MULTISCALE ANALYSIS IN SOBOLEV SPACES ON THE SPHERE*

Q. T. LE GIA†, I. H. SLOAN‡, AND H. WENDLAND§

Abstract. We consider a multiscale approximation scheme at scattered sites for functions in Sobolev spaces on the unit sphere $\mathbb{S}^n$. The approximation is constructed using a sequence of scaled, compactly supported radial basis functions restricted to $\mathbb{S}^n$. A convergence theorem for the scheme is proved, and the condition number of the linear system is shown to stay bounded by a constant from level to level, thereby establishing for the first time a mathematical theory for multiscale approximation with scaled versions of a single compactly supported radial basis function at scattered data points.

Key words. multiscale, approximation, interpolation, radial basis functions, unit sphere

AMS subject classifications. 65J10, 46E35

1. Introduction. Approximations based on spherical radial basis functions (SBFs) centered at scattered sites have many applications in the geosciences. Often the 'scale' of the SBF is an important parameter. A common situation is that the SBF Φ is the restriction to the unit sphere $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$ of a radial basis function (RBF) $\Psi(\mathbf{x}) = \psi(|\mathbf{x}|)$ for $\mathbf{x} \in \mathbb{R}^{n+1}$, where Ψ has support of radius 1 and $|\mathbf{x}|$ is the Euclidean distance in $\mathbb{R}^{n+1}$. The scaled RBF Ψ_δ is then defined by $\Psi_\delta(\mathbf{x}) = c_\delta \Psi(\frac{\mathbf{x}}{\delta})$ (which has support radius δ), and the scaled SBF Φ_δ is then the restriction of Ψ_δ to $\mathbb{S}^n$.

In this paper we consider an approximation scheme based on a single underlying RBF Ψ but using a sequence of scales $\delta_1, \delta_2, \ldots$ with limit zero. The motivating idea is that geophysical data typically occurs at many length scales: for example, the topography of the Sahara Desert varies slowly, while that of the Himalayas varies rapidly. Therefore it seems natural to use SBFs of different scales to accommodate the different length scales. But there is a significant theoretical difficulty in dealing with more than one scale at the same time, namely that the associated reproducing kernel Hilbert space (the '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 SBFs on the sphere, or indeed for scaled RBFs on Euclidean regions. New tools seem to be needed.

The approximation scheme studied here follows a path inspired by papers of Schaback (see [13]) and Floater & Iske [1], see also [18], which is intuitively plausible but until now unsubstantiated for either the sphere or Euclidean regions. The essential notion is that one begins with SBFs of large scale, and obtains a first approximation by interpolating at widely separated points; then forms the residual (or error) and finds a correction by interpolating that residual with RBFs of a finer scale and with points closer together; and so on, finally obtaining an approximation which is a sum of SBFs on a sequence with progressively finer scales. Floater and Iske [1] give numerical evidence to support the idea, but no analysis. We note that an analysis is available for a different multilevel scheme proposed by Narcowich, Schaback and Ward [9], in which the SBFs for the successive levels are obtained not by scaling, but rather by repeated convolution of the RBF for the finest scale. That scheme is intrinsically limited to

*This work was supported by the Australian Research Council. The authors acknowledge a helpful suggestion by Yuesheng Xu concerning the presentation of the results

†University of New South Wales, Sydney, Australia (qlegia@unsw.edu.au)

‡University of New South Wales, Sydney, Australia (i.sloan@unsw.edu.au)

§University of Oxford, Oxford, U.K. (holger.wendland@maths.ox.ac.uk)

a finite number of levels, and seems to present difficulties in implementation. The Floater and Iske [1] scheme is also limited to a finite number of levels, since it is formulated in terms of 'thinning' a given data set. The present scheme, by contrast, can be carried to as many levels as desired, and is easy to implement if the underlying RBF Ψ is available in closed form. We remark that, unlike the work of [1] and [9], we make no assumption that the successive point sets at the different scales are nested.

Different approaches to the multiscale problem for the sphere have been explored by Freeden and colleagues, [2, 3]. While the theory given, for example, in [3, Chapter 11] contains an elegant multiresolution analysis, it differs from the present work in a fundamental way, in that it does not include spatial discretization, except in the band-limited situation (in which case convolution integrals can be replaced by quadrature sums that are exact for appropriate spherical polynomials). In the present work, in contrast, spatial discretization is considered from the very beginning, and band-limited (i.e. polynomial) RBFs are excluded, since we require our RBFs to have compact support.

Yet another multilevel approach, this time for Euclidean spaces, is that of Hales and Levesley [5]. In that approach the point sets used at each level differ only by a change of scale, making this approach inherently unsuitable for the case of the sphere.

After a section devoted to preliminaries, we consider in Section 3 interpolation for SBFs of a single scale. In that section and the next we define the scaled SBFs through the behavior of the Fourier coefficients instead of as above, but later, in Section 6, we show that the compactly supported RBFs of Wendland do indeed have Fourier coefficients with the specified behavior.

The multiscale method itself is defined in Section 4. The main theorem, Theorem 4.1, establishes that the L_2 norm of the error, under appropriate assumptions, reduces linearly from scale to scale as the approximation progresses.

In Section 5 the multiscale method is expressed in the framework of multiresolution analysis.

The analysis in Section 6, in which we establish that the compactly supported RBFs associated with particular Sobolev spaces have Fourier coefficients on the sphere that transform in the expected way, turns out to be non-trivial. Theorem 6.2 is perhaps of independent interest, since for the first time it establishes a clear mathematical foundation for the study of scaled SBFs on the sphere.

In Section 7 we discuss the cost and the conditioning of the approximation scheme. It turns out that the condition number of the linear system stays constant from level to level, if we reduce the scale of the SBF and the separation radius of the points (see below) in proportion to each other. For the same reason the number of non-zero entries per row also stays roughly constant, making for efficient implementation.

The cost of the method is essentially the same as that of the computation on the finest scale.

A numerical example in Section 8 demonstrates the effectiveness of the scheme. In particular, Figures 8.3 and 8.4 compare the approximations on the one hand using the multiscale scheme, and on the other hand a "one-shot" approximation in which we proceed directly to an approximation with the finest scale and the smallest mesh norm. The multiscale results are unquestionably of better quality, yet the condition number and the work involved essentially the same.

2. Interpolation using spherical basis functions. Let $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$ be a set of M distinct scattered points on the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$. The mesh norm

h_X of X is defined by

$$h_X := \sup_{\mathbf{p}\in\mathbb{S}^n} \min_{\mathbf{x}\in X} \theta(\mathbf{p}, \mathbf{x}),$$

where $\theta(\mathbf{p}, \mathbf{x}) := \cos^{-1}(\mathbf{p}\cdot\mathbf{x})$ is the geodesic distance of two points $\mathbf{p}$ and $\mathbf{x}$ on the sphere $\mathbb{S}^n$. The separation radius is defined by

$$q_X := \frac{1}{2} \min_{j\neq k} \theta(\mathbf{x}_j, \mathbf{x}_k),$$

and the Euclidean separation radius in $\mathbb{R}^{n+1}$ by

$$\widetilde{q}_X := \frac{1}{2} \min_{j\neq k} |\mathbf{x}_j - \mathbf{x}_k| = \frac{1}{2}\sqrt{2(1-\cos q_X)}. \tag{2.1}$$

The mesh ratio

$$\rho_X := h_X/q_X \geq 1$$

provides a measure of how uniformly points in X are distributed on $\mathbb{S}^n$. In the case of the circle ($n = 1$), $\rho_X = 1$ means that the points are equally spaced. In all other cases we have $\rho_X > 1$. If ρ_X is bounded above by a uniform constant, we say that X is *quasi-uniform*.

Bizonal functions on $\mathbb{S}^n$ are functions that can be represented as $\phi(\mathbf{x}\cdot\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{S}^n$, where $\phi(t)$ is a continuous function on $[-1, 1]$. We shall be concerned exclusively with bizonal kernels of the type

$$\Phi(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}\cdot\mathbf{y}) = \sum_{\ell=0}^{\infty} a_\ell P_\ell(n+1; \mathbf{x}\cdot\mathbf{y}), \quad a_\ell > 0, \quad \sum_{\ell=0}^{\infty} a_\ell < \infty, \tag{2.2}$$

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 [14] and the later work of [19], we know that such a Φ is positive definite on $\mathbb{S}^n$, that is, the matrix $\mathbf{A} := [\Phi(\mathbf{x}_i, \mathbf{x}_j)]_{i,j=1}^{M}$ 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.

In particular, the following interpolation problem admits a unique solution: given a continuous function f on $\mathbb{S}^n$, a positive integer M, and a set of distinct points $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$ on $\mathbb{S}^n$, there exists a sequence of numbers $\{b_j\}_{j=1}^{M}$ such that the function

$$I_X f(\mathbf{x}) = \sum_{j=1}^{M} b_j \Phi(\mathbf{x}, \mathbf{x}_j) \tag{2.3}$$

satisfies the interpolation condition

$$I_X f(\mathbf{x}_k) = f(\mathbf{x}_k), \quad 1 \leq k \leq M.$$

The functions $\Phi(\cdot, \mathbf{x}_j) = \phi(\cdot\cdot\mathbf{x}_j)$, for $1 \leq j \leq M$, are called spherical basis functions (SBFs) in the literature.

For error analysis it is convenient to expand the kernel $\Phi(\mathbf{x}, \mathbf{y})$ into a series of spherical harmonics. A detailed discussion on spherical harmonics can be found in

[8]. In brief, spherical harmonics are the restriction to $\mathbb{S}^n$ of homogeneous polynomials $Y(\mathbf{x})$ in $\mathbb{R}^{n+1}$ which satisfy $\Delta Y(\mathbf{x}) = 0$, 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 H_ℓ, has an orthonormal basis

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

where

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

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

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

where dS is the surface measure of the unit sphere.

Using the addition theorem for spherical harmonics (see, for example, [8, page 18]), we can write

$$\Phi(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \widehat{\phi}(\ell) Y_{\ell k}(\mathbf{x}) \overline{Y_{\ell k}(\mathbf{y})}, \text{ where } \widehat{\phi}(\ell) = \frac{\omega_n}{N(n,\ell)} a_\ell, \tag{2.4}$$

where ω_n is the surface area of $\mathbb{S}^n$. Throughout the paper, we make a further assumption that, $\widehat{\phi}(\ell) \sim (1+\ell)^{-2\sigma}$, i.e. there exist constants $c_1, c_2 > 0$ and $\sigma > n/2$ such that

$$c_1(1+\ell)^{-2\sigma} \leq \widehat{\phi}(\ell) \leq c_2(1+\ell)^{-2\sigma}, \quad \ell \geq 0. \tag{2.5}$$

The *native space* induced by Φ is defined to be

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

where $\mathcal{D}'(\mathbb{S}^n)$ is the set of all tempered distributions defined on $\mathbb{S}^n$. Alternatively, N_Φ is the completion of $\operatorname{span}\{\Phi(\cdot, \mathbf{x}) : \mathbf{x} \in \mathbb{S}^n\}$ with respect to the norm $\|\cdot\|_\Phi$.

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

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

For $\sigma = 0$ the space is just $L_2(\mathbb{S}^n)$.

Under the condition (2.5), the norms $\|\cdot\|_\Phi$ and $\|\cdot\|_{H^\sigma}$ are equivalent, since

$$c_1^{1/2} \|f\|_\Phi \leq \|f\|_{H^\sigma} \leq c_2^{1/2} \|f\|_\Phi. \tag{2.6}$$

Because $\sigma > n/2$ it follows from the embedding theorem that functions in $H^\sigma(\mathbb{S}^n)$ are continuous.

3. Interpolation using scaled SBFs. For a given $\delta > 0$, we define the scaled version Φ_δ of the kernel Φ (with $\Phi_1 = \Phi$) by

$$\Phi_\delta(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \widehat{\phi_\delta}(\ell) Y_{\ell k}(\mathbf{x}) \overline{Y_{\ell k}(\mathbf{y})} \tag{3.1}$$

in which the Fourier coefficients satisfy the following asymptotic condition:

$$c_1(1 + \delta\ell)^{-2\sigma} \le \widehat{\phi_\delta}(\ell) \le c_2(1 + \delta\ell)^{-2\sigma}, \quad \ell \ge 0, \tag{3.2}$$

with the coefficients c_1 and c_2 from (2.5) possibly relaxed so that (3.2) holds for all $\delta \le 1$. In Section 6 we will show that the Fourier coefficients of the SBFs defined by restricting scaled Wendland's functions $\Psi_\delta(\mathbf{x}) = \delta^{-n}\Psi(\mathbf{x}/\delta)$ onto $\mathbb{S}^n$ behave as given by (3.1) and (3.2).

For a given function $f \in C(\mathbb{S}^n)$, the interpolant $I_{X,\delta} f$ of f is defined by

$$I_{X,\delta} f(\mathbf{x}) := \sum_{j=1}^{M} b_j \Phi_\delta(\mathbf{x}, \mathbf{x}_j) \text{ such that } I_{X,\delta} f(\mathbf{x}_k) = f(\mathbf{x}_k) \text{ for all } \mathbf{x}_k \in X. \tag{3.3}$$

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

$$\|f\|_{\Phi_\delta} = \left(\sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{|\widehat{f}_{\ell k}|^2}{\widehat{\phi_\delta}(\ell)} \right)^{1/2}. \tag{3.4}$$

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

Lemma 3.1. *For $\delta \le 1$ and all $g \in H^\sigma(\mathbb{S}^n)$ we have*

$$\|g\|_{\Phi_\delta} \le c_1^{-1/2} \|g\|_{H^\sigma} \le (c_2/c_1)^{1/2} \|g\|_\Phi, \quad \|g\|_\Phi \le (c_2/c_1)^{1/2} \delta^{-\sigma} \|g\|_{\Phi_\delta}.$$

Proof. Since $\delta \le 1$, on using (3.2) we have

$$\begin{aligned} \|g\|_{\Phi_\delta}^2 &\le \frac{1}{c_1} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} |\widehat{g}_{\ell k}|^2 (1 + \delta\ell)^{2\sigma} \le \frac{1}{c_1} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} |\widehat{g}_{\ell k}|^2 (1 + \ell)^{2\sigma} \\ &=: \frac{1}{c_1} \|g\|_{H^\sigma}^2 \le \frac{c_2}{c_1} \|g\|_\Phi^2. \end{aligned}$$

We also have $(1 + \delta\ell) = \delta(1/\delta + \ell) \ge \delta(1 + \ell)$. Hence, $(1 + \ell)^{2\sigma} \le \delta^{-2\sigma}(1 + \delta\ell)^{2\sigma}$, and

$$\begin{aligned} \|g\|_\Phi^2 &\le \frac{1}{c_1} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} |\widehat{g}_{\ell k}|^2 (1 + \ell)^{2\sigma} \\ &\le \frac{1}{c_1} \delta^{-2\sigma} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} |\widehat{g}_{\ell k}|^2 (1 + \delta\ell)^{2\sigma} \le \delta^{-2\sigma} \frac{c_2}{c_1} \|g\|_{\Phi_\delta}^2. \end{aligned}$$

$\square$

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 $c_3 > 0$ such that, for sufficiently small h_X and all $\delta \leq 1$,*

$$\|g\|_{L_2} \leq c_3 \left(\frac{h_X}{\delta}\right)^\sigma \|g\|_{\Phi_\delta}. \tag{3.5}$$

If the interpolant of $f \in H^\sigma(\mathbb{S}^n)$ is defined as in (3.3) *(with $\delta \leq 1$), then for sufficiently small h_X*

$$\|f - I_{X,\delta}f\|_{L_2} \leq c_3 \left(\frac{h_X}{\delta}\right)^\sigma \|f\|_{\Phi_\delta}. \tag{3.6}$$

Proof. By [7, Theorem 3.3], because $g|_X = 0$, there is a constant c such that

$$\|g\|_{L_2} \leq c h_X^\sigma \|g\|_{H^\sigma}. \tag{3.7}$$

Using Lemma 3.1 we have

$$\|g\|_{H^\sigma} \leq c_2^{1/2}\|g\|_\Phi \leq c_2^{1/2}(c_2/c_1)^{1/2}\delta^{-\sigma}\|g\|_{\Phi_\delta}. \tag{3.8}$$

Inequality (3.5) then follows from (3.7) and (3.8) with the constant $c_3 = cc_2/c_1^{1/2}$. Now take $g = f - I_{X,\delta}f$. Since $I_{X,\delta}f$ is the orthogonal projection of f into $\text{span}\{\Phi_\delta(\cdot, \mathbf{x}_j) : \mathbf{x}_j \in X\}$ with respect to the $\|\cdot\|_{\Phi_\delta}$ norm, it follows that

$$\|g\|_{\Phi_\delta} = \|f - I_{X,\delta}f\|_{\Phi_\delta} \leq \|f\|_{\Phi_\delta}.$$

Hence inequality (3.6) follows from (3.5). □

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

We select an SBF Φ whose Fourier coefficients satisfy (2.5) for some $\sigma > n/2$. For every $j = 1, 2, \ldots$ we choose δ_j and from it derive the scaled SBF $\Phi_j = \Phi_{\delta_j}$, where δ_j is a scaling parameter proportional to h_j, i.e. $\delta_j = \nu h_j$ for some $\nu > 0$. We denote the interpolation operator for the j-th scaled SBF Φ_j by

$$I_{X_j,\delta_j}f = \sum_{\mathbf{x}\in X_j} b_{\mathbf{x}}(f)\Phi_j(\cdot, \mathbf{x}), \qquad I_{X_j,\delta_j}f(\mathbf{x}) = f(\mathbf{x}) \quad \text{for } \mathbf{x} \in X_j.$$

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 $j = 1, 2, \ldots,$

$$\begin{aligned} s_j &= I_{X_j,\delta_j}e_{j-1} \\ f_j &= f_{j-1} + s_j, \\ e_j &= e_{j-1} - s_j. \end{aligned}$$

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 our main convergence result.

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 $c\mu h_j \le h_{j+1} \le \mu h_j$ 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 (3.2) with scale factor $\delta_j = \nu h_j$ and ν satisfying $1/h_1 \ge \nu \ge \beta/\mu$ 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*

$$\|e_j\|_{\Phi_{\delta_{j+1}}} \le \alpha \|e_{j-1}\|_{\Phi_{\delta_j}} \quad \text{for } j = 1, 2, \ldots, \tag{4.1}$$

with $\alpha = C\mu^\sigma$, and hence there exists $c_4 > 0$ such that

$$\|f - f_k\|_{L_2} \le c_4 \alpha^k \|f\|_{H^\sigma} \text{ for } k = 1, 2, \ldots.$$

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

Proof. Noting that $\delta_{k+1} \le \delta_1 = \nu h_1 \le 1$ and $e_k|_{X_k} = 0$, on using Theorem 3.2 with $\delta = \delta_{k+1}$ we obtain

$$\|f - f_k\|_{L_2} = \|e_k\|_{L_2} \le c_3 h_k^\sigma \delta_{k+1}^{-\sigma} \|e_k\|_{\Phi_{k+1}} \le c_3 (c\beta)^{-\sigma} \|e_k\|_{\Phi_{k+1}}, \tag{4.2}$$

since $h_k/\delta_{k+1} = h_k/(\nu h_{k+1}) \le 1/(c\mu\nu) \le 1/(c\beta)$.

The remaining part of the proof is devoted to proving the recurrence (4.1). Fixing $j \ge 1$ and writing Φ_{δ_j} as Φ_j, we have, from (3.2) and (3.4),

$$\begin{aligned}\|e_j\|_{\Phi_{j+1}}^2 &\le \frac{1}{c_1} \sum_{\ell \le 1/\delta_{j+1}} \sum_{k=1}^{N(n,\ell)} |\widehat{e_j}_{\ell k}|^2 (1 + \delta_{j+1}\ell)^{2\sigma} + \frac{1}{c_1} \sum_{\ell > 1/\delta_{j+1}} \sum_{k=1}^{N(n,\ell)} |\widehat{e_j}_{\ell k}|^2 (1 + \delta_{j+1}\ell)^{2\sigma} \\ &=: \frac{1}{c_1}(S_1 + S_2).\end{aligned}$$

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

$$S_1 \le \sum_{\ell \le 1/\delta_{j+1}} \sum_{k=1}^{N(n,\ell)} |\widehat{e_j}_{\ell,k}|^2 2^{2\sigma} \le 2^{2\sigma} \|e_j\|_{L_2}^2. \tag{4.3}$$

Since $e_j = e_{j-1} - I_{X_j,\delta_j} e_{j-1}$, by the last part of Theorem 3.2 we have

$$\|e_j\|_{L_2} = \|e_{j-1} - I_{X_j,\delta_j} e_{j-1}\|_{L_2} \le c_3 \left(\frac{h_j}{\delta_j}\right)^\sigma \|e_{j-1}\|_{\Phi_j} = c_3 \nu^{-\sigma} \|e_{j-1}\|_{\Phi_j}.$$

Therefore (4.3) gives

$$S_1 \le c_3^2 2^{2\sigma} \nu^{-2\sigma} \|e_{j-1}\|^2_{\Phi_j}.$$

For the sum S_2, since $\delta_{j+1}\ell > 1$ we have, using $\delta_{j+1}/\delta_j = h_{j+1}/h_j \le \mu$,

$$(1 + \delta_{j+1}\ell)^{2\sigma} < (2\delta_{j+1}\ell)^{2\sigma} \le (2\mu\delta_j\ell)^{2\sigma} \le (2\mu)^{2\sigma}(1 + \delta_j\ell)^{2\sigma}.$$

So with the aid of (3.2) and (3.4),

$$S_2 \le \sum_{\ell > 1/\delta_{j+1}} \sum_{k=1}^{N(n,\ell)} (2\mu)^{2\sigma} |\widehat{e_j}_{\ell k}|^2 (1 + \delta_j \ell)^{2\sigma} \le c_2 (2\mu)^{2\sigma} \|e_j\|^2_{\Phi_j} \le c_2 (2\mu)^{2\sigma} \|e_{j-1}\|^2_{\Phi_j},$$

where the key last step follows because $e_j = e_{j-1} - I_{X_j,\delta_j} e_{j-1}$ is an orthogonal projection of e_{j-1} on a subspace of N_{Φ_j}. (The subspace is the orthogonal complement of span $\{\Phi_j(\cdot, \mathbf{x}) : \mathbf{x} \in X_j\}$ in N_{Φ_j}.) Hence, since $\nu \ge \beta/\mu$,

$$\|e_j\|^2_{\Phi_{j+1}} \le \frac{1}{c_1}(c_3^2 \nu^{-2\sigma} + c_2 \mu^{2\sigma}) 2^{2\sigma} \|e_{j-1}\|^2_{\Phi_j} \le \frac{1}{c_1}(c_3^2 \beta^{-2\sigma} + c_2) 2^{2\sigma} \mu^{2\sigma} \|e_{j-1}\|^2_{\Phi_j},$$

from which follows

$$\|e_j\|_{\Phi_{j+1}} \le C\mu^\sigma \|e_{j-1}\|_{\Phi_j} = \alpha \|e_{j-1}\|_{\Phi_j}$$

with $C = 2^\sigma (c_3^2 \beta^{-2\sigma} + c_2)^{1/2}/\sqrt{c_1}$, thus proving (4.1).

Using (4.2) and repeating (4.1) k times, we obtain, with $c_5 = c_3(c\beta)^{-\sigma}$,

$$\|f - f_k\|_{L_2} \le c_5 \|e_k\|_{\Phi_{k+1}} \le c_5 \alpha \|e_{k-1}\|_{\Phi_k} \le c_5 \alpha^2 \|e_{k-2}\|_{\Phi_{k-1}} \le \ldots \le c_5 \alpha^k \|f\|_{\Phi_1},$$

and with the help of Lemma 3.1 the theorem now follows, with $c_4 = c_1^{-1/2} c_5 = c_1^{-1/2} c_3 (c\beta)^{-\sigma}$. □

We remark that the condition $\nu \ge \beta/\mu$ is needed in the analysis because in (4.2) we chose to introduce the seemingly unnatural $\|\cdot\|_{\Phi_{k+1}}$ norm of e_k. In turn, to prove (4.1) we bounded $\|e_j\|_{\Phi_{j+1}}$ (in the course of bounding S_2) by $(2\mu)^\sigma$ times the $\|\cdot\|_{\Phi_j}$ norm of e_j. We needed to reach the latter norm because it is only for that norm that we know the crucial fact that the norm of e_j is less than that of e_{j-1}.

5. Multiresolution analysis. For a given RBF Ψ 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 4 can be put into the following multiresolution framework.

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

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

and

$$V_j := \operatorname{span}\{\Phi_{\delta_i}(\cdot, \mathbf{x}) : \mathbf{x} \in X_i, \quad i \le j\},$$

where Φ_δ is as in (3.1). Thus

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

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,

$$V_j = \bigoplus_{1\le i\le j} W_i, \quad j \ge 1$$

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

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

That the sum is direct follows under appropriate conditions from the following lemma.

LEMMA 5.1. *Let Ψ be a compactly supported RBF as in [18] and let Φ_{δ_i} for $i = 1, 2, \ldots$ be scaled SBFs constructed as in Section 4 where $\delta_1, \delta_2, \ldots$ are distinct scales with $\delta_i \le 1$. Let $X_i = \{\mathbf{x}_{i,1}, \ldots, \mathbf{x}_{i,N_i}\} \subset \mathbb{S}^n$ for $i \ge 1$ be a set of N_i distinct points. Then for $j \ge 1$ and $a_{i,k} \in \mathbb{R}$*

$$\sum_{i=1}^{j}\sum_{k=1}^{N_i} a_{i,k}\Phi_{\delta_i}(\cdot, \mathbf{x}_{i,k}) = 0 \textit{ implies } a_{i,k} = 0 \textit{ for } i = 1, \ldots, j; k = 1, \ldots, N_i. \tag{5.1}$$

Proof. In (5.1), either all $a_{i,k}$ are zero, or one SBF, say $\Phi_{\delta_1}(\cdot, \mathbf{x}_{1,1})$, is a linear combination of the others,

$$\Phi_{\delta_1}(\cdot, \mathbf{x}_{1,1}) = \sum_{\substack{i=1 \\ (i,k)\ne(1,1)}}^{j}\sum_{k=1}^{N_i} b_{i,k}\Phi_{\delta_i}(\cdot, \mathbf{x}_{i,k}). \tag{5.2}$$

The SBF $\Phi_{\delta_i}(\mathbf{x}_{i,k}, \cdot)$ is singular (in the sense of a discontinuity in an appropriate derivative with respect to $\mathbf{x}$) on the $n-1$ dimensional sphere

$$S_{i,k} := \{\mathbf{x} \in \mathbb{S}^n : |\mathbf{x} - \mathbf{x}_{i,k}| = \delta_i\}$$

and at $\mathbf{x}_{i,k}$, and is infinitely smooth elsewhere. For $(i,k) \ne (1,1)$, $S_{i,k} \cap S_{1,1}$ is an $n-2$ dimensional sphere, hence

$$\left(\left(\bigcup_{(i,k)\ne(1,1)} S_{i,k}\right) \bigcup \{\mathbf{x}_{i,k} : 1 \le i \le j, 1 \le k \le N_i\}\right) \cap S_{1,1}$$

is of dimension $n-2$ and so cannot contain $S_{1,1}$. Therefore, there exists a point $\mathbf{x} \in S_{1,1}$ at which the left hand side of (5.2) is singular while the right hand side is infinitely smooth, which is a contradiction. Thus the sum in the lemma can vanish only if $a_{i,k} = 0$ for all i and k. □

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

THEOREM 5.2. *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 $c\mu h_j \le h_{j+1} \le \mu h_j$ 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 (3.2) *with scale factor $\delta_j = \nu h_j$ and ν satisfying $1/h_1 \geq \nu \geq \beta/\mu$ with a fixed $\beta > 0$. For all μ sufficiently small.*

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

Proof. Theorem 4.1 shows that $f \in H^\sigma(\mathbb{S}^n)$ can be approximated arbitrarily closely in the L_2 norm by linear combinations of functions in V_k, if k is sufficiently large. Since H^σ is dense in L_2, the result extends to arbitrary functions in L_2, proving the theorem. □

The algorithm presented in Section 4 can be generalized in the following way. Set $f_0 = 0$ and $e_0 = f$. For $j = 1, 2, \ldots$ compute

$$s_j = P_j(e_{j-1}) \in W_j,$$

where $P_j : H^\sigma(\mathbb{S}^n) \to W_j$ is a linear projection. Then we update the approximation and the residual by

$$\begin{aligned} f_j &= f_{j-1} + s_j \in V_j, \\ e_j &= e_{j-1} - s_j. \end{aligned}$$

Theorem 4.1 gives a convergence proof for the special case $P_j(g) = I_{X_j,\delta_j}(g)$. Note that the proof does not extend to general projections P_j because we made essential use of the fact that interpolation is an orthogonal projection in the corresponding native space.

6. Scaled SBFs defined from scaled RBFs. In this section we will show that SBFs constructed by scaling compactly supported RBFs $\Psi : \mathbb{R}^{n+1} \to \mathbb{R}$ have Fourier coefficients $\widehat{\phi}_\delta(\ell)$ that satisfy the decay condition (3.2).

We start with a compactly supported radial basis functions $\Psi(\mathbf{x}) = \psi(|\mathbf{x}|)$, typified by Wendland's radial basis functions [16]. We will assume that this radial basis function has $H^\tau(\mathbb{R}^{n+1})$ as its native space, which is equivalent to assuming that its $(n+1)$-variate Fourier transform $\widehat{\Psi}$ decays like

$$\widetilde{c}_1(1 + |\omega|^2)^{-\tau} \leq \widehat{\Psi}(\omega) \leq \widetilde{c}_2(1 + |\omega|^2)^{-\tau}, \qquad \omega \in \mathbb{R}^{n+1}. \tag{6.1}$$

Since $\Psi(\mathbf{x}) = \psi(|\mathbf{x}|)$ is radial, the Fourier transform of Ψ can be computed via a Hankel transform:

$$\begin{aligned} \widehat{\Psi}(\omega) &:= \int_{\mathbb{R}^{n+1}} \Psi(\mathbf{x}) e^{-i\mathbf{x}\cdot\omega} d\mathbf{x} \\ &= (2\pi)^{(n+1)/2} r^{-(n-1)/2} \int_0^\infty \psi(t) t^{(n+1)/2} J_{(n-1)/2}(rt) dt \\ &= (2\pi)^{(n+1)/2} r^{-(n+1)} \int_0^r \psi(t/r) t^{(n+1)/2} J_{(n-1)/2}(t) dt \\ &=: \mathcal{F}\psi(r), \end{aligned}$$

where $r = |\omega|$ and J_ν denotes the Bessel function of the first kind of order ν.

Then, we restrict the function $(\mathbf{x}, \mathbf{y}) \mapsto \Psi(\mathbf{x} - \mathbf{y})$ to the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$. Using $|\mathbf{x} - \mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2\mathbf{x} \cdot \mathbf{y} = 2 - 2\mathbf{x} \cdot \mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in \mathbb{S}^n$ we can conclude that the resulting function is indeed zonal:

$$\Psi(\mathbf{x} - \mathbf{y}) = \psi(|\mathbf{x} - \mathbf{y}|) = \psi(\sqrt{2 - 2\mathbf{x} \cdot \mathbf{y}}),$$

and hence we can define the kernel

$$\Phi(\mathbf{x},\mathbf{y}) := \Psi(\mathbf{x}-\mathbf{y})|_{\mathbb{S}^n}, \quad \mathbf{x},\mathbf{y} \in \mathbb{S}^n. \tag{6.2}$$

The connection between the $(n+1)$-variate Fourier transform $\widehat{\Psi}$ and the Fourier coefficients $\widehat{\phi}(\ell)$ is given by the following result.

LEMMA 6.1. *[11, Theorem 4.1] Let the kernel Φ be defined by restricting an RBF $\Psi(\mathbf{x}) = \psi(|\mathbf{x}|)$ to the sphere $\mathbb{S}^n$ as in* (6.2). *Then*

$$\widehat{\phi}(\ell) = \int_0^\infty r\mathcal{F}\psi(r)J_\nu^2(r)dr, \tag{6.3}$$

where $\nu = \ell + (n-1)/2$ and $J_\nu(r)$ is the Bessel function of the first kind.

A consequence of this lemma is that if Ψ generates the Sobolev space $H^\tau(\mathbb{R}^{n+1})$ with $\tau > (n+1)/2$, i.e. its Fourier transform satisfies (6.1), then Φ generates $H^{\tau-1/2}(\mathbb{S}^n)$, i.e. its Fourier coefficients satisfy (2.5) with $\sigma = \tau - 1/2$, see [10, Proposition 4.1].

The next step is to introduce scaling. For a given $\delta > 0$ and $\Psi : \mathbb{R}^{n+1} \to \mathbb{R}$ we define the scaled function $\Psi_\delta : \mathbb{R}^{n+1} \to \mathbb{R}$

$$\Psi_\delta(\mathbf{x}) = \delta^{-n}\Psi(\mathbf{x}/\delta). \tag{6.4}$$

Note that this scaling preserves volumes in the n-dimensional space $\mathbb{R}^n$, not in $\mathbb{R}^{n+1}$, otherwise the scaling factor would be δ^{-n-1}. Consequently, the Fourier transform of Ψ_δ scales as

$$\widehat{\Psi_\delta}(\omega) = \delta\widehat{\Psi}(\delta\omega). \tag{6.5}$$

This means that if Ψ generates $H^\tau(\mathbb{R}^{n+1})$ then so does Ψ_δ, and it follows from (6.1) that

$$\widetilde{c_1}\delta(1+\delta^2|\omega|^2)^{-\tau} \le \widehat{\Psi_\delta}(\omega) \le \widetilde{c_2}\delta(1+\delta^2|\omega|^2)^{-\tau}.$$

The scaled radial basis function Ψ_δ on $\mathbb{R}^{n+1}$ induces the corresponding scaled SBF Φ_δ on the unit sphere $\mathbb{S}^n$,

$$\Phi_\delta(\mathbf{x},\mathbf{y}) = \Psi_\delta(\mathbf{x}-\mathbf{y}) = \delta^{-n}\psi(|\mathbf{x}-\mathbf{y}|/\delta) = \delta^{-n}\psi(\delta^{-1}\sqrt{2-2\mathbf{x}\cdot\mathbf{y}}), \quad \mathbf{x},\mathbf{y} \in \mathbb{S}^n, \tag{6.6}$$

which has the Fourier coefficients

$$\widehat{\phi_\delta}(\ell) = \int_0^\infty t\mathcal{F}\psi_\delta(t)J_\nu^2(t)dt = \int_0^\infty \delta t\mathcal{F}\psi(\delta t)J_\nu^2(t)dt. \tag{6.7}$$

We will spend the rest of this section in proving that these Fourier coefficients possess the corresponding decay property (3.2), which is stated as the following theorem.

THEOREM 6.2. *If the radial function $\Psi = \psi(|\cdot|)$ has compact support in the unit ball and a Fourier transform satisfying the condition* (6.1) *with $\tau > (n+1)/2$ then there are constants C_1 and C_2 depending only on τ and n such that the associated scaled spherical basis function Φ_δ has Fourier coefficients satisfying*

$$C_1(1+\delta\ell)^{-2\tau+1} \le \widehat{\phi_\delta}(\ell) \le C_2(1+\delta\ell)^{-2\tau+1}.$$

To achieve the goal, we first have to analyse the proof of [10, Proposition 4.1] carefully. To this end, we introduce the integral

$$A(\nu,\tau)=\int_0^\infty t^{-2\tau+1}J_\nu^2(t)dt,\quad \nu>\tau-1, \tag{6.8}$$

which can be explicitly computed (cf [15, Formula (2) in 13.41] and [10, Formula (4.9)]) as

$$A(\nu,\tau)=\frac{\Gamma(2\tau-1)\Gamma(\nu-\tau+1)}{2^{2\tau-1}\Gamma^2(\tau)\Gamma(\tau+\nu)}=\frac{\Gamma(2\tau-1)}{2^{2\tau-1}\Gamma^2(\tau)}\frac{(\nu-\tau)\Gamma(\nu-\tau)}{\Gamma(\nu+\tau)}. \tag{6.9}$$

We will also make frequent use of Stirling's formula for the Γ-function in the form

$$\sqrt{2\pi}x^{x-1/2}e^{-x}\le\Gamma(x)\le\sqrt{2\pi}x^{x-1/2}e^{-x}e^{1/(12x)} \tag{6.10}$$

and of the fact that

$$\frac{2}{3}|x|\le|\log(1+x)|\le\frac{4}{3}|x| \tag{6.11}$$

for $|x|\le 1/4$. We start with the upper bound.

PROPOSITION 6.3. *If the radial function $\Psi=\psi(|\cdot|)$ has compact support in the unit ball and a Fourier transform satisfying $\widehat{\Psi}(\omega)\le\widetilde{c_2}(1+|\omega|^2)^{-\tau}$ with $\tau>(n+1)/2$ then there is a constant $C_2>0$ depending only on τ and n such that the associated scaled spherical basis function Φ_δ has Fourier coefficients satisfying*

$$\widehat{\phi_\delta}(\ell)\le C_2(1+\delta\ell)^{-2\tau+1}$$

for all $\ell\in\mathbb{N}_0$ and all $\delta\in(0,1]$.

Proof. In a first step we will show that there is a constant $c_4>0$ and an integer $\ell_0\in\mathbb{N}$ such that

$$\widehat{\phi_\delta}(\ell)\le c_4(\delta\ell)^{-2\tau+1} \tag{6.12}$$

for all $\ell\ge\ell_0$ and all $\delta>0$. The assumption on the decay of $\widehat{\Psi}$ together with formula (6.7) lead to

$$\widehat{\phi_\delta}(\ell)\le\widetilde{c_2}\int_0^\infty\frac{\delta tJ_\nu^2(t)}{(1+(\delta t)^2)^\tau}dt.$$

Since $1+(\delta t)^2>(\delta t)^2$ we can conclude that

$$\widehat{\phi_\delta}(\ell)\le\widetilde{c_2}\delta^{-2\tau+1}\int_0^\infty t^{-2\tau+1}J_\nu^2(t)dt=\widetilde{c_2}\delta^{-2\tau+1}A(\nu,\tau). \tag{6.13}$$

Using the explicit formula (6.9) together with Stirling's formula (6.10) and the fact that $e^{1/(12(\nu-\tau))}\le e^{1/(12\tau)}$ provided that $\nu\ge2\tau$, we find for $\nu\ge2\tau$ that

$$\begin{aligned}
A(\nu,\tau)&=\frac{\Gamma(2\tau-1)(\nu-\tau)\Gamma(\nu-\tau)}{2^{2\tau-1}\Gamma^2(\tau)\Gamma(\nu+\tau)}\\
&\le\frac{\Gamma(2\tau-1)}{2^{2\tau-1}\Gamma^2(\tau)}(\nu-\tau)\frac{\sqrt{2\pi}(\nu-\tau)^{\nu-\tau-1/2}e^{-(\nu-\tau)}e^{1/(12(\nu-\tau))}}{\sqrt{2\pi}(\nu+\tau)^{\nu+\tau-1/2}e^{-(\nu+\tau)}}\\
&\le\frac{\Gamma(2\tau-1)e^{2\tau}e^{1/(12\tau)}}{2^{2\tau-1}\Gamma^2(\tau)}\frac{(\nu-\tau)^{\nu-(\tau-1/2)}}{(\nu+\tau)^{\nu+\tau-1/2}}\\
&\le\frac{\Gamma(2\tau-1)e^{2\tau}e^{1/(12\tau)}}{2^{2\tau-1}\Gamma^2(\tau)}\frac{1}{(\nu^2-\tau^2)^{\tau-1/2}}\left(\frac{1-\frac{\tau}{\nu}}{1+\frac{\tau}{\nu}}\right)^\nu.
\end{aligned}$$

For $\nu \geq 2\tau$ we have also $\nu^2 - \tau^2 \geq \frac{3}{4}\nu^2$ and hence

$$\frac{1}{(\nu^2-\tau^2)^{\tau-1/2}} \leq \frac{2^{2\tau-1}}{3^{\tau-1/2}}\nu^{-2\tau+1}.$$

Assuming now $\nu \geq 4\tau$, we see that (6.11) yields

$$\begin{aligned}
\log\left(\frac{1-\frac{\tau}{\nu}}{1+\frac{\tau}{\nu}}\right)^{\nu} &= \nu\left[\log\left(1-\frac{\tau}{\nu}\right) - \log\left(1+\frac{\tau}{\nu}\right)\right] \\
&= -\nu\left[\left|\log\left(1-\frac{\tau}{\nu}\right)\right| + \left|\log\left(1+\frac{\tau}{\nu}\right)\right|\right] \\
&\leq -\nu \cdot 2 \cdot \frac{2}{3}\frac{\tau}{\nu} = -\frac{4}{3}\tau.
\end{aligned}$$

Taking this all together gives for $\nu \geq 4\tau$ the upper bound

$$A(\nu,\tau) \leq \frac{\Gamma(2\tau-1)e^{2\tau/3}e^{1/(12\tau)}}{3^{\tau-1/2}\Gamma^2(\tau)}\nu^{-2\tau+1},$$

which is consistent with the asymptotic result from [10, eq. (4.10)] for $A(\nu,\tau)$. This upper bound together with (6.13) and $\nu = \ell + \frac{n-1}{2} \geq \ell$ establishes our first step (6.12).

In the second step, we show that there is a constant $c_5 > 0$ such that

$$\widehat{\phi_\delta}(\ell) \leq c_5 \tag{6.14}$$

for all $\ell \in \mathbb{N}_0$ and all $\delta > 0$.

Using formula (3.1), the addition theorem for spherical harmonics and the following orthogonality

$$\int_{-1}^{1} P_\ell(n+1;t)P_{\ell'}(n+1;t)(1-t^2)^{\frac{n-2}{2}}dt = \frac{\omega_n}{\omega_{n-1}N(n,\ell)}\delta_{\ell,\ell'}, \tag{6.15}$$

we have

$$\widehat{\phi_\delta}(\ell) = \omega_{n-1}\int_{-1}^{1}\phi_\delta(t)P_\ell(n+1;t)(1-t^2)^{\frac{n-2}{2}}dt, \tag{6.16}$$

in which $\phi_\delta(\mathbf{x}\cdot\mathbf{y}) = \Phi_\delta(\mathbf{x},\mathbf{y})$. Since $\Psi(\mathbf{x}) = \psi(|\mathbf{x}|)$ is a compactly supported function with support in the unit ball, we have $\psi(r) = 0$ for $r \geq 1$. Hence, with the definition (6.6), the integral (6.16) is reduced to

$$\widehat{\phi_\delta}(\ell) = \delta^{-n}\omega_{n-1}\int_{1-\delta^2/2}^{1}\psi\left(\frac{\sqrt{2-2t}}{\delta}\right)P_\ell(n+1;t)(1-t^2)^{\frac{n-2}{2}}dt. \tag{6.17}$$

Using the fact that we integrate over the range $1-\delta^2/2 \leq t \leq 1$, we have

$$(1-t^2) = (1+t)(1-t) \leq 2(1-t) \leq 2\frac{\delta^2}{2} = \delta^2.$$

Since the Legendre polynomials satisfy $|P_\ell(n+1;t)| \leq 1$ for $t \in [-1,1]$, we can bound the integral in (6.17) by

$$\begin{aligned}
\widehat{\phi_\delta}(\ell) &\leq \omega_{n-1}\delta^{-2}\int_{1-\delta^2/2}^{1}|\psi(\delta^{-1}\sqrt{2-2t})|dt \\
&= \omega_{n-1}\int_{1/2}^{1}|\psi(\sqrt{2-2u})|du,
\end{aligned}$$

using the substitution $(1-t) = \delta^2(1-u)$. Thus (6.14) holds for all $\ell \in \mathbb{N}_0$.

In the final step we use both of the above bounds. If $\delta\ell \geq 1$ then we clearly have

$$(1+\delta\ell)^{2\tau-1} \leq 2^{2\tau-1}(\delta\ell)^{2\tau-1},$$

which together with (6.12) gives, for $\ell \geq \max\{\ell_0, 1/\delta\}$,

$$\widehat{\phi_\delta}(\ell) \leq 2^{2\tau-1} c_4 (1+\delta\ell)^{-2\tau+1}$$

For $\ell < \max\{\ell_0, 1/\delta\}$ we can use (6.14) and $\delta \leq 1$ to derive

$$\begin{aligned}
\widehat{\phi_\delta}(\ell) &\leq c_5 = c_5(1+\delta\ell)^{2\tau-1}(1+\delta\ell)^{-2\tau+1} \\
&\leq c_5 \max_{1\leq k \leq \max\{\ell_0, 1/\delta\}} (1+\delta k)^{2\tau-1}(1+\delta\ell)^{-2\tau+1} \\
&\leq c_5 \max\{(1+\delta\ell_0)^{2\tau-1}, 2^{2\tau-1}\}(1+\delta\ell)^{-2\tau+1} \\
&\leq c_5 \max\{(1+\ell_0)^{2\tau-1}, 2^{2\tau-1}\}(1+\delta\ell)^{-2\tau+1}
\end{aligned}$$

This give the stated upper bound with

$$C_2 = \max\{2^{2\tau-1} c_4, (1+\ell_0)^{2\tau-1} c_5\}.$$

□

We will now address the lower bound. However since this is more technical we break the actual proof up into several lemmas. We start by looking at the case where $\delta\nu$ is large. Here, we follow again [10] with special emphasis on the dependence on δ.

LEMMA 6.4. *Assume that the radial function $\Psi = \psi(|\cdot|)$ has a Fourier transform satisfying $\widehat{\Psi}(\omega) \geq \widetilde{c_1}(1+|\omega|^2)^{-\tau}$ with $\tau > (n+1)/2$. Then there is a constant $c_6 > 0$ depending only on τ such that, for $\nu = \ell + \frac{n-1}{2} \geq 1/\delta$ and $\delta \leq 1$, we have*

$$\widehat{\phi_\delta}(\ell) \geq c_6 (1+\delta\ell)^{-2\tau+1}.$$

Proof. Let us introduce a constant $c > 0$, which we specify later on. The assumptions together with (6.7) allow us to bound the Fourier coefficients by

$$\widehat{\phi_\delta}(\ell) \geq \widetilde{c_1} \int_0^\infty \frac{\delta t}{(1+\delta^2 t^2)^\tau} J_\nu^2(t)dt > \widetilde{c_1} \int_{c/\delta}^\infty \frac{\delta t}{(1+\delta^2 t^2)^\tau} J_\nu^2(t)dt,$$

where we used the constant $c > 0$ in the lower limit of the integral. In the latter integral we now have $t > c/\delta$, and hence, using $1 < \delta^2 t^2 / c^2$,

$$\frac{1}{(1+\delta^2 t^2)^\tau} > (\delta t)^{-2\tau} \left(\frac{c^2}{1+c^2}\right)^\tau,$$

and in turn

$$\widehat{\phi_\delta}(\ell) > \widetilde{c_1} \left(\frac{c^2}{1+c^2}\right)^\tau \delta^{-2\tau+1} \int_{c/\delta}^\infty t^{-2\tau+1} J_\nu^2(t)dt.$$

For $\nu > \tau - 1$, the latter integral can be decomposed, as in [10], as follows:

$$\begin{aligned}
\int_{c/\delta}^\infty t^{-2\tau+1} J_\nu^2(t)dt &= \int_0^\infty t^{-2\tau+1} J_\nu^2(t)dt - \int_0^{c/\delta} t^{-2\tau+1} J_\nu^2(t)dt \\
&= A(\nu,\tau) - B_\delta(\nu,\tau) = A(\nu,\tau)\left[1 - \frac{B_\delta(\nu,\tau)}{A(\nu,\tau)}\right],
\end{aligned}$$

where $A(\nu,\tau)$ is the integral from (6.8) and

$$B_\delta(\nu,\tau) = \int_0^{c/\delta} t^{-2\tau+1} J_\nu^2(t)dt.$$

The rest of the proof contains two steps. We will show first that $A(\nu,\tau)$ has the desired dependence on ν, and then show that $1 - B_\delta(\nu,\tau)/A(\nu,\tau)$ remains uniformly bounded away from zero under the given assumptions.

For the first task, the starting point is again the explicit formula (6.9). Using Stirling's formula we find in a similar way to the proof of Proposition 6.3 that

$$\begin{aligned} A(\nu,\tau) &= \frac{\Gamma(2\tau-1)}{2^{2\tau-1}\Gamma^2(\tau)} \frac{(\nu-\tau)\Gamma(\nu-\tau)}{\Gamma(\nu+\tau)} \\ &\geq \frac{\Gamma(2\tau-1)}{2^{2\tau-1}\Gamma^2(\tau)}(\nu-\tau)\frac{\sqrt{2\pi}(\nu-\tau)^{\nu-\tau-1/2}e^{-(\nu-\tau)}}{\sqrt{2\pi}(\nu+\tau)^{\nu+\tau-1/2}e^{-(\nu+\tau)}e^{1/(12(\nu+\tau))}} \\ &\geq \frac{\Gamma(2\tau-1)e^{2\tau}}{2^{2\tau-1}\Gamma^2(\tau)e^{1/(24\tau)}} \frac{1}{(\nu^2-\tau^2)^{\tau-1/2}} \left(\frac{1-\frac{\tau}{\nu}}{1+\frac{\tau}{\nu}}\right)^\nu, \end{aligned}$$

using the fact that $e^{1/(12(\nu+\tau))} < e^{1/(24\tau)}$ for $\nu \geq \tau$. We obviously have

$$\frac{1}{(\nu^2-\tau^2)^{\tau-1/2}} \geq \nu^{-2\tau+1}.$$

Furthermore, the bound in (6.11) yields this time

$$\log\left(\frac{1-\frac{\tau}{\nu}}{1+\frac{\tau}{\nu}}\right)^\nu = -\nu\left[\left|\log\left(1-\frac{\tau}{\nu}\right)\right| + \left|\log\left(1+\frac{\tau}{\nu}\right)\right|\right] \geq -\frac{8}{3}\tau,$$

provided that $\tau/\nu \leq 1/4$. Thus we have for $\nu \geq 4\tau$ the bound

$$A(\nu,\tau) \geq \frac{\Gamma(2\tau-1)e^{-2\tau/3}}{2^{2\tau-1}\Gamma^2(\tau)e^{1/(24\tau)}}\nu^{-2\tau+1}, \tag{6.18}$$

which has the desired dependence on ν.

It remains to look at the term $1 - B_\delta(\nu,\tau)/A(\nu,\tau)$. As in [10] we use the bound

$$|J_\nu(x)| \leq \frac{2^{-\nu}x^\nu}{\Gamma(\nu+1)},$$

which holds for $\nu > -1/2$ and $x > 0$, and Stirling's formula to derive a bound for $B_\delta(\nu,\tau)$: for $\nu > \tau - 1$ we have, with $\Gamma(\nu+1) = \nu\Gamma(\nu)$,

$$\begin{aligned} B_\delta(\nu,\tau) &= \int_0^{c/\delta} t^{-2\tau+1}J_\nu^2(t)dt \leq \frac{2^{-2\nu}}{\Gamma^2(\nu+1)}\int_0^{c/\delta} t^{2\nu-2\tau+1}dt \\ &= \frac{2^{-2\nu-1}}{\Gamma^2(\nu+1)(\nu-\tau+1)}\left(\frac{c}{\delta}\right)^{2\nu-2\tau+2} \\ &\leq \frac{2^{-2\nu-1}}{\nu^2(\nu-\tau+1)}\left(\frac{c}{\delta}\right)^{2\nu-2\tau+2}\frac{1}{2\pi\nu^{2\nu-1}e^{-2\nu}} \\ &= \frac{2^{-2\nu-2}}{\pi(\nu-\tau+1)}\left(\frac{c}{\delta}\right)^{2\nu-2\tau+2}\nu^{-2\nu-1}e^{2\nu}. \end{aligned}$$

This, together with the lower bound (6.18) on $A(\nu,\tau)$ gives after some simplifications, for $\nu \geq 4\tau$,

$$\frac{B_\delta(\nu,\tau)}{A(\nu,\tau)} \leq \frac{\Gamma^2(\tau)e^{1/(24\tau)}e^{\frac{8}{3}\tau+2}}{2\pi\Gamma(2\tau-1)}\frac{1}{\nu-\tau+1}\left(\frac{2\delta\nu}{ce}\right)^{2\tau-2\nu-2}.$$

If we now choose $c = 2/e$ and use $\delta\nu \geq 1$, then we see that

$$\frac{B_\delta(\nu,\tau)}{A(\nu,\tau)} \leq \frac{\Gamma^2(\tau)e^{1/(24\tau)}e^{\frac{8}{3}\tau+2}}{2\pi\Gamma(2\tau-1)}\frac{1}{\nu-\tau+1} =: \frac{c_\tau}{\nu-\tau+1}.$$

This expression becomes less than $1/2$ for $\nu \geq 2c_\tau + \tau - 1$. Thus, putting everything together gives

$$\widehat{\phi_\delta}(\ell) \geq \widetilde{c_1}(\delta\nu)^{-2\tau+1}$$

for $\nu = \ell + \frac{n-1}{2} \geq \max\{k_\tau, 1/\delta\}$, where $k_\tau := \max\{4\tau, 2c_\tau + \tau - 1\}$. If necessary, we can enlarge k_τ so that without restriction $k_\tau \geq n-1$. Then, $\nu \geq k_\tau \geq n-1$ means in particular $\ell \geq (n-1)/2 = \nu - \ell$ and hence $\nu \leq 2\ell$, which gives

$$\widehat{\phi_\delta}(\ell) \geq \widetilde{c_1}(\delta\nu)^{-2\tau+1} \geq \widetilde{c_1}2^{-2\tau+1}(\delta\ell)^{-2\tau+1} \geq \widetilde{c_1}2^{-2\tau+1}(1+\delta\ell)^{-2\tau+1},$$

for $\nu = \ell + \frac{n-1}{2} \geq \max\{k_\tau, 1/\delta\}$, which settles our statement if $k_\tau \leq 1/\delta$. In the case of $k_\tau > 1/\delta$ we have to make sure that the statement also holds for $1/\delta \leq \nu \leq k_\tau$. This can be seen as follows. We have for all $\nu \leq k_\tau$, using $1 \leq k_\tau\delta$,

$$\begin{aligned}
\widehat{\phi_\delta}(\ell) &\geq \widetilde{c_1}\int_0^{k_\tau}\frac{\delta t}{(1+\delta^2t^2)^\tau}J_\nu^2(t)dt \\
&\geq \widetilde{c_1}\delta^{-2\tau+1}2^{-\tau}k_\tau^{-2\tau}\int_0^{k_\tau} tJ_\nu^2(t)dt \\
&\geq \widetilde{c_1}\delta^{-2\tau+1}2^{-\tau}k_\tau^{-2\tau}\inf_{\nu\leq k_\tau}\int_0^{k_\tau} tJ_\nu^2(t)dt \\
&=: c_6\delta^{-2\tau+1} \geq c_6(1+\ell\delta)^{-2\tau+1},
\end{aligned}$$

where $c_6 > 0$ because there are only a finite number of values of $\nu = \ell + \frac{n-1}{2}$ less than or equal to k_τ. Hence, we have the stated lower bound for $\delta\nu \geq 1$. □

It remains to investigate the behaviour of $\widehat{\phi_\delta}(\ell)$ for small values of $\delta\nu$. For this, the following result, which is interesting on its own will be useful.

Lemma 6.5. *Let $\nu \geq 1/2$. Let λ be any positive zero of J_ν. Then,*

$$\lambda[J_\nu'(\lambda)]^2 > \frac{2}{\pi}\frac{(\lambda^2-\nu^2)^{1/2}}{\lambda}.$$

Proof. We will use the Weber functions, also known as Bessel functions of the second kind defined for non-integer index $\nu \notin \mathbb{Z}$ by

$$Y_\nu(x) = \frac{J_\nu(x)\cos(\nu\pi) - J_{-\nu}(x)}{\sin(\nu\pi)}$$

and by $Y_n := \lim_{\nu\to n} Y_\nu$ for integer index, see [15, section 3.54, pages 63/64]. They give another fundamental solution to the Bessel differential equation. Furthermore, according to [15, formula (1), section 3.63, page 76] the Wronskian of the system $\{J_\nu, Y_\nu\}$ is given by

$$J_\nu(x)Y_\nu'(x) - J_\nu'(x)Y_\nu(x) = \frac{2}{\pi x} \tag{6.19}$$

for all ν and all $x \neq 0$. We will also employ the bounds

$$\frac{2}{\pi x} < J_\nu^2(x) + Y_\nu^2(x) < \frac{2}{\pi}\frac{1}{(x^2-\nu^2)^{1/2}}, \tag{6.20}$$

which hold for $x \geq \nu \geq 1/2$, see [15, formula (1), section 13.74, page 447].

According to [15, formula (1), section 15.3, page 485] the smallest positive root λ_1 of J_ν satisfies $\lambda_1 > \nu$, thus we can set $x = \lambda$ in (6.20) where λ is any positive root of J_ν to derive

$$\lambda[J_\nu'(\lambda)]^2 = \frac{4}{\pi^2\lambda[Y_\nu(\lambda)]^2}. \tag{6.21}$$

From (6.19) we have

$$Y_\nu^2(\lambda) = J_\nu^2(\lambda) + Y_\nu^2(\lambda) < \frac{2}{\pi}\frac{1}{(\lambda^2-\nu^2)^{1/2}},$$

thus (6.21) gives

$$\lambda[J_\nu'(\lambda)]^2 > \frac{2}{\pi}\frac{(\lambda^2-\nu^2)^{1/2}}{\lambda}.$$

□

LEMMA 6.6. *Let* $\nu = \ell + \frac{n-1}{2} \geq 1/2$. *For every* $x \geq 4\nu$ *there is at least one zero of* J_ν *in the interval* $[x, 2x]$.

Proof. Let $j_{\nu,k}$ denote the kth positive zero of J_ν. From [6, page 251] we have, for $\nu \geq 1/2$,

$$j_{\nu,k+1} - j_{\nu,k} \leq \frac{\pi}{\sqrt{1 - \frac{\nu^2-\frac{1}{4}}{j_{\nu,k}^2}}},$$

which gives, since the sequence of zeros is increasing in k,

$$\frac{j_{\nu,k+1} - j_{\nu,k}}{j_{\nu,k}} \leq \frac{\pi}{\sqrt{j_{\nu,k}^2 - \nu^2 + 1/4}} \leq \frac{\pi}{\sqrt{j_{\nu,1}^2 - \nu^2 + 1/4}} \leq \frac{\pi}{\sqrt{2\nu + 1/4}}, \tag{6.22}$$

where the last inequality comes from the lower bound $j_{\nu,1} \geq \sqrt{\nu(\nu+2)}$, see [15, formula (5), section 15.3, page 486]. From (6.22) it follows easily that

$$\frac{j_{\nu,k+1} - j_{\nu,k}}{j_{\nu,k}} \leq 1: \tag{6.23}$$

for $\nu \geq 5$ this follows from the last bound in (6.22), while for $\nu = \frac{1}{2}, 1, \frac{3}{2}, \ldots, \frac{9}{2}$ the result follows from the penultimate expression in (6.22) by direct computation. The bound is sharp for $\nu = 1/2$ and is easily satisfied for $1 \leq \nu \leq \frac{9}{2}$. The inequality (6.23) is equivalent to

$$j_{\nu,k+1} \leq 2j_{\nu,k}. \tag{6.24}$$

Consider first the case $x > j_{\nu,1}$, and fix $k \geq 1$ by letting $j_{\nu,k}$ be the biggest of the zeros of J_ν which is smaller than x. Then we have, by (6.24),

$$j_{\nu,k} < x \leq j_{\nu,k+1} \leq 2j_{\nu,k} < 2x,$$

thus $j_{\nu,k+1} \in [x, 2x]$. To complete the argument, we consider the case $4\nu \leq x \leq j_{\nu,1}$. We know by [15, formula (5), section 15.3, page 486] that $j_{\nu,1} < \sqrt{2(\nu+1)(\nu+3)}$, and hence

$$\nu < j_{\nu,1} \leq \nu\sqrt{2}\sqrt{\left(1+\frac{1}{\nu}\right)\left(1+\frac{3}{\nu}\right)} \leq \sqrt{42}\nu < 8\nu \text{ for } \nu \geq \frac{1}{2}.$$

For $4\nu \leq x \leq j_{\nu,1}$ we have

$$4\nu \leq x \leq j_{\nu,1} < 8\nu \leq 2x,$$

and hence $j_{\nu,1} \in [x, 2x]$, completing the proof. $\square$

This now leads to a lower bound for the Fourier coefficients in the case of $\delta\nu \leq 1$.

Lemma 6.7. *For $\delta\nu \leq 1$ we have*

$$\widehat{\phi_\delta}(\ell) \geq c_7.$$

Proof. For any $\lambda \geq 1/\delta$ we have the lower bound

$$\begin{aligned}
\widehat{\phi_\delta}(\ell) &\geq \widetilde{c_1}\int_0^\lambda \frac{\delta t}{(1+\delta^2t^2)^\tau} J_\nu^2(t)dt \\
&\geq \widetilde{c_1}2^{-\tau}\delta(\delta\lambda)^{-2\tau}\int_0^\lambda tJ_\nu^2(t)dt \\
&= \widetilde{c_1}2^{-\tau}\delta(\delta\lambda)^{-2\tau}\frac{\lambda^2}{2}\left[[J'_\nu(\lambda)]^2 + \left(1-\frac{\nu^2}{\lambda^2}\right)J_\nu^2(\lambda)\right],
\end{aligned}$$

where the explicit form of the integral comes from [4, formula (24), page 70]. If we pick λ as a zero of J_ν, this reduces with the help of Lemma 6.5 to

$$\widehat{\phi_\delta}(\ell) \geq \widetilde{c_1}2^{-\tau-1}(\delta\lambda)^{-2\tau+1}\lambda[J'_\nu(\lambda)]^2 \tag{6.25}$$

$$\geq \widetilde{c_1}2^{-\tau-1}(\delta\lambda)^{-2\tau+1}\frac{2}{\pi}\frac{(\lambda^2-\nu^2)^{1/2}}{\lambda}. \tag{6.26}$$

Now choose $\lambda = \lambda_{\nu,4/\delta}$, where for $x \geq \nu$

$$\lambda_{\nu,x} := \min\{\lambda : \lambda \geq x \text{ and } J_\nu(\lambda) = 0\}.$$

Since

$$\lambda_{\nu,\frac{4}{\delta}} \geq \frac{4}{\delta} \geq 4\nu,$$

and since the function $f(\lambda) := (\lambda^2-\nu^2)^{1/2}/\lambda$ is strictly monotonically increasing on $[\nu,\infty)$, we have

$$\frac{(\lambda_{\nu,4/\delta}^2-\nu^2)^{1/2}}{\lambda_{\nu,4/\delta}} = f(\lambda_{\nu,4/\delta}) \geq f(4\nu) = \frac{\sqrt{15}}{4}.$$

Since lemma 6.6 gives

$$x \le \lambda_{\nu,x} \le 2x \text{ for } x \ge 4\nu,$$

we also have an upper bound

$$\lambda_{\nu,\frac{4}{\delta}} \le 2\frac{4}{\delta} \implies \delta\lambda_{\nu,\frac{4}{\delta}} \le 8.$$

Substituting the two bounds into (6.26), we find

$$\widehat{\phi_\delta}(\ell) \ge \widetilde{c}_1 \frac{2^{-7\tau+1}\sqrt{15}}{\pi} =: c_7,$$

proving the lemma. □

Proposition 6.3 establishes the upper bound while Lemmas 6.4 and 6.7 with $C_1 := \min\{c_6, c_7\}$ give the lower bound stated in Theorem 6.2.

7. Condition number and cost. We will now analyse the condition number and cost of the multiscale algorithm. For simplicity, we will drop the index indicating the current level. Thus in the present section $X = X_j$ is the point set at the step j.

As in the previous section, we will assume that the spherical basis function Φ_δ is defined by the restriction of a positive definite radial function $\Psi_\delta = \delta^{-n}\Psi(\cdot/\delta)$, where the unscaled basis function Ψ generates $H^\tau(\mathbb{R}^{n+1})$, i.e. it satisfies (6.1).

In each step of the multiscale algorithm, we have to compute an interpolant of the form (3.3), where the coefficients b_j for $j = 1, \ldots, M$ are determined by a linear system

$$\mathbf{A}_{X,\delta}\mathbf{b} = \mathbf{f}. \tag{7.1}$$

Here, $\mathbf{A}_{X,\delta} = (a_{i,j})$ is the interpolation matrix with entries $a_{i,j} = \Phi_\delta(\mathbf{x}_i, \mathbf{x}_j) = \delta^{-n}\Psi((\mathbf{x}_i - \mathbf{x}_j)/\delta)$ and $\mathbf{f} = (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_M))^T$. The nature of the multiscale algorithm guarantees that we can assume that $\delta = \nu h_X$ with $\nu > 1$ a fixed number, i.e. δ is proportional to h_X and this number ν is independent of the level.

Since the matrix $\mathbf{A}_{X,\delta}$ is symmetric and positive definite, an iterative method such as the conjugate gradient method can be used to solve (7.1) efficiently. The complexity of the conjugate gradient method depends on the condition number of the matrix $\mathbf{A}_{X,\delta}$ and on the cost of a matrix-vector multiplication.

We start with a discussion of the condition number. It is well known, see for example [18, Section 12.2] that the lower bound on the Fourier transform in (6.1) gives for the RBF with scale δ and the point set X a lower bound on the minimal eigenvalue of the form

$$\lambda_{\min}(\mathbf{A}_{X,1}) \ge c\widetilde{q}_X^{2\tau-(n+1)}, \tag{7.2}$$

with a constant $c > 0$ independent of X (but depending on δ_1), where we are using the Euclidean separation radius $\widetilde{q}_X$ in $\mathbb{R}^{n+1}$ from (2.1). We now introduce the data set $X/\delta = \{\mathbf{x}_1/\delta, \ldots, \mathbf{x}_M/\delta\}$, which obviously has Euclidean separation radius

$$\widetilde{q}_{X/\delta} = \widetilde{q}_X/\delta,$$

and make the simple observation that

$$\mathbf{A}_{X,\delta} = (\Phi_\delta(\mathbf{x}_i, \mathbf{x}_j)) = \left(\delta^{-n}\Psi\left(\frac{\mathbf{x}_i - \mathbf{x}_j}{\delta}\right)\right) = \delta^{-n}\mathbf{A}_{X/\delta,1}.$$

From this and (7.2) we obtain

$$\lambda_{\min}(\mathbf{A}_{X,\delta}) \geq c\delta^{-n} \left(\frac{\widetilde{q}_X}{\delta}\right)^{2\tau-(n+1)}.$$

Thus, we have the following result.

LEMMA 7.1. *The minimum eigenvalue of the matrix $\mathbf{A}_{X,\delta}$ can be bounded by*

$$\lambda_{\min}(\mathbf{A}_{X,\delta}) \geq c\delta^{-n} \left(\frac{\widetilde{q}_X}{\delta}\right)^{2\tau-(n+1)}$$

with a constant $c > 0$ independent of X and δ.

We state an upper bound of the maximum eigenvalue of $\mathbf{A}_\delta$ in the following lemma.

LEMMA 7.2. *Suppose $\Psi : \mathbb{R}^{n+1} \to \mathbb{R}$ has support in the unit ball and is bounded by 1. If the Euclidean separation radius $\widetilde{q}_X$ is less than $1/2$ then*

$$\lambda_{\max}(\mathbf{A}_{X,\delta}) \leq C\delta^{-n} \left(\frac{\widetilde{q}_X}{\delta}\right)^{-n-1}.$$

Proof. According to [17, Theorem 4.2] we have

$$\lambda_{\max}(\mathbf{A}_{X,1}) \leq \left(\frac{4}{\widetilde{q}_X}\right)^{n+1}$$

and the same result together with the previous scaling argument leads to

$$\lambda_{\max}(\mathbf{A}_{X,1}) \leq 4^{n+1}\delta^{-n} \left(\frac{\widetilde{q}_X}{\delta}\right)^{-n-1},$$

which is the desired estimate. □

From the previous lemmas, we have the following theorem.

THEOREM 7.3. *The condition number $\kappa(\mathbf{A}_{X,\delta})$ of the interpolation matrix $\mathbf{A}_{X,\delta}$ is bounded by*

$$\kappa(\mathbf{A}_{X,\delta}) \leq C \left(\frac{\delta}{\widetilde{q}_X}\right)^{2\tau},$$

where $C > 0$ is independent of X and the scaling parameter δ.

Thus Theorem 7.3 leads to the conclusion that if $\delta/\widetilde{q}_X \leq \widetilde{c}$ then the condition number should remain constant. However, since we already know that $\delta = \nu h_X$, these two assumptions lead to

$$\nu q_X \leq \nu h_X = \delta \leq \widetilde{c}\widetilde{q}_X.$$

Bearing in mind that the Euclidean and geodesic norm are comparable, this means that the data set has to be *quasi-uniform.*

We finish with a result on the cost of the multilevel algorithm.

THEOREM 7.4. *Suppose that the data sets used in each step of the multiscale algorithm are quasi-uniform. Then, for a given precision the conjugate gradient method can solve the linear system (7.1) on each level in $\mathcal{O}(M \log M)$ time, where M denotes the number of points at that level.*

Proof. The most expensive operation in the conjugate gradient method is the matrix vector product. However, in our case the interpolation matrix $\mathbf{A}_{X,\delta}$ is sparse and the number of nonzero entries in each row is bounded by a constant, which is independent of the level. Hence, each matrix-vector multiplication can be done in linear time.

According to Theorem 7.3 the condition number on each level is bounded by the same constant, which means that the number of iterations of the conjugate gradient method is also bounded by a constant independent of the level.

Hence, if the nonzero entries of the matrix $\mathbf{A}_{X,\delta}$ are known, the conjugate gradient method can solve the system in linear time. However, with scattered data points the nonzero entries are at unknown locations, and have to be sought in a preprocessing step. This boils down to finding, for every $\mathbf{x}_i$, all data sites $\mathbf{x}_j$ within the support of $\Psi_\delta(\mathbf{x}_i - \cdot)$. This is a classical problem in computational geometry and efficient methods based on tree-structures are known to solve this in $\mathcal{O}(M \log M)$ time, if M is the number of points used for the finest level. □

In general, we will keep the number of steps in the multiscale algorithm small compared to the number of data sites on the final level, such that this number can be seen as a constant. Hence, the overall complexity for solving the original interpolation problem is again of the form $\mathcal{O}(M \log M)$, if M is the number of points at the finest level.

8. Numerical experiments. In the experiments, the sets $X_1, \ldots, X_5$ are generated using a modified version of Saff's equal area algorithm [12] so that latitudes and longitudes (in degree) of the points are odd multiples of half a degree. The cardinalities, mesh norms and separation radii of these sets are listed in Table 8.1. The

	X_1	X_2	X_3	X_4	X_5
M	32	125	500	2000	8000
h_X	0.4372	0.2288	0.1165	0.0613	0.0367
q_X	0.3020	0.1281	0.0637	0.0263	0.0091

TABLE 8.1
Cardinalities, mesh norms and separation radii of the sets of points

function we wish to interpolate is the topographic data from NOAA (US Department of Commerce), which is available in MATLAB as an array of size $(180, 360)$ (which can be invoked by the command `load topo`) as shown in Figure 8.1. The function values represent the average heights in meters above and below the mean sea level at cells of size 1 degree times 1 degree. The RBF used in the experiment is the Wendland function

$$\psi(r) = \psi_{3,1}(r) = (1-r)_+^4(4r+1)$$

and the scaled version is

$$\psi_\delta(r) = \delta^{-2}(1-r/\delta)_+^4(4r/\delta+1)$$

where at stage j, we set $\delta = \delta_j$. We remark that the target function does not belong to the native space.

In Table 8.2 we show the errors and the condition numbers as reported by MATLAB at each stage of three independent multilevel experiments, using always the same 5 point sets given in Table 8.1, but with different starting values for the scale of the SBF (the three starting values of δ_1 being $2, 1, 1/2$ respectively). The condition

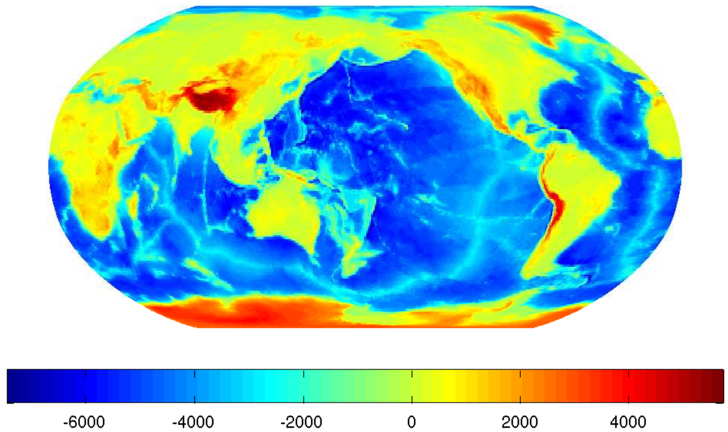

FIG. 8.1. *Exact function*

numbers are in good agreement with Theorem 7.3, and by the usual standards are all small. In particular, Experiment 3 has condition numbers reflecting the fact that $\mathbf{A}_\delta$ is essentially a constant matrix, but at the cost of a very large errors. The errors $\|e_j\|$ in the table are approximations to the true errors in the mean-square sense: in principle we define

$$\begin{aligned}\|e_j\| &:= \left(\frac{1}{4\pi}\int_{\mathbb{S}^2}|f(\mathbf{x})-f_j(\mathbf{x})|^2 d\mathbf{x}\right)^{1/2}\\ &= \left(\frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi}|f(\theta,\phi)-f_j(\theta,\phi)|^2\sin\theta d\phi d\theta\right)^{1/2},\end{aligned}$$

and in practice approximate this by the product midpoint rule at 1 degree intervals,

$$\left(\frac{1}{4\pi}\frac{2\pi^2}{|\mathcal{G}|}\sum_{\mathbf{x}(\theta,\phi)\in\mathcal{G}}|f(\theta,\phi)-f_j(\theta,\phi)|^2\sin\theta\right)^{1/2},$$

where $\mathcal{G}$ is the grid containing the centres of rectangles of size 1 degree times 1 degree and $|\mathcal{G}| = 180\cdot 360 = 64800$.

The approximations f_3, f_4, f_5, the details s_4, s_5 and the signed error $f_5 - f$ for Experiment 2 are shown in Figure 8.2. In Table 8.3 we show the results obtained with different scales by the "one-shot" approximation. Finally, in Figures 8.3 and 8.4 we show for comparison the final multiscale approximation obtained with Experiment 2, and therefore finishing at scale 1/16, and the one-shot approximation at scale 16. The latter shows a marked graininess, reflecting the poor overlap of the scale 16 radial basis functions on the 8,000 point set.

REFERENCES

[1] M. S. FLOATER AND A. ISKE, *Multistep scattered data interpolation using compactly supported radial basis functions*, J. Comput. Appl. Math., 73 (1996), pp. 65–78.

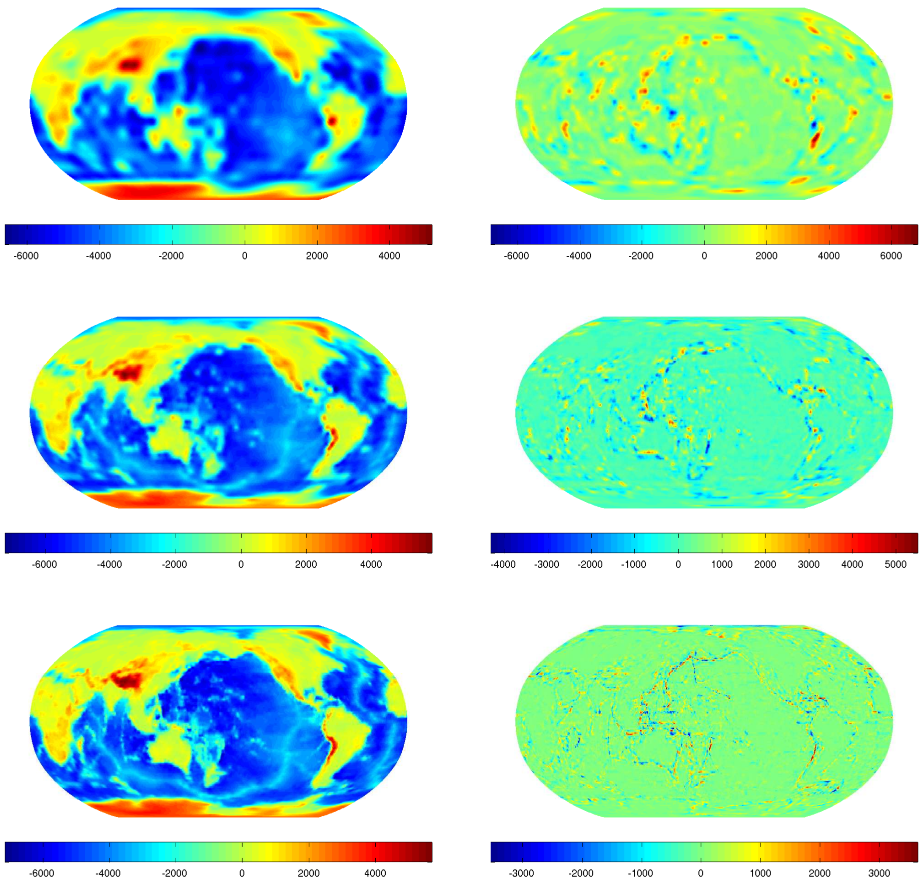

FIG. 8.2. *Approximations f_3, f_4, f_5 and details s_4, s_5 and signed error $f_5 - f$ for Experiment* 2.

[2] W. FREEDEN, *Multiscale modelling of spaceborne geodata*, B. G. Teubner, Stuttgart, 1999.

[3] W. FREEDEN, T. GERVENS, AND M. SCHREINER, *Constructive Approximation on the Sphere with Applications to Geomathematics*, Oxford University Press, Oxford, 1998.

[4] A. GRAY AND G.B. MATHEWS, *A treatise on Bessel functions and their applications to physics*, McMillan, London, 1952.

[5] S. J. HALES AND J. LEVESLEY, *Error estimates for multilevel approximation using polyharmonic splines*, Numer. Algorithms, 30 (2002), pp. 1–10.

[6] H. HOCHSTADT, *The functions of mathematical physics*, Dover, New York, 1986.

[7] Q. T. LE GIA, F. J. NARCOWICH, J. D. WARD, AND H. WENDLAND, *Continuous and discrete least-squares approximation by radial basis functions on spheres*, J. Approx. Theory, 143 (2006), pp. 124–133.

[8] C. MÜLLER, *Spherical Harmonics*, Springer, Berlin, 1966.

[9] F. J. NARCOWICH, R. SCHABACK, AND J. D. WARD, *Multilevel interpolation and approximation*, Appl. Comput. Harmon. Anal., 7 (1999), pp. 243–261.

[10] F. J. NARCOWICH, X. SUN, AND J. D. WARD, *Approximaton power of RBFs and their associated SBFs: A connection*, Adv. Comput. Math., 27 (2007), pp. 107–124.

[11] F. J. NARCOWICH AND J. D. WARD, *Scattered data interpolation on spheres: Error estimates and locally supported basis functions*, SIAM J. Math. Anal., 33 (2002), pp. 1393–1410.

[12] E. B. SAFF AND A. B. J. KUIJLAARS, *Distributing many points on a sphere*, Math. Intell., 19 (1997), pp. 5–11.

[13] R. SCHABACK, *Multivariate interpolation and approximation by translates of a basis function*,

	Level j	1	2	3	4	5
Experiment 1	δ_j	2	1	1/2	1/4	1/8
	$\kappa(\mathbf{A}_{\delta_j})$	42.22	53.54	49.45	83.47	174.21
	$\|e_j\|$	1955.40	1490.14	999.75	637.25	371.81
Experiment 2	δ_j	1	1/2	1/4	1/8	1/16
	$\kappa(\mathbf{A}_{\delta_j})$	1.71	2.14	2.16	4.09	7.80
	$\|e_j\|$	1961.31	1432.02	965.86	615.98	364.83
Experiment 3	δ_j	1/2	1/4	1/8	1/16	1/32
	$\kappa(\mathbf{A}_{\delta_j})$	1.00	1.00	1.00	1.01	1.29
	$\|e_j\|$	2831.27	2321.06	1852.14	1466.27	1169.33

Table 8.2
Errors and condition numbers of multiscale approximation

δ	2	1	1/2	1/4	1/8	1/16
$\kappa(\mathbf{A}_\delta)$	1.18e+8	3.70e+6	1.19e+5	4.14e+3	174.21	7.80
$\|e\|$	373.61	373.65	373.54	373.19	375.39	894.21

Table 8.3
One shot approximations with different scales

in Approximation Theory VIII, Vol. 1: Approximation and Interpolation, C. K. Chui and L. L. Schumaker, eds., Singapore, 1995, World Scientific Publishing, pp. 491–514.
[14] I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J., 9 (1942), pp. 96–108.
[15] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, 1966.
[16] H. Wendland, *Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree*, Adv. Comput. Math., 4 (1995), pp. 389–396.
[17] ———, *Error estimates for interpolation by compactly supported radial basis functions of minimal degree*, J. Approx. Theory, 93 (1998), pp. 258–272.
[18] ———, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, UK, 2005.
[19] Y. Xu and E. W. Cheney, *Strictly positive definite functions on spheres*, Proc. Am. Math. Soc., 116 (1992), pp. 977–981.

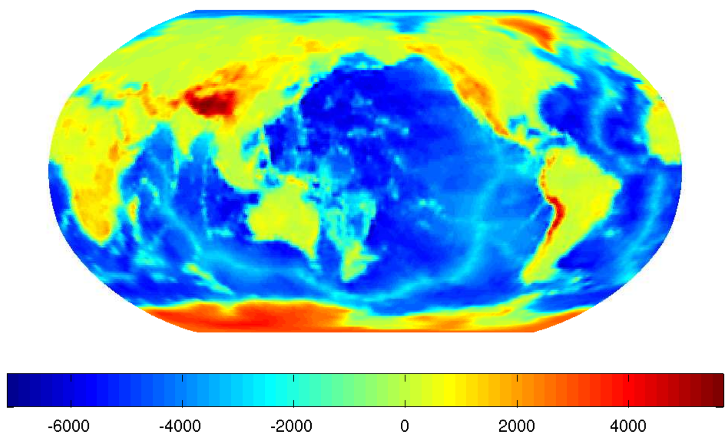

FIG. 8.3. *Multiscale approximation f_5 with finishing scale $\delta = 1/16$.*

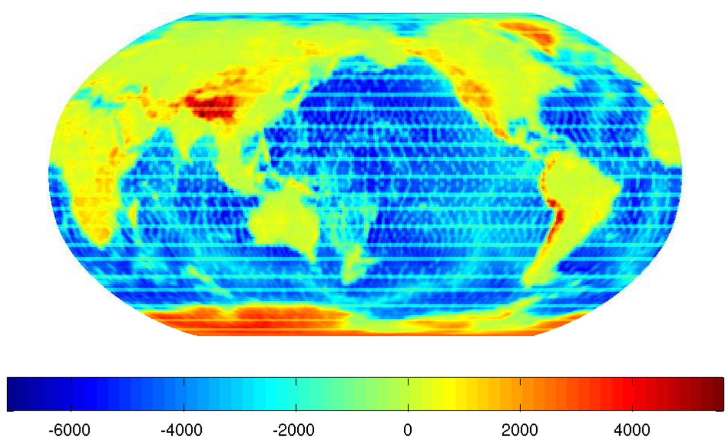

FIG. 8.4. *One shot approximation with $\delta = 1/16$.*