

# Multiscale RBF collocation for solving PDEs on spheres

Q. T. Le Gia · I. H. Sloan · H. Wendland

Received: date / Accepted: date

**Abstract** In this paper, we discuss multiscale radial basis function collocation methods for solving certain elliptic partial differential equations on the unit sphere. The approximate solution is constructed in a multi-level fashion, each level using compactly supported radial basis functions of smaller scale on an increasingly fine mesh. Two variants of the collocation method are considered (sometimes called symmetric and unsymmetric, although here both are symmetric). A convergence theory is given, which builds on recent theoretical advances for multiscale approximation using compactly supported radial basis functions.

**Keywords** partial differential equation · multiscale · collocation · unit sphere

**Mathematics Subject Classification (2000)** 65N35 · 65N55

## 1 Introduction

Partial differential equations (PDEs) on the unit sphere have many applications in the geosciences. Since analytic solutions are difficult or impossible to find, good algorithms for finding approximate solutions are essential. Radial basis functions (RBFs) present a simple and effective way to construct approximate solutions to PDEs on spheres, via a collocation method [16] or a Galerkin method [13]. They have been used successfully for solving transport-like equations on the sphere [3,4].

The quality of the approximation depends on the distribution of the centers of the RBFs used to define the approximate solution. However, in practice the solution usually represents some physical quantities, which are available in many physical scales. A solution using RBFs with a single scale may fail to capture these features, unless the radial basis function has a large support and the number of centers is also large, a combination which can be computationally prohibitive. To overcome this, we propose a multiscale approximation scheme,

---

Q. T. Le Gia  
School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia.

I. H. Sloan  
School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia.

H. Wendland  
Mathematical Institute, University of Oxford, 24-29 St Giles' Oxford, OX1 3LB England.

in which the approximate solution is constructed using a multi-stage process, in which the residual of the current stage is the target function for the next stage, and in each stage, RBFs with smaller support and with more closely spaced centers will be used as basis functions.

While meshless methods using RBFs have been employed to derive numerical solutions for PDEs on the sphere only recently [16, 13, 3, 4], it should be mentioned that approximation methods using RBFs for PDEs on bounded domains have been around for the last two decades. Originally proposed by Kansa [11, 12] for fluid dynamics, approximation methods for many types of PDEs defined on bounded domains in  $\mathbb{R}^n$  using RBFs have since been used widely. Examples include [1, 6, 9, 10].

For boundary value problems, the technique predominantly used in the literature, with the exception of [23], where a Galerkin method was used, has been collocation, mainly because of the simplicity and the fact that there is no requirement for numerical integration, which is still a problematic issue in all meshfree methods. There are two popular approaches to deriving the approximation scheme, usually called unsymmetric and symmetric collocation. In the first approach the collocation matrix is unsymmetric on bounded domains because of the two different operators involved, the differential operator and the boundary operator. This however can lead to nonsolvable systems [10]. Nonetheless, the method is widely used since the solution is just a linear combination of the RBFs and usually shows good approximation properties. In our case, where our differential operator is independent of position and we do not have a boundary, the nonsymmetric approach turns out actually to be an alternative symmetric approach, which we will refer to as the *standard collocation method*.

In the classical symmetric approach the operators are incorporated into the approximation space and hence the collocation matrix is symmetric and always positive definite, but the numerical solution is slightly more complicated to construct. Since the solution minimizes a certain Hilbert space norm amongst all possible functions satisfying the collocation condition, we will refer to this approach as *norm-minimal collocation*. A comparison of the two variants can be found in [19].

A common drawback of using RBFs in approximation schemes for PDEs is that the collocation matrix arising from the approximation problem tends to be ill-conditioned. There are two main approaches to overcome this: either to use a preconditioner, or to use a multilevel approximation approach. With a multilevel method, the condition number of the matrix at each level can be relatively small, and has only slow growth. There are papers [1, 2] dealing with multilevel approximation methods for PDEs using compactly supported radial basis functions on bounded domains in  $\mathbb{R}^n$ , however the theory there is incomplete. In this work, we will prove convergence results in Sobolev spaces for both kinds of collocation for a class of elliptic PDEs defined on the unit sphere  $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ .

The present paper builds upon theoretical advances in multiscale approximation for the sphere [15], which were subsequently extended to bounded regions in  $\mathbb{R}^n$  [25].

In Section 2 we will review necessary background on spherical harmonics, positive definite kernels, and Sobolev spaces on the unit sphere. In Section 3 we will present two variants of the collocation method for solving PDEs on the unit sphere using RBFs of a single scale. In Section 4 we present the corresponding multiscale methods. A convergence theorem for the multiscale methods will be proved. Finally, Section 5 presents some numerical examples.

## 2 Preliminaries

In our work, we will use zonal functions to construct approximate solutions for the PDEs. Zonal 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 zonal kernels of the type

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

where  $\{P_{\ell}(n+1; t)\}_{\ell=0}^{\infty}$  is the sequence of  $(n+1)$ -dimensional Legendre polynomials normalized to  $P_{\ell}(n+1; 1) = 1$ . Thanks to the seminal work of Schoenberg [22] and the later work of [26], we know that such a  $\phi$  is (strictly) positive definite on  $\mathbb{S}^n$ , that is, the matrix  $\mathbf{A} := [\phi(\mathbf{x}_i \cdot \mathbf{x}_j)]_{i,j=1}^M$  is positive definite for every set of distinct points  $\{\mathbf{x}_1, \dots, \mathbf{x}_M\}$  on  $\mathbb{S}^n$  and every positive integer  $M$ .

For mathematical analysis it is sometimes 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 [17]. 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  $\Delta$  is the Laplacian operator in  $\mathbb{R}^{n+1}$ . The space of all spherical harmonics of degree  $\ell$  on  $\mathbb{S}^n$ , denoted by  $\mathcal{H}_{\ell}$ , has an  $L_2$  orthonormal basis

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

where

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

thus

$$\int_{\mathbb{S}^n} Y_{\ell k} Y_{\ell' k'} dS = \delta_{\ell \ell'} \delta_{kk'},$$

where  $dS$  is the surface measure of the unit sphere. The space of spherical harmonics of degree  $\leq L$  will be denoted by  $\mathcal{P}_L := \bigoplus_{\ell=0}^L \mathcal{H}_{\ell}$ ; it has dimension  $N(n+1, L)$ . Every function  $g \in L_2(\mathbb{S}^n)$  can be expanded in terms of spherical harmonics,

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

Using the addition theorem for spherical harmonics (see, for example, [17, page 10]),

$$\sum_{k=1}^{N(n, \ell)} Y_{\ell k}(\mathbf{x}) Y_{\ell k}(\mathbf{y}) = \frac{N(n, \ell)}{\omega_n} P_{\ell}(n+1; \mathbf{x} \cdot \mathbf{y}), \quad (2)$$

we can write

$$\Phi(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n, \ell)} \widehat{\phi}(\ell) Y_{\ell k}(\mathbf{x}) Y_{\ell k}(\mathbf{y}), \quad \text{where } \widehat{\phi}(\ell) = \frac{\omega_n}{N(n, \ell)} a_{\ell}, \quad (3)$$

where  $\omega_n$  is the surface area of  $\mathbb{S}^n$ . We shall assume that, for some  $\sigma > n/2$ ,

$$c_1(1+\ell)^{-2\sigma} \leq \widehat{\phi}(\ell) \leq c_2(1+\ell)^{-2\sigma}, \quad \ell \geq 0. \quad (4)$$

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

Assume that we are given a positive definite kernel on  $\mathbb{S}^n$  defined from a compactly supported radial basis function  $R : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$

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

where  $|\cdot|$  is the Euclidean distance in  $\mathbb{R}^{n+1}$ . We may then define a scaled version,

$$\Phi_\delta(\mathbf{x}, \mathbf{y}) = \frac{1}{\delta^n} R\left(\frac{\mathbf{x} - \mathbf{y}}{\delta}\right), \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^n, \quad (5)$$

where  $\delta > 0$  is a scaling parameter. In the following, we expand  $\Phi_\delta$  as

$$\Phi_\delta(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \widehat{\phi}_\delta(\ell) Y_{\ell k}(\mathbf{x}) Y_{\ell k}(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^n. \quad (6)$$

We assume, strengthening (4), that for some  $\sigma > n/2$ ,

$$c_1(1 + \delta\ell)^{-2\sigma} \leq \widehat{\phi}_\delta(\ell) \leq c_2(1 + \delta\ell)^{-2\sigma}, \quad \ell \geq 0. \quad (7)$$

In fact, we have shown previously ([15, Theorem 6.2]) that condition (7) is satisfied if  $R$  is a compactly supported RBF of Wendland's type [24].

The *native space* associated with  $\Phi$  is defined to be

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

where  $\mathcal{D}'(\mathbb{S}^n)$  is the space of distributions on  $\mathbb{S}^n$ . More details on native spaces can be found in [24]. It can be shown that  $\mathcal{N}_\Phi$  is a Hilbert space with respect to the following inner product

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

Moreover, we can show that  $\Phi$  is a reproducing kernel for  $\mathcal{N}_\Phi$ , i.e. for all  $g \in \mathcal{N}_\Phi$ ,

$$\langle g(\cdot), \Phi(\mathbf{x}, \cdot) \rangle_\Phi = g(\mathbf{x}), \quad \mathbf{x} \in \mathbb{S}^n. \quad (8)$$

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

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

Under the condition  $\sigma > n/2$ , the norms  $\|\cdot\|_\Phi$  and  $\|\cdot\|_{H^\sigma}$  are equivalent if (4) holds, with the norms related by

$$c_1^{1/2} \|g\|_\Phi \leq \|g\|_{H^\sigma} \leq c_2^{1/2} \|g\|_\Phi. \quad (9)$$

Using the scaled version of the kernel  $\Phi_\delta(\mathbf{x}, \mathbf{y})$ , for a function  $g \in \mathcal{N}_\Phi$  we define the following norm:

$$\|g\|_{\Phi_\delta} := \left( \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{|\widehat{g}_{\ell k}|^2}{\widehat{\phi}_\delta(\ell)} \right)^{1/2}. \quad (10)$$

This norm, too, is equivalent to the norm  $\|g\|_{H^\sigma}$ . The following lemma gives information about that equivalence.

**Lemma 1** For  $\sigma > n/2$  and  $\delta \leq 2$ , for all  $g \in H^\sigma(\mathbb{S}^n)$ , we have

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

**Proof.** This is essentially Lemma 3.1 in [15].  $\square$

### 3 Single-scale Collocation for Solving PDEs

We consider the following PDE

$$Lu = f \quad \text{on } \mathbb{S}^n, \quad (11)$$

where  $L$  is an elliptic self-adjoint differential operator with constant symbol  $\widehat{L}(\ell)$  and order  $t$ , for some  $t > 0$ . That is, we can expand  $Lu$  as a Fourier series

$$Lu = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \widehat{L}(\ell) \widehat{u}_{\ell k} Y_{\ell k}$$

in which

$$c_3(1+\ell)^t \leq \widehat{L}(\ell) \leq c_4(1+\ell)^t, \quad \ell \geq 0, \quad (12)$$

where  $c_3, c_4$  are two positive constants independent of  $\ell$ . For example, we may take  $L = -\Delta^* + \omega^2$ , where  $\Delta^*$  is the Laplace–Beltrami operator and  $\omega > 0$ , in which case  $\widehat{L}(\ell) = \ell(\ell + n - 1) + \omega^2$  and  $t = 2$ .

Since  $\widehat{L}(\ell) > 0$  for  $\ell \geq 0$ , we can define  $L^{1/2}$  by

$$L^{1/2}u = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \sqrt{\widehat{L}(\ell)} \widehat{u}_{\ell k} Y_{\ell k}.$$

Note that we have an intrinsic relation between the smoothness of the given function  $f$  and the solution  $u$  of (11), as follows:  $f \in H^\sigma$  if and only if  $u \in H^{\sigma+t}$ . Hence, in the future we can make assumptions on the smoothness of the solution  $u$ , which translate immediately to assumptions on  $f$ .

In this section, we will discuss collocation methods to solve (11) approximately. Initially, our collocation methods will be based upon radial basis functions of a single scale. We will discuss two single scale approaches, known in the literature as symmetric and unsymmetric collocation.

Suppose  $X := \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subseteq \mathbb{S}^n$  is a given discrete set of scattered points on the  $n$ -dimensional sphere  $\mathbb{S}^n$ . Then solving equation (11) by collocation on the set  $X$  means to find a function  $u_h$  from a given approximation space which satisfies the collocation equations

$$Lu_h(\mathbf{x}_j) = Lu(\mathbf{x}_j) = f(\mathbf{x}_j), \quad 1 \leq j \leq N. \quad (13)$$

### 3.1 Standard Collocation

When working with radial basis functions, an obvious approach to finding such an approximate solution is as follows.

Suppose that  $\Phi$  is a kernel that satisfies condition (4) for some  $\sigma > (t+n)/2$ . This assumption guarantees that we may apply  $L$  to one of the arguments of  $\Phi$  and still have a continuous function. Hence, we may pick our approximation  $u_h$  from the approximation space

$$V_X := \text{span} \{ \Phi(\cdot, \mathbf{x}_j) : \mathbf{x}_j \in X \}. \quad (14)$$

To find  $u_h \in V_X$ , we represent  $u_h$  as a linear combination of the basis functions in  $V_X$ :

$$u_h = \sum_{j=1}^N b_j \Phi(\cdot, \mathbf{x}_j),$$

and the condition (13) leads now to the linear system

$$\mathbf{A} \mathbf{b} = \mathbf{f}, \quad (15)$$

where  $\mathbf{A}$  is the collocation matrix with entries  $a_{i,j} = L\Phi(\mathbf{x}_i, \mathbf{x}_j)$  and the right-hand side is given by  $\mathbf{f} = (f(\mathbf{x}_j))$ .

Here, the function  $L\Phi$  is defined to be the kernel having the Fourier expansion

$$L\Phi(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \widehat{\phi}(\ell) \widehat{L}(\ell) Y_{\ell k}(\mathbf{x}) Y_{\ell k}(\mathbf{y}), \quad (16)$$

which can be computed by applying  $L$  to either of the arguments of  $\Phi$ . In particular, the new kernel  $L\Phi$  is symmetric and positive definite, since its Fourier coefficients  $\widehat{L}(\ell) \widehat{\phi}(\ell)$  are positive. This ensures that the system (15) is always uniquely solvable.

**Lemma 2** *Let  $\sigma > (t+n)/2$  and assume that  $\Phi$  satisfies (4) and  $V_X$  is given by (14). There exists a unique function  $u_h \in V_X$  satisfying the collocation conditions (13). The solution  $u_h$  belongs to  $H^{\sigma+t/2}$ .*

**Proof.** Existence and uniqueness have been established already. The second part follows from the fact that each  $\varphi_j := \Phi(\cdot, \mathbf{x}_j)$  belongs to  $H^{\sigma+t/2}$ . To see this, we note from (3) that  $\widehat{\varphi}_j(\ell) = \widehat{\phi}(\ell) Y_{\ell k}(\mathbf{x}_j)$ , which leads via (4) to

$$\begin{aligned} \|\varphi_j\|_{H^{\sigma+t/2}}^2 &= \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \widehat{\phi}(\ell)^2 Y_{\ell k}(\mathbf{x}_j)^2 (1+\ell)^{2\sigma+t} \\ &\leq c_2^2 \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} Y_{\ell k}(\mathbf{x}_j)^2 (1+\ell)^{t-2\sigma} \\ &= c_2^2 \sum_{\ell=0}^{\infty} \frac{N(n,\ell)}{\omega_n} P_{\ell}(n+1;1) (1+\ell)^{t-2\sigma} \\ &\leq c \sum_{\ell=0}^{\infty} (1+\ell)^{t-2\sigma+n-1}, \end{aligned}$$

where we have used the addition theorem (2), the fact that the Legendre polynomials are normalized to  $P_{\ell}(n+1;1) = 1$ , and the fact that the dimension  $N(n,\ell)$  behaves like  $(1+\ell)^{n-1}$ . This sum is finite if  $t-2\sigma+n-1 < -1$  which is equivalent to our assumption on  $\sigma$ .  $\square$

This result stands in sharp contrast to standard RBF collocation for boundary value problems. In the situation of solving a PDE on a bounded domain with given boundary values, this approach does not always lead to an invertible collocation matrix, see [10]. Nonetheless, it is widely used, usually under the name *unsymmetric collocation*.

To understand the solution process, we introduce a new kernel  $\Psi$  defined by

$$\Psi = L^{-1}\Phi.$$

This kernel has Fourier coefficients  $\widehat{\psi}(\ell) = \widehat{\phi}(\ell)/\widehat{L}(\ell)$  and hence defines an inner product  $\langle \cdot, \cdot \rangle_{\Psi}$ ,

$$\langle f, g \rangle_{\Psi} = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{\widehat{L}(\ell) \widehat{f}_{\ell k} \widehat{g}_{\ell k}}{\widehat{\phi}(\ell)}, \quad f, g \in H^{\sigma+t/2}, \quad (17)$$

and the corresponding norm by  $\|g\|_{\Psi}^2 = \langle g, g \rangle_{\Psi}$ . Under our assumptions (4) and (12) on the kernel  $\Phi$  and the operator  $L$ , respectively, we easily see that the  $\|\cdot\|_{\Psi}$  norm is equivalent to the Sobolev norm  $\|\cdot\|_{H^{\sigma+t/2}}$ , and that  $H^{\sigma+t/2}$  with the inner product (17) is a reproducing kernel Hilbert space with kernel  $\Psi$ . As in the case of the original kernel  $\Phi$ , we can define a scaled version of  $\Psi$ , that is

$$\Psi_{\delta} := L^{-1}\Phi_{\delta}.$$

This kernel defines an inner product  $\langle \cdot, \cdot \rangle_{\Psi_{\delta}}$ , which is defined by

$$\langle f, g \rangle_{\Psi_{\delta}} = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{\widehat{L}(\ell) \widehat{f}_{\ell k} \widehat{g}_{\ell k}}{\widehat{\phi}_{\delta}(\ell)}, \quad f, g \in H^{\sigma+t/2}, \quad (18)$$

and the corresponding norm is  $\|g\|_{\Psi_{\delta}} = \sqrt{\langle g, g \rangle_{\Psi_{\delta}}}$ . This norm is also equivalent to the Sobolev norm  $\|\cdot\|_{H^{\sigma+t/2}}$ , as given in the following lemma.

**Lemma 3** For  $\sigma > n/2$  and  $\delta \leq 2$ , for all  $g \in H^{\sigma+t/2}(\mathbb{S}^n)$ , we have

$$c_5 \|g\|_{\Psi_{\delta}} \leq \|g\|_{H^{\sigma+t/2}} \leq c_6 \delta^{-\sigma} \|g\|_{\Psi_{\delta}}.$$

**Proof.** Since  $\delta \leq 2$ , using (7), (12) and (18) we have

$$\begin{aligned} \|g\|_{\Psi_{\delta}}^2 &= \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \frac{\widehat{L}(\ell) |\widehat{g}_{\ell k}|^2}{\widehat{\phi}_{\delta}(\ell)} \\ &\leq \frac{c_4}{c_1} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} (1+\ell)^t |\widehat{g}_{\ell k}|^2 (1+\delta\ell)^{2\sigma} \\ &\leq \frac{2^{2\sigma} c_4}{c_1} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} |\widehat{g}_{\ell k}|^2 (1+\ell)^{2\sigma+t} = \frac{2^{2\sigma} c_4}{c_1} \|g\|_{H^{\sigma+t/2}}^2. \end{aligned}$$

We also have  $(1+\delta\ell) = \delta(1/\delta + \ell) \geq \delta(1/2 + \ell/2)$ . Hence,

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

and again using (7) and (18),

$$\begin{aligned} \|g\|_{H^{\sigma+t/2}}^2 &= \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} |\widehat{g}_{\ell k}|^2 (1+\ell)^{2\sigma+t} \\ &\leq 2^{2\sigma} \delta^{-2\sigma} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} (1+\ell)^t |\widehat{g}_{\ell k}|^2 (1+\delta\ell)^{2\sigma} \leq 2^{2\sigma} \delta^{-2\sigma} c_2 c_3^{-1} \|g\|_{\Psi_{\delta}}^2. \end{aligned}$$

Setting  $c_5 := (c_1/c_4)^{1/2} 2^{-\sigma}$  and  $c_6 := 2^{\sigma} (c_2/c_3)^{1/2}$  we obtain the result of the lemma.  $\square$

**Lemma 4** Let  $u \in H^{\sigma+t/2}$  with  $\sigma > (t+n)/2$ . Let  $u_h \in V_X$  be the solution of the collocation equation (13) by the standard approach. Then

$$\langle u - u_h, \chi \rangle_\Psi = 0 \text{ for all } \chi \in V_X, \quad (19)$$

and hence

$$\|u - u_h\|_\Psi \leq \|u\|_\Psi. \quad (20)$$

**Proof.** From the collocation equation (13) we have  $Lu(\mathbf{x}_j) = Lu_h(\mathbf{x}_j)$  for all  $\mathbf{x}_j \in X$ . This means that the error function  $e_h = u - u_h$  satisfies  $Le_h(\mathbf{x}_j) = 0$  for all  $\mathbf{x}_j \in X$ , and hence, since  $L$  is invertible and self-adjoint,

$$\begin{aligned} \langle e_h, \Phi(\cdot, \mathbf{x}_j) \rangle_\Psi &= \langle Le_h, L^{-1}\Phi(\cdot, \mathbf{x}_j) \rangle_\Psi \\ &= \langle Le_h, \Psi(\cdot, \mathbf{x}_j) \rangle_\Psi \\ &= Le_h(\mathbf{x}_j) = 0. \end{aligned}$$

Since  $V_X$  is the finite dimensional space spanned by  $\Phi(\cdot, \mathbf{x}_j)$  we obtain (19), which immediately implies Pythagoras' theorem,

$$\|u - u_h\|_\Psi^2 + \|u_h\|_\Psi^2 = \|u\|_\Psi^2,$$

from which inequality (20) follows.  $\square$

We will now discuss the error  $u - u_h$ . As usual, we will use the mesh norm  $h_X$ , which is defined by

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

where  $\theta(\mathbf{x}, \mathbf{y}) = \cos^{-1}(\mathbf{x} \cdot \mathbf{y})$  is the geodesic distance between  $\mathbf{x}$  and  $\mathbf{y}$ .

**Theorem 1** Assume that the exact solution  $u$  belongs to  $H^{\sigma+t/2}$  with  $\sigma > (t+n)/2$ . Let  $h_X$  be the mesh norm of the scattered set  $X$ , let  $\Phi$  be a positive definite kernel satisfying (4) and let  $u_h \in V_X$  be the approximate solution obtained by the collocation equation (13). Then,

$$\|u - u_h\|_{L_2} \leq c \|L^{1/2}(u - u_h)\|_{L_2} \leq ch_X^{\sigma-t/2} \|u - u_h\|_{H^{\sigma+t/2}} \leq ch_X^{\sigma-t/2} \|u\|_{H^{\sigma+t/2}}.$$

**Proof.** Since the function  $Lu - Lu_h$  vanishes on  $X$  and belongs to  $H^{\sigma-t/2}$ , the ‘‘sampling inequality’’, [14, Theorem 3.3], guarantees the existence of a constant  $c > 0$  such that

$$\|Lu - Lu_h\|_{L_2} \leq ch_X^{\sigma-t/2} \|Lu - Lu_h\|_{H^{\sigma-t/2}} \leq ch_X^{\sigma-t/2} \|u - u_h\|_{H^{\sigma+t/2}},$$

where, in the last step, we have used the condition (12) and the definition of Sobolev norms. Using (12) again and the equivalence between the  $\|\cdot\|_\Psi$  norm and the norm  $\|\cdot\|_{H^{\sigma+t/2}}$ , we obtain

$$\begin{aligned} \|u - u_h\|_{L_2} &\leq c \|L^{1/2}(u - u_h)\|_{L_2} \\ &\leq c \|Lu - Lu_h\|_{L_2} \\ &\leq ch_X^{\sigma-t/2} \|u - u_h\|_{H^{\sigma+t/2}} \\ &\leq ch_X^{\sigma-t/2} \|u - u_h\|_\Psi \\ &\leq ch_X^{\sigma-t/2} \|u\|_\Psi \\ &\leq ch_X^{\sigma-t/2} \|u\|_{H^{\sigma+t/2}}, \end{aligned}$$

where we used Lemma 4 in the last but one step.  $\square$



### 3.2 Norm-minimal Collocation

We will now discuss another collocation technique. Assume that we know that the exact solution  $u$  belongs to  $H^\sigma$  and assume that  $\Phi$  is a reproducing kernel of  $H^\sigma$  in the sense that its Fourier coefficients satisfy (4). Then, it seems natural to choose an approximate solution  $u_h$  as the solution of

$$\min \{ \|s\|_\Phi : s \in H^\sigma \text{ with } Ls(\mathbf{x}_j) = f(\mathbf{x}_j), \quad 1 \leq j \leq N \}, \quad (21)$$

i.e.,  $u_h$  is minimizing the native space norm amongst all possible functions that collocate the given data. It turns out that this corresponds to what is known as the *symmetric collocation method*, see [24]. However, we have now to assume that  $\sigma > t + n/2$ .

It can be shown, in a much more general context, that the solution of (21) must necessarily come from the finite dimensional space

$$W_X := \text{span}\{L\Phi(\cdot, \mathbf{x}_j) : 1 \leq j \leq N\},$$

and that the concrete solution  $u_h \in W_X$  can be computed by imposing the collocation conditions (13). To find  $u_h$ , we can represent  $u_h$  as a linear combination of the basis functions in  $W_X$ :

$$u_h = \sum_{j=1}^N b_j L\Phi(\cdot, \mathbf{x}_j)$$

and condition (13) will then lead to the linear system

$$\mathbf{A}\mathbf{b} = \mathbf{f}, \quad (22)$$

where  $\mathbf{A}$  is now the collocation matrix with entries  $a_{ij} = LL\Phi(\mathbf{x}_i, \mathbf{x}_j)$  and the right-hand side is again given by  $\mathbf{f} = (f(\mathbf{x}_j))$ .

It is also well-known that this symmetric collocation solution is the best approximation from  $W_X$  in the native space norm, see [24].

**Lemma 5** *Suppose  $u \in H^\sigma$  with  $\sigma > t + n/2$  is the exact solution of (11). Let  $u_h \in W_X$  be the solution of (21). Then,*

$$\langle u - u_h, \chi \rangle_\Phi = 0 \quad \text{for all } \chi \in W_X, \quad (23)$$

and hence

$$\|u - u_h\|_\Phi \leq \|u\|_\Phi. \quad (24)$$

**Proof.** For  $j = 1, \dots, N$  we have, using the reproducing property of  $\Phi$  and the self-adjoint property of  $L$ ,

$$\begin{aligned} \langle u - u_h, L\Phi(\cdot, \mathbf{x}_j) \rangle_\Phi &= \langle L(u - u_h), \Phi(\cdot, \mathbf{x}_j) \rangle_\Phi \\ &= Lu(\mathbf{x}_j) - Lu_h(\mathbf{x}_j) = 0, \end{aligned}$$

from which the result follows immediately.  $\square$

Error estimates for the norm-minimal collocation method have, even in the more complicated situation of bounded domains, been derived in [5, 8].

The error between the solution and the approximate solution depends on the mesh norm  $h_X$  of the set  $X$ , as given in the following convergence theorem.

**Theorem 2** Assume that  $u \in H^\sigma$ , for  $\sigma > t + n/2$ , is the exact solution of (11) and  $u_h \in W_X$  is the solution of (21). Then, provided  $h_X$  is sufficiently small,

$$\|u - u_h\|_{L_2} \leq ch_X^{\sigma-t} \|u - u_h\|_{H^\sigma} \leq ch_X^{\sigma-t} \|u\|_{H^\sigma}.$$

**Proof.** Since the function  $Lu - Lu_h$  vanishes on  $X$ , we can again employ the ‘‘sampling inequality’’ from [14, Theorem 3.3], which gives a constant  $c > 0$  such that

$$\begin{aligned} \|Lu - Lu_h\|_{L_2} &\leq ch_X^{\sigma-t} \|Lu - Lu_h\|_{H^{\sigma-t}} \leq ch_X^{\sigma-t} \|u - u_h\|_{H^\sigma} \\ &\leq ch_X^{\sigma-t} \|u - u_h\|_\Phi, \end{aligned}$$

where in the last two steps we have used condition (12), the definition of Sobolev norms and the equivalence (9). Using condition (12) again and the equivalence between the native space norm  $\|\cdot\|_\Phi$  and the Sobolev norm  $\|\cdot\|_{H^\sigma}$  we obtain

$$\begin{aligned} \|u - u_h\|_{L_2} &\leq c \|Lu - Lu_h\|_{L_2} \leq ch_X^{\sigma-t} \|u - u_h\|_\Phi \leq ch_X^{\sigma-t} \|u\|_\Phi \\ &\leq ch_X^{\sigma-t} \|u\|_{H^\sigma}, \end{aligned}$$

where in the second to last step we have used (24) from Lemma 5.  $\square$

### 3.3 Sharpness of the Results

In both Theorems 1 and 2 we used an inequality of the form

$$\|u - u_h\|_{L_2} \leq \|Lu - Lu_h\|_{L_2}.$$

This clearly is a coarse estimate, which one might think would leave some leeway for better estimates. Interestingly, the following 1-dimensional example shows that the estimates in Theorems 1 and 2 are the best we can hope for.

**Lemma 6** Consider  $Lu = f$  on  $\mathbb{S}^1$  where  $L$  is defined by  $\widehat{L}(\ell) = (1 + \ell)^t$ ,  $t > 0$ . Let  $X = \{x_j = j\pi/m : 1 \leq j \leq 2m\}$  and let  $u_h$  be any collocation solution satisfying  $Lu_h(x_j) = f(x_j)$ . Assume that the collocation solution is constructed from a kernel  $\Phi$  which satisfies (4) for some  $\sigma > t/2 + 1/2$  for the standard method, or  $\sigma > t + 1/2$  for the norm-minimal method. Then, for the standard collocation method

$$\sup_{u \in H^{\sigma+t/2}} \frac{\|u - u_h\|_{L_2}}{\|u\|_{H^{\sigma+t/2}}} \geq Ch_X^{\sigma-t/2},$$

while for the norm-minimal collocation method

$$\sup_{u \in H^\sigma} \frac{\|u - u_h\|_{L_2}}{\|u\|_{H^\sigma}} \geq Ch_X^{\sigma-t}.$$

**Proof.** Let  $f(x) = \sin^2(mx) = \frac{1}{2}(1 - \cos 2mx)$ ,  $m \in \mathbb{N}$ , and let  $k = 2m$ . The pseudo-differential operator  $L$  of order  $t$  is defined by its effect on the eigenfunctions,

$$L(\cos(\ell x)) = (1 + \ell)^t \cos(\ell x), \quad L(\sin(\ell x)) = (1 + \ell)^t \sin(\ell x).$$

Then the pseudo-differential equation  $Lu = f$  admits the exact solution

$$u(x) = \frac{1}{2} \left( 1 - \frac{\cos(kx)}{(1+k)^t} \right).$$

For the chosen collocation points we have  $f(x_j) = 0$  and hence the approximate solution  $u_h$  is identically zero. It is easily seen that the mesh norm of  $X$  is  $h_X = \pi/k$ . We therefore have

$$\|u - u_h\|_{L_2}^2 = \|u\|_{L_2}^2 = \frac{\pi}{2} + \frac{\pi}{4(1+k)^{2t}} \geq \frac{\pi}{2}.$$

The  $H^\sigma$  norm of  $u$  can also be computed,

$$\|u\|_{H^\sigma}^2 = \frac{\pi}{2} + \frac{\pi}{4}(1+k)^{2(\sigma-t)}.$$

For the norm-minimal case we have  $\sigma > t$ , and hence  $\|u\|_{H^\sigma}^2 \leq ck^{2(\sigma-t)}$ , and

$$\frac{\|u - u_h\|_{L_2}}{\|u\|_{H^\sigma}} = \frac{\|u\|_{L_2}}{\|u\|_{H^\sigma}} \geq ck^{-(\sigma-t)} = Ch_X^{\sigma-t}.$$

Similarly, the  $H^{\sigma+t/2}$  norm of  $u$  can also be computed,

$$\|u\|_{H^{\sigma+t/2}}^2 = \frac{\pi}{2} + \frac{\pi}{4}(1+k)^{2\sigma-t}.$$

Thus, for the standard collocation case we have (since  $\sigma > t/2$ )  $\|u\|_{H^{\sigma+t/2}}^2 \leq ck^{2\sigma-t}$ , and hence

$$\frac{\|u - u_h\|_{L_2}}{\|u\|_{H^{\sigma+t/2}}} = \frac{\|u\|_{L_2}}{\|u\|_{H^{\sigma+t/2}}} \geq ck^{-(\sigma-t/2)} = Ch_X^{\sigma-t/2}.$$

□

#### 4 Multiscale Collocation for Solving PDEs

While the (fixed) single scale approach for solving PDEs via collocation yields good approximation in both the standard and the norm-minimizing approaches, it suffers from two major drawbacks. On the one hand, the condition number grows rapidly with decreasing fill distance. On the other hand, even when using compactly supported RBFs, the matrices quickly become dense, and as a result the computational cost becomes prohibitive.

Recently (see [15, 25]), in the case of interpolation, a multiscale technique has been proven to have the advantages of good approximation and computational efficiency. We are now going to analyze how this method can be carried over to solving PDEs. We know already from numerical examples that, at least in the boundary value PDE case (see [1]), the approach has to be modified to be successful. The theory below will guide us to an appropriate modification for PDEs on the sphere.

The general idea of the multiscale approach can be described as follows.

We start with a widely spread set of points  $X_1$  and use a basis function with a large scale  $\delta_1$  to recover the global behavior of the solution  $u$ , by solving the collocation equation  $Ls_1|_{X_1} = f|_{X_1}$ . We then set  $u_1 = s_1$  as the first approximation, so that the residual at the first step is  $f_1 = f - Ls_1$ . To reduce the residual, at the next step we use a finer set of points  $X_2$  and a smaller scale  $\delta_2$ , and compute a correction  $s_2$  from an appropriate finite dimensional space by solving  $Ls_2|_{X_2} = f_1|_{X_2}$ . We then obtain a new approximation  $u_2 = u_1 + s_2$ , so that the new residual is  $f_2 = f_1 - Ls_2$ ; and so on. Stated as an algorithm, this takes the following form. We first set  $u_0 = 0$  and  $f_0 = f$ . Then, we do for  $j = 1, 2, \dots$ :

- Determine a correction  $s_j$  as the solution in a prescribed finite-dimensional space of

$$Ls_j(\mathbf{x}) = f_{j-1}(\mathbf{x}) \text{ for all } \mathbf{x} \in X_j.$$

- Update the solution and the residual according to

$$\begin{aligned} u_j &= u_{j-1} + s_j \\ f_j &= f_{j-1} - Ls_j. \end{aligned}$$

As a consequence, for  $j \geq 1$ , we have

$$Lu_j + f_j = Lu_{j-1} + f_{j-1} = \dots = Lu_0 + f_0 = f.$$

Hence, the residual at level  $j$  is  $f_j = f - Lu_j$ . Since  $L$  is injective, let  $e_j := L^{-1}f_j$ , for  $j \geq 1$ . We note that

$$e_j = L^{-1}f_j = L^{-1}f - u_j = u - u_j,$$

thus  $e_j$  is the error at step  $j$ , and also we have

$$e_j = e_{j-1} - s_j. \quad (25)$$

In the following, we will analyze this multiscale algorithm using either standard collocation or norm-minimal collocation for the local reconstruction step. The proofs are similar to each other and follow proofs from [15, 25] but require careful consideration of the details.

#### 4.1 Standard Multiscale Collocation

We begin with standard collocation for the local reconstruction. Hence, the setting is as follows. Suppose  $X_1, X_2, \dots$  is a sequence of point sets on  $\mathbb{S}^n$  with decreasing mesh norms  $h_1, h_2, \dots$  respectively. For every  $j = 1, 2, \dots$  we choose a basis function  $\Phi_j = \Phi_{\delta_j}$ , where  $\delta_j$  is a scaling parameter depending on  $h_j$ . Specifically, we will choose

$$\delta_j = \nu h_j^{1-t/(2\sigma)}$$

for some fixed constant  $\nu > 1$ , which means that  $\delta_j/h_j = \nu h_j^{-t/(2\sigma)}$  is not a constant but grows mildly with  $h_j$  decreasing to zero. We also define for each  $j = 1, 2, \dots$  finite dimensional spaces

$$V_j = \text{span} \{ \Phi_j(\cdot, \mathbf{x}) : \mathbf{x} \in X_j \}.$$

Hence, we pick the local solution  $s_j$  from  $V_j$  such that  $Ls_j(\mathbf{x}) = f_{j-1}(\mathbf{x})$  for all  $\mathbf{x} \in X_j$ . This means that  $s_j \in V_j$  is the standard collocation approximation to  $Le_{j-1} = f_{j-1}$  on  $X_j$  using the kernel  $\Phi_j$ .

To analyze the convergence of the algorithm, we introduce, as in the case of the single scale method, kernels  $\Psi_j = L^{-1}\Phi_j$ .

With this notation, we are able to formulate and prove our first convergence result.

**Theorem 3** Assume that  $u \in H^{\sigma+t/2}$  is the exact solution of (11) with  $\sigma > (t+n)/2$ . Suppose that  $X_1, X_2, \dots$  is a sequence of point sets on  $\mathbb{S}^n$  with decreasing mesh norms  $h_1, h_2, \dots$  respectively. The mesh norms are assumed to satisfy  $\gamma\mu \leq h_{j+1}/h_j \leq \mu$  for some fixed constants  $\gamma$  and  $\mu$  in  $(0, 1)$ . Let  $\Phi$  be a kernel satisfying (4) and let  $\Phi_j = \Phi_{\delta_j}$  be a sequence of scaled kernels, where the scales are defined by  $\delta_j = (h_j/\mu)^{1-t/(2\sigma)}$ . Let  $\Psi_j = L^{-1}\Phi_j$ . Then there exists a constant  $C$  independent of  $\mu, j$  and  $f$  such that

$$\|u - u_j\|_{\Psi_{j+1}} \leq \beta \|u - u_{j-1}\|_{\Psi_j} \quad \text{for } j = 1, 2, \dots,$$

with  $\beta = C\mu^{2\sigma-t}$ , and hence there exists  $c > 0$  such that

$$\|u - u_k\|_{L_2} \leq c\beta^k \|u\|_{H^{\sigma+t/2}} \quad \text{for } k = 1, 2, \dots$$

Thus the standard multiscale collocation  $u_k$  converges linearly to  $u$  in the  $L_2$  norm if  $\mu < C^{-1/(2\sigma-t)}$ .

**Proof.** From (18) and (7) we have, with  $e_j = u - u_j$ ,

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

Since  $\delta_{j+1}\ell \leq 1$  in the first term, we have

$$S_1 \leq 2^{2\sigma} \|L^{1/2} e_j\|_{L_2}^2.$$

We note that  $s_j \in V_j$  is the approximate solution with the standard collocation method of  $Le_{j-1} = f_{j-1}$ . Thus by (25) and Theorem 1 and Lemma 3 we have

$$\begin{aligned} \|L^{1/2} e_j\|_{L_2} &= \|L^{1/2} (e_{j-1} - s_j)\|_{L_2} \leq ch_j^{\sigma-t/2} \|e_{j-1}\|_{H^{\sigma+t/2}} \\ &\leq ch_j^{\sigma-t/2} \delta_j^{-\sigma} \|e_{j-1}\|_{\Psi_j} = c\mu^{\sigma-t/2} \|e_{j-1}\|_{\Psi_j}, \end{aligned}$$

and hence

$$S_1 \leq c' 2^{2\sigma} \mu^{2\sigma-t} \|e_{j-1}\|_{\Psi_j}^2,$$

where we have used  $\delta_j = (h_j/\mu)^{1-t/(2\sigma)}$ .

For the second sum  $S_2$ , we note that

$$\delta_{j+1}/\delta_j = (h_{j+1}/h_j)^{1-t/(2\sigma)} \leq \mu^{1-t/(2\sigma)},$$

and since  $\delta_{j+1}\ell > 1$  we have

$$(1 + \delta_{j+1}\ell)^{2\sigma} < (2\delta_{j+1}\ell)^{2\sigma} \leq (2\mu^{1-t/(2\sigma)} \delta_j \ell)^{2\sigma} \leq 2^{2\sigma} \mu^{2\sigma-t} (1 + \delta_j \ell)^{2\sigma}.$$

Therefore, from (7) and (18)

$$S_2 \leq c_2 2^{2\sigma} \mu^{2\sigma-t} \|e_j\|_{\Psi_j}^2 = c_2 2^{2\sigma} \mu^{2\sigma-t} \|e_{j-1} - s_j\|_{\Psi_j}^2 \leq c_2 2^{2\sigma} \mu^{2\sigma-t} \|e_{j-1}\|_{\Psi_j}^2,$$

where in the last step we used Lemma 4 with  $u$  replaced by  $e_{j-1}$  and  $\Psi$  by  $\Psi_j$ . Therefore

$$\|e_j\|_{\Psi_{j+1}}^2 \leq \frac{2^{2\sigma}}{c_1} (c' + c_2) \mu^{2\sigma-t} \|e_{j-1}\|_{\Psi_j}^2$$

So if we write  $\beta = C\mu^{\sigma-t/2}$  where  $C := 2^\sigma(c' + c_2)^{1/2}/c_1^{1/2}$  then

$$\|e_j\|_{\Psi_{j+1}} \leq \beta \|e_{j-1}\|_{\Psi_j}. \quad (26)$$

Using (25), Theorem 1 and Lemma 3 and then repeating (26)  $k$  times gives

$$\begin{aligned} \|u - u_k\|_{L_2} &= \|e_k\|_{L_2} = \|e_{k-1} - s_k\|_{L_2} \leq ch_k^{\sigma-t/2} \|e_{k-1} - s_k\|_{H^{\sigma+t/2}} \\ &= ch_k^{\sigma-t/2} \|e_k\|_{H^{\sigma+t/2}} \\ &\leq ch_k^{\sigma-t/2} \delta_{k+1}^{-\sigma} \|e_k\|_{\Psi_{k+1}} \\ &\leq c \|e_k\|_{\Psi_{k+1}} \\ &\leq c\beta^k \|u\|_{\Psi_1} \leq c\beta^k \|u\|_{H^{\sigma+t/2}}, \end{aligned}$$

where we have used the fact that

$$h_k^{\sigma-t/2} \delta_{k+1}^{-\sigma} = (\mu h_k / h_{k+1})^{\sigma-t/2} \leq \gamma^{t/2-\sigma}.$$

□

## 4.2 Norm-minimal Multiscale Collocation

We will now analyze the multiscale method using norm-minimal collocation in the local reconstruction step. Again, we have a sequence of point sets  $X_1, X_2, \dots$  on  $\mathbb{S}^n$  with decreasing mesh norms  $h_1, h_2, \dots$ . For every  $j = 1, 2, \dots$  we choose a scaled basis function  $\Phi_j = \Phi_{\delta_j}$ , where  $\delta_j$  is a scaling parameter depending on  $h_j$ . This time, however, we choose

$$\delta_j = \nu h_j^{1-t/\sigma}$$

for some fixed constant  $\nu > 1$ , which means that  $\delta_j/h_j = \nu h_j^{-t/\sigma}$  is again not a constant but grows mildly with  $h_j$  decreasing to zero. Again, in a similar way to the single scale case, we also define, for  $j = 1, 2, \dots$ , finite dimensional spaces

$$W_j = \text{span} \{L\Phi_j(\cdot, \mathbf{x}) : \mathbf{x} \in X_j\}$$

and pick the local reconstruction  $s_j$  from  $W_j$  as the norm-minimal collocation solution to  $Le_{j-1} = f_{j-1}$  based on the set  $X_j$  and the kernel  $\Phi_j$ .

Our convergence result this time takes the following form.

**Theorem 4** *Assume that  $u \in H^\sigma$  is the exact solution of (11) with  $\sigma > t + n/2$ . Let  $X_1, X_2, \dots$  be a sequence of point sets on  $\mathbb{S}^n$  with mesh norms  $h_1, h_2, \dots$  satisfying  $\gamma\mu h_j \leq h_{j+1} \leq \mu h_j$  for all  $j = 1, 2, \dots$ , for some fixed  $\mu \in (0, 1)$  and  $\gamma \in (0, 1)$ .*

*Let  $\delta_j = (h_j/\mu)^{1-t/\sigma}$ , for  $j = 1, 2, \dots$ , be a sequence of scale factors. Let  $\Phi_j = \Phi_{\delta_j}$  be a kernel satisfying (7). Then there exists a constant  $C$  independent of  $\mu, j$  and  $f$  such that*

$$\|u - u_j\|_{\Phi_{j+1}} \leq \alpha \|u - u_{j-1}\|_{\Phi_j} \quad \text{for } j = 1, 2, \dots,$$

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

$$\|u - u_k\|_{L_2} \leq c\alpha^k \|u\|_{H^\sigma} \text{ for } k = 1, 2, \dots$$

Thus the norm-minimal multiscale collocation  $u_k$  converges linearly to  $u$  in the  $L_2$  norm if  $\mu < C^{-1/(\sigma-t)}$ .

**Proof.** From (10) and (7) we have, with  $e_j = u - u_j$ ,

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

Since  $\delta_{j+1}\ell \leq 1$  in the first term, we have

$$S_1 \leq 2^{2\sigma} \|e_j\|_{L_2}^2.$$

We note that  $s_j \in W_j$  is the approximate solution of  $Le_{j-1} = f_{j-1}$  with the norm-minimal collocation method. Hence, by using Theorem 2 and Lemma 1 we can conclude that

$$\begin{aligned} \|e_j\|_{L_2} &= \|e_{j-1} - s_j\|_{L_2} \leq ch_j^{\sigma-t} \|e_{j-1}\|_{H^\sigma} \\ &\leq 2^\sigma c_2^{1/2} ch_j^{\sigma-t} \delta_j^{-\sigma} \|e_{j-1}\|_{\Phi_j} = c2^\sigma \mu^{\sigma-t} \|e_{j-1}\|_{\Phi_j}, \end{aligned}$$

where in the last step we have used  $\delta_j = (h_j/\mu)^{1-t/\sigma}$ . This means that

$$S_1 \leq 2^{2\sigma} c' \mu^{2\sigma-2t} \|e_{j-1}\|_{\Phi_j}^2.$$

For  $S_2$ , note that  $\delta_{j+1}/\delta_j = (h_{j+1}/h_j)^{1-t/\sigma} \leq \mu^{1-t/\sigma}$ . Since  $\delta_{j+1}\ell > 1$  we have

$$(1 + \delta_{j+1}\ell)^{2\sigma} < (2\delta_{j+1}\ell)^{2\sigma} \leq (2\mu^{1-t/\sigma} \delta_j \ell)^{2\sigma} \leq 2^{2\sigma} \mu^{2\sigma-2t} (1 + \delta_j \ell)^{2\sigma}.$$

Thus, we have the upper bound

$$S_2 \leq c_2 2^{2\sigma} \mu^{2\sigma-2t} \|e_j\|_{\Phi_j}^2 = c_2 2^{2\sigma} \mu^{2\sigma-2t} \|e_{j-1} - s_j\|_{\Phi_j}^2 \leq c_2 2^{2\sigma} \mu^{2\sigma-2t} \|e_{j-1}\|_{\Phi_j}^2,$$

where in the last step we used Lemma 5 with  $\Phi$  replaced by  $\Phi_j$ .

Therefore

$$\|e_j\|_{\Phi_{j+1}}^2 \leq \frac{2^{2\sigma}}{c_1} (c' + c_2) \mu^{2\sigma-2t} \|e_{j-1}\|_{\Phi_j}^2$$

Hence, if we choose  $\alpha = C\mu^{\sigma-t}$  where  $C = 2^\sigma (c' + c_2)^{1/2} / c_1^{1/2}$  then

$$\|e_j\|_{\Phi_{j+1}} \leq \alpha \|e_{j-1}\|_{\Phi_j}. \quad (27)$$

Using (25), Theorem 2, Lemma 1 and then (27) repeatedly  $k$  times, gives

$$\begin{aligned} \|u - u_k\|_{L_2} &= \|e_k\|_{L_2} = \|e_{k-1} - s_k\|_{L_2} \\ &\leq ch_k^{\sigma-t} \|e_{k-1} - s_k\|_{H^\sigma} = ch_k^{\sigma-t} \|e_k\|_{H^\sigma} \\ &\leq ch_k^{\sigma-t} \delta_{k+1}^{-\sigma} \|e_k\|_{\Phi_{k+1}} \\ &\leq c\gamma^{\tau-\sigma} \|e_k\|_{\Phi_{k+1}} \\ &\leq c\alpha^k \|u\|_{\Phi_1} \leq c\alpha^k \|u\|_{H^\sigma}. \end{aligned}$$

□

### 4.3 Condition Numbers of Collocation Matrices

In each step of the multiscale algorithm, we have to solve a linear system arising from the collocation condition on a set  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ :

$$\mathbf{A}_\delta \mathbf{b} = \mathbf{f}, \quad (28)$$

where the collocation matrix  $\mathbf{A}_\delta$  is the collocation matrix with entries either  $L\Phi_\delta(\mathbf{x}_i, \mathbf{x}_j)$  (standard collocation) or  $LL\Phi_\delta(\mathbf{x}_i, \mathbf{x}_j)$  (norm-minimal collocation).

Since the matrix  $\mathbf{A}_\delta$  is symmetric and positive definite in both cases, an iterative method such as the conjugate gradient method can be used to solve (28) efficiently. The complexity of the conjugate gradient method depends on the condition number of the matrix  $\mathbf{A}_\delta$  and on the cost of a matrix-vector multiplication.

The collocation equation (13) can be viewed as an interpolation problem with the kernel  $L\Phi_\delta(\mathbf{x}, \mathbf{y})$  (or  $LL\Phi_\delta(\mathbf{x}, \mathbf{y})$  in the norm-minimal case). It is well known, see for example [24, Section 12.2] that the lower bound of the interpolation matrix depends on the smoothness of the kernel and the separation radius  $q_X$  of the set  $X$ ,

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

where  $\theta(\mathbf{x}, \mathbf{y}) := \cos^{-1}(\langle \mathbf{x}, \mathbf{y} \rangle)$  is the geodesic distance between two points  $\mathbf{x}$  and  $\mathbf{y}$  on the unit sphere  $\mathbb{S}^n$ . This geodesic separation radius is comparable to the Euclidean separation radius  $\tilde{q}_X$  of the set  $X$  when being viewed as a subset of  $\mathbb{R}^{n+1}$ ,

$$\tilde{q}_X := \frac{1}{2} \min_{i \neq j} |\mathbf{x}_i - \mathbf{x}_j|.$$

We can use a result for condition numbers of multiscale interpolation [15, Theorem 7.3] to arrive at the following conclusion.

**Theorem 5** *The condition number  $\bar{\kappa}(\mathbf{A}_\delta)$  in the standard approach is bounded by*

$$\bar{\kappa}(\mathbf{A}_\delta) \leq C \left( \frac{\delta}{\tilde{q}_X} \right)^{1+2\sigma-t}, \quad (29)$$

while the condition number  $\tilde{\kappa}(\mathbf{A}_\delta)$  of the collocation matrix  $\mathbf{A}_\delta$  in the norm-minimal approach is bounded by

$$\tilde{\kappa}(\mathbf{A}_\delta) \leq C \left( \frac{\delta}{\tilde{q}_X} \right)^{1+2(\sigma-t)}. \quad (30)$$

**Proof.** The kernel  $L\Phi(\mathbf{x}, \mathbf{y})$  can be expanded as

$$L\Phi(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n,\ell)} \widehat{L}(\ell) \widehat{\phi}(\ell) Y_{\ell k}(\mathbf{x}) Y_{\ell k}(\mathbf{y}).$$

Using the assumptions (4) and (12) on the unscaled kernel  $\Phi$  and the differential operator  $L$  we obtain  $\widehat{L}(\ell) \widehat{\phi}(\ell) \sim (1+\ell)^{-2\sigma+t}$ . (Here,  $A \sim B$  means that there are two positive constants  $c$  and  $c'$  such that  $cB \leq A \leq c'B$ ). Thus we can apply [15, Theorem 7.3] with  $2\tau = 1 + 2\sigma - t$  to derive (29).

Similarly, using the Fourier expansion of  $LL\Phi(\mathbf{x}, \mathbf{y})$ , since  $[\widehat{L}(\ell)]^2 \widehat{\phi}(\ell) \sim (1+\ell)^{-2\sigma+2t}$  we can apply [15, Theorem 7.3] with  $2\tau = 1 + 2\sigma - 2t$  to derive (30).  $\square$



This indicates that a choice of  $\delta$  proportional to  $q_X$  would lead to a level independent condition number. However, to derive convergence, we are not allowed to choose  $\delta$  proportional to  $q_X$ . The specific choices of  $\delta$  in our situation lead to the following result. Here  $q_j := q_{X_j}$ .

**Corollary 1** *In the standard approach, the choice  $\delta_j = \nu h_j^{1-t/(2\sigma)}$  leads to a level-dependent condition number of the form*

$$\bar{\kappa}_j \leq C \left( \frac{h_j}{q_j} \right)^{1+2\sigma-t} h_j^{-\frac{t}{2\sigma}(1+2\sigma-t)},$$

which, in the case of quasi-uniform data sets and  $t \geq 1$  reduces to

$$\bar{\kappa}_j \leq C h_j^{-\frac{t}{2\sigma}(1+2\sigma-t)} \leq C h_j^{-t}.$$

*In the norm-minimal approach, the choice  $\delta_j = \nu h_j^{1-t/\sigma}$  leads to a level-dependent condition number of the form*

$$\tilde{\kappa}_j \leq C \left( \frac{h_j}{q_j} \right)^{1+2(\sigma-t)} h_j^{-\frac{t}{\sigma}(1+2(\sigma-t))},$$

which, in the case of quasi-uniform data sets and  $t \geq 1/2$  reduces to

$$\tilde{\kappa}_j \leq C h_j^{-\frac{t}{\sigma}(1+2(\sigma-t))} \leq C h_j^{-2t}.$$

It is important to see that, though the condition number grows with  $1/h_j$ , the order of this growth is bounded by the order of the operator (or twice the order of the operator in the norm-minimal case). This has to be compared to the larger order of growth of a single scale collocation method, which follows from Theorem 5 by setting  $\delta$  to constant:

$$\kappa(\mathbf{A}_\delta) \leq C h_X^{-(1+2\sigma-t)}.$$

Thus for the single-scale method and quasi-uniformity the order of growth of the condition number is larger by a factor of  $2\sigma/t > 1 + n/t$  (since  $\sigma > (t+n)/2$ ) for the standard case, and larger by a factor of  $\sigma/t > 1 + n/(2t)$  for the norm-minimal case (since  $\sigma > t + n/2$ ).

## 5 Numerical Experiments

In this section, we consider the following PDE of order  $t = 2$  on the unit sphere  $\mathbb{S}^2 \subset \mathbb{R}^3$ :

$$Lu(\mathbf{x}) := -\Delta^* u(\mathbf{x}) + \omega^2 u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{S}^2,$$

where  $\Delta^*$  is the Laplace-Beltrami operator on  $\mathbb{S}^2$  and  $\omega$  is a positive constant. The PDE arises from discretizing the heat equation on the sphere.

Let  $\omega = 1$  and let  $f$  be defined so that the exact solution is given by the Franke function [7] defined on the unit sphere  $\mathbb{S}^2$ . To be more precise, let

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

**Table 1** Mesh norms and separation radii of sets of point

	$X_1$	$X_2$	$X_3$
$M$	1500	6000	24000
$h_X$	0.0647	0.0325	0.0162
$q_X$	0.0423	0.0212	0.0106

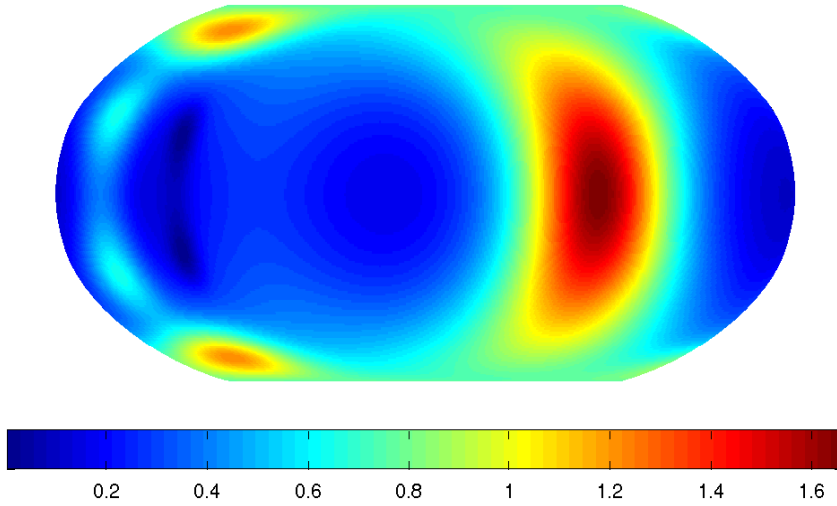
Then we define

$$u(\mathbf{x}) = 0.75 \exp\left(-\frac{(9x-2)^2 + (9y-2)^2}{4}\right) + 0.75 \exp\left(-\frac{(9x+1)^2}{49} - \frac{9y+1}{10}\right) \\ + 0.5 \exp\left(-\frac{(9x-7)^2 + (9y-3)^2}{4}\right) - 0.2 \exp(-(9x-4)^2 - (9y-7)^2)$$

and compute the function  $f$  via the formula

$$f(\mathbf{x}(\theta, \phi)) = -\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial u}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2 u}{\partial \phi^2} + \omega^2 u(\mathbf{x}(\theta, \phi)).$$

A plot of the exact solution  $u$  is given in Figure 1. Even though the algorithm allows the

**Fig. 1** Exact solution

collocation points to be scattered freely on the sphere, choosing sets of collocation points distributed roughly uniformly over the whole sphere significantly improves the quality of the approximate solutions and condition numbers. To this end, the sets of points used to construct the approximate solutions are generated using the equal area partitioning algorithm [21,20]. The mesh norms and separation radii of these sets are listed in Table 1. The RBF

used is

$$\psi(r) = (1-r)_+^6(3+18r+35r^2), \quad \psi_\delta(r) = \delta^{-2}\psi(r/\delta),$$

and

$$\Phi_\delta(\mathbf{x}, \mathbf{y}) = \psi_\delta(|\mathbf{x} - \mathbf{y}|) = \psi_\delta(\sqrt{2 - 2\mathbf{x} \cdot \mathbf{y}}).$$

It can be shown that  $\Phi_\delta$  is a kernel which satisfies condition (7) with  $\sigma = 7/2$  ([18]).

The kernel  $\Phi_\delta$  is a zonal function, i.e.  $\Phi_\delta(\mathbf{x}, \mathbf{y}) = \phi_\delta(\mathbf{x} \cdot \mathbf{y})$  where  $\phi_\delta(t)$  is a univariate function. For zonal functions, the Laplace-Beltrami operator can be computed via

$$\Delta^* \phi_\delta(\mathbf{x} \cdot \mathbf{y}) = \mathcal{L} \phi_\delta(t), \quad t = \mathbf{x} \cdot \mathbf{y},$$

where

$$\mathcal{L} = \frac{d}{dt}(1-t^2) \frac{d}{dt}$$

In our case,

$$\mathcal{L} \phi_\delta(t) = \frac{112}{\delta^{10}} (\sqrt{2-2t} - \delta)^4 \left( 25t^2 - 10t + \delta^2 t - 15 + \frac{-8\delta t^2 + 8\delta t}{\sqrt{2-2t}} \right).$$

At each level the normalized  $L_2$  error  $\|e_j\|$  is approximated by an  $\ell_2$  error, thus in principle we define

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

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

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

where  $\mathcal{G}$  is a longitude-latitude grid containing the centers of rectangles of size 1 degree times 1 degree and  $|\mathcal{G}| = 180 \cdot 360 = 64800$ . We also record the condition numbers  $\bar{\kappa}_j$  of the collocation matrix

$$\mathbf{A}_{\delta_j} = [L\Phi_j(\mathbf{x}, \mathbf{y})]_{\mathbf{x}, \mathbf{y} \in X_j}$$

for the standard approach and the condition numbers  $\tilde{\kappa}_j$  of the collocation matrix

$$\mathbf{A}_{\delta_j} = [LL\Phi_j(\mathbf{x}, \mathbf{y})]_{\mathbf{x}, \mathbf{y} \in X_j}$$

for the norm-minimal approach. The errors and condition numbers of the collocation matrices at each step of the multiscale algorithm for two variants of the collocation method are listed in Tables 2 and 3, respectively. In the upper part of each table we use the results for the scale  $\delta_j$  taken proportional to  $h_j$ , whereas in the lower part we use the scale in accordance with Theorem 3 or 4. As can be seen from the tables, if the scaling parameters  $\delta_j$  decrease linearly with respect to the mesh norms  $h_j$ , we may not get a good convergence rate, at least in the norm-minimal case, whereas in both cases we get a good convergence rate if we follow the theoretical predictions. Figure 2 shows the approximate solutions using the standard approach at each level corresponding to  $\delta_j = 2; 1.2230; 0.7438$ , in accordance with Theorem 3. If we use one-shot approximation on the final set of 24000 points with various scales  $\delta$  then we obtain the errors listed in Tables 4 and 5. As can be seen from the tables, the multiscale approach provides a more accurate approximation with a collocation matrix of a smaller condition number.

**Table 2** The approximation errors and condition numbers of multiscale approximation using the standard approach

Level	1	2	3
$M$	1500	6000	24000
$\delta_j$	2.0	1.0	0.5
$\ e_j\ $	2.1173E-04	3.9021E-06	1.1509E-07
$\tilde{\kappa}_j$	4.4028E+04	4.4172E+04	4.4562E+04
$\delta_j$	2.0000	1.2230	0.7438
$\ e_j\ $	2.1173E-04	3.8357E-06	1.0944E-07
$\tilde{\kappa}_j$	4.4028E+04	1.2154E+05	3.2279E+05

**Table 3** The approximation errors and condition numbers of multiscale approximation using the norm-minimal approach

Level	1	2	3
$M$	1500	6000	24000
$\delta_j$	2.0	1.0	0.5
$\ e_j\ $	5.1048E-02	3.4713E-02	3.3648E-02
$\tilde{\kappa}_j$	1.7471E+02	2.1987E+02	3.1605E+02
$\delta_j$	2.0000	1.4820	1.1033
$\ e_j\ $	5.1048E-02	7.2364E-03	2.6972E-04
$\tilde{\kappa}_j$	1.7471E+02	5.7738E+02	1.8929E+03

**Table 4** Errors by one-shot approximation of  $u$  with various scales using the final set of 24000 points (standard approach)

$\delta$	2.0000	1.2230	1.0000	0.7438	0.5000
$\ e\ $	1.1372E-07	2.8564E-07	1.0765E-06	8.4260E-06	1.3642E-04
$\tilde{\kappa}$	4.6265E+07	3.9109E+06	1.4203E+06	3.2273E+05	4.4562E+04

**Table 5** Errors by one-shot approximation of  $u$  with various scales using the final set of 24000 points (norm-minimal approach)

$\delta$	2.000	1.4820	1.1033	1.000	0.500
$\ e\ $	3.9640E-04	8.0738E-03	6.2875E-02	1.1298E-01	5.7892E-01
$\tilde{\kappa}$	1.1200E+04	4.6228E+03	1.8927E+03	1.8927E+03	3.1605E+02

## 6 Concluding remarks

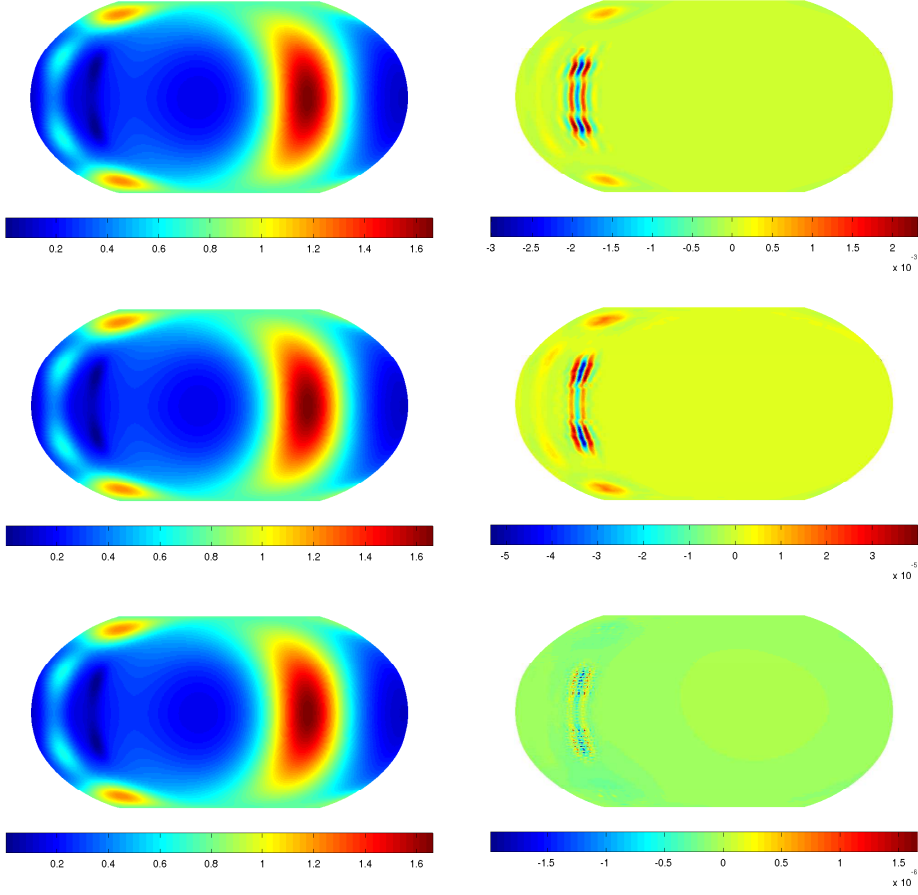
In this paper we have developed a multiscale approximation for elliptic PDEs on spheres, by adapting the method and the proof of the multiscale approximation scheme in [15].

A natural question is whether the scheme so developed is equivalent to applying the multiscale approximation scheme to the right hand side of equation (11), with an appropriate redefinition of the kernel.

For the single-scale collocation schemes with  $\delta = 1$  there is indeed a close connection, in that the standard and symmetric collocation schemes of Sections 3.1 and 3.2 can be re-expressed this way: the standard choice is equivalent to

$$f \approx f^{\text{stand}} := \sum_{i=1}^N b_i L \Phi(\cdot, \mathbf{x}_i), \quad f^{\text{stand}}(\mathbf{x}_j) = f(\mathbf{x}_j), \quad 1 \leq j \leq N,$$

$$u_h^{\text{stand}} := L^{-1} f^{\text{stand}} = \sum_{i=1}^N b_i \Phi(\cdot, \mathbf{x}_i),$$



**Fig. 2** The left column are approximate solutions  $u_j$ ,  $j = 1, 2, 3$ . The right column are the details  $s_j = u_j - u_{j-1}$ ,  $j = 2, 3$  and the errors  $u_3 - u_1$

and the symmetric choice to

$$f \approx f^{\text{symm}} := \sum_{i=1}^N b_i L^2 \Phi(\cdot, \mathbf{x}_i), \quad f^{\text{symm}}(\mathbf{x}_j) = f(\mathbf{x}_j), \quad 1 \leq j \leq N,$$

$$u_h^{\text{symm}} := L^{-1} f^{\text{symm}} = \sum_{i=1}^N b_i L \Phi(\cdot, \mathbf{x}_i).$$

But for the multiscale versions there is no such simple correspondence. The multiscale approximation in [15] is based upon scaled versions of a single kernel, scaled as in (5). Suppose we approximate the right hand side of (11) by the multiscale approximation, using the kernel (at face value appropriate for the standard case)

$$\zeta(\mathbf{x}, \mathbf{y}) := L \Phi(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(n, \ell)} \widehat{L}(\ell) \widehat{\phi}(\ell) Y_{\ell k}(\mathbf{x}) Y_{\ell k}(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^n,$$

We can define a scaled version of the new kernel after extending the definition to  $\mathbb{R}^{n+1}$ , but the difficulty is that the scaled kernel  $\zeta_\delta$  is not related in any simple way to  $L\Phi_\delta$ , even for the case  $L = -\Delta + 1$ . To state the problem differently, an effective multiscale method based on the scaled kernel  $\zeta_\delta$  would require an ability to solve analytically the equation  $Lg = \zeta_\delta(\cdot, \mathbf{x}_i)$ , something that does not seem possible in general.

The multiscale method presented in Section 4 is therefore significantly different from the multiscale approximation in [15].

**Acknowledgements** The support of the Australian Research Council is gratefully acknowledged.

## References

1. G. E. Fasshauer. Solving differential equations with radial basis functions: multilevel methods and smoothing. *Advances in Comput. Math.*, 11:139–159, 1999.
2. G. E. Fasshauer and J. W. Jerome. Multistep approximation algorithms: improved convergence rates through postconditioning with smooth kernels. *Advances in Comput. Math.*, 10:1–27, 1999.
3. N. Flyer and G. Wright. Transport schemes on a sphere using radial basis functions. *J. Comp. Phys.*, 226:1059–1084, 2007.
4. N. Flyer and G. Wright. A radial basis function method for the shallow water equations on a sphere. *Proc. R. Soc. A*, 465:1949–1976, 2009.
5. C. Franke and R. Schaback. Convergence order estimates of meshless collocation methods using radial basis functions. *Adv. in Comp. Math.*, 8:381–399, 1998.
6. C. Franke and R. Schaback. Solving partial differential equations by collocation using radial basis functions. *Appl. Math. Comput.*, 93:73–82, 1998.
7. R. Franke. A critical comparison of some methods for interpolation of scattered data. Technical Report NPS-53-79-003, Naval Postgraduate School, 1979.
8. P. Giesl and H. Wendland. Meshless collocation: Error estimates with application to dynamical systems. *SIAM J. Num. Analysis*, 45:1723–1741, 2007.
9. Y. C. Hon and X. Z. Mao. An efficient numerical scheme for Burgers’ equation. *Appl. Math. Comput.*, 95:37–50, 1998.
10. Y. C. Hon and R. Schaback. On unsymmetric collocation by radial basis functions. *Appl. Math. Comput.*, 119:177–186, 2001.
11. E. J. Kansa. Multiquadrics - A scattered data approximation scheme with applications to computational fluid-dynamics i. *Comput. Math.*, 19:127–145, 1990.
12. E. J. Kansa. Multiquadrics - A scattered data approximation scheme with applications to computational fluid-dynamics ii: solutions to parabolic, hyperbolic and elliptic partial differential equations. *Comput. Math.*, 19:147–161, 1990.
13. Q. T. Le Gia. Galerkin approximation for elliptic PDEs on spheres. *J. Approx. Theory*, 130:123–147, 2004.
14. Q. T. Le Gia, F. J. Narcowich, J. D. Ward, and H. Wendland. Continuous and discrete least-square approximation by radial basis functions on spheres. *J. Approx. Theory*, 143:124–133, 2006.
15. Q. T. Le Gia, I. H. Sloan, and H. Wendland. Multiscale analysis in Sobolev spaces on the sphere. *SIAM J. Numerical Analysis*, 48:2065–2090, 2010.
16. T. M. Morton and M. Neamtu. Error bounds for solving pseudodifferential equations on spheres. *J. Approx. Theory*, 114:242–268, 2002.
17. C. Müller. *Spherical Harmonics*, volume 17 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1966.
18. F. J. Narcowich and J. D. Ward. Scattered data interpolation on spheres: error estimates and locally supported basis functions. *SIAM J. Math. Anal.*, 33:1393–1410, 2002.
19. H. Power and V. Barraco. A comparison analysis between unsymmetric and symmetric radial basis function collocation methods for the numerical solution of partial differential equations. *Comput. Math. Appl.*, 43:551–583, 2002.
20. E. B. Saff and A. B. J. Kuijlaars. Distributing many points on a sphere. *Math. Intelligencer*, 19:5–11, 1997.
21. E. B. Saff, E. A. Rakhmanov, and Y. M. Zhou. Minimal discrete energy on the sphere. *Mathematical Research Letters*, 1:647–662, 1994.
22. I. J. Schoenberg. Positive definite function on spheres. *Duke Math. J.*, 9:96–108, 1942.

- 
23. H. Wendland. Meshless Galerkin methods using radial basis functions. *Math. Comp.*, 68:1521–1531, 1999.
  24. H. Wendland. *Scattered Data Approximation*. Cambridge University Press, Cambridge, 2005.
  25. H. Wendland. Multiscale analysis in Sobolev spaces on bounded domains. *Numerische Mathematik*, page (to appear), 2010.
  26. Y. Xu and E. W. Cheney. Strictly positive definite functions on spheres. *Proc. Amer. Math. Soc.*, 116:977–981, 1992.