\hat{\phi}(\ell) < \infty \right\}. \tag{12}
\]
Alternatively, $N_\Phi$ is the completion of span$\{\Phi(\cdot, x) : x \in S^d\}$ with respect to the norm $\|\cdot\|_\Phi$. The norm is associated with the following inner product
\[
(f, g)_\Phi = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \hat{f}_{\ell,k}\hat{g}_{\ell,k}\hat{\phi}(\ell), \quad f, g \in N_\Phi. \tag{13}
\]
It follows from Eq. 11 that the norms in $H^\sigma(S^d)$ and $N_\Phi$ are equivalent.

2.3 Kernels defined from radial basis functions

Let $\Pi : \mathbb{R}^{d+1} \to \mathbb{R}$ be a compactly supported radial basis function (RBF) with associated RKHS $H^\tau(\mathbb{R}^{d+1})$ with $\tau > (d + 1)/2$. Examples of such RBFs are the Wendland functions (see [28]).

By restricting the function $\Pi$ to the unit sphere $S^d \subset \mathbb{R}^{d+1}$, we have a positive definite, zonal kernel on the unit sphere
\[
\Phi(x, y) = \Pi(x - y), \quad x, y \in S^d.
\]

Lemma 2.2 (Native spaces) Let $\Pi : \mathbb{R}^{d+1} \to \mathbb{R}$ be a positive definite function with native space $N_\Pi(\mathbb{R}^{d+1}) = H^\tau(\mathbb{R}^{d+1})$ with $\tau > (d + 1)/2$. Then $N_\Phi(S^d) = H^\sigma(S^d)$ with $\sigma = \tau - \frac{1}{2}$.

Proof Using [20, Proposition 4.2], we deduce that
\[
c(1 + \ell)^{-2\sigma} \leq \hat{\phi}(\ell) \leq C(1 + \ell)^{-2\sigma}.
\]
So the result follows from the definition of the Sobolev spaces (2) and the native spaces (12) on $S^d$. \hfill \Box

For a given $\delta > 0$, we define the scaled version $\Phi_\delta$ of the kernel $\Phi$ by
\[
\Phi_\delta(x, y) = \delta^{-d}\Pi((x - y)/\delta). \tag{14}
\]
We can expand $\Phi_\delta$ into a series of spherical harmonics

$$
\Phi_\delta(x, y) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d, \ell)} \hat{\phi}_\delta(\ell) Y_{\ell,k}(x) Y_{\ell,k}(y),
$$
in which the Fourier coefficients satisfy the following condition (see [12, Theorem 6.2])

$$
c_1^2 (1 + \delta \ell)^{-2\sigma} \leq \hat{\phi}_\delta(\ell) \leq c_2^2 (1 + \delta \ell)^{-2\sigma}, \quad (15)
$$

with the coefficients $c_1$ and $c_2$ from Eq. 11) possibly relaxed so that Eq. 15 holds for all $0 < \delta \leq 1$.

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

$$
\| f \|_{\Phi_\delta} = \left( \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d, \ell)} \frac{|\hat{f}_{\ell,k}|^2}{\hat{\phi}_\delta(\ell)} \right)^{1/2}, \quad (16)
$$

and the corresponding inner product is

$$
(f, g)_{\Phi_\delta} = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d, \ell)} \frac{\hat{f}_{\ell,k} \hat{g}_{\ell,k}}{\hat{\phi}_\delta(\ell)}, \quad f, g \in \mathcal{N}_\Phi. \quad (17)
$$

Clearly the norms $\| \cdot \|_{\Phi_\delta}$ for different $\delta$ are all equivalent, as given in the following lemma.

**Lemma 2.3 (Norm-equivalence)** Let $\Pi : \mathbb{R}^{d+1} \to \mathbb{R}$ be a reproducing kernel of $H^\tau(\mathbb{R}^{d+1})$ with $\tau > (d + 1)/2$. Let $\Phi_\delta(x, y) = \delta^{-d} \Pi((x - y)/\delta)$ with $x, y \in \mathbb{S}^d$. Then with $\sigma = \tau - 1/2$,

$$
c_1 \| u \|_{\Phi_\delta} \leq \| u \|_{H^\sigma(\mathbb{S}^d)} \leq c_2 \delta^{-\sigma} \| u \|_{\Phi_\delta}
$$

for all $u \in H^\sigma(\mathbb{S}^d)$.

**Proof** See [12, Lemma 3.1].

From Eq. 17 it follows that the reproducing property (13) extends to general $\delta$, that is

$$
f(x) = (f, \Phi_\delta(x, \cdot))_{\Phi_\delta}, \quad x \in \mathbb{S}^d, \quad f \in H^\sigma(\mathbb{S}^d). \quad (18)
$$

### 3 From global to local multiscale RBF interpolation

In this section we first consider RBF interpolation with a single scale, then turn to multiscale interpolation, both global and local.
3.1 Interpolation using spherical RBFs

Let $X = \{x_1, \ldots, x_N\} \subset \Omega \subseteq \mathbb{S}^d$ be a finite set of distinct points on $\Omega$. We define the mesh norm $h_{X,\Omega}$ and the separation radius $q_X$ of this point set by

$$h_{X,\Omega} = \sup_{x \in \Omega} \min_{x_j \in X} \theta(x, x_j), \quad q_X = \frac{1}{2} \min_{i \neq j} \theta(x_i, x_j),$$

where $\theta(x, y) = \cos^{-1}(x \cdot y)$ is the geodesic distance on $\mathbb{S}^d$. If $\Omega$ is a proper subset of $\mathbb{S}^d$ then we say that $h_{X,\Omega}$ is a local mesh norm. If $\Omega = \mathbb{S}^d$ then the mesh norm is global, and we simply write $h_X$.

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

$$I_{X,\delta}f(x) = \sum_{j=1}^N b_j \Phi_\delta(x, x_j), \quad I_{X,\delta}f(x_j) = f(x_j) \text{ for all } x_j \in X. \quad (19)$$

If $\delta = 1$ then we simply write $I_X f$ instead of $I_{X,1} f$. From the interpolation condition and Eq. 18 we deduce that

$$(f - I_{X,\delta} f, \Phi_\delta(\cdot, x_j))_{\Phi_\delta} = f(x_j) - I_{X,\delta} f(x_j) = 0, \text{ for all } x_j \in X.$$

Hence $I_{X,\delta} f$ is the orthogonal projection of $f$ into $\text{span}\{\Phi_\delta(\cdot, x_j) : x_j \in X\}$, from which it follows that

$$\|f - I_{X,\delta} f\|_{\Phi_\delta} \leq \|f\|_{\Phi_\delta}. \quad (20)$$

From Lemma 2.3 we then have

$$\|f - I_{X,\delta} f\|_{H^\sigma(\mathbb{S}^d)} \leq c_2 \delta^{-\sigma} \|f\|_{\Phi_\delta}. \quad (21)$$

**Lemma 3.1 (Zeros Theorem)** Let $\Omega \subseteq \mathbb{S}^d$ be either an open connected region with Lipschitz boundary or $\Omega = \mathbb{S}^d$. Assume that a finite set $X \subset \Omega$ has a sufficiently small (local) mesh norm $h_{X,\Omega}$. Then, for any function $u \in H^\sigma(\Omega)$, $\sigma > d/2$, with $u|_X = 0$, for all $0 \leq \nu \leq \sigma$ we have

$$\|u\|_{H^\nu(\Omega)} \leq C h_{X,\Omega}^{\sigma-\nu} \|u\|_{H^\sigma(\Omega)}. \quad (22)$$

**Proof** For $\Omega \subset \mathbb{S}^d$ an open and connected set with Lipschitz boundary, the proof follows from the zeros lemma for Lipschitz domains on a Riemannian manifold in [7, Theorem A.11]. The case $\Omega = \mathbb{S}^d$ was proved earlier in [10].
Theorem 3.2 Let $\Omega \subseteq \mathbb{S}^d$ be either a spherical cap that satisfies the conditions of Theorem 2.1 or $\Omega = \mathbb{S}^d$. Assume that a finite set $X \subset \Omega$ has a sufficiently small (local) mesh norm $h_{X,\Omega}$. Then,
\[ \| f - I_{X,\delta} f \|_{L_2(\Omega)} \leq C \delta^{-\sigma} h_{X,\Omega}^\sigma \| f \|_{H^\sigma(\Omega)}. \]
In particular, when $\delta = 1$, we have
\[ \| f - IX f \|_{L_2(\Omega)} \leq C h_{X,\Omega} \| f \|_{H^\sigma(\Omega)}. \]

Proof Let $u := f - I_{X,\delta} f$, then $u|_X = 0$. Using Lemma 3.1, we have
\[ \| f - I_{X,\delta} f \|_{L_2(\Omega)} \leq C h_{X,\Omega} \| f - I_{X,\delta} f \|_{H^\sigma(\Omega)}. \]

If $\Omega$ is a spherical cap our assumptions on $\Omega$ allow us to extend the function $f \in H^\sigma(\Omega)$ to a function $Ef \in H^\sigma(\mathbb{S}^d)$. Moreover, since $X \subset \Omega$ and $Ef|_{\Omega} = f|_{\Omega}$, the interpolant $I_{X,\delta} f$ coincides with the interpolant $I_{X,\delta}(Ef)$ on $\Omega$. Therefore,
\[ \| f - I_{X,\delta} f \|_{H^\sigma(\Omega)} = \| Ef - I_{X,\delta}(Ef) \|_{H^\sigma(\Omega)} \leq C \| Ef - I_{X,\delta}(Ef) \|_{H^\sigma(\mathbb{S}^d)} \]
\[ \leq C \delta^{-\sigma} \| Ef \|_{H^\sigma(\mathbb{S}^d)} \quad \text{by (21)} \]
\[ \leq C \delta^{-\sigma} \| f \|_{H^\sigma(\mathbb{S}^d)} \quad \text{by Lemma 2.3} \]
\[ \leq C \delta^{-\sigma} \| f \|_{H^\sigma(\Omega)} \quad \text{by Theorem 2.1}. \]

The case $\Omega = \mathbb{S}^d$ was proved in [12, Theorem 3.2].

3.2 The global and local multiscale algorithm

Suppose $X_1, X_2, \ldots, X_m \subset \mathbb{S}^d$ is a sequence of finite point sets with mesh norms $h_1, h_2, \ldots, h_m$ respectively. The mesh norms are assumed to satisfy $h_{j+1} \approx \mu h_j$ for some fixed $\mu \in (0, 1)$. After that, suppose $X_{m+1}, X_{m+2}, \ldots, X_n \subset \Omega$ is a sequence of point sets with local mesh norms $h_{m+1,\Omega}, \ldots, h_n,\Omega$, where $\Omega \subset \mathbb{S}^d$ is some open connected subset. In future we will write $h_j$ for $h_{X_j,\Omega}$ for all $j = 1, \ldots, n$.

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 in this way 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 $\Phi_j := \Phi_{\delta_j}$, and also define the scaled approximation space $W_j = \text{span}\{\Phi_j(\cdot, x) : x \in X_j\}$.

We start with a widely spread set of points on the global scale and use a basis function with scale $\delta_1$ 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 $\delta_2$, 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 $e_2 = f - f_2 = e_1 - I_{X_2,\delta_2} e_1$; and so on. After $m$ global steps we switch to local refinement, i.e. from step $(m+1)$ onwards the set $X_{m+1}$ is localized to a small region $\Omega$ on the sphere, and the new correction $s_{m+1}$ is constructed from the
local space $W_{m+1}$ and the new approximation is $f_{m+1} = f_m + s_{m+1}$. The multiscale algorithm then is continued for a further local $n - m$ steps.

\begin{algorithm}
\textbf{Algorithm 1} Multiscale global/local algorithm

\textbf{Data:} Right hand side $f$, number of levels $n$

\begin{algorithmic}
\State Set $f_0 = 0$, $e_0 = f$.
\For {$j = 1, 2, \ldots, n$}
\State Determine the (global or local) interpolant $s_j \in W_j$ to $e_{j-1}$
\State Set $f_j = f_{j-1} + s_j$
\State Set $e_j = e_{j-1} - s_j$
\EndFor
\end{algorithmic}

\textbf{Result:} Approximation solution $f_n \in W_1 + \cdots + W_n$
\end{algorithm}

Remark Clearly, we could continue the algorithm by choosing an even smaller region $\Omega' \subset \Omega$ and a sequence of point sets $X_{n+1}, X_{n+2}, \ldots \subset \Omega'$, and so on, until a desired resolution is reached. For simplicity of presentation, we restrict ourselves to the situation of one zooming-in region $\Omega$ in the subsequent error analysis (though not in the numerical example). Extension of the convergence theory to the general case is trivial.

We will show convergence for the scheme within a spherical cap $\Omega$.

**Theorem 3.3 (Convergence for functions in $H^\sigma(\mathbb{S}^d)$)** Let $X_1, \ldots, X_m$ be a sequence of point sets in $\mathbb{S}^d$ and let $X_{m+1}, \ldots, X_n$ be a sequence of point sets in $\Omega \subset \mathbb{S}^d$, where $\Omega$ satisfies the requirements in Theorem 2.1. Assume that we are performing $m$ steps of the global multilevel algorithm on $\mathbb{S}^d$ and then $n - m$ steps of the local multilevel algorithm, localised to $\Omega$. Let $h_1, \ldots, h_m$ be the global mesh norms and $h_{m+1}, \ldots, h_n$ be the local mesh norms of the sets $X_1, \ldots, X_m$ and $X_{m+1}, \ldots, X_n$, respectively, and assume that, for some $\mu \in (0, 1)$, $h_{j+1} = \mu h_j$, for each $j = 1, \ldots, n - 1$.

Let $\Phi$ be a kernel generating $H^\sigma(\mathbb{S}^d)$ and let $\Phi_j := \Phi_{\delta_j}$ be defined by Eq. 14 with scale factor $\delta_j = \nu h_j$ where $1/h_1 \geq \nu \geq \gamma/\mu \geq 1$ with a fixed $\gamma > 0$. Assume that the target function $f$ belongs to $H^\sigma(\mathbb{S}^d)$.

Then the algorithm converges in the $L_2(\Omega)$ sense linearly in the number of levels. To be more precise, there is a constant $C > 0$ and a constant $\alpha > 0$, where $\alpha < 1$ for $\mu$ sufficiently small, such that

$$\|f - f_n\|_{L_2(\Omega)} \leq C\alpha^n \|f\|_{H^\sigma(\mathbb{S}^d)}$$

for all $f \in H^\sigma(\mathbb{S}^d)$.

The theorem is a generalisation of the main result in [12].

In preparation for the proof of the theorem we first prove the following technical lemma.

**Lemma 3.4** Let $e_j$ for $j = 0, \ldots, n$ be as in Algorithm 1, and let $E$ be the extension operator from $\Omega$ to $\mathbb{S}^d$ as defined in Theorem 2.1. Then
Proof For \( j = 1, \ldots, m \) we have, using (21),
\[
\| e_j \|_{H^\sigma(\mathbb{R}^d)} = \| e_j - I X_j, \delta_j e_j - 1 \|_{H^\sigma(\mathbb{R}^d)} \leq C \delta_j^{-\sigma} \| e_j - 1 \|_{\Phi_j}.
\]

For \( j = m + 2, \ldots, n \) we have, by using the property of the extension operator and Eq. 21
\[
\| e_j \|_{H^\sigma(\mathbb{R}^d)} = \| e_m - I X_{m+1}, \delta_{m+1} e_m \|_{H^\sigma(\mathbb{R}^d)} \leq C \delta_{m+1}^{-\sigma} \| e_m \|_{\Phi_{m+1}}.
\]

For the intermediate case, when \( j = m + 1 \), we avoid the extension operator by arguing as follows
\[
\| e_{m+1} \|_{H^\sigma(\mathbb{R}^d)} = \| e_m - I X_{m+1}, \delta_{m+1} e_m \|_{H^\sigma(\mathbb{R}^d)} \leq C \delta_{m+1}^{-\sigma} \| e_m \|_{\Phi_{m+1}}.
\]

Proof of Theorem 3.3 In the proof we use repeatedly the fact that \( e_j | X_j = 0 \), allowing us to use the zeros theorem (Lemma 3.1), and we also make essential use of the extension theorem (Theorem 2.1) and Lemma 3.4.

We start by noting that
\[
\| f - f_n \|_{L^2(\Omega)} = \| e_n \|_{L^2(\Omega)} \leq C h_n^\sigma \| e_n \|_{H^\sigma(\Omega)} = C h_n^\sigma \| E e_n \|_{H^\sigma(\Omega)}
\]
\[
\leq C h_n^\sigma \| E e_n \|_{H^\sigma(\mathbb{R}^d)}
\]
\[
\leq C h_n^\sigma \delta_{n+1}^{-\sigma} \| E e_n \|_{\Phi_{n+1}}
\]
\[
= C \| E e_n \|_{\Phi_{n+1}},
\]
where in the second-last step we use Lemma 2.3 with \( \delta = \delta_{n+1} \), and in the last step \( h_n/\delta_{n+1} = 1/(\mu \nu) \leq 1/\gamma \).

The result will then follow by establishing the recursions
\[
\| e_j \|_{\Phi_{j+1}} \leq \alpha \| e_{j-1} \|_{\Phi_j}, \quad j = 1, \ldots, m
\]
\[
\| E e_m \|_{\Phi_{m+1}} \leq \alpha \| e_m \|_{\Phi_{m+1}}
\]
\[
\| E e_j \|_{\Phi_{j+1}} \leq \alpha \| E e_{j-1} \|_{\Phi_j}, \quad j = m + 2, \ldots, n,
\]
where \( \alpha \) is some real number satisfying \( 0 < \alpha < 1 \).

The first of these is exactly as in [12]. We shall prove the third recursion (25), noting that Eq. 23 can then be recovered by replacing \( \Omega \) by \( \mathbb{R}^d \) and omitting the extension operator \( E \).
For $j = m + 2, \ldots, n$ we have by definition of the norm (16) together with Eq. 15

$$\|Ee_j\|_{\Phi_{j+1}}^2 = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \frac{|(\hat{E}e_j)_{\ell,k}|^2}{\phi_{\delta_{j+1}}(\ell)}$$

$$\leq C \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} |(\hat{E}e_j)_{\ell,k}|^2 (1 + \delta_{j+1} \ell)^{2\sigma}$$

$$=: I_1 + I_2,$$

where

$$I_1 = C \sum_{\ell \leq 1/\delta_{j+1}} \sum_{k=1}^{N(d,\ell)} |(\hat{E}e_j)_{\ell,k}|^2 (1 + \delta_{j+1} \ell)^{2\sigma}$$

$$\leq C 2^{2\sigma} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} |(\hat{E}e_j)_{\ell,k}|^2 = C \|Ee_j\|_{L^2(\Sigma^d)}^2$$

$$\leq C \|e_j\|_{L^2(\Omega)}^2$$

$$\leq C h_j^{2\sigma} \|e_j\|_{H^\sigma(\Omega)}^2$$

by Lemma 3.1

$$\leq C \left( \frac{h_j}{\delta_j} \right)^{2\sigma} \|Ee_{j-1}\|_{\Phi_j}^2$$

by Lemma 3.4 iii)

$$= C \nu^{-2\sigma} \|Ee_{j-1}\|_{\Phi_j}^2 \leq C \mu^{2\sigma} \|Ee_{j-1}\|_{\Phi_j}^2,$$

and

$$I_2 = C \sum_{\ell > 1/\delta_{j+1}} \sum_{k=1}^{N(d,\ell)} |(\hat{E}e_j)_{\ell,k}|^2 (1 + \delta_{j+1} \ell)^{2\sigma}$$

$$\leq C (2\delta_{j+1})^{2\sigma} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} |(\hat{E}e_j)_{\ell,k}|^2 \ell^{2\sigma}$$

$$\leq C \delta_{j+1}^{2\sigma} \|Ee_j\|_{H^\sigma(\Sigma^d)}^2 \leq C \delta_{j+1}^{2\sigma} \|e_j\|_{H^\sigma(\Omega)}^2$$

by Theorem 2.1

$$\leq C \left( \frac{\delta_{j+1}}{\delta_j} \right)^{2\sigma} \|Ee_{j-1}\|_{\Phi_j}^2$$

by Lemma 3.4 iii)

$$= C \mu^{2\sigma} \|Ee_{j-1}\|_{\Phi_j}^2.$$

Thus we have proved $\|Ee_j\|_{\Phi_{j+1}} \leq C \mu^{-\sigma} \|Ee_{j-1}\|_{\Phi_j}$. With $\mu$ small enough, we can choose $\alpha = C \mu^{2\sigma} < 1$, so proving the recursion (25).

Next, we discuss the switch over from global to local. We have

$$e_{m+1} = e_m - I_{X_{m+1}, \delta_{m+1}} e_m, \quad \text{with} \ X_m \subset \Sigma^d, X_{m+1} \subset \Omega.$$

As before, we decompose

$$\|Ee_{m+1}\|_{\Phi_{m+2}}^2 = I_1 + I_2.$$
Now we have
\[ I_1 \leq C \| E_{m+1} \|_{L_2(\mathbb{S}^d)}^2 \]
\[ \leq C \| e_{m+1} \|_{L_2(\Omega)}^2 \]
\[ \leq C h_{m+1}^{2\sigma} \| e_{m+1} \|_{H^\sigma(\Omega)}^2 \]
\[ \leq C \left( \frac{h_{m+1}}{\delta_{m+1}} \right)^{2\sigma} \| e_m \|_{\Phi_{m+1}}^2 \]
by Lemma 3.1
\[ \leq C \left( \frac{\delta_{m+2}}{\delta_{m+1}} \right)^{2\sigma} \| e_m \|_{\Phi_{m+1}}^2 \]
by Lemma 3.4 ii).

The second term can be bounded by
\[ I_2 \leq C \delta_{m+2}^{2\sigma} \| E_{m+1} \|_{H^\sigma(\mathbb{S}^d)}^2 \]
\[ \leq C \delta_{m+2}^{2\sigma} \| e_{m+1} \|_{H^\sigma(\Omega)}^2 \]
\[ \leq C \left( \frac{\delta_{m+2}}{\delta_{m+1}} \right)^{2\sigma} \| e_m \|_{\Phi_{m+1}}^2 \]
by Lemma 3.4 ii).

Hence we find
\[ \| E_{m+1} \|_{\Phi_{m+2}}^2 \leq C \mu^{2\sigma} \| e_m \|_{\Phi_{m+1}}^2, \]
and this can be no larger than \( \alpha^2 \| e_m \|_{\Phi_{m+1}}^2 \) for \( \mu \) sufficiently small.

The first recursion (23) follows by the same proof if \( E \) is omitted and \( \Omega \) is replaced by \( \mathbb{S}^d \).

Taken in the reverse order, the recursive steps (25), (24) and (23) give
\[ \| E_n \|_{\Phi_{n+1}} \leq \alpha^{n-m} \| E_{m+1} \|_{\Phi_{m+2}} \]
\[ \leq \alpha^{n-m} \| e_{m+1} \|_{\Phi_{m+1}} \]
\[ \leq \alpha^n \| e_0 \|_{\Phi_{1}} = \alpha^n \| f \|_{H^\sigma(\mathbb{S}^d)}, \]
which together with Eq. 22 proves the desired result.

The following result on the condition numbers of the matrices is adapted from [12, Theorem 7.3].

**Theorem 3.5** Assume that the conditions in Theorem 3.3 hold, together with
\[ q_j \leq h_j \leq c_q q_j \text{ for } j = 1, 2, \ldots, n \text{ with } c_q > 1. \]
There exists \( C > 0 \) such that the condition number of the interpolation matrices at each level of the multiscale approximation in Algorithm 1 are bounded by
\[ \kappa \leq C, \quad j = 1, \ldots, n. \]

### 4 Escaping the native space

In this section, our target function \( f \) will be assumed to be in \( H^\beta(\mathbb{S}^d) \) for some \( \beta \in (d/2, \sigma) \). The extension of an approximation result to spaces rougher than the native space is often referred to as “escaping the native space”.

Let \( K \) be the reproducing kernel of the Sobolev space \( H^{\beta+1/2}(\mathbb{R}^{d+1}) \). We define the kernel \( \Psi \) by restricting \( K \) to the sphere,
\[ \Psi(x, y) = K(x - y), \quad x, y \in \mathbb{S}^d. \]
For $0 < \delta \leq 1$, the scaled version of $\Psi$ is defined by

$$\Psi_\delta(x, y) = \delta^{-d} K((x - y)/\delta), \quad x, y \in \mathbb{S}^d.$$  

It can be expanded into a series of spherical harmonics as

$$\Psi_\delta(x, y) = \sum_{\ell=0}^{\infty} \sum_{k=0}^{N(d, \ell)} \tilde{\psi}_\delta(\ell) Y_{\ell,k}(x) Y_{\ell,k}(y).$$  

(26)

It is known [13, Lemma 2.1] that there are positive constants $c_3, c_4$ independent of $\delta$ and $\ell$ so that

$$c_2^2 (1 + \delta \ell)^{-2\beta} \leq \tilde{\psi}_\delta(\ell) \leq c_2^2 (1 + \delta \ell)^{-2\beta}, \quad \ell \geq 0.$$  

(27)

We can define the RKHS with the reproducing kernel $\Psi_\delta$ and its norm $\| \cdot \|_{\Psi_\delta}$ as in Eqs. 12 and 16. By Lemma 2.3, the norm $\| \cdot \|_{\Psi_\delta}$ defined on $N_{\Psi_\delta}$ is equivalent to $\| \cdot \|_{H^{1/2}(\mathbb{S}^d)}$.

For the multiscale convergence theory, the sole thing that prevents us from using the proof of Theorem 3.3 with $\Phi_\delta$ replaced by $\Psi_\delta$ is that a key stability property is missing: the orthogonal projection property (21) no longer holds. We therefore approximate a function in $\Psi_\delta$ by a polynomial (which of course lies in all Sobolev spaces), and apply the orthogonal projection property to that polynomial.

For a given smooth function $f$, the following lemma [13, Lemma 4.3] asserts the existence of a spherical polynomial that interpolates $f$ on a set of scattered points $X$ and, simultaneously, has a $\Psi_\delta$ norm comparable to that of $f$.

**Lemma 4.1** Let $f \in H^{\beta} (\mathbb{S}^d)$, $\beta > d/2$ and let $X$ be a finite subset of $\mathbb{S}^d$ with separation radius $q_X$. Let $\delta \in (0, 1]$ be given. There exists a constant $\kappa$, which depends only on $d$ and $\beta$, such that if $L \geq \kappa \max(\delta/q_X, 1/\delta)$, then there is a spherical polynomial $p \in \mathcal{P}_L$ such that $p|_X = f|_X$ and

$$\| f - p \|_{\Psi_\delta} \leq 5 \| f \|_{\Psi_\delta}.$$  

**Remark** The dependence of the lower bound for $L$ on the mesh radius $q_X$ in the last lemma makes it necessary to impose a weak condition on $q_X$ in the following theorem.

**Theorem 4.2** (Convergence outside the native space) Let $X_1, \ldots, X_m$ be a sequence of point sets in $\mathbb{S}^d$ with separation radius $q_X$. Let $\delta \in (0, 1]$ be given. There exists a constant $\kappa$, which depends only on $d$ and $\beta$, such that if $L \geq \kappa \max(\delta/q_X, 1/\delta)$, then there is a spherical polynomial $p \in \mathcal{P}_L$ such that $p|_X = f|_X$ and

$$\| f - p \|_{\Psi_\delta} \leq 5 \| f \|_{\Psi_\delta}.$$  

Let $\Phi$ be a kernel generating $H^{1/2} (\mathbb{S}^d)$ and let $\Phi_j := \Phi_{\delta_j}$ be defined by Eq. 14 with scale factor $\delta_j = \nu h_j$ where $1/h_1 \geq \gamma \geq \gamma/\mu \geq 1$ with a fixed $\gamma > 0$. Let $\Psi$
be a kernel generating $H^\beta(S^d)$ with $\sigma > \beta > d/2$ and let $\Psi_j := \Psi_{\delta_j}$ be the scaled version (26) using the scale factor $\delta_j$. Assume that the target function $f$ belongs to $H^\beta(S^d)$.

Then, Algorithm 1 converges in the $L_2(\Omega)$ sense linearly in the number of levels. To be more precise, there is a constant $C > 0$ and a constant $\alpha > 0$, which for $\mu$ sufficiently small is < 1, such that

$$\|f - f_n\|_{L_2(\Omega)} \leq C\alpha^n \|f\|_{H^\beta(S^d)}$$

for all $f \in H^\beta(S^d)$.

Similarly to the case of Theorem 3.3, the proof of the theorem rests upon the following technical lemma. But in this case the proof is necessarily different, because the orthogonal projection property (21) is not available.

**Lemma 4.3** Let $e_j$ for $j = 0, \ldots, n$ be as in Algorithm 1, and let $E$ be the extension operator from $\Omega$ to $S^d$ as defined in Theorem 2.1. Let the assumptions on $h_j$ and $q_j$ be satisfied as in Theorem 4.2. Then

1. $\|e_j\|_{H^\beta(S^d)} \leq C\delta_j^{-\beta}\|e_{j-1}\|_{\Psi_j}$ for $j = 1, \ldots, m$,
2. $\|e_{m+1}\|_{H^\beta(\Omega)} \leq C\delta_{m+1}^{-\beta}\|e_m\|_{\Psi_{m+1}}$,
3. $\|e_j\|_{H^\beta(\Omega)} \leq C\delta_j^{-\beta}\|Ee_{j-1}\|_{\Psi_j}$ for $j = m + 2, \ldots, n$.

**Proof** We prove part iii) since part i) follows easily by replacing $\Omega$ by $S^d$ and omitting the extension operator, and part (ii) is in an obvious sense intermediate. See also the proof of [13, Lemma 4.4].

We use the extension operator to extend $e_{j-1}$ to $Ee_{j-1}$ defined on the whole sphere, for $j = m + 2, \ldots, n$. Then, with $L_j := \lceil\kappa \max\{\delta_j/q_j, 1/\delta_j\}\rceil$, by Lemma 4.1, there is a polynomial $p \in P_{L_j}$ that interpolates and approximates $Ee_{j-1}$, in the sense that

$$p|_{X_j} = Ee_{j-1}|_{X_j} \text{ and } \|p - Ee_{j-1}\|_{\Psi_j} \leq 5\|Ee_{j-1}\|_{\Psi_j}. \quad (28)$$

We note that the RBF interpolant for $e_{j-1}$ coincides with the RBF interpolant for $Ee_{j-1}$ on $X_j$. Therefore,

$$\|e_j\|_{H^\beta(\Omega)} = \|e_{j-1} - I_{X_j, \delta_j}e_{j-1}\|_{H^\beta(\Omega)} = \|Ee_{j-1} - I_{X_j, \delta_j}Ee_{j-1}\|_{H^\beta(\Omega)} \leq C\|Ee_{j-1} - I_{X_j, \delta_j}Ee_{j-1}\|_{H^\beta(S^d)} \leq C\left(\|Ee_{j-1} - p\|_{H^\beta(S^d)} + \|p - I_{X_j, \delta_j}Ee_{j-1}\|_{H^\beta(S^d)}\right). \quad (29)$$

The first term of Eq. 29 can be bounded using Lemmas 2.3 and 4.1,

$$\|Ee_{j-1} - p\|_{H^\beta(S^d)} \leq c_4\delta_j^{-\beta}\|Ee_{j-1} - p\|_{\Psi_j} \leq 5c_4\delta_j^{-\beta}\|Ee_{j-1}\|_{\Psi_j}. \quad (30)$$
For the second term, since $p|_{X_j} = E e_{j-1}|_{X_j}$, the interpolant $I_{X_j, \delta_j} E e_{j-1}$ is identical to $I_{X_j, \delta_j} p$, hence by using Lemma 3.1 and Eq. 21, we have

$$
\| p - I_{X_j, \delta_j} E e_{j-1} \|_{H^\beta(S^d)} = \| p - I_{X_j, \delta_j} p \|_{H^\beta(S^d)} \leq C h_j^{\sigma-\beta} \| p - I_{X_j, \delta} p \|_{H^\sigma(S^d)} \\
\leq C h_j^{\sigma-\beta} \delta_j^{-\sigma} \| p \|_{\Phi_j} \leq C \delta_j^{-\beta} \| p \|_{\Phi_j}.
$$

For the polynomial $p$ of degree $L_j$, using the definition (16), condition (15) and the fact that $\beta < \sigma$, we have

$$
\| p \|_{\Phi_j}^2 \leq C \sum_{\ell=0}^{L_j} \sum_{k=1}^{N(d, \ell)} (1 + \delta_j \ell)^{2\sigma} |\tilde{p}_{\ell k}|^2 \\
\leq C (1 + \delta_j L_j)^{2(\sigma-\beta)} \sum_{\ell=0}^{L_j} \sum_{k=1}^{N(d, \ell)} (1 + \delta_j \ell)^{2\beta} |\tilde{p}_{\ell k}|^2 \\
\leq C \| p \|_{\Psi_j}^2,
$$

where in the last step we used $L_j \leq C/\delta_j$. (Since $h_j \leq c q^{1/2} q_j$ and since $\delta_j = \nu h_j$, we see that $\delta_j/q_j \leq c/\delta_j$ and hence $L_j \leq C/\delta_j$). Thus, combining these above estimates together with the fact that $\| p \|_{\Psi_j} \leq 6 \| E e_{j-1} \|_{\Psi_j}$ we obtain

$$
\| p - I_{X_j, \delta_j} E e_{j-1} \|_{H^\beta(S^d)} \leq C \delta_j^{-\beta} \| E e_{j-1} \|_{\Psi_j}. \tag{31}
$$

Combining (29), (30) and (31), we obtain the desired result.

Proof of Theorem 4.2 The proof is identical to that for Theorem 3.3 once we have established Lemma 4.3: the only difference is that $\sigma$ is replaced by $\beta$ and $\Phi_j$ by $\Psi_j$. We leave the details to the reader.

5 Computational cost

Since the interpolation matrix is symmetric positive definite, we can use the conjugate gradient (CG) method to obtain the coefficients for each level.

The computational cost for solving a linear system $A c = b$ using the CG method is $O(M \sqrt{\kappa})$, where $M$ is the number of non-zero entries in the matrix $A$ and $\kappa$ is the condition number of $A$ [26, Chapter 10].

In our context, the matrix $A$ is the interpolation matrix at each level, which is defined by

$$
A_{i, j} = \Phi_\delta(x_i, x_j), \quad x_i, x_j \in X.
$$

Since the RBF $\Phi_\delta$ has as its support a spherical cap, to bound the number of non-zero entries of the matrix $A$, we need to estimate the number of points from the set $X$ that lie inside the spherical cap. To this end, we need the following lemma, which is taken from [14, Lemma 4.1]. The proof is based on volume packing arguments.
Lemma 5.1 Suppose $X$ is a set of scattered points on the unit sphere $\mathbb{S}^d$ with separation radius $q_X$. Let $q_X \leq R \leq \pi/2$ be given and let

$$\tilde{M} := \sup_{x \in \mathbb{S}^d} \# \{ x_k \in X : \theta(x, x_k) < R \}. $$

Then

$$\tilde{M} \leq \frac{A(2R)}{A(q_X)}, \tag{32}$$

where $A(\alpha)$ denotes the area of a spherical cap of geodesic radius $\alpha$, given by

$$A(\alpha) = \omega_{d-1} \int_{0}^{\alpha} \sin^{d-1} \theta d\theta,$$

where $\omega_{d-1}$ is the surface area of the unit sphere $\mathbb{S}^{d-1}$.

Theorem 5.2 Suppose $X_1, \ldots, X_n$ be a quasi-uniform sequence of point sets on $\mathbb{S}^2$. Ignoring the pre-processing steps, the computational cost of the multiscale algorithm described in Section 3.2 is

$$O(N_1 + N_2 + \ldots + N_n),$$

where $N_k$ is the cardinality of the set $X_k$ for $k = 1, \ldots, n$.

Proof On $\mathbb{S}^2$, the area of a spherical cap of radius $\alpha$ is explicitly given by

$$A(\alpha) = 2\pi (1 - \cos \alpha).$$

The support of $\Phi_\delta(x, y)$ on the sphere is determined by the condition

$$\sqrt{2 - 2x \cdot y} \leq \delta,$$

thus the support radius of the RBF is

$$R(\delta) = \cos^{-1} \left( 1 - \frac{\delta^2}{2} \right). \tag{33}$$

With $R = R(\delta)$ as above, the right hand side of Eq. 32 is bounded by

$$\frac{A(2R(\delta))}{A(q_X)} = \frac{\delta^2 (1 - \delta^2/4)}{\sin^2(q_X/2)} \leq \frac{\delta^2}{\sin^2(q_X/2)} \leq \pi^2 \left( \frac{\delta}{q_X} \right)^2, \tag{34}$$

where in the last step we used $\sin x \geq 2x/\pi$ for $x \in [0, \pi/2]$.

In the multiscale algorithm, the support radius $\delta$ is chosen proportional to the separation radius $q_X$, hence the right hand side of Eq. 34 is a constant. So it follows from Lemma 5.1 that the number of points of $X$ inside the support of the RBF, which is $\tilde{M}$, is bounded by a constant. Since each row of the corresponding interpolation matrix $A$ has at most $\tilde{M}$ non-zero entries, the number of non-zero entries of the interpolation matrix $A$ of size $N \times N$ is

$$M = N \tilde{M} = O(N).$$
Since the condition number $\kappa$ is bounded (see Theorem 3.5), the computational cost for the multiscale algorithm with $n$ levels is the total cost of solving $n$ linear systems using the CG method, which is

$$O(N_1 + N_2 + \ldots + N_n),$$

where $N_k$ is the cardinality of each point set $X_k$, for $k = 1, \ldots, N$.

If on the other hand, a one-shot interpolation algorithm is used (without scaling the RBFs) with $N$ points, and if the separation radius of the set of point is $q_X$, then the number of non-zero elements of the matrix $A$ is $O(N/q^2_X)$ (by Eq. 34) and the condition number $\kappa = O(1/q^{2+\tau}_X)$ (see [12, Theorem 7.3]). Here $\tau$ is a real and positive parameter which depends on the smoothness of the RBF. Hence the computational cost for solving the linear system using the conjugate gradient method is

$$O(N/q^2_X + \tau).$$ (35)

From the results of Theorem 5.2 and Eq. 35, it can be seen that the multiscale algorithm is significantly more cost effective, especially when $q_X$ is small.

6 Numerical experiment

In this section, we describe a numerical experiment to illustrate the multiscale algorithm described in previous sections.

Let $p = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^T$ and $q = (-0.7476, 0.5069, 0.4289)^T$ be two given points on $S^2$, and let $\alpha := \pi/12$ and $\rho := \pi/96$. Let $\Omega_1 = G(q, \alpha)$ and $\Omega_2 = G(q, \rho)$ be concentric spherical caps centered at $q$, with geodesic radii $\alpha$ and $\rho$ respectively. Note that the successive areas of $S^2$, $\Omega_1$ and $\Omega_2$ are decreasing by a factor of roughly 60.

A point on $S^2$ is parametrized by polar coordinates $\theta, \phi$, with

$$x = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \text{ for } \theta \in [0, \pi] \text{ and } \phi \in [0, 2\pi).$$

Let $t = \cos^{-1}(p \cdot x)$ and let $s = \cos^{-1}(q \cdot x)$. The target function $f$ is given by

$$f(x) = 2 + \left[\sin(3t) \cos(100t) + (1 - 3s/2\rho)^2 \cos(2000\theta) \right] S(\theta),$$

where $S(\theta)$ is a cubic spline which takes the values of 1 for $\theta \in [0, \pi/2]$ and 0 for $\theta \in [2\pi/3, \pi]$. The function $f$, shown in Fig. 1, is designed to have both global features and finer features. On the global scale, the effect of the spline multiplying the second term is that $f$ has the constant value 2 below a latitude of 30° south. This feature was chosen because we want to be sure that the approximation scheme approximates a constant satisfactorily. (We remind the reader that approximating a constant with compactly supported radial basis functions is non-trivial, specially if the scale is comparable to the mesh norm). The function $f$ also contains a slow oscillation (seen in Fig. 1) and a localized fast oscillation inside the spherical cap $\Omega_2$, as shown in left panel of Fig. 7. Note that the period of the oscillation, given by the last term in the expression for $f$, corresponds to approximately 20 km if mapped to Earth’s surface. This finer oscillation is too localized to be seen in Fig. 1.
In the experiment we use 9 multiscale levels, zooming in to the cap $\Omega_1$ after three global levels, and zooming in again to the smaller cap $\Omega_2$ after a further three levels. In the first three (global) levels, the sets of points $X_1$, $X_2$, and $X_3$ are each centers of equal area regions generated by a partitioning algorithm \[23\]. The number of points in each set $X_1$, $X_2$, $X_3$ is increasing by a factor of 4 (see Table 1 below); the sets are not nested. The configuration of the set $X_1$ is plotted in Fig. 2.

The sets $X_4$, $X_5$ and $X_6$ are also centers of equal area regions, but the regions are partitioned from $\Omega_1$ rather than the whole sphere (see Fig. 3 for an illustration). For simplicity of language we call levels 4 to 6 the “local” levels. Similarly, $X_7$, $X_8$ and $X_9$ are the results of partitioning $\Omega_2$ into equal area regions. We call levels 7 to 9 the “superlocal” levels. At every stage the scale is halved exactly and the mesh norm halved approximately. The parameter details for the successive levels are given in Table 1.

**Fig. 1** Exact function from a global view

**Table 1** Parameters and local errors using 9 levels of global and local interpolation

| Level | $N$  | $\delta_j$ | $h_j$  | $\|e_j\|_{L^2(\Omega_2)}$ | $\kappa_j$ |
|-------|------|------------|--------|--------------------------|------------|
| 1     | 500  | 1/4        | 0.1129 | 4.24e-02                 | 1.68       |
| 2     | 2000 | 1/8        | 0.0569 | 4.07e-02                 | 1.68       |
| 3     | 8000 | 1/16       | 0.0281 | 3.45e-02                 | 1.69       |
| 4     | 500  | 1/32       | 0.0186 | 1.56e-02                 | 3.25       |
| 5     | 2000 | 1/64       | 0.0089 | 9.83e-03                 | 3.39       |
| 6     | 8000 | 1/128      | 0.0041 | 8.94e-03                 | 3.30       |
| 7     | 500  | 1/256      | 0.0018 | 7.87e-03                 | 3.24       |
| 8     | 2000 | 1/512      | 0.0009 | 2.87e-03                 | 3.37       |
| 9     | 8000 | 1/1024     | 0.0005 | 7.97e-04                 | 3.28       |
The RBF used in the experiment is the Wendland function
\[ \Pi(||x||) = (1 - ||x||)^4 \frac{1}{3} (4||x|| + 1) \]
and its scaled version is
\[ \Pi_\delta(||x||) = \delta^{-2} (1 - ||x||/\delta)^4 \frac{1}{3} (4||x||/\delta + 1), \]

Fig. 2 The set \( X_1 \) of 500 points generated by the partitioning algorithm [23]

Fig. 3 A set of 100 points inside the spherical cap \( \Omega_1 \)
where at level \( j \), we set \( \delta = \delta_j \). It is known that \( \Pi \) generates \( H^3(\mathbb{R}^3) \) (see [28]) and hence the kernel \( \Phi(x, y) = \Pi(x - y) \) for \( x, y \in S^2 \) generates \( H^{5/2}(S^2) \) (see [21]).

In Fig. 4, we show the approximation after the three global levels, using the point sets \( X_1, X_2 \) and \( X_3 \). We also show on this figure the spherical cap \( \Omega_1 \), to show the

---

**Fig. 4** The global view after three global levels, with the cap \( \Omega_1 \) shown

---

**Fig. 5** The local view after 3 global and 3 local levels, showing both the large cap \( \Omega_1 \) and the small (“superlocal”) cap \( \Omega_2 \).
first region where we intend to zoom in. At this stage it is clear visually that the approximation scheme not yet resolved the slower oscillations, but the broad features, including the constant value in southern latitudes, are already apparent.

In Fig. 5, we show the approximation on the spherical cap $\Omega_1$ after 6 levels (3 global and 3 local). We also show the smaller spherical cap $\Omega_2$, inside which it is clear that after 6 multiscale levels the slow oscillations have largely been resolved but fine scale features have not.

Finally, in Fig. 6, we show the approximation on the small spherical cap $\Omega_2$ after 9 levels (3 global, 3 local and 3 superlocal). By this stage even the fine scale features are well resolved (Fig. 7).

For comparison, we carry out a more modest multiscale approximation in which we use just the last three (superlocal) levels, and separately also a single scale (‘one-shot’) approximation, in the second case using the final scale $\delta = 2^{-10}$ and the 8000 sampling points inside the cap $\Omega_2$. Poorer approximation quality of the one-shot interpolation can be seen by eye in the right panel of Fig. 8. For the multiscale result in the left panel of Fig. 8 that uses just the last three levels the visual result is of intermediate quality: not as good as the full multiscale result, but certainly better than the one-shot result.

In Tables 1 and 2 approximate $L_2(\Omega_2)$ errors are given, in the first case for the full 9-level multiscale approximation, in the second case for the 3-level superlocal
version. These were computed over a rectangular grid $\mathcal{G}$ of size $1/64 \text{ degree} \times 1/64 \text{ degree}$ restricted to the spherical cap $\Omega_2$,

$$\|e_j\|_{L^2(\Omega_2)} := \left( \frac{|\Omega_2|}{|\mathcal{G} \cap \Omega_2|} \sum_{x(\theta, \phi) \in \mathcal{G} \cap \Omega_2} |f(\theta, \phi) - f_j(\theta, \phi)|^2 \right)^{1/2},$$

where the area $|\Omega_2|$ of the cap $\Omega_2$ is included so that the computed quantity is an approximation to the $L^2(S^2)$ norm of the error. With the grid $\mathcal{G}$ as above the number of points in the cap $\Omega_2$ is $|\mathcal{G} \cap \Omega_2| = 50063$. The condition number of the interpolation matrix at level $j$ is denoted by $\kappa_j$.

The $\|e\|_{L^2(\Omega_2)}$ error for the one-shot approximation is $2.00 \cdot 10^{-2}$, which is much larger than errors from the level 9-level multiscale approach, and also larger than the error from the 3-level multiscale approach in Table 2. Indeed, it is even an order of magnitude larger than the approximate $L^2(\Omega_2)$ norm of the function $f$ itself, which is $7.20 \cdot 10^{-3}$. The reason for this bad result is that the one-shot approximation, with its relatively small scale compared to the mesh norm, fails to resolve well even the slowly varying background features – witness the “pepper and salt” nature of the image on the slowly varying part of the right-hand image in Fig. 8. Even the 3-level multiscale approximation is struggling to resolve the slowly varying background.
Table 2 Local errors table when using multiscale approximation only at the last 3 superlocal levels

| Level | \(N\)  | \(\delta_j\) | \(h_j\) | \(\|e_j\|_{L^2(\Omega_2)}\) | \(\kappa_j\) |
|-------|-------|-------------|---------|-----------------|-------|
| 1     | 500   | 1/256       | 0.0018  | 2.75e-02        | 3.24  |
| 2     | 2000  | 1/512       | 0.0009  | 1.49e-02        | 3.37  |
| 3     | 8000  | 1/1024      | 0.0005  | 9.18e-03        | 3.28  |

Before finishing, we discuss principles that might apply in choosing the particular sequence of spherical caps \(\Omega_1 \supseteq \Omega_2 \supseteq \cdots\). The most obvious such principle is that of “user’s choice”: the user may have his or her own reason to focus special attention on some region. The second principle, which is clearly seen in Fig. 5, is that of “discovered features”, where after some number of levels of refinement, an interesting feature begins to emerge, requiring closer study.

A final conclusion might be that the “zooming in” multiscale approximation is successful at all levels. It could be continued indefinitely to smaller and smaller regions, giving a consistent approximation scheme at all levels if the data is available.

Acknowledgments The authors gratefully acknowledge the support of the Australian Research Council. We also thank the reviewers for their constructive comments.

References

1. Adams, R.A.: Sobolev Spaces. Academic Press, New York (1975)
2. Antoine, J.P., Vandergheynst, P.: Wavelets on the 2-sphere: a group-theoretical approach. Appl. Comput. Harmon. Anal. 7, 262–291 (1999)
3. Chen, D., Menegatto, V.A., Sun, X.: A necessary and sufficient condition for strictly positive definite functions on spheres. Proc. Amer. Math. Soc. 131, 2733–2740 (2003)
4. Filbir, F., Themistoclakis, W.: Polynomial approximation on the sphere using scattered data. Math. Nachr. 281, 650–668 (2008)
5. Freeden, W., Gervens, T., Schreiner, M.: Constructive Approximation on the Sphere with Applications to Geomathematics. Oxford University Press, Oxford (1998)
6. Gräf, M., Kunis, S., Potts, D.: On the computation of nonnegative quadrature weights on the sphere. Appl. Comput. Harmon. Anal. 27, 124–132 (2009)
7. Hangelbroek, T., Narcowich, F.J., Ward, J.D.: Polyharmonic and related kernels on manifolds: interpolation and approximation. Found. Comput. Math. 12, 625–670 (2012)
8. Hesse, K., Sloan, I.H., Womersley, R.S.: Numerical integration on the sphere. In: Freeden, W., Nashed, Z., Sonar, T. (eds.) Handbook of Geomathematics, pp. 1187–1220. Springer Verlag (2010)
9. Hubbert, S., Morton, T.M.: A Duchon framework for the sphere. J Approx. Theory 129, 28–57 (2004)
10. Le Gia, Q.T., Narcowich, F.J., Ward, J.D., Wendland, H.: Continuous and discrete least-square approximation by radial basis functions on spheres. J. Approx. Theory 143, 124–133 (2006)
11. Le Gia, Q.T., Mhaskar, H.: Quadrature formulas and localized linear polynomial operators on the sphere. SIAM Numer. Anal. 47, 440–466 (2008)
12. Le Gia, Q.T., Sloan, I.H., Wendland, H.: Multiscale analysis in Sobolev spaces on the sphere. SIAM J. Numer. Anal. 48, 2065–2090 (2010)
13. Le Gia, Q.T., Sloan, I.H., Wendland, H.: Multiscale approximation for functions in arbitrary Sobolev spaces by scaled radial basis functions on the unit sphere. Appl. Comput. Harmon. Anal. 32, 401–412 (2012)
14. Le Gia, Q.T., Sloan, I.H., Wendland, H.: Data compression on the sphere using multiscale radial basis functions. Adv. Comp. Math. 40, 923–943 (2014)
15. Lions, J.L., Magenes, E.: Non-Homogeneous Boundary Value Problems and Applications I. Springer-Verlag, New York (1972)
16. Mhaskar, H.N.: On the representation of smooth functions on the sphere using finitely many bits. Appl. Comput. Harmon. Anal. 18, 215–233 (2005)
17. Mhaskar, H.N., Narcowich, F.J., Ward, J.D.: Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature. Math. Comp. 70, 1113–1130 (2001)
18. Müller, C.: Spherical Harmonics, Volume 17 of Lecture Notes in Mathematics. Springer-Verlag, Berlin (1966)
19. Narcowich, F.J., Petrushev, P., Ward, J.D.: Localized tight frames on spheres. SIAM J. Math Anal. 38, 574–594 (2006)
20. Narcowich, F.J., Sun, X., Ward, J.D.: Approximating power of RBFs and their associated SBFs: a comparison. Adv. Comp. Math. 27, 107–124 (2007)
21. Narcowich, F.J., Ward, J.D.: Scattered data interpolation on spheres: error estimates and locally supported basis functions. SIAM J. Math. Anal. 33, 1393–1410 (2002)
22. Ratcliffe, J.G.: Foundations of Hyperbolic Manifolds. Springer, New York (1994)
23. Saff, E.B., Rakhmanov, E.A., Zhou, Y.M.: Minimal discrete energy on the sphere. Math. Res. Lett. 1, 647–662 (1994)
24. Schoenberg, I.J.: Positive definite function on spheres. Duke Math. J. 9, 96–108 (1942)
25. Schröder, P., Sweldens, W.: Spherical wavelets: efficiently representing functions on the sphere. Computer Graphics Proceedings (SIGGRAPH ’95) (1995)
26. Shewchuk, J.R.: An introduction to the conjugate gradient method without the agonizing pain. School of Computer Science, Carnegie Mellon University, Pittsburgh. http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf
27. Triebel, H.: Interpolation Theory, Function Spaces and Differential Operators. North-Holland, Amsterdam (1978)
28. Wendland, H.: Scattered Data Approximation. Cambridge University Press, Cambridge (2005)
29. Wendland, H.: Multiscale analysis in Sobolev spaces on bounded domains. Numer. Math. 116, 493–517 (2010)
30. Xu, Y., Cheney, E.W.: Strictly positive definite functions on spheres. Proc. Amer. Math Soc. 116, 977–981 (1992)