SIAM J. NUMER. ANAL.
Vol. 56, No. 1, pp. 274–295

A PETROV–GALERKIN KERNEL APPROXIMATION ON THE SPHERE*

DAVOUD MIRZAEI†

Abstract. In this paper, a numerical solution of partial differential equations on the unit sphere is given by using a kernel trial approximation in combination with a special Petrov–Galerkin test discretization. The solvability of the scheme is proved, and the error bounds are obtained for functions in appropriate Sobolev spaces. The condition number of the final system is estimated in terms of discretization parameters. The method is meshless because in the trial side the numerical solution parameterizes entirely in terms of scattered points and in the test side everything breaks down to simple numerical integrations over independent spherical caps. This means that no connected background mesh is required for either approximation or integration.

Key words. radial basis functions, spherical basis functions, Petrov–Galerkin method, convolution data, error analysis, condition numbers

AMS subject classifications. 65M12, 65M15, 65M60, 41A17, 41A30

DOI. 10.1137/16M1106626

1. Introduction.

Scattered data approximation in Euclidian spaces has a rather long and rich history. In recent years researchers have become increasingly interested in using tools from approximation theory to develop numerical methods for problems on spheres and other manifolds [11]. Fitting a surface to scattered data arising from sampling an unknown function defined on a manifold, or solving a partial differential equation (PDE) where the underlying domain is a manifold with information available at scattered points, comes up frequently in applied problems. There are applications to geodesy, meteorology, astrophysics, geophysics, and other areas when the underlying manifold is a sphere [7].

Spherical basis functions (SBF) are known as one of the most promising and interesting tools for solving spherical problems [11]. They are closely related to *radial basis functions* (RBF), which are well-established for pure function approximation and PDE problems on regions in Euclidean spaces [3, 32]. There exists a reasonable number of publications for solving data fitting problems on spheres using SBFs [13, 12, 24, 15]. Applications for numerical solution of PDEs can be found, for example, in [16, 18, 14, 21]. See also [2, 25] for spherical spline solutions to PDEs. In this direction, the collocation method is used in [18, 16], and the Gelerkin method is employed in [14, 21, 2, 25]. In Galerkin methods entries in the stiffness matrices have to be numerically computed via quadrature. For instance, the methods used in [2, 25] require spherical triangulation and special quadratures, and the method of [14] requires solving an optimization problem to find the quadrature weights. This will sometimes increase the computational costs.

In this paper we present an alternative approach based on a special Petrov–Galerkin test discretization. The new method utilizes the interpolation by a given

*Received by the editors December 5, 2016; accepted for publication (in revised form) August 17, 2017; published electronically January 23, 2018.
http://www.siam.org/journals/sinum/56-1/M110662.html
Funding: The work of the author was supported by a grant from the Institute for Research in Fundamental Science (95650077).
†Department of Mathematics, University of Isfahan, 81746-73441 Isfahan, Iran (d.mirzaei@sci.ui.ac.ir).

SBF as a *trial function* and local integrated forms against certain test functions on spherical caps as *test functionals*. The Petrov–Galerkin test discretization makes the new technique different from already well-recognized collocation and Galerkin methods. However, the approach itself is different from the usual Petrov–Galerkin method, which usually involves bilinear forms directly. Test functionals have special *convolution* form giving us suitable insight to analyze and implement the method simply and interestingly. Error analysis is given for functions that lie in Sobolev spaces defined on the unit sphere. As well as the Galerkin methods, the new technique needs a numerical quadrature, but here everything breaks down to simple numerical integration on spherical caps.

The remainder of this paper is organized as follows. In section 2 some preliminary results about spherical harmonics and Fourier series expansions on the sphere are outlined. In section 3, inspired by our final extracted test functionals, the well-known convolution on the sphere is reviewed and its main properties are addressed. In section 4 a simple local integrated form of the underlying PDE problem is extracted. In section 5 the trial space used in this work is discussed. In section 6 the properties of some specific kernels obtained by restricting a Euclidian kernel on the sphere are given. In section 7 the test space is constructed by the restricted compactly supported kernels. In section 8 a detailed description of the method and an error analysis are given. The condition number of the final linear system is estimated in section 9, and finally, section 10 is devoted to some discussions about constructing a numerical quadrature on spherical caps as needed for numerical implementation done in section 11.

With regard to notation, constants are denoted by c, $\widetilde{c}$, or C with or without subscript. In asymptotic expressions, c and $\widetilde{c}$ are used in lower and upper bounds, respectively. Constants without subscript are understood as "generic," i.e., they can change their value when used at different places, but we shall sometimes state explicitly on which problem parameters they depend. But constants with subscript are "specific" and their value will be unchanged through the paper.

2. Spherical harmonics. Spherical harmonics are restrictions to the unit sphere $\mathbb{S}^d$ of polynomials Y which satisfy

$$\Delta Y = 0,$$

where Δ is the Laplacian operator in $\mathbb{R}^{d+1}$. The space of all spherical harmonics of degree ℓ on $\mathbb{S}^d$ is denoted by $\mathcal{H}_\ell^d$ and has an L_2 orthonormal basis

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

where

$$N(d,0) = 1, \quad N(d,\ell) = \frac{(2\ell + d - 1)\Gamma(\ell + d - 1)}{\Gamma(\ell + 1)\Gamma(d)}, \quad \ell \geqslant 1,$$

where Γ is the known Gamma function. The orthonormality is expressed as

$$\int_{\mathbb{S}^d} Y_{\ell k}(x) Y_{\ell' k'}(x) d\sigma(x) = \delta_{\ell\ell'}\delta_{kk'},$$

where $d\sigma$ is the surface measure of the unit sphere. The space of spherical harmonics of order m or less will be denoted by

$$\mathcal{P}_m^d := \bigoplus_{\ell=0}^{m} \mathcal{H}_\ell^d$$

with dimension $N(d+1,m)$. It is known that the spherical harmonics are the eigenfunctions of the Laplace–Beltrami operator Δ_0, and every function $f \in L_2 = L_2(\mathbb{S}^d)$ can be expanded as

$$f = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widehat{f}_{\ell k} Y_{\ell k}, \quad \widehat{f}_{\ell k} = \frac{1}{\omega_d} \int_{\mathbb{S}^d} f Y_{\ell k} d\sigma, \tag{2.1}$$

where ω_d denotes the surface area of $\mathbb{S}^d$,

$$\omega_d := \int_{\mathbb{S}^d} d\sigma = \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)}.$$

The L_2-norm of f given by the formula

$$\|f\|_{L_2}^2 := \int_{\mathbb{S}^d} |f|^2 d\sigma$$

can also be expressed, via Parseval's identity, as

$$\|f\|_{L_2}^2 = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} |\widehat{f}_{\ell k}|^2.$$

Finally we note that

$$\sum_{k=1}^{N(d,\ell)} Y_{\ell k}(x) Y_{\ell k}(y) = \frac{N(d,\ell)}{\omega_d} P_\ell(d+1; x^T y) \tag{2.2}$$

is the *addition formula* for spherical harmonics. For a detailed discussion about the subject of this section see [4, 19].

3. Convolution on spheres. In this section the definition of convolution on the sphere and some of its properties are briefly explained. We refer the reader to [5] as an earlier reference and to [26, 4] as some newer sources. Let

$$w_\alpha(t) := (1-t^2)^{\alpha-1/2}, \quad \alpha > -\frac{1}{2}, \quad t \in (-1,1).$$

The space of univariate functions for which the pth power of the absolute value is integrable with respect to the weight w_α on $[-1,1]$ is denoted by $L_p(w_\alpha;[-1,1])$, and the norm of this space is denoted by $\|\cdot\|_{\alpha,p}$ and is defined to be

$$\|g\|_{\alpha,p} := \left(c_\alpha \int_{-1}^{1} |g(t)|^p w_\alpha(t) dt \right)^{1/p}, \quad g \in L_p(w_\alpha;[-1,1]),$$

where c_α is a normalization constant such that $c_\alpha \int_{-1}^{1} w_\alpha(t) dt = 1$. For $\alpha = \frac{d-1}{2}$ one can prove $c_\alpha = \frac{\omega_{d-1}}{\omega_d}$. Now, a convolution on the unit sphere $\mathbb{S}^d$ can be defined as follows.

DEFINITION 3.1. *Let $f \in L_1(\mathbb{S}^d)$, the space of integrable functions on $\mathbb{S}^d$, and $g \in L_1(w_\alpha;[-1,1])$ with $\alpha = \frac{d-1}{2}$. The convolution of f and g is defined by*

$$(f * g)(x) := \frac{1}{\omega_d} \int_{\mathbb{S}^d} f(y) g(x^T y)\, d\sigma(y),$$

where $d\sigma$ is the surface measure and ω_d is the surface area of $\mathbb{S}^d$, respectively.

If f is expressed as (2.1), then the following lemma can be viewed as an analogue of the fact that the Fourier transform of $f * g$ is equal to the product of the Fourier transforms of f and g.

LEMMA 3.2. *For $f \in L_2(\mathbb{S}^d)$ and $g \in L_1(w_\alpha, [-1,1])$ with $\alpha = \frac{d-1}{2}$ we have*

$$\widehat{(f * g)}_{\ell k} := \widehat{g}_\ell \widehat{f}_{\ell k}.$$

This fundamental result plays an important role in analysis of the numerical algorithm presented in section 8.

4. A Petrov–Galerkin formulation. We consider the PDE

$$Lu = f \text{ on } \mathbb{S}^d, \tag{4.1}$$

where L is an elliptic self-adjoint differential operator of order κ, for some $\kappa > 0$.

Instead of the usual global weak formulation, we use a *local weak* form in this paper. For this purpose we first define

$$B(x, \varrho) := \{y \in \mathbb{S}^d : \operatorname{dist}(x, y) \leqslant \varrho\}$$

as a spherical cap centered at $x \in \mathbb{S}^d$ with geodesic radius ϱ for some $0 < \varrho < \pi$. Here $\operatorname{dist}(x, y)$ is the geodesic distance between two points x, y on $\mathbb{S}^d$ which is defined by $\operatorname{dist}(x, y) = \arccos(x^T y)$.

Let $\Psi(x, \cdot)$ be a compactly supported *test function* on $B(x, \varrho)$ for $x \in \mathbb{S}^d$. Sufficient smoothness will be assumed for the test function as needed. By integrating (4.1) against $\Psi(x, \cdot)$ we have

$$\int_{B(x,\varrho)} Lu(y)\Psi(x, y)d\sigma(y) = \int_{B(x,\varrho)} f(y)\Psi(x, y)d\sigma(y), \quad x \in \mathbb{S}^d. \tag{4.2}$$

This is similar to the standard Galerkin method when one integrates against a basis in a given trial space. However, we will choose Ψ from a space different from the underlying trial space. It is clear that the integrals in (4.2) both can be written over $\mathbb{S}^d$ instead of $B(x, \varrho)$ because Ψ is assumed to be compactly supported. If we assume that the test function is zonal, i.e., there exists a continuous function $v_\rho : [-1, 1] \to \mathbb{R}$ such that

$$\Psi(x, y) = v_\rho(x^T y) \text{ for all } x, y \in \mathbb{S}^d,$$

then the left-hand side of (4.2) gets the convolution form as defined in section 3. Here $\rho \leqslant \varrho$ is the scaling parameter of the test function. In this case we can rewrite (4.2) as

$$(Lu * v_\rho)(x) = (f * v_\rho)(x), \quad x \in \mathbb{S}^d.$$

To approximate the unknown solution u, two different finite dimensional SBF subspaces will be assigned as *trial* and *test* spaces. Any SBF, local or global, with certain smoothness, can be used to construct the trial space. However, in this paper we are mainly interested in positive definite functions that possess a Fourier transform that decays only algebraically. The test function v_ρ should be necessarily compactly supported to allow us to form a set of independent test functionals and in parallel to relax the costs of numerical integrations. This type of Petrov–Galerkin method is implemented and analyzed on the sphere for the first time in this paper. The idea comes form the local Petrov–Galerkin methods on bounded domains in $\mathbb{R}^d$, introduced in a mechanical engineering community [1].

5. Trial functions. Zonal kernels are employed to construct the approximate solution u_N of u. Recall that zonal kernels on $\mathbb{S}^d$ are functions that can be represented as $\phi(x^T y)$ for all $x, y \in \mathbb{S}^d$, where $\phi(t)$ is a continuous function on $[-1, 1]$. We are especially interested in zonal kernels of the type

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

where $\{P_\ell(d+1;t)\}_{\ell=0}^{\infty}$ is the sequence of $(d+1)$-dimensional Legendre polynomials normalized to $P_\ell(d+1;1) = 1$. In [30] and [33] it was proved that such ϕ is positive definite on $\mathbb{S}^d$. One can expand the kernel $\Phi(x, y)$ in terms of spherical harmonics. Using the addition formula (2.2) we have

$$\Phi(x, y) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widehat{\phi}(\ell) Y_{\ell k}(x) Y_{\ell k}(y), \tag{5.1}$$

where

$$\widehat{\phi}(\ell) = \frac{\omega_d}{N(d, \ell)} a_\ell.$$

If we assume that for some $\sigma > d/2$,

$$c_\phi (1+\ell)^{-2\sigma} \leqslant \widehat{\phi}(\ell) \leqslant \widetilde{c}_\phi (1+\ell)^{-2\sigma}, \quad \ell \geqslant 0, \tag{5.2}$$

holds for specific positive constants c_ϕ and $\widetilde{c}_\phi$, then the *native space* associated to Φ is norm equivalent to $H^\sigma = H^\sigma(\mathbb{S}^d)$, the Sobolev space of order σ on $\mathbb{S}^d$. If fact, the native space $\mathcal{N}_\Phi = \mathcal{N}_\Phi(\mathbb{S}^d)$ is defined by

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

where $\mathcal{D}'(\mathbb{S}^d)$ is the space of distributions on $\mathbb{S}^d$. It can be shown that $\mathcal{N}_\Phi$ is a Hilbert space with respect to the inner product

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

Moreover, Φ is reproducing kernel for $\mathcal{N}_\Phi$, i.e., for all $f \in \mathcal{N}_\Phi$,

$$\langle f, \Phi(x, \cdot) \rangle_\Phi = f(x), \quad x \in \mathbb{S}^d.$$

On the other hand, the Sobolev space H^σ with real parameter σ is defined by

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

Thus, under condition $\sigma > d/2$, due to the definitions of $\mathcal{N}_\Phi$ and H^σ and condition (5.2), we deduce that

$$c_\phi \|f\|_\Phi^2 \leqslant \|f\|_{H^\sigma}^2 \leqslant \widetilde{c}_\phi \|f\|_\Phi^2,$$

which means that $\mathcal{N}_\Phi$ and H^σ are norm equivalent.

A class of basis functions which satisfy condition (5.2) for some σ will be briefly addressed in the next section.

6. Restricted kernels. The restriction of a positive definite kernel from $\mathbb{R}^{d+1}$ to any submanifold $\mathbb{M}$ is a seemingly naive way for obtaining a positive definite kernel on $\mathbb{M}$. If the original kernel is positive definite, so is its restriction to $\mathbb{M}$, making it well-suited for scattered data interpolation problems. The case $\mathbb{M} = \mathbb{S}^d$ has been studied in [22, 34], while the general case has been investigated in [9]. Before all, a variation of compactly supported RBFs of Wendland's type [31] on the sphere was introduced in [24].

Assume that S is an RBF on $\mathbb{R}^{d+1}$, i.e., there exists a univariate function ψ such that $S(x) = \psi(\|x\|_2)$, where $\|\cdot\|_2$ is Euclidian norm in $\mathbb{R}^{d+1}$. Since for points $x, y \in \mathbb{S}^d$ we have $\|x - y\|_2 = \sqrt{2 - 2x^T y}$, we may therefore define

$$\Phi(x, y) = \phi(x^T y) := \psi(\sqrt{2 - 2x^T y}) = S(x - y), \quad x, y \in \mathbb{S}^d.$$

It is clear that Φ inherits the property of positive definiteness from S. In [22] it was proved that if Φ is represented in the form (5.1), then the Fourier coefficients $\widehat{\phi}(\ell)$ satisfy the decay condition (5.2) for some $\sigma > 0$. To be more precise, if we assume that the RBF S has $H^s(\mathbb{R}^{d+1})$ as its native space, which is equivalent to this fact that its $(d+1)$-variate Fourier transform $\widehat{S}$ behaves like

$$c(1 + \|\omega\|_2^2)^{-s} \leqslant \widehat{S}(\omega) \leqslant \widetilde{c}(1 + \|\omega\|_2^2)^{-s}, \quad \omega \in \mathbb{R}^{d+1}, \tag{6.1}$$

for $s > \frac{d+1}{2}$, then Φ (the restriction of S on $\mathbb{S}^d$) generates $H^{s-1/2}(\mathbb{S}^d)$, i.e., its Fourier coefficients satisfy (5.2) for $\sigma = s - \frac{1}{2}$. This loss of "$\frac{1}{2}$ a derivative" is familiar from the theory of Sobolev spaces and traces of functions. In the general case, when S is restricted to a k-dimensional smooth submanifold $\mathbb{M} \subset \mathbb{R}^{d+1}$, then the native space of the restricted kernel is $H^{s-(d+1-k)/2}(\mathbb{M})$. See [9, Theorem 5].

7. Test functions. In Petrov–Galerkin schemes trial and test functions come from different spaces. Here we use a scaled compactly supported positive definite kernel as a reference test function and then we rotate it to some different test points to construct a set of independent test functionals. Assume that we are given a zonal kernel v on $\mathbb{S}^d$ defined from a compactly supported RBF ψ on $\mathbb{R}^{d+1}$ by

$$v(x^T y) := \psi(\|x - y\|_2), \quad x, y \in \mathbb{S}^d.$$

Then we define for $x, y \in \mathbb{S}^d$

$$v_\rho(x^T y) := \rho^{-d} \psi\left(\frac{\|x - y\|_2}{\rho}\right),$$

where $\rho > 0$ is a scaling parameter. If we expand v_ρ as

$$v_\rho(x^T y) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widehat{v}_\rho(\ell) Y_{\ell k}(x) Y_{\ell k}(y), \quad x, y \in \mathbb{S}^d,$$

and if we assume that ψ is a compactly supported RBF of Wendland's type [31], then we can show that for some $\nu > d/2$,

$$c_v(1 + \rho\ell)^{-2\nu} \leqslant \widehat{v}_\rho(\ell) \leqslant \widetilde{c}_v(1 + \rho\ell)^{-2\nu}, \quad \ell \geqslant 0, \tag{7.1}$$

where c_v and $\widetilde{c}_v$ are two specific positive constants. More precisely, we have the following lemma from [15].

LEMMA 7.1. *If the radial function $\psi(\|\cdot\|_2) = S(\cdot)$ has compact support in the unit ball and a Fourier transform satisfying the condition* (6.1) *with $s > (d+1)/2$, then there are constants c_v and $\widetilde{c}_v$ depending only on s and d such that the associated scaled SBF v_ρ has Fourier coefficients satisfying* (7.1) *with $\nu = s - \frac{1}{2}$.*

8. Numerical method and error analysis. Assume that we can expand Lu, for a κ-order elliptic self-adjoint differential operator L, as a Fourier series

$$Lu = \sum_{k=1}^{\infty} \sum_{k=1}^{N(d,\ell)} \widehat{L}(\ell)\widehat{u}_{\ell k} Y_{\ell k}, \tag{8.1}$$

in which

$$c_L(1+\ell)^\kappa \leqslant \widehat{L}(\ell) \leqslant \widetilde{c}_L(1+\ell)^\kappa, \quad \ell \geqslant 0, \tag{8.2}$$

where c_L, $\widetilde{c}_L$ are two positive constants independent of ℓ. For example, we may take $L = -\Delta_0 + \omega^2 I$, where Δ_0 is the Laplace–Beltrami operator and $\omega > 0$, in which case $\widehat{L}(\ell) = \ell(\ell+d-1) + \omega^2$ and $\kappa = 2$.

Assumptions (8.1) and (8.2) mean that $L = P(-\Delta_0)$, where P is a polynomial with $P(x) > 0$ for all $x \geqslant 0$. This excludes, for example, PDEs with operators of the form $Lu = \sum_{i,j} a_{ij} \frac{\partial^2 u}{\partial x^i \partial x^j}$ + lower order terms, unless a is the metric tensor for the unit sphere. This also implies that κ in (8.2) is an even number. Otherwise L would be a pseudodifferential operator [7].

Assume that Φ is a kernel that satisfies condition (5.2) for some $\sigma > d/2$. Suppose $X = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{S}^d$ is a given discrete set of scattered points on the unit sphere $\mathbb{S}^d$. Our numerical solution u_N comes from the trial space

$$V_{\Phi,X} := \operatorname{span}\{\Phi(\cdot, x_j) : x_j \in X\}$$

by linear combination

$$u_N = \sum_{j=1}^{N} b_j \Phi(\cdot, x_j).$$

Recall the test functionals

$$(Lu * v_\rho)(x) = (f * v_\rho)(x), \quad x \in \mathbb{S}^d, \tag{8.3}$$

from section 4. Replacing u by u_N in (8.3) and then imposing at $x = x_k$, $k = 1, \ldots, N$, we will finally get the linear system

$$Ab = F, \tag{8.4}$$

where $A = (a_{kj})$ is an $(N \times N)$-matrix with

$$a_{kj} = (L\Phi(\cdot, x_j) * v_\rho)(x_k), \quad k, j = 1, \ldots, N, \tag{8.5}$$

and $F = (f_k)$ is an N-vector with

$$f_k = (f * v_\rho)(x_k), \quad k = 1, \ldots, N. \tag{8.6}$$

Although it is not explicitly expressed in the notation, the coefficient vector b depends on ρ and so does u_N.

The following theorem establishes the solvability of the proposed scheme.

THEOREM 8.1. *Suppose that $\psi(r)$ is a compactly supported RBF supported in $[0,1]$. Define*

$$v_\rho(t) := \rho^{-d}\psi\left(\frac{\sqrt{2-2t}}{\rho}\right), \quad t \in [-1,1]. \tag{8.7}$$

In additions, let L be an elliptic self-adjoint differential operator of order κ, and assume that Φ is a positive definite RBF on the unit sphere which satisfies (5.2) *for $\sigma > d/2 + \kappa/2$. Then there exists a unique function $u_N \in V_{\Phi,X}$, which depends on ρ, that fulfills the conditions*

$$(Lu_N * v_\rho)(x_k) = (f * v_\rho)(x_k) \tag{8.8}$$

for $k = 1, 2, \ldots, N$.

Proof. The lower bound of σ guarantees that L can be applied to one of the arguments of Φ. Since the entries of the final linear system are given by (8.5), it is enough to prove that

$$\Lambda(x,y) := (L\Phi(\cdot,y) * v_\rho)(x) \tag{8.9}$$

is a positive definite kernel on the unit sphere. If Φ is a zonal kernel represented by the Fourier series (5.1), then $L\Phi$ is a zonal kernel having the Fourier expansion

$$L\Phi(x,y) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widehat{\phi}(\ell)\widehat{L}(\ell) Y_{\ell k}(x) Y_{\ell k}(y). \tag{8.10}$$

Using Lemma 3.2 and the series representation (8.10) we conclude that the Fourier coefficients of kernel $\Lambda(x,y)$ are

$$\widehat{\lambda}(\ell) = \widehat{\phi}(\ell)\widehat{L}(\ell)\widehat{v}_\rho(\ell).$$

Both $\widehat{\phi}(\ell)$ and $\widehat{v}_\rho(\ell)$ are positive because the trial kernel Φ and the test function v_ρ are positive definite. On the other hand, by assumption, $\widehat{L}(\ell)$ are positive numbers. Consequently, $\widehat{\lambda}(\ell)$ are positive and thus Λ is positive definite. This guarantees that (8.4) is always uniquely solvable. ∎

The remaining parts of this section will be devoted to convergence and stability properties of the numerical solution. Things start with the following theorem.

THEOREM 8.2. *Assume that $\rho \in (0,2]$ is given and v_ρ is defined via* (8.7) *where* (7.1) *holds for its corresponding Fourier coefficients with $\nu > d/2$ and $\nu \geqslant \kappa/2$. If $u \in H^\tau$ for $\tau \geqslant \kappa$, then $Lu * v_\rho \in H^{\tau-\kappa+2\nu}$ and*

$$c\|u\|_{H^\mu} \leqslant \|Lu * v_\rho\|_{H^{\mu-\kappa+2\nu}} \leqslant \widetilde{c}\,\rho^{-2\nu}\|u\|_{H^\mu} \tag{8.11}$$

hold for any μ with $0 \leqslant \mu \leqslant \tau$.

Proof. By assumptions, $Lu \in H^{\tau-\kappa} \subseteq L_2(\mathbb{S}^d)$ and the convolution is well-defined. Using the definition of norms by Fourier series we have

$$\begin{aligned}\|Lu * v_\rho\|^2_{H^{\mu-\kappa+2\nu}} &= \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2(\mu-\kappa+2\nu)} |(\widehat{Lu * v_\rho})_{\ell k}|^2 \\ &= \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2(\mu-\kappa+2\nu)} |\widehat{L}(\ell)\widehat{v}_\rho(\ell)\widehat{u}_{\ell k}|^2 \\ &\leqslant \widetilde{c}_v^2 \widetilde{c}_L^2 \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2(\mu+2\nu)} (1+\rho\ell)^{-4\nu} |\widehat{u}_{\ell k}|^2,\end{aligned}$$

where we have used conditions (7.1) and (8.2) to bound $\widehat{v}_\rho(\ell)$ and $\widehat{L}(\ell)$, respectively. Since $(1+\rho\ell) = \rho(1/\rho + \ell) \geqslant \frac{\rho}{2}(1+\ell)$, we have

$$(1+\rho\ell)^{-4\nu} \leqslant 2^{4\nu}\rho^{-4\nu}(1+\ell)^{-4\nu}.$$

Thus

$$\begin{aligned}\|Lu * v_\rho\|^2_{H^{\mu-\kappa+2\nu}} &\leqslant \widetilde{c}_v^2 \widetilde{c}_L^2 2^{4\nu}\rho^{-4\nu} \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2\mu} |\widehat{u}_{\ell k}|^2 \\ &= \widetilde{c}^2\, \rho^{-4\nu} \|u\|^2_{H^\mu}\end{aligned}$$

for $\widetilde{c} = \widetilde{c}_v \widetilde{c}_L 2^{2\nu}$. On the other hand we can write

$$\begin{aligned}\|u\|^2_{H^\mu} &= \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2\mu} |\widehat{u}_{\ell k}|^2 \\ &\leqslant \frac{1}{c_v^2 c_L^2} \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2(\mu-\kappa)} (1+\rho\ell)^{4\nu} |\widehat{L}(\ell)\widehat{v}_\rho(\ell)\widehat{u}_{\ell k}|^2 \\ &\leqslant \frac{2^{4\nu}}{c_v^2 c_L^2} \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2(\mu-\kappa+2\nu)} |(\widehat{Lu * v_\rho})_{\ell k}|^2 \\ &= \frac{1}{c^2} \|Lu * v_\rho\|^2_{H^{\mu-\kappa+2\nu}},\end{aligned}$$

where we have used conditions (7.1) and (8.2) and the fact that $(1+\rho\ell)^{4\nu} \leqslant 2^{4\nu}(1+\ell)^{4\nu}$ for $\rho \in (0,2]$. This proves the left-hand-side inequality in (8.11) for $c = 2^{-2\nu} c_v c_L$. ∎

Now, following [16], we introduce a new positive definite kernel. We can interpret the kernel $\Lambda(x,y)$ as the action of an operator K on Φ,

$$\Lambda(x,y) = (L\Phi(\cdot,y) * v_\rho)(x) =: K\Phi(x,y), \quad x,y \in \mathbb{S}^d. \tag{8.12}$$

The new kernel Θ is then defined by

$$\Theta := K^{-1}\Phi.$$

This kernel has Fourier coefficients

$$\widehat{\theta}(\ell) = \frac{\widehat{\phi}(\ell)}{\widehat{L}(\ell)\widehat{v}_\rho(\ell)}$$

and defines an inner product

$$\langle f, g\rangle_\Theta = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \frac{\widehat{L}(\ell)\widehat{v}_\rho(\ell)}{\widehat{\phi}(\ell)} \widehat{f}_{\ell k}\widehat{g}_{\ell k}, \quad f, g \in H^{\sigma+\kappa/2-\nu}, \tag{8.13}$$

with the corresponding norm

$$\|f\|_\Theta^2 := \langle f, f\rangle_\Theta.$$

Under conditions (5.2), (7.1), and (8.2) on trial kernel Φ, test kernel v_ρ, and operator L, respectively, we can prove that

$$c\|u\|_{H^{\sigma+k/2-\nu}} \leqslant \|u\|_\Theta \leqslant \widetilde{c}\rho^{-\nu}\|u\|_{H^{\sigma+k/2-\nu}}, \tag{8.14}$$

with $c = \sqrt{c_L c_v/\widetilde{c}_\phi}2^{-\nu}$ and $\widetilde{c} = \sqrt{\widetilde{c}_v\widetilde{c}_L/c_\phi}2^{\nu}$, which shows that the $\|\cdot\|_\Theta$ norm is equivalent to the Sobolev norm $\|\cdot\|_{H^{\sigma+\kappa/2-\nu}}$, and $H^{\sigma+\kappa/2-\nu}$ with the inner product (8.13) is a reproducing kernel Hilbert space with kernel Θ, provided that $\sigma > d/2 + \nu - \kappa/2$.

Next, we need to specify the exact order of smoothness of the approximate solution u_N. The smoothness of u_N, of course, depends on the behavior of Fourier coefficients $\widehat{\phi}(\ell)$. The following lemma makes it precise.

LEMMA 8.3. *If Φ satisfies (5.2) for $\sigma > d/2 + \beta$ for $\beta \geqslant 0$, then u_N belongs to $H^{\sigma+\beta}$.*

Proof. If we define $\varphi_j := \Phi(\cdot, x_j)$, then from (5.1) we have $\widehat{\varphi}_j(\ell) = \widehat{\phi}(\ell)Y_{\ell k}(x_j)$, which leads via (5.2) to

$$\begin{aligned}
\|\varphi_j\|_{H^{\sigma+\beta}} &= \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} \widehat{\phi}(\ell)^2 Y_{\ell k}(x_j)^2(1+\ell)^{2\sigma+2\beta} \\
&\leqslant \widetilde{c}_\phi^2 \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} Y_{\ell k}(x_j)^2(1+\ell)^{-2\sigma+2\beta} \\
&= \widetilde{c}_\phi^2 \sum_{\ell=0}^{\infty} \frac{N(d,\ell)}{\omega_d} P_\ell(d+1;1)(1+\ell)^{-2\sigma+2\beta} \\
&\leqslant \widetilde{c}\sum_{\ell=0}^{\infty}(1+\ell)^{-2\sigma+2\beta+d-1}.
\end{aligned}$$

In the third line above we have used the addition formula (2.2) for $x = y = x_j$, and in the forth line we have used $P_\ell(d+1;1) = 1$ and the fact that $N(d,\ell) = \mathcal{O}((1+\ell)^{d-1})$. Since by assumption $-2\sigma + 2\beta + d < 0$, the last sum is finite. This shows that φ_j belong to $H^{\sigma+\beta}$ and completes the proof because $u_N \in V_{\Phi,X}$ and $V_{\Phi,X}$ is a finite dimensional space spanned by functions φ_j. ∎

The following theorem gives an orthogonality property helping us to prove some error bounds for our numerical solution.

THEOREM 8.4. *Suppose that the test function v_ρ, $\rho \in (0, 2]$, satisfies (7.1) for $\nu > d/2$, and operator L satisfies (8.2) for positive integer κ. Let $\nu \geqslant \kappa/2$ and assume that the trial kernel Φ satisfies (5.2) for some σ with*

$$\sigma > d/2 + \nu - \kappa/2 \quad \textit{and} \quad \sigma \geqslant \nu + \kappa/2. \tag{8.15}$$

Finally assume that $u \in H^{\sigma+\kappa/2-\nu}$, and $u_N \in V_{\Phi,X}$ is the solution of the Petrov–Galerkin scheme. Then

$$\langle u - u_N, s\rangle_\Theta = 0 \text{ for all } s \in V_{\Phi,X}.$$

Proof. The second lower bound of σ together with $\nu > d/2$ imply that $\sigma > d/2 + \kappa/2$. This with the first lower bound of σ gives $\sigma > d/2+\beta$ for $\beta := \max\{\kappa/2, \nu-\kappa/2\}$. Lemma 8.3 ensures that $u_N \in H^{\sigma+\beta}$. Condition $\nu \geqslant \kappa/2$ guarantees that

$$e_N := u - u_N$$

belongs to $H^{\sigma+\kappa/2-\nu}$. Theorem 8.2 implies that $Ke_N \in H^{\sigma+\nu-\kappa/2}$. Note that K can operate on e_N because by assumption $\sigma \geqslant \nu + \kappa/2$, which implies $\sigma + \kappa/2 - \nu \geqslant \kappa$. It is clear that $Ke_N \in H^{\sigma+k/2-\nu}$ because $\nu \geqslant \kappa/2$. On the other hand, from (8.8) and (8.12) we have

$$Ku(x_k) = Ku_N(x_k), \quad k = 1, 2, \ldots, N,$$

which means that the error function e_N satisfies $Ke_N(x_k) = 0$ for $k = 1, 2, \ldots, N$. Using (8.1) and the definition of the Θ-inner product in (8.13) one can easily show that K is self-adjoint. Operator K is also invertible, thus we can write

$$\begin{aligned}\langle e_N, \Phi(\cdot, x_k)\rangle_\Theta &= \langle Ke_N, K^{-1}\Phi(\cdot, x_k)\rangle_\Theta \\ &= \langle Ke_N, \Theta(\cdot, x_k)\rangle_\Theta \\ &= Ke_N(x_k) \\ &= 0.\end{aligned}$$

This completes the proof because $V_{\Phi,X}$ is a finite dimensional space spanned by functions $\Phi(\cdot, x_k)$. ∎

Theorem 8.4 immediately implies Pythagoras' theorem

$$\|u - u_N\|_\Theta^2 + \|u_N\|_\Theta^2 = \|u\|_\Theta^2,$$

which proves

$$\|u - u_N\|_\Theta \leqslant \|u\|_\Theta. \tag{8.16}$$

The norm equivalence property (8.14) together with (8.16) gives the stability bound

$$\begin{aligned}\|u - u_N\|_{H^{\sigma+\kappa/2-\nu}} &\leqslant C\|u - u_N\|_\Theta \\ &\leqslant C\|u\|_\Theta \\ &\leqslant C\rho^{-\nu}\|u\|_{H^{\sigma+\kappa/2-\nu}}.\end{aligned} \tag{8.17}$$

The order of convergence of kernel methods is mainly based on the density and the quality of trial and test points. Recall the set $X = \{x_1, x_2, \ldots, x_N\}$ of scattered points on $\mathbb{S}^d$. There are three geometrical quantities associated with X. The *separation distance* q_X is the radius of the largest ball that can be placed around every point in X such that no two balls overlap, i.e.,

$$q_X := \frac{1}{2}\min_{j\neq k} \operatorname{dist}(x_j, x_k).$$

On the other hand, the *fill distance* corresponds to the radius of the largest empty possible ball that can be placed between the points in X. It is defined to be

$$h_X := \max_{x\in\mathbb{S}^d} \min_{x_j\in X} \operatorname{dist}(x, x_j).$$

Finally the *mesh ratio* r_X is defined by

$$r_X := \frac{h_X}{q_X},$$

which measures how uniformly the points are placed. When it is close to 1, the distribution of the points in X is said to be *quasi uniform*. For $R \geqslant 1$, let $\mathcal{X}_R = \mathcal{X}_R(\mathbb{S}^d)$ be the family of all sets of centers X with $r_X \leqslant R$; we will say that the family $\mathcal{X}_R$ is *R-uniform*.

The following "sampling inequality" is very important in our analysis. The authors of [21] have applied Theorem 5.5 of [23] to prove this lemma. Going through the details of the proof, one can find that the assumption of quasi uniformity for set X is not actually required.

LEMMA 8.5. *Let $\alpha, \beta \in \mathbb{R}$ satisfy $\beta > d/2$ and $0 \leqslant \alpha \leqslant \beta$. Suppose that $X \subset \mathbb{S}^d$ is a set of scattered points with fill distance h_X. If $u \in H^\beta$ satisfies $u|_X = 0$, then for h_X sufficiently small, we have*

$$\|u\|_{H^\alpha} \leqslant C h_X^{\beta-\alpha} \|u\|_{H^\beta}.$$

Finally, the following theorem gives an error bound for the presented method. We assume that the integrals in (8.5) and (8.6) are computed exactly and thus no error is produced by numerical integration.

THEOREM 8.6. *Suppose that all assumptions of Theorem 8.4 are satisfied. Let $\tau := \sigma + \kappa/2 - \nu$. Then the error bound*

$$\|u - u_N\|_{H^\beta} \leqslant C h_X^{\tau-\beta} \rho^{-3\nu} \|u\|_{H^\tau} \tag{8.18}$$

holds for $0 \leqslant \beta \leqslant \tau$ and for sufficiently small fill distance h_X of set X on $\mathbb{S}^d$. Here ρ is the radius of support of a test function.

Proof. By applying Theorem 8.2, Lemma 8.5, and inequalities (8.17) we have

$$\begin{aligned}
\|u - u_N\|_{H^\beta} &\leqslant C\|Ku - Ku_N\|_{H^{\beta+2\nu-\kappa}} \\
&\leqslant C h_X^{\sigma+\kappa/2-\nu-\beta} \|Ku - Ku_N\|_{H^{\sigma+\nu-\kappa/2}} \\
&\leqslant C h_X^{\sigma+\kappa/2-\nu-\beta} \rho^{-2\nu} \|u - u_N\|_{H^{\sigma+\kappa/2-\nu}} \\
&\leqslant C h_X^{\sigma+\kappa/2-\nu-\beta} \rho^{-3\nu} \|u\|_{H^{\sigma+\kappa/2-\nu}}.
\end{aligned}$$

This completes the proof. ∎

Note that, with $\tau = \sigma + \kappa/2 - \nu$, the lower bounds (8.15) on σ are equivalent to the following lower bounds on τ:

$$\tau > d/2 \quad \text{and} \quad \tau \geqslant \kappa.$$

The error bound (8.18) estimates the error for a function u that lies in Sobolev space H^τ for $\tau = \sigma + \kappa/2 - \nu$. It is also interesting to estimate the error for approximating functions smoother than those in the native space H^τ of kernel Θ. In the

following we use the "doubling trick" in the case where $u \in H^{\tau+\alpha}$ for $\alpha \in [0,\tau]$. Results of this kind have been developed by Schaback [29] for positive definite functions on $\mathbb{R}^d$ and on manifolds. See also [9, Proposition 11] and [21, Appendix A.1]. Our argument is partly based on [21]. First, we measure the error in the H^τ-norm and then we extend it to the H^β-norm for $\beta \in [0,\tau]$.

LEMMA 8.7. *Let $\tau := \sigma + \kappa/2 - \nu$ and $\alpha \in [0,\tau]$. Under the assumptions of Theorem* 8.4 *we have*

$$\|u - u_N\|_{H^\tau} \leqslant C\rho^{-2\nu} h_X^\alpha \|u\|_{H^{\tau+\alpha}},$$

provided that h_X is sufficiently small and $u \in H^{\tau+\alpha}$.

Proof. By Theorem 8.4 we have $\langle u - u_N, s\rangle_\Theta = 0$ for all $s \in V_X$. Consequently, $\langle u - u_N, u_N\rangle_\Theta = 0$ and so $\langle u - u_N, u_N\rangle_{H^\tau} = 0$. This implies $\|u - u_N\|_{H^\tau}^2 = \langle u - u_N, u\rangle_{H^\tau}$. Let $w := u - u_N$. We have

$$\begin{aligned}
\|u - u_N\|_{H^\tau}^2 = \langle w, u\rangle_{H^\tau} &= \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2\tau}\widehat{u}_{\ell k}\overline{\widehat{w}_{\ell k}} \\
&\leqslant \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{\tau+\alpha}|\widehat{u}_{\ell k}|(1+\ell)^{\tau-\alpha}|\widehat{w}_{\ell k}| \\
&\leqslant \left(\sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2(\tau+\alpha)}|\widehat{u}_{\ell k}|^2\right)^{1/2} \left(\sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} (1+\ell)^{2(\tau-\alpha)}|\widehat{w}_{\ell k}|^2\right)^{1/2} \\
&= \|u\|_{H^{\tau+\alpha}}\|w\|_{H^{\tau-\alpha}} \\
&= \|u\|_{H^{\tau+\alpha}}\|u - u_N\|_{H^{\tau-\alpha}}.
\end{aligned}$$

On the other hand we can write

$$\begin{aligned}
\|u - u_N\|_{H^{\tau-\alpha}} &\leqslant C\|Ku - Ku_N\|_{H^{\tau-\alpha-\kappa+2\nu}} && \text{(using Theorem 8.2)} \\
&\leqslant Ch_X^\alpha\|Ku - Ku_N\|_{H^{\tau-\kappa+2\nu}} && \text{(using Lemma 8.5)} \\
&\leqslant Ch_X^\alpha\rho^{-2\nu}\|u - u_N\|_{H^\tau}. && \text{(using Theorem 8.2).}
\end{aligned}$$

Combining these and then dividing both sides by $\|u - u_N\|_{H^\tau}$ yield the desired bound. ∎

THEOREM 8.8. *Let $\tau := \sigma + \kappa/2 - \nu$, $\beta \in [0,\tau)$, and $\gamma \in (\tau, 2\tau]$. Under the assumptions of Theorem* 8.4 *we have*

$$\|u - u_N\|_{H^\beta} \leqslant C\rho^{-4\nu} h_X^{\gamma-\beta}\|u\|_{H^\gamma}, \tag{8.19}$$

provided that h_X is sufficiently small and $u \in H^\gamma$.

Proof. By applying Theorem 8.2 and Lemma 8.5 we have

$$\begin{aligned}
\|u - u_N\|_{H^\beta} &\leqslant C\|Ku - Ku_N\|_{H^{\beta-\kappa+2\nu}} \\
&\leqslant Ch_X^{\tau-\beta}\|Ku - Ku_N\|_{H^{\tau-\kappa+2\nu}} \\
&\leqslant C\rho^{-2\nu}h_X^{\tau-\beta}\|u - u_N\|_{H^\tau}.
\end{aligned}$$

Then Lemma 8.7 with $\alpha = \gamma - \tau$ gives the desired bound. ∎

Theorems 8.6 and 8.8 estimate the error function when $u \in H^\gamma$ for $\gamma \in [\tau, 2\tau]$. The following theorem concerns the case $\gamma \in [\mu, \tau)$ for $\mu > d/2$ and $\mu \geqslant \kappa$, i.e., the case where u lies outside the native space H^τ of kernel Θ.

THEOREM 8.9. *Let $\tau := \sigma + \kappa/2 - \nu$. Assume that $\gamma > d/2$, $\gamma \geqslant \kappa$, $\gamma < \tau$, and $0 \leqslant \beta \leqslant \gamma$. Then for all $u \in H^\gamma$ we have*

$$\|u - u_N\|_{H^\beta} \leqslant C\rho^{-2\nu} r_X^{\tau-\gamma} h_X^{\gamma-\beta} \|u\|_{H^\gamma}, \tag{8.20}$$

provided that h_X is sufficiently small. Here r_X is the mesh ratio of set X.

Proof. Let $\bar{\tau} := \sigma + \nu - \kappa/2$. Remember that $Ku_N = Lu_N * v_\rho$ can be viewed as the Λ-interpolant of function Ku. Since $\widehat{\lambda}(\ell) = \widehat{\phi}(\ell)\widehat{L}(\ell)\widehat{v}_\rho(\ell)$, by using (5.2), (7.1), (8.2), and inequality $(1 + \rho\ell) \leqslant 2(1 + \ell)$ for $\rho \in (0, 2]$, we have

$$c_\phi c_v c_L 2^{-2\nu}(1+\ell)^{-2\bar{\tau}} \leqslant \widehat{\lambda}(\ell) \leqslant \widetilde{c}_\phi \widetilde{c}_v \widetilde{c}_L 2^{-2\nu} \rho^{-2\nu} (1+\ell)^{-2\bar{\tau}}.$$

This means that the native space of Λ, i.e., $\mathcal{N}_\Lambda$, is $H^{\bar{\tau}}$ and their norms are equivalent, namely,

$$c\rho^\nu \|u\|_{H^{\bar{\tau}}} \leqslant \|u\|_\Lambda \leqslant \widetilde{c}\|u\|_{H^{\bar{\tau}}}. \tag{8.21}$$

Let $\bar{\gamma} > d/2 + 2\nu - \kappa$, $\bar{\gamma} \geqslant 2\nu$, $\bar{\gamma} \leqslant \bar{\tau}$, and $0 \leqslant \bar{\beta} \leqslant \bar{\gamma}$. Since $\nu \geqslant \kappa/2$ and $\nu > d/2$, thus $\bar{\gamma} > d/2$ and [23, Theorem 5.5] yields

$$\|Ku - Ku_N\|_{H^{\bar{\beta}}} \leqslant C r_X^{\bar{\tau}-\bar{\gamma}} h_X^{\bar{\gamma}-\bar{\beta}} \|Ku\|_{H^{\bar{\gamma}}},$$

where r_X is the mesh ratio of set X. Now, by setting $\bar{\beta} := \beta - \kappa + 2\nu$ and $\bar{\gamma} := \gamma - \kappa + 2\nu$ and by using Theorem 8.2 and the above inequality we have

$$\begin{aligned}
\|u - u_N\|_{H^\beta} &\leqslant C\|Ku - Ku_N\|_{H^{\bar{\beta}}} \\
&\leqslant C r_X^{\bar{\tau}-\bar{\gamma}} h_X^{\bar{\gamma}-\bar{\beta}} \|Ku\|_{H^{\bar{\gamma}}} \\
&\leqslant C\rho^{-2\nu} r_X^{\bar{\tau}-\bar{\gamma}} h_X^{\bar{\gamma}-\bar{\beta}} \|u\|_{H^\gamma} \\
&= C\rho^{-2\nu} r_X^{\tau-\gamma} h_X^{\gamma-\beta} \|u\|_{H^\gamma},
\end{aligned}$$

because $\bar{\tau} - \bar{\gamma} = \tau - \gamma$ and $\bar{\gamma} - \bar{\beta} = \gamma - \beta$. ∎

If we work with an R-uniform family of centers, i.e., $X \in \mathcal{X}_R$, then in inequality (8.20) we can replace r_X by R to get the bound

$$\|u - u_N\|_{H^\beta} \leqslant C_{\mathcal{X}_R} \rho^{-2\nu} h_X^{\gamma-\beta} \|u\|_{H^\gamma},$$

where $C r_X^{\tau-\gamma} \leqslant C R^{\tau-\gamma} =: C_{\mathcal{X}_R}$.

The error bounds (8.18), (8.19), and (8.20), which contain the scaling test parameter ρ, suggest the *stationary* and the *nonstationary* test discretizations. These concepts are well-known in the kernel approximation when the trial space is formed by translates of a scaled kernel [32]. However, they are new here in the test space. The stationary and the nonstationary test discretizations refer to the cases when $\rho = \mathcal{O}(h_X)$ and $\rho = \mathcal{O}(1)$, respectively. In the stationary case, the error bounds (8.18), (8.19), and (8.20) show that the order of $\|u - u_N\|_{H^\beta}$ is reduced by (at most) -4ν, although the numerical results of section 11 predict a better order of convergence. On the other hand, the order of convergence in the nonstationary case is

more than what one should expect since the solution u in (4.1) is in general κ times smoother than the data function f.

In [16, Theorem 1] the error bound for the collocation solution u_N^C of (4.1) with the same kernel Φ for $u \in H^\gamma$, $\gamma = \sigma + \kappa/2$, has been obtained as

$$\|u - u_N^C\|_{L_2} \leqslant \|Lu - Lu_N^C\|_{L_2} \leqslant Ch_X^{\gamma-\kappa}\|u\|_{H^\gamma}. \tag{8.22}$$

The first inequality above is sharp and [16] proves that (in general) there is no leeway for a better estimate.

In [21, Theorem 6.2] the error bound for the Galerkin solution u_N^G of (4.1) for $\kappa = 2$ and $d = 2$ with the same kernel Φ has been obtained using the Nietsche trick as

$$\|u - u_N^G\|_{L_2} \leqslant Ch_X^\gamma \|u\|_{H^\gamma},$$

provided that $u \in H^\gamma$ for $2 \leqslant \gamma \leqslant 2\sigma$.

Compared with the above estimates, the order of convergence of the nonstationary Petrov–Galerkin method improves that of the collocation method at least for the special case $\beta = 0$ and $u \in H^{\sigma+\kappa/2}$. On the other hand, for the special case $d = 2$, $\kappa = 2$, and $\beta = 0$ and for a common range of smoothness index γ, the same order of convergence is observed for the Galerkin and the nonstationary Petrov–Galerkin methods.

9. Condition numbers. The final matrix A in (8.4) is symmetric and positive definite, thus

$$\mathrm{cond}_2(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.$$

To bound the condition number it is enough to find suitable lower and upper bounds for $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$, respectively. The lower bound of $\lambda_{\min}(A)$ depends on the smoothness of kernel Λ and the geodesic separation distance q_X of set X. The smoothness of Λ goes back to the smoothness of Φ and v_ρ and the order of differential operator L.

Our approach to bound $\lambda_{\min}(A)$ is the use of an inverse inequality in the trial space to turn the conditioning of the PDE matrix back to one of the approximation theory. Thus we first review the conditioning of the pure interpolation by kernel Φ, and then we give some inverse inequalities on the sphere.

LEMMA 9.1. *Assume Φ satisfies (5.2) for $\sigma > d/2$ and $B_{\Phi,X} = (\Phi(x_j, x_k))_{j,k=1}^N$ is the kernel interpolation matrix on $\mathbb{S}^d$. Then the minimum eigenvalue of $B_{\Phi,X}$ can be bounded by*

$$\lambda_{\min}(B_{\Phi,X}) \geqslant Cq_X^{2\sigma-d}.$$

Proof. First define the Euclidean separation distance $\widetilde{q}_X$ by

$$\widetilde{q}_X := \frac{1}{2}\min_{k\neq j}\|x_k - x_j\|_2,$$

and then apply Theorem 3.1 of [28] (see also [32, Theorem 12.3]) when X is viewed as a subset of $\mathbb{R}^{d+1}$. Note that q_X and $\widetilde{q}_X$ are comparable. ∎

The Bernstein-type inequalities control stronger Sobolev norms of trial functions by the weaker ones via the separation distance q_X. A global form of such inequalities on $\mathbb{R}^d$ was proved in [20] which bounds a stronger norm by the L_2-norm. An analogous inequality on $\mathbb{S}^d$ was established in [23] which reads as

$$\|u\|_{H^\alpha} \leqslant Cq_X^{-\alpha}\|u\|_{L_2} \tag{9.1}$$

for all $u \in V_{\Phi,X}$ and $0 \leqslant \alpha \leqslant \sigma$. Moreover, a Bernstein inequality which estimates the L_p Bessel-potential Sobolev norms of functions in this space in terms of q_X and the L_p norm of the function itself was proved in [17].

Here we need a generalization of (9.1) holding for Sobolev norms of order μ, $\mu \leqslant \alpha$, instead of the L_2-norm on the right-hand side.

LEMMA 9.2. *Assume* Φ *satisfies* (5.2) *for* $\sigma > d/2+\mu/2$, *where* $\mu \leqslant \alpha \leqslant \sigma$. *Then*

$$\|u\|_{H^\alpha} \leqslant C q_X^{\mu-\alpha} \|u\|_{H^\mu} \tag{9.2}$$

holds for all $u \in V_{\Phi,X}$.

Proof. Assume that

$$u = u_{\Phi,b} = \sum_{j=1}^{N} b_j \Phi(\cdot, x_j).$$

If u is expanded in the Fourier series (2.1), then $\widehat{u}_{\ell k} = \widehat{\phi}(\ell) \sum_{j=1}^{N} b_j Y_{\ell k}(x_j) =: \widehat{\phi}(\ell)\widetilde{b}_{\ell k}$. On the other hand assume Ψ is a zonal positive definite kernel satisfying

$$c_\psi (1+\ell)^{-2\sigma+\mu} \leqslant \widehat{\psi}(\ell) \leqslant \widetilde{c}_\psi (1+\ell)^{-2\sigma+\mu}$$

for $\sigma - \mu/2 > d/2$. Then we get

$$\begin{aligned}
\|u_{\Phi,b}\|_{H^\mu}^2 &= \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widehat{u}_{\ell k}^2 (1+\ell)^{2\mu} \\
&= \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widetilde{b}_{\ell k}^2 \widehat{\phi}(\ell)^2 (1+\ell)^{2\mu} \\
&\geqslant c_\phi^2 \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widetilde{b}_{\ell k}^2 (1+\ell)^{2\mu-4\sigma} \\
&\geqslant c_\phi^2 \widetilde{c}_\psi^{-2} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widetilde{b}_{\ell k}^2 \widehat{\psi}(\ell)^2 \\
&= c_\phi^2 \widetilde{c}_\psi^{-2} \|u_{\Psi,b}\|_{L_2}^2.
\end{aligned}$$

Now we invoke (9.1) for kernel Ψ under the condition $0 \leqslant \beta \leqslant \sigma - \mu/2$ to find

$$\|u_{\Psi,b}\|_{H^\beta} \leqslant C q_X^{-\beta} \|u_{\Psi,b}\|_{L_2}.$$

The left-hand side can be treated as

$$\begin{aligned}
\|u_{\Psi,b}\|_{H^\beta}^2 &= \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widetilde{b}_{\ell k}^2 (1+\ell)^{-4\sigma+2\mu+2\beta} \\
&\geqslant \widetilde{c}_\phi^{-2} \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widetilde{b}_{\ell k}^2 \widehat{\phi}(\ell)^2 (1+\ell)^{2\mu+2\beta} \\
&= \widetilde{c}_\phi^{-2} \|u_{\Phi,b}\|_{H^{\mu+\beta}}^2,
\end{aligned}$$

which is feasible for $\mu + \beta < 2\sigma - d/2$. Summarizing all, we have

$$\begin{aligned}\|u_{\Phi,b}\|^2_{H^{\mu+\beta}} &\leqslant \widetilde{c}_\phi^2 \|u_{\Psi,b}\|^2_{H^\beta} \\ &\leqslant C^2 \widetilde{c}_\phi^2 q_X^{-2\beta} \|u_{\Phi,b}\|^2_{L_2} \\ &\leqslant C^2 \widetilde{c}_\phi^2 c_\phi^{-2} \widetilde{c}_\psi^2 q_X^{-2\beta} \|u_{\Phi,b}\|^2_{H^\mu}.\end{aligned}$$

Setting $\alpha = \mu + \beta$ completes the proof. ∎

If we proceed with the proof of Theorem 6.6 of [15] we see that for a given $\rho < 1$ and for sufficiently large values of ℓ, say, $\ell > 1/\rho - (d-1)/2$, the lower bound of $\widehat{v}_\rho(\ell)$ is indeed given by $c(\rho\ell)^{-2\nu}$, which leads to

$$C\rho^{-2\nu}\|u\|_{H^{\sigma-\nu+\kappa/2}} \leqslant \|Lu * v_\rho\|_{H^{\sigma+\nu-\kappa/2}} \tag{9.3}$$

by modifying the proof of Theorem 8.2.

THEOREM 9.3. *Under the assumptions of Theorem* 8.4*, for a given* $\rho < 1$*, the minimum eigenvalue of* A *can be bounded by*

$$\lambda_{\min}(A) \geqslant C\rho^{-2\nu} q_X^{2\sigma+2\nu-\kappa-d}.$$

Proof. An appropriate formula for $\lambda_{\min}(A)$ is

$$\lambda_{\min}(A) = \min_{0\neq b\in\mathbb{R}^N} \frac{b^T A b}{\|b\|_2^2}.$$

For a given $b \in \mathbb{R}^N$ assume that $u_{\Lambda,b} := \sum_{j=1}^N b_j \Lambda(\cdot, x_j)$ and $u_{\Phi,b} := \sum_{j=1}^N b_j \Phi(\cdot, x_j)$. It is clear from (8.9) that $u_{\Lambda,b} = L u_{\Phi,b} * v_\rho$. Now we can write

$$\begin{aligned} b^T A b &= \|u_{\Lambda,b}\|^2_\Lambda \\ &\geqslant C\rho^{2\nu}\|u_{\Lambda,b}\|^2_{H^{\sigma+\nu-\kappa/2}} && \text{(using (8.21))} \\ &= C\rho^{2\nu}\|Lu_{\Phi,b} * v_\rho\|^2_{H^{\sigma+\nu-\kappa/2}} \\ &\geqslant C\rho^{-2\nu}\|u_{\Phi,b}\|^2_{H^{\sigma-\nu+\kappa/2}} && \text{(using (9.3))} \\ &\geqslant C\rho^{-2\nu} q_X^{2\nu-\kappa}\|u_{\Phi,b}\|^2_{H^\sigma} && \text{(using Lemma 9.2)} \\ &\geqslant C\rho^{-2\nu} q_X^{2\nu-\kappa}\|u_{\Phi,b}\|^2_\Phi \\ &= C\rho^{-2\nu} q_X^{2\nu-\kappa} b^T B_{\Phi,X} b \\ &\geqslant C\rho^{-2\nu} q_X^{2\nu-\kappa} q_X^{2\sigma-d}\|b\|_2^2 && \text{(using Lemma 9.1),} \end{aligned}$$

which completes the proof. In the fifth line above the inverse inequality in Lemma 9.2 for $\alpha = \sigma$ and $\mu = \sigma - \nu + \kappa/2$ was applied. This application requires, due to the assumption of Lemma 9.2, $\sigma - \mu/2 > d/2$ or $\sigma > d - \nu + \kappa/2$. But this imposes no additional condition because we have $\sigma \geqslant \nu + \kappa/2$ from Theorem 8.4 and, on the other hand, $\nu + \kappa/2 > d - \nu + \kappa/2$ holds since $\nu > d/2$. ∎

To bound $\lambda_{\max}(A)$, using the Gershgorin's theorem there exists an index $j \in \{1, 2, \ldots, N\}$ such that

$$|\lambda_{\max}(A) - \Lambda(x_j, x_j)| \leqslant \sum_{k=1, k\neq j}^N |\Lambda(x_j, x_k)|,$$

which becomes

$$\lambda_{\max}(A) \leqslant N \max_{x,y\in\mathbb{S}^d} |\Lambda(x, y)|.$$

Due to the definition of Λ via the convolution with factor ρ^{-d} instead of $\rho^{-(d+1)}$ the maximum value of Λ is of $\mathcal{O}(\rho^{-1})$. Moreover, if $X \in \mathcal{X}_R$, then $N = \mathcal{O}(h_X^{-d})$. Thus the upper bound for $\lambda_{\max}(A)$ is given by

$$\lambda_{\max}(A) \leqslant C\rho^{-1}h_X^{-d}.$$

By inserting the bounds of minimum and maximum eigenvalues into the formula of the condition number, we have the following corollary.

COROLLARY 9.4. *The condition number of the final linear system of the Petrov–Galerkin method is bounded by*

$$\operatorname{cond}_2(A) \leqslant C\rho^{2\nu-1}h_X^{-2\sigma+\kappa-2\nu}, \tag{9.4}$$

where h_X is the geodesic fill distance of set $X \in \mathcal{X}_R$ on $\mathbb{S}^d$ and σ, ν, and κ are those given in Theorem 8.4.

One can see that the stability of the scheme depends on the distribution of points and the smoothness of Λ, where the latter is determined by the smoothness of the trial and the test functions via σ and ν, respectively. The condition number grows polynomially when h_X goes to zero.

The condition number of the collocation method with the same kernel Φ can be bounded as [16, Theorem 5]

$$\operatorname{cond}_2(A) \leqslant Ch_X^{-2\sigma+\kappa-1}.$$

Compared with (9.4), this bound is the same as the bound of the condition number of the stationary Petrov–Galerkin method. However, the condition number of the nonstationary Petrov–Galerkin method is increased by factor $h_X^{1-2\nu}$.

10. Numerical integration on spherical caps. Although the analysis in the previous sections is presented for exact computations of involved integrals, in practice the scheme requires numerical integration over spherical caps $B(x_k, \rho)$ for all test points x_k from scattered data set X. Spherical caps may overlap each other, but we integrate independently over all of them. In fact, each spherical cap is the support of a test functional introduced in section 4 and it is independent of the other functionals. This is in contrast with the Galerkin methods where either integration over whole $\mathbb{S}^d$ is required for computing each element of the final matrix or a connected and non-overlapping spherical triangulation and then integration over triangles are needed.

Thanks to the work of Hesse and Womersly [10], numerical integrations over caps can be easily performed. Due to the following lemma it is sufficient to construct rules for spherical caps for one given fixed test point, and then rotate it to the other points. The proof uses the fact that the space of spherical harmonics on $\mathbb{S}^d$ is invariant under rotation.

LEMMA 10.1 (see [10]). *Let $B(z;\rho) \subset \mathbb{S}^d$ be the spherical cap with center $z \in \mathbb{S}^d$ and radius $\rho \in (0,\pi)$. Let $Q_{B(z;\rho),m}$ given by*

$$Q_{B(z;\rho),m}(f) := \sum_{j=1}^{M} w_j f(z_j), \quad f \in C(B(z;\rho)),$$

where $z_1, z_2, \ldots, z_M \in B(z;\rho)$ and $w_1, w_2, \ldots, w_M \in \mathbb{R}$, be a rule for numerical integration over the spherical cap $B(z;\rho)$ that is exact on $\mathcal{P}_m^d(B(z;\rho))$. Let $B(z';\rho) \subset \mathbb{S}^d$ be another spherical cap with center $z' \in \mathbb{S}^d$ and the same radius ρ, and let R denote any rotation on $\mathbb{R}^{d+1}$ such that $z' = Rz$. Then the rule $Q_{B(z';\rho),m}$ defined by

$$Q_{B(z';\rho),m}(f) := \sum_{j=1}^{M} w_j f(Rz_j), \quad f \in C(B(z';\rho)),$$

is a rule for numerical integration over $B(z';\rho)$, with nodes $Rz_1, Rz_2, \dots, Rz_M \in B(z';\rho)$, and this rule is exact on $\mathcal{P}_m^d(B(z';\rho))$.

Lemma 10.1 shows that it is sufficient to construct only rules for numerical integration over the north polar cap. We refer the reader to [10] for explicit formulas.

Finally, we note that the numerical integration will be a more difficult task if f is known only at a finite number of scattered points. This can be considered as a shortcoming of this and the Galerkin methods.

11. Numerical results. In this section we present the results of a numerical experiment for approximating the solution of differential equation (4.1) on $\mathbb{S}^2$, where $L = -\Delta_0 + I$. To construct a finitely smooth true solution, let $\{\xi_1, \dots, \xi_n\}$ be a set of n points on $\mathbb{S}^2$ and define

$$u(x) := \sum_{k=1}^{n} b_k \psi_4(\sqrt{2 - 2x^T\xi_k}), \quad x \in \mathbb{S}^2,$$

for some known coefficients b_k, where

$$\psi_\beta(r) = (\varepsilon r)^{\beta-3/2} K_{\beta-3/2}(\varepsilon r) \tag{11.1}$$

is the well-known Matérn kernel. Here K_β is the modified Bessel function of the second kind of order β. Since ψ_4 produces $H^4(\mathbb{R}^3)$, its restriction to $\mathbb{S}^2$ produces $H^{3.5}(\mathbb{S}^2)$. Theorem 8.3 can be applied to show that $u \in H^\gamma(\mathbb{S}^2)$ for any $\gamma < 6$. In this experiment the shape parameter $\varepsilon = 2$ and a set $\{\xi_1, \dots, \xi_{100}\}$ of scattered points on $\mathbb{S}^2$ [27] are used. Moreover, we set

$$\tilde{b} = (0.1, -0.2, 0.4, 0.3, -0.1, -0.4, 0.3, -0.5, 0.1, 0.2), \quad b = (\underbrace{\tilde{b}, -\tilde{b}, \tilde{b}, \dots, -\tilde{b}}_{10 \text{ times}}).$$

The right-hand-side function f is calculated using the fact that $\Delta_0\varphi(x^T\xi) = \mathcal{L}\varphi(t)$ for $t = x^T\xi$ with

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

The test function v_ρ is chosen to be the restriction of the compactly supported RBF $(1 - r/\rho)_+^2$ to $\mathbb{S}^2$. This kernel satisfies (7.1) with $\nu = 1.5$. The Matérn kernel $\Phi(x) = \psi_{5.5}(\|x\|_2)$, where ψ_β is defined in (11.1), with $\varepsilon = 5$ is employed to form the trial space. This kernel satisfies (5.2) with $\sigma = 5$. The equal area partitioning algorithm [27] is used to generate the sets of scattered quasi-uniform points on $\mathbb{S}^2$.

The results of the collocation, the Galerkin, the stationary Petrov–Galerkin, and the nonstationary Petrov–Galerkin methods are given in Tables 1, 2, 3, and 4, respectively. The L_2 errors are computed via a quadrature (with many points) on $\mathbb{S}^2$ by evaluating the discrete solution on the set of quadrature points. Since for X in a family of R-uniform sets on $\mathbb{S}^2$ we have $h_X = \mathcal{O}(N^{-1/2})$, the numerical orders are computed (for example, for errors) via

$$\log\left(\frac{\|e_{\text{old}}\|_{L_2}}{\|e_{\text{new}}\|_{L_2}}\right) \Big/ \log\left(\sqrt{\frac{N_{\text{new}}}{N_{\text{old}}}}\right).$$

The results of the collocation method (Table 1) are approximately the same as the results of the stationary Petrov–Galerkin method (Table 3). Although the orders

TABLE 1

The collocation method: the relative L_2-norm of $e = u - u_N$, the minimum and the maximum eigenvalues of the final matrix A, together with the numerical orders.

	Errors		Min. eigs.		Max. eigs.	
N	$\|e\|_{L_2}$	Orders	$\lambda_{\min}(A)$	Orders	$\lambda_{\max}(A)$	Orders
100	4.30e − 1	−	3.08e + 1	−	1.76e + 3	−
200	8.50e − 2	4.68	5.91e − 0	4.77	3.41e + 3	−1.90
400	1.06e − 2	6.01	9.08e − 1	5.40	6.73e + 3	−1.96
800	2.34e − 3	4.35	1.25e − 1	5.72	1.34e + 4	−1.98
1600	5.90e − 4	3.97	1.64e − 2	5.86	2.67e + 4	−1.99
3200	1.27e − 4	4.42	1.10e − 3	5.93	5.33e + 4	−2.00

TABLE 2

The Galerkin method: the relative L_2-norm of $e = u - u_N$, the minimum and the maximum eigenvalues of the final matrix A, together with the numerical orders.

	Errors		Min. eigs.		Max. eigs.	
N	$\|e\|_{L_2}$	Orders	$\lambda_{\min}(A)$	Orders	$\lambda_{\max}(A)$	Orders
100	1.21e − 1	−	8.91e + 2	−	2.56e + 5	−
200	2.40e − 2	4.66	1.21e + 2	5.76	5.05e + 5	−1.96
400	4.25e − 3	4.99	1.31e + 1	6.40	1.00e + 6	−1.98
800	6.97e − 4	5.22	1.22e − 0	6.86	2.00e + 6	−1.99
1600	1.03e − 4	5.51	1.12e − 1	6.91	4.00e + 6	−2.00
3200	1.78e − 5	5.08	2.59e − 3	10.85	8.00e + 6	−2.00

TABLE 3

The stationary Petrov–Galerkin method with $\rho = \pi/8h_X$: the relative L_2-norm of $e = u - u_N$, the minimum and the maximum eigenvalues of the final matrix A, together with the numerical orders.

	Errors		Min. eigs.		Max. eigs.	
N	$\|e\|_{L_2}$	Orders	$\lambda_{\min}(A)$	Orders	$\lambda_{\max}(A)$	Orders
100	3.94e − 1	−	6.39e + 2	−	3.76e + 4	−
200	8.01e − 2	4.59	1.73e + 2	3.77	1.03e + 5	−2.90
400	1.02e − 2	5.95	3.76e + 1	4.40	2.87e + 5	−2.97
800	2.14e − 3	4.50	7.34e − 0	4.72	8.08e + 5	−2.98
1600	5.48e − 4	3.93	1.36e − 0	4.86	2.28e + 6	−2.99
3200	1.18e − 4	4.43	2.46e − 1	4.93	6.44e + 6	−3.00

TABLE 4

The nonstationary Petrov–Galerkin method with $\rho = \pi/8$: the relative L_2-norm of $e = u - u_N$, the minimum and the maximum eigenvalues of the final matrix A, together with the numerical orders.

	Errors		Min. eigs.		Max. eigs.	
N	$\|e\|_{L_2}$	Orders	$\lambda_{\min}(A)$	Orders	$\lambda_{\max}(A)$	Orders
100	8.14e − 1	−	1.10e + 2	−	9.24e + 3	−
200	1.45e − 1	4.97	1.61e + 1	5.55	1.79e + 4	−1.91
400	1.82e − 2	6.00	1.76e − 0	6.39	3.54e + 4	−1.97
800	2.90e − 3	5.30	1.66e − 1	6.81	7.04e + 4	−1.98
1600	4.41e − 4	5.43	1.53e − 2	6.87	1.40e + 5	−1.99
3200	5.93e − 5	5.79	1.35e − 3	7.01	2.80e + 5	−2.00

of $\lambda_{\min}$ and $\lambda_{\max}$ are different, the order of $\text{cond}(A) = \lambda_{\max}/\lambda_{\min}$ is the same for both methods. We can roughly say, at the price of a more computational cost, the stationary Petrov–Galerkin method adds nothing to the simple and easily handled collocation method.

On the other hand the results of the Galerkin (Table 2) and the nonstationary Petrov–Galerkin (Table 4) methods are comparable. Both methods improve the errors

and the order of convergence of the previous methods, while they have worse numerical conditioning. Numerical results (approximately) confirm the theoretical order of convergence of the collocation method which is $\gamma - \kappa \approx 4$ from (8.22) and the nonstationary Petrov–Galerkin method which is $\gamma \approx 6$ from (8.19). On the other hand, the numerical order of convergence of the stationary Petrov–Galerkin method is much better than the theoretical order $\gamma - 4\nu$. This would leave some leeway for a better estimate in a future work.

In [21] the Galerkin method is implemented via spatially well-localized "small footprint" bases for the associated kernel space on $\mathbb{S}^2$. This leads to a well-conditioned numerical method with a condition number of order h_X^{-2}. The idea can, of course, be applied to the method of this paper. We do not pursue this further but leave it for a future study.

12. Final remarks and future research. An alternative approach for solving PDEs on the unit sphere was presented and analyzed in this work. Compared with the well-established collocation and Galerkin methods, a Petrov–Galerkin test discretization was applied to generate a set of convolution-form test functionals. The trial space was constructed by rotations of a restricted kernel on the unit sphere. The final error bounds were given for functions in Sobolev space $H^\gamma(\mathbb{S}^d)$ for $\max\{d/2-\epsilon, \kappa\} \leqslant \gamma \leqslant 2\tau$ with $\tau = \sigma + \kappa/2 - \nu$, where σ and ν are smoothness parameters of the trial and the test functions, respectively, and κ is the order of differential operator L. On the other hand, by applying an inverse inequality, the condition number of the final linear system was estimated. It was proved that the condition number depends on the smoothness of the underlying trial and test kernels, the scaling test parameter, and the distribution of scattered points, as it should be. New research can focus on theoretical and computational tricks to improve the condition number. For example, one can implement the present method in a multiscale setting, as has been done for the collocation method on spheres in [16] and on bounded Euclidean domains in [6]. Or, one can use the approach introduced in [8] which leads to a sparse final linear system by employing spatially well-localized "small footprint" bases for the associated kernel space.

Acknowledgment. The proof of Lemma 9.2 borrows an idea from an unpublished draft by Robert Schaback in Euclidean spaces and private communications with him.

REFERENCES

[1] S. N. Atluri and S. Shen, *The Meshless Local Petrov-Galerkin (MLPG) Method*, Tech Science Press, Encino, CA, 2002.

[2] V. Baramidze and M. J. Lai, *Spherical spline solution to a PDE on the sphere*, in Splines and Wavelets, G. Chen and M. J. Lai, eds., Nashboro Press, Brentwood, TN, 2006.

[3] M. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge University Press, Cambridge, UK, 2003.

[4] F. Dai and Y. Xu, *Approximation Theory and Harmonic Analysis on Spheres and Balls*, Springer, New York, 2013.

[5] C. F. Dunkl, *Operators and harmonic analysis on the sphere*, Math. Comp., 125 (1966), pp. 250–263.

[6] P. Farrell and H. Wendland, *RBF multiscale collocation for second order elliptic boundary value problems*, SIAM J. Numer. Anal., 51 (2013), pp. 2403–2425.

[7] W. Freeden, T. Gervens, and M. Schreiner, *Constructive Approximation on the Sphere with Applications to Geomathematics*, Clarendon Press, Oxford, UK, 1998.

[8] E. Fuselier, T. Hangelbroek, F. J. Narcowich, J. D. Ward, and G. B. Wright, *Localized bases for kernel spaces on the unit sphere*, SIAM J. Numer. Anal., 51 (2013), pp. 2538–2562.

[9] E. Fuselier and G. B. Wright, *Scattered data interpolation on embeded submanifolds with restricted positive definite kernels: Sobolev error estimates*, SIAM J. Numer. Anal., 50 (2012), pp. 1753–1776.
[10] K. Hesse and R. S. Womersley, *Numerical integration with polynomial exactness over a spherical cap*, Adv. Comput. Math., 36 (2012), pp. 451–483.
[11] S. Hubbert, Q. T. Le Gia, and T. M. Morton, *Spherical Radial Basis Functions, Theory and Applications*, Springer, New York, 2015.
[12] S. Hubbert and T. M. Morton, *A Duchon framework for the sphere*, J. Approx. Theory, 129 (2004), pp. 28–57.
[13] K. Jetter, J. Stöckler, and J. D. Ward, *Error estimates for scattered data interpolation on spheres*, Math. Comp., 68 (1999), pp. 733–747.
[14] Q. T. Le Gia, *Galerkin approximation for elliptic PDEs on spheres*, J. Approx. Theory, 130 (2004), pp. 123–147.
[15] Q. T. Le Gia, I. H. Sloan, and H. Wendland, *Multiscale analysis in Sobolev spaces on the sphere*, SIAM J. Numer. Anal., 48 (2010), pp. 2065–2090.
[16] Q. T. Le Gia, I. H. Sloan, and H. Wendland, *Multiscale RBF collocation for solving PDEs on spheres*, Numer. Math., 121 (2012), pp. 99–125.
[17] H. N. Mhaskar, F. J. Narcowich, J. Prestin, and J. D. Ward, *L_p Bernstein estimates and approximation by spherical basis functions*, Math. Comp., 79 (2010), pp. 1647–1679.
[18] T. M. Morton and M. Neamtu, *Error bonds for solving pseudodifferential equations on spheres*, J. Approx. Theory, 114 (2002), pp. 242–268.
[19] C. Müller, *Spherical Harmonics*, Lecture Notes in Math., 17, Springer-Verlag, Berlin, 1966.
[20] F. J. Narcowich, J. D. Ward, and H. Wendland, *Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions*, Constr. Approx., 24 (2006), pp. 175–186.
[21] F. J. Narcowich, S. T. Rowe, and J. D. Ward, *A novel Galerkin method for solving PDEs on the sphere using highly localized kernel bases*, Math. Comp., 86 (2017), pp. 197–231.
[22] F. J. Narcowich, X. Sun, and J. D. Ward, *Approximation power of RBFs and their associated SBFs: A connection*, Adv. Comput. Math., 27 (2007), pp. 107–124.
[23] F. J. Narcowich, X. Sun, J. D. Ward, and H. Wendland, *Direct and inverse Sobolev error estimates for scattered data interpolation via spherical basis functions*, Found. Comput. Math., 7 (2007), pp. 369–390.
[24] 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.
[25] T. D. Pham, T. Tran, and A. Chernov, *Pseudodifferential equations on the sphere with spherical splines*, Math. Models Methods Appl. Sci., 21 (2011), pp. 19–33.
[26] J. R. Driscoll and D. M. Healy, *Computing Fourier transforms and convolutions on the 2-sphere*, Adv. Appl. Math., 15 (1994), pp. 202–250.
[27] E. B. Saff and A. B. J. Kuijlaars, *Distributing many points on a sphere*, Math. Intell., 19 (1997), pp. 5–11.
[28] R. Schaback, *Error estimates and condition numbers for radial basis function interpolation*, Adv. Comput. Math., 3 (1995), pp. 251–264.
[29] R. Schaback, *A unified theory of radial basis functions. Native Hilbert spaces for radial basis functions* II, J. Comput. Appl. Math., 121 (2000), pp. 165–177.
[30] I. J. Schoenberg, *Positive definite function on spheres*, Duke Math. J., 9 (1942), pp. 96–108.
[31] H. Wendland, *Piecewise polynomial, positive definite and compactly supported radial basis functions of minimal degree*, Adv. Comput. Math., 4 (1995), pp. 389–396.
[32] H. Wendland, *Scattered Data Approximation*, Cambridge University Press, Cambridge, UK, 2005.
[33] Y. Xu and E. W. Cheney, *Strictly positive definite functions on spheres*, Proc. Amer. Math. Soc., 116 (1992), pp. 977–981.
[34] W. zu Castell and F. Filbir, *Radial basis functions and corresponding zonal series expansions on the sphere*, J. Approx. Theory, 134 (2005), pp. 65–79.