1-1-2013

Localized Bases for Kernel Spaces on the Unit Sphere

E. Fuselier
High Point University

T. Hangelbroek
University of Hawaii

F. J. Narcowich
Texas A & M University

J. D. Ward
Texas A & M University

G. B. Wright
Boise State University

First published in *SIAM Journal on Numerical Analysis* in Volume 51, Issue 5, published by the Society of Industrial and Applied Mathematics (SIAM).

SIAM J. NUMER. ANAL.
Vol. 51, No. 5, pp. 2538–2562

LOCALIZED BASES FOR KERNEL SPACES ON THE UNIT SPHERE*

E. FUSELIER†, T. HANGELBROEK‡, F. J. NARCOWICH§, J. D. WARD§, AND G. B. WRIGHT¶

Abstract. Approximation/interpolation from spaces of positive definite or conditionally positive definite kernels is an increasingly popular tool for the analysis and synthesis of scattered data and is central to many meshless methods. For a set of N scattered sites, the standard basis for such a space utilizes N *globally* supported kernels; computing with it is prohibitively expensive for large N. Easily computable, well-localized bases with "small-footprint" basis elements—i.e., elements using only a small number of kernels—have been unavailable. Working on $\mathbb{S}^2$, with focus on the restricted surface spline kernels (e.g., the thin-plate splines restricted to the sphere), we construct easily computable, spatially well-localized, small-footprint, robust bases for the associated kernel spaces. Our theory predicts that each element of the local basis is constructed by using a combination of only $\mathcal{O}((\log N)^2)$ kernels, which makes the construction computationally cheap. We prove that the new basis is L_p stable and satisfies polynomial decay estimates that are stationary with respect to the density of the data sites, and we present a quasi-interpolation scheme that provides optimal L_p approximation orders. Although our focus is on $\mathbb{S}^2$, much of the theory applies to other manifolds—$\mathbb{S}^d$, the rotation group, and so on. Finally, we construct algorithms to implement these schemes and use them to conduct numerical experiments, which validate our theory for interpolation problems on $\mathbb{S}^2$ involving over 150,000 data sites.

Key words. interpolation, thin-plate splines, sphere, kernel approximation

AMS subject classifications. 41A05, 41A30, 41A63, 65D05

DOI. 10.1137/120876940

1. Introduction.

Approximation/interpolation with positive definite or conditionally positive definite kernels is an increasingly popular tool for analyzing and synthesizing of scattered data and is central to many meshless methods. A basic challenge in using this tool is that well-localized bases with "small-footprint" elements—i.e., elements using only a small number of kernels—have been unavailable. With this in mind, we have two main goals for this paper.

The first is the theoretical development of small-footprint bases that are well-localized spatially, for a variety of kernels. For important classes of kernels on $\mathbb{S}^2$, the theory itself predicts that a basis element requires only $\mathcal{O}(\log(N)^2)$ kernels, where N is the number of data sites.

Previous numerical experiments on data sets, with N on the order of 1000, proved quite successful, but the method for determining the localized basis elements does not scale. The predictions of our theory, on the other hand, have been verified numerically on $\mathbb{S}^2$ for data sets with over 100,000 sites.

*Received by the editors May 14, 2012; accepted for publication (in revised form) June 17, 2013; published electronically September 10, 2013.
http://www.siam.org/journals/sinum/51-5/87694.html

†Department of Mathematics, High Point University, High Point, NC 27262 (efuselie@highpoint.edu).

‡Department of Mathematics, University of Hawaii, Honolulu, HI 96822 (hangelbr@math.hawaii.edu). This author's research was supported by grant DMS-1232409 from the National Science Foundation.

§Department of Mathematics, Texas A&M University, College Station, TX 77843 (fnarc@math.tamu.edu, jward@math.tamu.edu). These authors were supported by grant DMS-1211566 from the National Science Foundation.

¶Department of Mathematics, Boise State University, Boise, ID 83725 (gradywright@boisestate.edu). This author's research was supported by grants DMS-0934581 and DMS-1160379 from the National Science Foundation.

Our second goal is to show how to easily and efficiently compute these small-footprint, robust (i.e., well-localized, L_p stable) bases for spaces associated with restricted surface-spline kernels on the sphere $\mathbb{S}^2$. The kernels in question are spherical basis functions having the form

$$k_m(x,\alpha) := (-1)^m (1 - x\cdot\alpha)^{m-1} \log(1 - x\cdot\alpha) \tag{1.1}$$

for $m = 2, 3, \ldots$ (cf. [18, equation 3.3]). The kernel spaces are denoted $S_m(\Xi)$—these are finite dimensional spaces of functions obtained as linear combinations of k_m, sampled at some (finite) set of nodes $\Xi \subset \mathbb{S}^2$, plus a spherical polynomial p of degree $m-1$, i.e., $\sum_{\xi\in\Xi} a_\xi k_m(\cdot,\xi) + p(\cdot)$. The coefficients involved satisfy the simple side conditions given in 3.1.

The Lagrange functions χ_ξ, which interpolate cardinal sequences $\chi_\xi(\zeta) = \delta_{\xi,\zeta}$, $\zeta \in \Xi$, form a basis for $S_m(\Xi)$. Recently, it has been shown in [12], for restricted surface splines as well as many other kernels, that these functions decay extremely rapidly away from ξ. Thus, $\{\chi_\xi\}_{\xi\in\Xi}$ forms a basis that is theoretically quite good (sufficient to demonstrate that the Lebesgue constant is uniformly bounded, among many other things). However, determining a Lagrange basis function generally requires solving a full linear system with at least $N := \#\Xi$ unknowns, so working with this basis directly is computationally expensive. In this paper we consider an alternative basis, one that shares many of the nice properties of the Lagrange basis, yet whose construction is computationally cheap.

Here is what we would desire in an easily computed, robust basis $\{b_\xi\}_{\xi\in\Xi}$ for $S_m(\Xi)$. Each basis function should be highly localized with respect to the mesh norm $h := \max_{x\in\mathbb{S}^2} \operatorname{dist}(x,\Xi)$ of Ξ. Moreover, each should have a nearly stationary construction. By this we mean that each basis element b_ξ is of the form $\sum_{\eta\in\Upsilon(\xi)} A_{\xi,\eta} k_m(\cdot,\eta) + p_\xi$, where the coefficients $A_{\xi,\eta}$ and the degree $m-1$ polynomial p_ξ are completely determined by k_m and a small subset of centers $\Upsilon(\xi)$. Specifically, we wish b_ξ to satisfy the following requirements:

$$\begin{aligned} &\text{(i)} \quad \#\Upsilon(\xi) = \mathrm{c}(N), \\ &\text{(ii)} \quad |b_\xi(x)| \le \sigma(r/h),\ r := \operatorname{dist}(x,\xi), \end{aligned}$$

where the number of points influencing each basis function $\mathrm{c}(N)$ is constant or slowly growing with N, and the function $\sigma(\cdot)$ decays rapidly—at an exponential rate $\sigma(t) \le Ce^{-\nu|t|}$ or at least at a fast polynomial rate $\sigma(t) \le C(1+|t|)^{-J}$. The B-spline basis, constructed from the family of truncated power functions (i.e., using $(x-y)_+^m$ in place of $k_m(x,y)$), is a model solution to the problem we consider.

Main results. The solution we present is to consider a basis of "local Lagrange" functions, which are constructed below in section 3. It has the following properties:

- *Numerical stability.* For any $J > 2$, one can construct a numerically stable basis with decay $\sigma(t) \le C(1+|t|)^{-J}$.
- *Small footprint.* Each basis function is determined by a relatively small set of centers: $\mathrm{c}(N) \le M(\log N)^2$, where the constant M is proportional to the square of the rate of decay J: $M \propto J^2$.
- *L_p stability.* The basis is stable in L_p: sequence norms $\|c\|_{\ell_p}$ of the coefficients are comparable to L_p norms of the expansion $\sum_{\xi\in\Xi} c_\xi b_\xi$.
- *Near-best L_∞ approximation.* For sufficiently large J, the operator $Q_\Xi f = \sum_{\xi\in\Xi} f(\xi) b_\xi$ provides near-best L_∞ approximation.

Preconditioners. Over the years practical implementation of kernel approximation has progressed despite the ill-conditioning of kernel bases. This has happened

with the help of clever numerical techniques like multipole methods and other fast methods of evaluation [1, 4, 13] and often with the help of preconditioners [3, 7, 15, 25]. Many results already exist in the radial basis function (RBF) literature concerning preconditioners and "better" bases. For a good list of references and further discussion, see [6]. Several of these papers use local Lagrange functions in their efforts to efficiently construct interpolants, but the number of points chosen to localize the Lagrange functions are based on experimental evidence. For example, Faul and Powell, in [8], devise an algorithm which converges to a given RBF interpolant that is based on local Lagrange interpolants using about thirty nearby centers. Beatson, Cherrie, and Mouat, in [2, Table 1, p. 260], use fifty local centers in their construction along with a few "far away" points to control the growth of the local interpolant at a distance from the center. In other work, Ling and Kansa [15] and co-workers have studied approximate cardinal basis functions based on solving least squares problems.

An offshoot of our results is a strategy for selecting centers for preconditioning (as in [8] and [2]) that scales correctly with the total number of centers N. We demonstrate the power of this approach in section 7, where the local basis is used to successfully precondition kernel interpolation problems varying in size by several orders of magnitude.

Organization. We now sketch the outline of the remainder of the article. In section 2 we give some background on kernel approximation and analysis on spheres. Section 3 presents the construction of the local Lagrange basis. Much of the remainder of the article is devoted to proving that this basis has the desired properties mentioned above. However, doing this will first require a thorough understanding of the (full) Lagrange basis $\{\chi_\xi\}_{\xi\in\Xi}$, which we study in detail in sections 4 and 5.

In section 4 we consider the full Lagrange basis: the stable, local bases constructed in [12]. We demonstrate, numerically, the decay of these functions as well as the coefficients used in their construction. These numerical observations confirm the theory in section 5, where it is proved that the Lagrange coefficients indeed decay quickly and stationarily with respect to h as ζ moves away from ξ.

Section 6 treats the main arguments of the paper. Section 6.1 introduces the *truncated* Lagrange functions (essentially $\widetilde{\chi}_\xi = \sum_{\zeta\in\Upsilon(\xi)} A_{\xi,\zeta} k_m(\cdot,\zeta)$), obtained by thresholding most of the coefficients of the Lagrange function. It demonstrates the basis properties of these functions and discusses an extra adjustment to the coefficients which is necessary to make each $\widetilde{\chi}_\xi$ satisfy moment conditions associated to the restricted surface splines.

2. Background.

2.1. The sphere. We denote by $\mathbb{S}^2$ the unit sphere in $\mathbb{R}^3$, and by μ we denote Lebesgue measure. The distance between two points, x and ξ, on the sphere is written $\operatorname{dist}(x,\xi) := \arccos(x\cdot\xi)$. The basic neighborhood is the spherical "cap" $B(\alpha,r) := \{x\in\mathbb{S}^2 : \operatorname{dist}(x,\alpha) < r\}$. The volume of a spherical cap is $\mu(B(\alpha,r)) = 2\pi(1-\cos r)$.

Throughout this article, Ξ is assumed to be a finite set of distinct nodes on $\mathbb{S}^2$, and we denote the number of elements in Ξ by $\#\Xi$. The *mesh norm* or *fill distance*, $h := h(\Xi,\mathbb{S}^2) := \max_{x\in\mathbb{S}^2}\operatorname{dist}(x,\Xi)$, measures the density of Ξ in $\mathbb{S}^2$. The *separation radius* is $q := \frac{1}{2}\min_{\xi\neq\zeta}\operatorname{dist}(\zeta,\xi)$, where $\xi,\zeta\in\Xi$, and the *mesh ratio* is $\rho_\Xi := h/q$.

Our results will be coordinate independent. Nevertheless, it is important to provide various geometric quantities in spherical coordinates. Given a north pole on $\mathbb{S}^2$, we will use the longitude $\theta_1\in[0,2\pi)$ and the colatitude $\theta_2\in[0,\pi]$ as coordinates. The metric tensor g_{ij} in these coordinates is a diagonal matrix with $g_{11} = \sin^2(\theta_2)$ and $g_{22} = 1$. As is customary in differential geometry, the inverse of g_{ij} is denoted by

g^{ij}. This is of course diagonal and has the form $g^{11} = (\sin^2(\theta_2))^{-1}$ and $g^{22} = 1$. Also, the measure $d\mu = \sin(\theta_2)d\theta_1 d\theta_2$, and the Laplace–Beltrami operator is then given by

$$\Delta = \frac{1}{\sin^2(\theta_2)}\frac{\partial^2}{\partial\theta_1^2} + \frac{1}{\sin(\theta_2)}\frac{\partial}{\partial\theta_2}\sin(\theta_2)\frac{\partial}{\partial\theta_2}.$$

Spectral properties of the Laplace–Beltrami operator. For each $\ell \in \mathbb{N}$, the eigenvalues of the negative of the Laplace–Beltrami operator, $-\Delta$, have the form $\nu_\ell := \ell(1+\ell)$; these have multiplicity $2\ell+1$. For each fixed ℓ, the eigenspace $\mathcal{H}_\ell$ has an orthonormal basis of $2\ell+1$ eigenfunctions, $\{Y_\ell^\mu\}_{\mu=-\ell}^{\ell}$, the *spherical harmonics* of degree ℓ. The space of spherical harmonics of degree $\ell \le \sigma$ is $\Pi_\sigma = \bigoplus_{\ell\le\sigma}\mathcal{H}_\ell$ and has dimension $(\sigma+1)^2$. These are the basic objects of Fourier analysis on the sphere. In order to simplify notation, we often denote a generic basis for Π_σ as $(\phi_j)_{j=1\ldots(\sigma+1)^2}$. We deviate from this only when a specific basis of spherical harmonics is required.

Covariant derivatives and smoothness spaces. The covariant derivatives $\{\nabla^k\}_{k=1}^{\infty}$ play an important role in defining smoothness spaces and in proving results about surface splines, but they play no role in the actual implementation of the algorithms. For a detailed discussion of these operators, where the relevant concepts are developed for Riemannian manifolds (including $\mathbb{S}^2$), see [11, section 2]. Using these derivatives and the standard inner product on the space of tensors, we have

$$\langle\nabla^m f, \nabla^m g\rangle_x = \sum_{i_1,\ldots,i_m}(\nabla^m f(x))_{i_1,\ldots,i_m}(\nabla^m g(x))_{i_1,\ldots,i_m}g^{i_1,i_1}(x)\ldots g^{i_m,i_m}(x)$$

with the norm being $|\nabla^m f(x)| := \sqrt{\langle\nabla^m f, \nabla^m f\rangle_x}$. For each m and each measurable subset of a manifold, the L_2 Sobolev norm may be defined this way:

$$\|f\|_{W_2^m(\Omega)} := \left(\sum_{k\le m}\int_\Omega |\nabla^k f(x)|^2 \mathrm{d}\mu(x)\right)^{1/2}.$$

2.2. Conditionally positive definite kernels and interpolation. Many of the useful computational properties of restricted surface splines stem from the fact that they are *conditionally positive definite*.

DEFINITION 2.1. *A kernel k is conditionally positive definite with respect to a finite dimensional space Π if, for any set of N distinct centers Ξ, the matrix $\mathrm{K}_\Xi := (k(\xi,\zeta))_{\xi,\zeta\in\Xi}$ is positive definite on all vectors $\mathbf{a} \ne \mathbf{0} \in \mathbb{C}^N$ satisfying $\sum_{\xi\in\Xi} a_\xi p(\xi) = 0$ for $p \in \Pi$.*

Here is an example pertinent to the kernels discussed in this paper. Let $(\varphi_j)_{j\in\mathbb{N}}$ be the complete set of eigenfunctions of the Laplace–Beltrami for some compact Riemannian manifold $\mathbb{M}$. Consider a kernel

$$k(x,y) := \sum_{j\in\mathbb{N}}\tilde{k}(j)\varphi_j(x)\overline{\varphi_j(y)} \tag{2.1}$$

having all but finitely many coefficients $\tilde{k}(j)$ positive and $\sum_{j\in\mathbb{N}}|\tilde{k}(j)|\|\varphi_j\|_\infty^2$ being finite; k is thus continuous and the series is uniformly convergent. We claim that k is conditionally positive definite with respect to the finite dimensional space $\Pi := \mathrm{span}(\varphi_j \mid j \in \mathcal{J})$, where $\mathcal{J} = \{j \mid \tilde{k}(j) \le 0\}$. Indeed, if $\sum_\xi a_\xi\varphi_j(\xi) = 0$ for $j \in \mathcal{J}$, then

$$\sum_{\xi\in\Xi}\sum_{\zeta\in\Xi}a_\xi k(\xi,\zeta)\overline{a_\zeta} = \sum_{j\in\mathbb{N}}\tilde{k}(j)\sum_{\xi,\zeta\in\Xi}a_\xi\varphi_j(\xi)\overline{a_\zeta\varphi_j(\zeta)} = \sum_{j\notin\mathcal{J}}\tilde{k}(j)|\mathbf{a}\varphi_j|_\Xi|^2 \ge 0.$$

We must show that if this is 0, then $\mathbf{a} = \mathbf{0}$. Suppose the quadratic form above is 0. Then all of the terms in the series vanish, and $\sum_\xi a_\xi\varphi_j(\xi) = 0$ for *all* $j \in \mathbb{N}$, not just

those in $\mathcal{J}$. Equivalently, for all $j \in \mathbb{N}$, $u(\varphi_j) = 0$, where $u := \sum_\xi a_\xi \delta_\xi(\cdot)$. However, in the proof of [19, Theorem 3.6] it was shown that this implies that the distribution $u \equiv 0$ and, consequently, that $\mathbf{a} = \mathbf{0}$.

Conditionally positive definite kernels are important for the following interpolation problem. Suppose $\Xi \subset \mathbb{S}^2$ is a set of nodes on the sphere, $f : \mathbb{S}^2 \to \mathbb{R}$ is some target function, and $f|_\Xi$ are the samples of f at the nodes in Ξ. We look for a function that interpolates this data from the space

$$S(k,\Xi) := S(k,\Xi,\Pi) := \left\{ \sum_{\xi\in\Xi} a_\xi k(\cdot,\xi) \mid \sum_{\xi\in\Xi} a_\xi p(\xi) = 0,\ \forall p \in \Pi \right\} + \Pi.$$

Provided $\Xi \subset \mathbb{S}^2$ is unisolvent with respect to Π (meaning that $p(\xi) = 0$ for all $\xi \in \Xi$ implies that $p = 0$ for any $p \in \Pi$), the unique interpolant from $S(k,\Xi)$ can be written

$$s(\cdot) = \sum_{\xi\in\Xi} a_\xi k(\cdot,\xi) + \sum_{j\in\mathcal{J}} c_j \varphi_j(\cdot),$$

where the expansion coefficients satisfy the (nonsingular) linear system of equations:

$$\begin{pmatrix} \mathrm{K}_\Xi & \Phi \\ \Phi^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{c} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{0} \end{pmatrix}, \tag{2.2}$$

where $\mathrm{K}_\Xi = (k(\xi_i,\xi_j))$, $i,j = 1,\dots,N$, and $\Phi = (\varphi_j(\xi_i))$, $i = 1,\dots,N$, $j \in \mathcal{J}$. This interpolant plays a dual role as the minimizer of the seminorm $|\cdot|_k$ induced from the "native space" semiinner product

$$\langle u, v\rangle_k = \left\langle \sum_{j\in\mathbb{N}} \hat{u}(j)\varphi_j, \sum_{j\in\mathbb{N}} \hat{v}(j)\varphi_j \right\rangle_k := \sum_{j\notin\mathcal{J}} \frac{\hat{u}(j)\overline{\hat{v}(j)}}{\tilde{k}(j)}. \tag{2.3}$$

Namely, it is the interpolant to $\mathbf{f}$ having minimal seminorm $|u|_k = \sqrt{\langle u,u\rangle_k}$.

3. Constructing the local Lagrange basis. The restricted surface splines k_m (see (1.1)) are conditionally positive definite with respect to the space of spherical harmonics of degree up to $m-1$, i.e., Π_{m-1}. The finite dimensional spaces associated with these kernels are denoted as in the previous section:

$$\begin{aligned} S_m(\Xi) &:= S(k_m,\Xi,\Pi_{m-1}) \\ &= \left\{ \sum_{\xi\in\Xi} a_\xi k_m(\cdot,\xi) \mid \sum_{\xi\in\Xi} a_\xi \phi(\xi) = 0,\ \forall \phi \in \Pi_{m-1} \right\} + \Pi_{m-1}. \end{aligned} \tag{3.1}$$

The goal of this section is to provide an easily constructed, robust basis for $S_m(\Xi)$. The fundamental idea behind building this basis is to associate with each $\xi \in \Xi$, a new basis function that interpolates over a relatively small set of nodes a function that is cardinal at ξ.

Specifically, let $\Upsilon(\xi)$ be a set of $n \ll N$ nearest neighbors to the node ξ, including the node ξ; see Figure 3.1 for an illustration. Then the new basis function associated with ξ is given by

$$\check{\chi}_\xi(\cdot) = \sum_{\zeta\in\Upsilon(\xi)} A_{\xi,\zeta} k_m(\cdot,\zeta) + \sum_{1\le j\le m^2} c_{\xi,j}\phi_j(\cdot), \tag{3.2}$$

where ϕ_j are a basis for the spherical harmonics of degree $\le m-1$. The coefficients

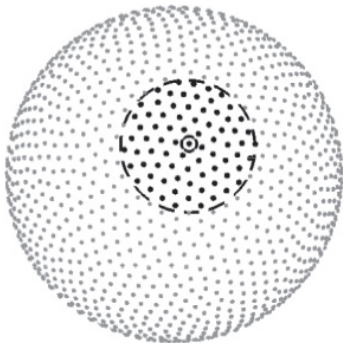

FIG. 3.1. *Illustration of the centers that make up the local Lagrange basis. The solid gray and black spheres mark the set of N nodes making up* Ξ. *The solid black sphere with a circle around it marks the node* ξ, *where a local Lagrange function* $\check{\chi}_\xi$ *is to be computed. The the solid black spheres enclosed in the dashed circular line mark the set of* $n = M(\log N)^2$ *centers* $\Upsilon(\xi)$ *used to compute* $\check{\chi}_\xi$. *For each* $\xi \in \Xi$, *a similar set* $\Upsilon(\xi)$ *is determined for computing* $\check{\chi}_\xi$.

$A_{\xi,\zeta}$ and $c_{\xi,j}$ are determined from the cardinal conditions

$$\check{\chi}_\xi(\zeta) = \begin{cases} 1 & \text{if } \zeta = \xi, \\ 0 & \text{if } \zeta \in \Upsilon(\xi) \setminus \xi, \end{cases} \quad \text{and} \quad \sum_{\zeta \in \Upsilon(\xi)} A_{\xi,\zeta}\phi_j(\zeta) = 0,\ 1 \le j \le m^2. \tag{3.3}$$

These coefficients can be determined by solving the (small) linear system

$$\begin{pmatrix} \mathrm{K}_{\Upsilon(\xi)} & \Phi \\ \Phi^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{A}_\xi \\ \mathbf{c}_\xi \end{pmatrix} = \begin{pmatrix} \mathbf{y}_\xi \\ \mathbf{0} \end{pmatrix}, \tag{3.4}$$

where $\mathbf{y}_\xi$ represents the cardinal data and the entries of the matrix follow from (2.2). We call $\check{\chi}_\xi$ a *local Lagrange function* about ξ.

The new basis for $S_m(\Xi)$ will consist of the collection of all the local Lagrange functions for the nodes in Ξ. It will be shown in section 6.3 that choosing the number of nearest neighbors to each ξ as $n = M(\log N)^2$ will give a basis with sufficient locality. The choice of M is related to the polynomial rate of decay of $\check{\chi}_\xi$ away from its center, and a priori estimates are given for M in section 6.3. However, in practice it will be sufficient to choose M by tuning it appropriately to get the desired rate of decay.

The exact details of the algorithm for constructing this basis then proceed as follows: For each $\xi \in \Xi$

1. find the $n = M(\log N)^2$ nearest neighbors to ξ, $\Upsilon(\xi)$,
2. construct $\check{\chi}_\xi$ according to the conditions (3.3), which amounts to solving the associated linear system (3.4) and storing the coefficients $\mathbf{A}_\xi$, $\mathbf{c}_\xi$.

We note that each set $\Upsilon(\xi)$ can be determined in $\mathcal{O}(\log N)$ operations by using a KD-tree algorithm for sorting and searching through the nodes Ξ. After the initial construction of the KD-tree, which requires $\mathcal{O}(N(\log N)^2)$, the construction of all the sets $\Upsilon(\xi)$ thus takes $\mathcal{O}(N(\log N)^2)$ operations.

Before continuing, we note that our main results, given in Theorem 6.5 and its corollaries, depend heavily on properties that this local Lagrange basis inherits from the full Lagrange basis $\{\chi_\xi\}_{\xi\in\Xi}$. Thus, much of what follows is spent on developing a working understanding of the full Lagrange basis and its connections to the local Lagrange basis. Even though the local Lagrange basis is the focus of our work, we will delay any further mention of $\{\check{\chi}_\xi\}_{\xi\in\Xi}$ until section 6.3.

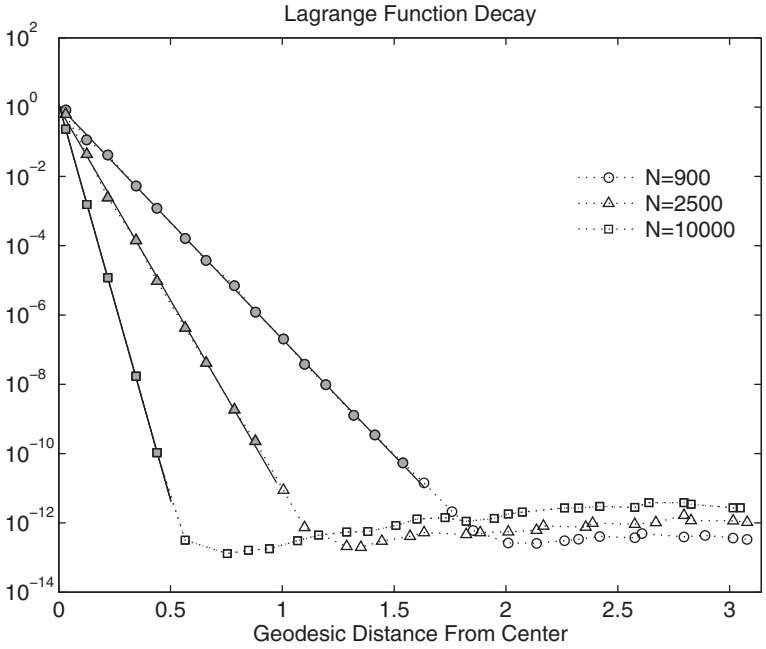

FIG. 4.1. *Maximum latitudinal values of the Lagrange function for the kernel $k_2(x,\alpha)$. This experiment was carried out in double precision arithmetic, and the plateau at roughly 10^{-11} occurs due to ill-conditioning of the collocation matrices and truncation error.*

4. The full Lagrange basis: Numerical observations. In this section we numerically examine a full Lagrange basis function χ_ξ and its associated coefficients for the kernel $k_2(x,\alpha) = (1 - x\cdot\alpha)\log(1 - x\cdot\alpha)$, the second order restricted surface spline (also known as the thin plate spline) on $\mathbb{S}^2$. First, we demonstrate numerically that χ_ξ decays exponentially away from its center. Second, we provide the initial evidence that the Lagrange coefficients decay at roughly the same rate, which is proved later in Theorem 5.3.

The full Lagrange function centered at ξ takes the form $\chi_\xi = \sum_{\zeta\in\Xi} A_{\xi,\zeta}k(\cdot,\zeta)+p_\xi$, where p_ξ is a degree 1 spherical harmonic. In this example, we use the "minimal energy points" of Womersley for the sphere—these are described and distributed at the website [27].[1] Because of the quasi-uniformity of the minimal energy point sets, it is sufficient to consider the Lagrange function χ_ξ centered at the north pole $\xi = (0,0,1)$.

Figure 4.1 displays the maximal colatitudinal values[2] of $|\chi_\xi|$. Until a terminal value of roughly 10^{-11}, we clearly observe the exponential decay of the Lagrange function, which follows

$$|\chi_\xi(x)| \le C_L \exp\left(-\nu_L \frac{\operatorname{dist}(x,\xi)}{h}\right). \tag{4.1}$$

(This "plateau" at 10^{-11} is caused by roundoff error—see Figure 4.3.) The estimate (4.1) has in fact been proved in [12, Theorem 5.3], where this and other analytic

[1] These point sets are used as benchmarks: each set of centers has a nearly identical mesh ratio, and the important geometric properties (e.g., fill distance and separation distance) are explicitly documented.

[2] The function χ_ξ is evaluated on a set of points (θ_1,θ_2) with 152 equispaced longitudes $\theta_1 \in [0,2\pi]$ and 179 equispaced colatitudes $\theta_2 \in [0,\pi]$.

TABLE 4.1

Estimates of the decay constants ν and C for Lagrange functions and coefficients on the sphere using the kernel $k_2(x,\alpha)$ with relevant geometric measurements of the minimum energy node sets used.

N	h_X	ρ_X	ν_L	C_L	ν_c	C_c
400	0.1136	1.2930	1.1119	0.8382	1.0997	0.5402
900	0.0874	1.5302	1.3556	1.0982	1.3445	0.7554
1600	0.0656	1.5333	1.3513	1.2170	1.3216	0.5946
2500	0.0522	1.5278	1.3345	0.9618	1.3117	0.5494
5041	0.0365	1.5304	1.3395	1.1080	1.3158	0.6188
10000	0.0260	1.5421	1.3645	1.1934	1.3369	0.7291

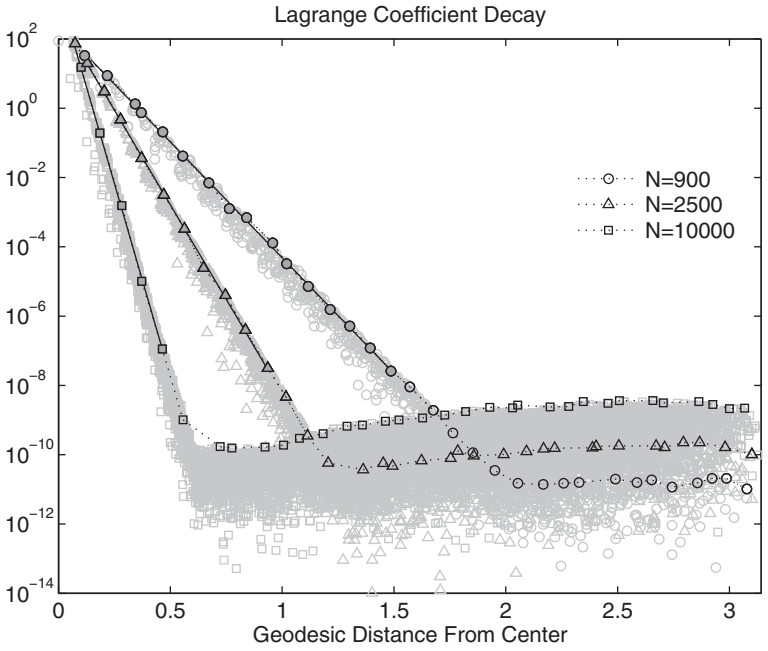

FIG. 4.2. *Plot of coefficients for a Lagrange function in the kernel space $S(k_2,\Xi)$. This experiment was carried out in double precision arithmetic.*

properties of bases for $S_m(\Xi)$ were studied in detail. By fitting a line to the data in Figure 4.1 where the exponential decay is evident, one can estimate the constants ν_L and C_L, which in this case are quite reasonable. For example, the value of ν_L, which measures the rate of exponential decay, is observed to be close to 1.35 (see Table 4.1).

We can visualize the decay of the corresponding coefficients in the same way. We again take the Lagrange function centered at the north pole: for each $\zeta' \in \Xi$, the coefficient $|A_{\xi,\zeta'}|$ in the expansion $\chi_\xi = \sum A_{\xi,\zeta} k(\cdot,\zeta) + p_\xi$ is plotted with horizontal coordinate $\mathrm{dist}(\xi,\zeta')$. The results for sets of centers of size $N = 900, 2500$, and 10,000 are given in Figure 4.2. The exponential decay seems to follow

$$|A_{\zeta,\xi}| \le C_c q^{-2} \exp\left(-\nu_c \frac{\mathrm{dist}(\xi,\zeta)}{h}\right).$$

Indeed, this is established later in Theorem 5.3. As before, we can estimate the constants ν_c and C_c for the decay of the coefficients. Comparing Figures 4.1 and 4.2,

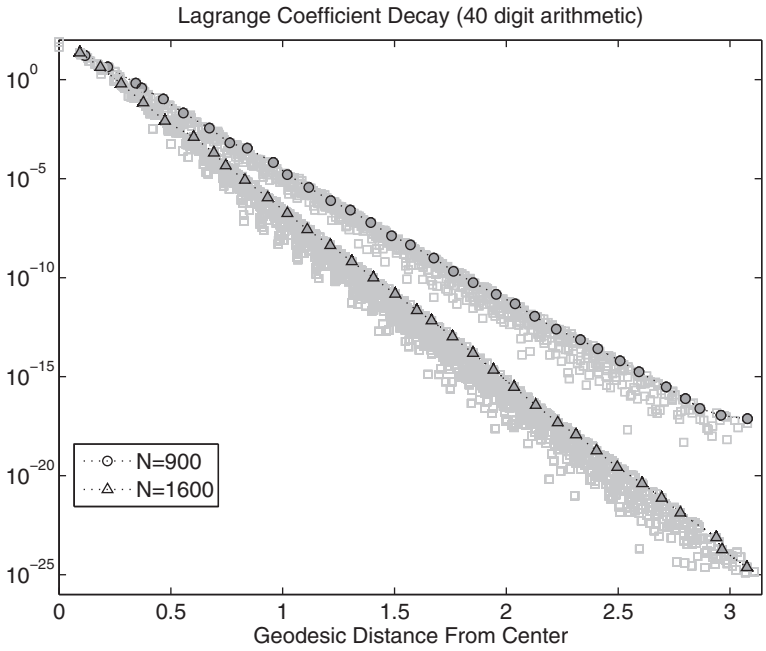

FIG. 4.3. *Size of the coefficients for a Lagrange function in the kernel space* $S(k_2, \Xi)$. *This experiment was carried out in Maple with* 40 *digit arithmetic.*

we note that the coefficient plot is shifted vertically. This is a consequence of the factor of q^{-2} in the estimate (5.6) below. Table 4.1 gives estimates for the constants ν_c and C_c, along with the constants involved in the decay of the Lagrange functions.

The perceived plateau present in the Lagrange function values as well as the coefficients shown in Figures 4.1 and 4.2 is due purely to round-off error related to the conditioning of kernel collocation and evaluation matrices. These results were produced using double-precision (approximately 16 digits) floating point arithmetic. To illustrate this point, we plot the decay rate of the Lagrange coefficients for the 900 and 1600 point node sets as computed using high-precision (40 digits) floating point arithmetic in Figure 4.3. The figure clearly shows that the exponential decay does not plateau, but continues as the theory predicts (see Theorem 5.3).

5. Coefficients of the full Lagrange functions. In this section we give theoretical results for the coefficients in the kernel expansion of Lagrange functions. In the first part we give a formula relating the size of coefficients to native space inner products of the Lagrange functions themselves (this is Proposition 5.1). We then obtain estimates for the restricted surface splines on $\mathbb{S}^2$, demonstrating the rapid, stationary decay of these coefficients.

5.1. Interpolation with conditionally positive definite kernels. In this section we demonstrate that the Lagrange function coefficients $A_{\xi,\zeta}$ can be expressed as a certain kind of inner product of different Lagrange functions χ_ξ and χ_ζ. Because this is a fundamental result, we work in generality in this subsection: the kernels we consider here are conditionally positive of the type considered in section 2.2.

When $u, v \in S(k, \Xi)$—meaning that they have the expansion $u = \sum_{\xi\in\Xi} a_{1,\xi} k(\cdot, \xi) + p_u$ and $v = \sum_{\xi\in\Xi} a_{2,\xi} k(\cdot, \xi) + p_v$ with coefficients $(a_{j,\xi})_{\xi\in\Xi} \perp (\Pi)|_\Xi$ for $j = 1, 2$—the

semiinner product is

$$\langle u, v\rangle_k = \left\langle \sum_{\xi\in\Xi} a_{1,\xi} k(\cdot,\xi), \sum_{\xi\in\Xi} a_{2,\xi} k(\cdot,\xi) \right\rangle_k = \sum_{\xi\in\Xi}\sum_{\zeta\in\Xi} a_{1,\xi}\overline{a_{2,\zeta}} k(\xi,\zeta).$$

(This follows directly from the definition (2.3) coupled with the observation that for $j \notin \mathcal{J}$, $\hat{u}(j) = \sum_{\xi\in\Xi} a_{1,\xi}\tilde{k}(j)\phi_j(\xi)$ and $\hat{v}(j) = \sum_{\xi\in\Xi} a_{2,\xi}\tilde{k}(j)\phi_j(\xi)$.) We can use this expression of the inner product to investigate the kernel expansion of the Lagrange function.

PROPOSITION 5.1. *Let $k(\cdot,\xi) = \sum_{j\in\mathbb{N}} \tilde{k}(j)\phi_j(\cdot)\overline{\phi_j(\xi)}$ be a conditionally positive definite kernel with respect to the space $\Pi = \operatorname{span}_{j\in\mathcal{J}} \phi_j$, and let Ξ be unisolvent for Π. Then $\chi_\eta \in S(k,\Xi)$ (the Lagrange function centered at η) has the kernel expansion $\chi_\eta(x) = \sum_{\xi\in\Xi} A_{\eta,\xi} k(x,\xi) + p_\zeta$ with coefficients*

$$\mathbf{A}_\eta = (A_{\eta,\xi})_{\xi\in\Xi} = \big(\langle \chi_\zeta(x), \chi_\eta(x)\rangle_k\big)_{\xi\in\Xi}.$$

Proof. Select two centers $\zeta, \eta \in \Xi$ with corresponding Lagrange functions χ_ζ and $\chi_\eta \in S(k,\Xi)$. Denote the collocation and auxiliary matrices, introduced in section 2.2, by $\mathrm{K}_\Xi = (k(\xi,\zeta))_{\zeta,\xi}$ and $\Phi = (\phi_j(\xi))_{\xi,j}$. Because $\mathbf{A}_\zeta$ and $\mathbf{A}_\eta$ are both orthogonal to $(\Pi)|_\Xi$, we have

$$\langle \chi_\zeta, \chi_\eta\rangle_k = \sum_{\xi_1\in\Xi}\sum_{\xi_2\in\Xi} A_{\zeta,\xi_1}\overline{A_{\eta,\xi_2}} k(\xi_1,\xi_2) = \langle \mathrm{K}_\Xi \mathbf{A}_\zeta, \mathbf{A}_\eta\rangle_{\ell_2(\Xi)}.$$

Now define $P := \Phi(\Phi^*\Phi)^{-1}\Phi^* : \ell_2(\Xi) \to (\Pi_\mathcal{J})|_\Xi \subset \ell_2(\Xi)$ to be the orthogonal projection onto the subspace of samples of Π on Ξ and let $P^\perp = \mathrm{Id} - P$ be its complement. Then for any data $\mathbf{y}$, (2.2) yields coefficient vectors $\mathbf{A}$ and $\mathbf{c}$ satisfying $P^\perp\mathbf{A} = \mathbf{A}$ and $P^\perp\Phi\mathbf{c} = \mathbf{0}$, and hence $P^\perp \mathrm{K}_\Xi P^\perp \mathbf{A} = P^\perp \mathrm{K}_\Xi \mathbf{A} = P^\perp \mathbf{y}$. Because $P^\perp : \ell_2(\Xi) \to \ell_2(\Xi)$ is also an orthogonal projector, and therefore self-adjoint, it follows that

$$\begin{aligned}\langle \chi_\zeta(x), \chi_\eta(x)\rangle_k &= \langle \mathrm{K}_\Xi \mathbf{A}_\zeta, \mathbf{A}_\eta\rangle_{\ell_2(\Xi)} = \langle \mathrm{K}_\Xi \mathbf{A}_\zeta, P^\perp \mathbf{A}_\eta\rangle_{\ell_2(\Xi)} = \langle P^\perp \mathrm{K}_\Xi \mathbf{A}_\zeta, \mathbf{A}_\eta\rangle_{\ell_2(\Xi)} \\ &= \langle P^\perp \mathbf{e}_\zeta, \mathbf{A}_\eta\rangle_{\ell_2(\Xi)}.\end{aligned}$$

In the last line, we have introduced the sequence $\mathbf{e}_\zeta = (\delta_{\zeta,\xi})_{\xi\in\Xi}$ for which $\mathrm{K}_\Xi \mathbf{A}_\zeta + p_\zeta|_\Xi = \mathbf{e}_\zeta$, which implies that $P^\perp \mathrm{K}_\Xi \mathbf{A}_\zeta = P^\perp \mathbf{e}_\zeta$. Using once more the fact that $P^\perp$ is self-adjoint and that $\mathbf{A}_\eta$ is in its range, we have

$$\langle \chi_\zeta(x), \chi_\eta(x)\rangle_{k,\mathcal{J}} = \langle P^\perp \mathbf{e}_\zeta, \mathbf{A}_\eta\rangle = \langle \mathbf{e}_\zeta, P^\perp \mathbf{A}_\eta\rangle = \langle \mathbf{e}_\zeta, \mathbf{A}_\eta\rangle$$

and the lemma follows. ∎

The next result involves estimating the norms $\|\mathbf{a}\|_{\ell_2(\Xi)}$ and $\|\mathbf{c}\|_{\ell_2(\mathcal{J})}$, where $\mathbf{a}$ and $\mathbf{c}$ are as in (2.2). It will be useful later, when we discuss local Lagrange functions. The notation is the same as that used in the proof above. In addition, because k is a conditionally positive definite kernel for Π, the matrix $P^\perp \mathrm{K}_\Xi P^\perp$ is positive definite on the orthogonal complement of the range of Φ. We will let ϑ be the minimum eigenvalue of this matrix, that is,

$$\vartheta := \min_{\|P^\perp\alpha\|=1} \langle P^\perp \mathrm{K}_\Xi P^\perp \alpha, \alpha\rangle > 0.$$

PROPOSITION 5.2. *Suppose* $\mathbf{a}$ *and* $\mathbf{c}$ *satisfy* (2.2). *Let* $G_\Xi = \Phi^*\Phi$. *Then,*

$$\|\mathbf{a}\|_{\ell_2(\Xi)} \le \vartheta^{-1}\|\mathbf{y}\|_{\ell_2(\Xi)} \le \vartheta^{-1}\sqrt{\#\Xi}\|\mathbf{y}\|_{\ell_\infty(\Xi)},$$
$$\|\mathbf{c}\|_{\ell_2(\mathcal{J})} \le 2\|k\|_\infty\|G_\Xi^{-1}\|^{1/2}\vartheta^{-1}\#\Xi\|\mathbf{y}\|_{\ell_2(\Xi)} \le 2\|k\|_\infty\|G_\Xi^{-1}\|^{1/2}\vartheta^{-1}(\#\Xi)^{3/2}\|\mathbf{y}\|_{\ell_\infty(\Xi)}.$$

Proof. From (2.2) and the fact that $P^\perp$ projects onto the orthogonal complement of the range of Φ, we have that $P^\perp \mathrm{K}_\Xi P^\perp \mathbf{a} = P^\perp \mathbf{y}$ and that $P^\perp \mathbf{a} = \mathbf{a}$. Consequently,

$$\vartheta\|\mathbf{a}\|^2_{\ell_2(\Xi)} = \vartheta\|P^\perp\mathbf{a}\|^2_{\ell_2(\Xi)} \le \langle P^\perp \mathrm{K}_\Xi P^\perp\mathbf{a}, \mathbf{a}\rangle \le \|\mathbf{a}\|_{\ell_2(\Xi)}\|P^\perp\mathbf{y}\|_{\ell_2(\Xi)}.$$

The bound on $\|\mathbf{a}\|_{\ell_2(\Xi)}$ follows immediately from this and $\|\mathbf{y}\|_{\ell_2(\Xi)} \le \sqrt{\#\Xi}\|\mathbf{y}\|_{\ell_\infty(\Xi)}$. To get the bound on $\|\mathbf{c}\|_{\ell_2(\Xi)}$, note that $\Phi\mathbf{c} = P\mathbf{y} - P\mathrm{K}_\Xi\mathbf{a}$ and, hence, that

$$\|\Phi\mathbf{c}\|_{\ell_2(\Xi)} \le \|P\mathbf{y}\|_{\ell_2(\Xi)} + \|P\mathrm{K}_\Xi P^\perp\mathbf{a}\|_{\ell_2(\Xi)} \le \|P\mathbf{y}\|_{\ell_2(\Xi)} + \vartheta^{-1}\|P\mathrm{K}_\Xi P^\perp\|\|P^\perp\mathbf{y}\|_{\ell_2(\Xi)}.$$

We also have that $\|\Phi\mathbf{c}\|^2_{\ell_2(\Xi)} = \langle\Phi^*\Phi\mathbf{c},\mathbf{c}\rangle \ge \lambda_{\min}(\Phi^*\Phi)\|\mathbf{c}\|^2_{\ell_2(\mathcal{J})}$. However, $\lambda_{\min}(\Phi^*\Phi) = \|(\Phi^*\Phi)^{-1}\|^{-1}$, which implies that

$$\|\mathbf{c}\|_{\ell_2(\mathcal{J})} \le \|(\Phi^*\Phi)^{-1}\|^{1/2}\|\Phi\mathbf{c}\|_{\ell_2(\Xi)} = \|G_\Xi^{-1}\|^{1/2}\|\Phi\mathbf{c}\|_{\ell_2(\Xi)}.$$

Next, note that the following hold: $\|P\mathrm{K}_\Xi P^\perp\| \le \|\mathrm{K}_\Xi\| \le \#\Xi\|k\|_\infty$, $\|P\mathbf{y}\|_{\ell_2(\Xi)}$, $\|P^\perp\mathbf{y}\|_{\ell_2(\Xi)} \le \|\mathbf{y}\|_{\ell_2(\Xi)} \le \sqrt{\#\Xi}\|\mathbf{y}\|_{\ell_\infty(\Xi)}$, and $\frac{\#\Xi\|k\|_\infty}{\vartheta} \ge 1$. Applying these to the inequality

$$\|\mathbf{c}\|_{\ell_2(\mathcal{J})} \le \|G_\Xi^{-1}\|^{1/2}\big(\|P\mathbf{y}\|_{\ell_2(\Xi)} + \#\Xi\|k\|_\infty\vartheta^{-1}\|P^\perp\mathbf{y}\|_{\ell_2(\Xi)}\big)$$

then yields the desired bound on $\|\mathbf{c}\|_{\ell_2(\mathcal{J})}$, completing the proof. ∎

5.2. Estimating Lagrange function coefficients. In [12], it has been shown that Lagrange functions for restricted surface splines decay exponentially fast away from the center. We can use these decay estimates in conjunction with Proposition 5.1 to estimate the decay of the coefficients $|A_{\xi,\zeta}|$.

Recall that the eigenvalues of $-\Delta$ are $\lambda_\ell = \ell(\ell+1)$. Let $Q(z) := \Pi_{\nu=1}^m(z-\lambda_{\nu-1}) = \sum_{\nu=1}^m b_\nu z^\nu$. The kernel $k_m : \mathbb{S}^2\times\mathbb{S}^2 \to \mathbb{R}$ has the expansion

$$k_m(x,\alpha) = \sum_{\ell=0}^\infty \tilde{k}(\ell)\sum_{\mu=-\ell}^{\ell} Y_\ell^\mu(x)Y_\ell^\mu(\alpha),$$

where, for $\ell \ge m$, $\tilde{k}(\ell) = C_m Q(\lambda_\ell)^{-1}$ with $C_m = 2^{m+1}\pi\Gamma(m)^2$ [18, equation 3.3]. From the expansion, one sees that k_m is conditionally positive definite with respect to Π_{m-1}. Kernels such as k_m are said to be of *polyharmonic or related type*; they have been studied in [12]. The kernel k_m acts as the Green's function for the elliptic operator $\mathcal{L}_m := C_m^{-1}Q(-\Delta)$ (cf. [12, Example 3.3]), in the sense that

$$f = \int_{\mathbb{M}} k_m(\cdot,\alpha)\mathcal{L}_m\big[f(\alpha) - p_f(\alpha)\big]\mathrm{d}\alpha + p_f,$$

where p_f is the orthogonal projection of f onto Π_{m-1}.

The native space "inner product" on subsets. In [12] it was shown that for any $k \in \mathbb{N}$, the operator $(\nabla^k)^*\nabla^k$ (which involves $(\nabla^k)^*$ the adjoint—with respect to the $L_2(\mathbb{S}^2)$ inner product—of the covariant derivative operator ∇^k, which was

introduced in section 2.1) can be expressed as $\sum_{\nu=0}^{k} d_\nu \Delta^\nu$ with $d_k = (-1)^k$. Consequently, any operator of the form $\sum_{j=0}^{k} c_j (\nabla^j)^* \nabla^j$ can be expressed as $\sum_{\nu=0}^{k} d_\nu \Delta^\nu$ with $d_k = (-1)^k c_k$ and vice-versa:

(5.1)
$$\forall (d_0, \dots d_m)\ \exists (c_0, \dots, c_m) \text{ with } d_m = (-1)^m c_m \text{ and } \sum_{\nu=0}^{m} d_\nu \Delta^\nu = \sum_{j=1}^{m} c_j (\nabla^j)^* \nabla^j.$$

Because $\mathcal{L}_m = C_m^{-1} Q(-\Delta)$, it follows that $\mathcal{L}_m = \sum_{j=0}^{m} c_j (\nabla^j)^* \nabla^j$ with $c_m = C_m^{-1}$, and so the native space semiinner product, introduced in (2.3), can be expressed as

$$\langle u, v \rangle_{k_m} = \langle \mathcal{L}_m u, v \rangle_{L_2(\mathbb{S}^2)} = \int_{\mathbb{S}^2} \beta(u, v)_x \mathrm{d}\mu(x)$$

with $\beta(u, v)_x = \sum_{k=0}^{m} c_k \langle \nabla^k u, \nabla^k v \rangle_x$ and $c_0, \dots, c_m$ the appropriate constants guaranteed by (5.1). The latter expression allows us to extend naturally the native space inner product to measurable subsets Ω of $\mathbb{S}^2$. Namely,

$$\langle u, v \rangle_{\Omega, k_m} := \int_{\Omega} \beta(u, v)_x \mathrm{d}\mu(x).$$

This has the desirable property of set additivity: for sets A and B with $\mu(A \cap B) = 0$, we have $\langle u, v \rangle_{A \cup B, k_m} = \langle u, v \rangle_{A, k_m} + \langle u, v \rangle_{B, k_m}$. Unfortunately, since some of the coefficients c_k may be negative, $\beta(u, u)$ and $\langle u, u \rangle_{\Omega, k_m}$ may assume negative values for some u: in other words, the bilinear form $(u, v) \mapsto \langle u, v \rangle_{\Omega, k_m}$ is only an *indefinite* inner product.

A Cauchy–Schwarz type inequality. When restricted to the cone of functions in $W_2^m(\Omega)$ having a sufficiently dense set of zeros, the quadratic form $\langle u, u \rangle_{\Omega, k_m}$ is positive definite. We now briefly discuss this.

When Ω has Lipschitz boundary and u has many zeros, we can relate the quadratic form $|\!|\!| u |\!|\!|^2_{\Omega, k_m} := \langle u, u \rangle_{\Omega, k_m}$ to a Sobolev norm $\|u\|^2_{W_2^m(\Omega)}$. Arguing as in [12, (4.2)], we see that

$$c_m |u|^2_{W_2^m(\Omega)} - \left(\max_{j \le m-1} |c_j| \right) \|u\|^2_{W_2^{m-1}(\Omega)} \le \int_{\Omega} \beta(u, u)_x \mathrm{d}\mu(x) \le \left(\max_{j \le m} |c_j| \right) \|u\|^2_{W_2^m(\Omega)}.$$

If $u|_{\Xi} = 0$ on a set Ξ with $h(\Xi, \Omega) \le h_0$ with h_0 determined only by the boundary of Ω (specifically the radius and aperture of an interior cone condition satisfied by $\partial\Omega$), Theorem A.11 of [12] guarantees that $\|u\|^2_{W_2^{m-1}(\Omega)} \le C h^2 |u|^2_{W_2^m(\Omega)}$ with C depending only on the order m and the roughness of the boundary. (In this case, depending only on the aperture of the interior cone condition.) Thus, by choosing $h \le h^*$, where h^* satisfies the two conditions

$$h^* \le h_0 \quad \text{and} \quad C(h^*)^2 \times \left(\max_{j \le m} |c_j| \right) \le \frac{|c_m|}{2}, \tag{5.2}$$

we have

$$\frac{c_m}{2} \|u\|^2_{W_2^m(\Omega)} \le |\!|\!| u |\!|\!|^2_{\Omega, k_m} \le \left(\max_{j \le m} |c_j| \right) \|u\|^2_{W_2^m(\Omega)}.$$

The threshold value h^* depends on the coefficients c_j as well as the radius R_Ω and aperture ϕ_Ω of the cone condition for Ω. When Ω is an annulus of sufficiently small

inner radius, the cone parameters can be replaced by a single global constant, and h_* can be taken to depend only on $c_0, \ldots, c_m$, in other words, only on k_m—cf. [12, Corollary A.16].

A direct consequence of this is positive definiteness for such functions, $\|u\|_{\Omega,k_m} \geq 0$ with equality only if $u|_\Omega = 0$. From this, we have a version of the Cauchy–Schwarz inequality: if u and v share a set of zeros Z (i.e., $u|_Z = v|_Z = \{0\}$) that is sufficiently dense in Ω, then

$$|\langle u, v\rangle_{\Omega,k_m}| \leq \|u\|_{\Omega,k_m} \|v\|_{\Omega,k_m} \tag{5.3}$$

follows (sufficient density means that $h(Z,\Omega) < h^*$ as above).

Decay of Lagrange functions. [12, Lemma 5.1] guarantees that the Lagrange function χ_ξ satisfies the *bulk chasing* estimate: there is a fixed constant $0 \leq \epsilon < 1$ so that for radii r the estimate

$$\|\chi_\xi\|_{W_2^m(B^c(\xi,r))} \leq \epsilon \|\chi_\xi\|_{W_2^m(B^c(\xi,r-\frac{h}{4h_0}))}$$

holds. In other words, a fraction (roughly $1-\epsilon$) of the bulk of the tail $\|\chi_\xi\|_{W_2^m(B^c(\xi,r))}$ is to be found in the annulus $B(\xi,r)\setminus B(\xi,r-\frac{h}{4h_0})$ of width $\frac{h}{4h_0} \propto h$ (with a constant of proportionality $\frac{1}{4h_0}$ that depends only on m). For $r > 0$, it is possible to iterate this n times, provided $n\frac{h}{4h_0} \leq r$. It follows that there is $\nu = -4h_0 \log \epsilon > 0$ so that

$$\|\chi_\xi\|_{W_2^m(B^c(\xi,r))} \leq \epsilon^n \|\chi_\xi\|_{W_2^m(\mathbb{S}^2)} \leq Ce^{-\nu r/h} \|\chi_\xi\|_{W_2^m(\mathbb{S}^2)}.$$

By [12, (5.1)][3] we have

$$\|\chi_\xi\|_{W_2^m(B^c(\xi,r))} \leq Cq^{1-m} e^{-\nu\frac{r}{h}}. \tag{5.4}$$

This leads us to our main result.

THEOREM 5.3. *Let $\rho > 0$ be a fixed mesh ratio. There exist constants h^*, ν, and C depending only on m and ρ so that if $h \leq h^*$, then the Lagrange function $\chi_\zeta = \sum_{\xi\in\Xi} A_{\zeta,\xi} k_m(\cdot,\xi) + p_\zeta \in S_m(\Xi)$ has the following properties:*

$$|\chi_\xi| \leq C \exp\left(-\nu \frac{\operatorname{dist}(x,\xi)}{h}\right), \tag{5.5}$$

$$|A_{\zeta,\xi}| \leq Cq^{2-2m} \exp\left(-\nu \frac{\operatorname{dist}(\xi,\zeta)}{h}\right), \tag{5.6}$$

$$c_1 q^{2/p} \|\mathbf{a}\|_{\ell_p(\Xi)} \leq \left\| \sum_{\xi\in\Xi} a_\xi \chi_\xi \right\|_{L_p(\mathbb{S}^2)} \leq c_2 q^{2/p} \|\mathbf{a}\|_{\ell_p(\Xi)}. \tag{5.7}$$

Proof. The bounds (5.5) and (5.7) are given in [12, Theorems 5.3 and 5.7]. Only (5.6) requires proof. By Proposition 5.1 and set additivity, we have that

$$A_{\zeta,\xi} = \langle \chi_\xi, \chi_\zeta\rangle_{k_m} = \langle \chi_\xi, \chi_\zeta\rangle_{\Omega_\zeta,k_m} + \langle \chi_\xi, \chi_\zeta\rangle_{\Omega_\xi,k_m},$$

where we use the decomposition into hemispheres: $\Omega_\zeta = \{\alpha \in \mathbb{S}^2 \mid \operatorname{dist}(\alpha,\zeta) < \operatorname{dist}(\alpha,\xi)\}$, $\Omega_\xi = \{\alpha \in \mathbb{S}^2 \mid \operatorname{dist}(\alpha,\xi) < \operatorname{dist}(\alpha,\zeta)\}$, and (modulo a set of measure zero) $\mathbb{S}^2 \setminus \Omega_\zeta = \Omega_\xi$.

[3] This is simply a comparison of χ_ξ to a smooth "bump" ϕ_ξ of radius q—also an interpolant to the delta data (δ_ξ), but worse in the sense that $\|\chi_\xi\|_{k_m} \leq \|\phi\|_{k_m}$. This idea is repeated in the proof of Theorem 5.3.

We apply the Cauchy–Schwarz type inequality (5.3) to obtain

$$\begin{aligned}|A_{\zeta,\xi}| &\le \|\chi_\zeta\|_{\Omega_\zeta,k_m}\|\chi_\xi\|_{\Omega_\zeta,k_m} + \|\chi_\zeta\|_{\Omega_\xi,k_m}\|\chi_\xi\|_{\Omega_\xi,k_m}\\ &\le \sqrt{\max_{j\le m}|c_j|}\left(\|\chi_\zeta\|_{W_2^m(\Omega_\zeta)}\|\chi_\xi\|_{\Omega_\zeta,k_m} + \|\chi_\zeta\|_{\Omega_\xi,k_m}\|\chi_\xi\|_{W_2^m(\Omega_\xi)}\right).\end{aligned}$$

Since $\Omega_\zeta \subset B^c(\zeta,r) := \mathbb{S}^2 \setminus B(\zeta, \frac{1}{2}\mathrm{dist}(\xi,\zeta))$ and $\Omega_\xi \subset B^c(\xi,r)$, we can again employ set additivity and positive definiteness (this time $\|\chi_\xi\|_{\Omega_\zeta,k_m} \le \|\chi_\xi\|_{\mathbb{S}^2,k_m}$, which follows from the fact that $\mathbb{S}^2 = \Omega_\zeta \cup \overline{\Omega_\xi}$ and that χ_ξ vanishes to high order in Ω_ξ—the same holds for χ_ζ) to obtain

$$|A_{\zeta,\xi}| \le \sqrt{\max_{j\le m}|c_j|}\left(\|\chi_\zeta\|_{W_2^m(B^c(\zeta,r))}\|\chi_\xi\|_{k_m} + \|\chi_\zeta\|_{k_m}\|\chi_\xi\|_{W_2^m(B^c(\xi,r))}\right).$$

The full energy of the Lagrange function can be bounded by comparing it to the energy of a bump function—for χ_ξ this is ϕ_ξ, which can be defined by using a smooth cutoff function σ. In spherical coordinates (colatitude, longitude) around ξ, $\phi_\xi(\theta,\varphi) = \sigma(\theta/q)$. This is done in [12, (5.1)] and we have that $\|\chi_\xi\|_{k_m}$ and $\|\chi_\zeta\|_{k_m}$ are bounded by Cq^{1-m}.

On the other hand, to treat $\|\chi_\zeta\|_{W_2^m(B^c(\zeta,r))}$ and $\|\chi_\xi\|_{W_2^m(B^c(\zeta,r))}$, we can use (5.4), which gives

$$\|\chi_\xi\|_{W_2^m(B^c(\zeta,r))}, \|\chi_\zeta\|_{W_2^m(B^c(\zeta,r))} \le Cq^{1-m}e^{-\nu\frac{r}{h}} = Cq^{1-m}e^{-\nu\frac{\mathrm{dist}(\xi,\zeta)}{2h}}.$$

The bound (5.6) follows immediately from this. ∎

Remark 5.4. Because the proof doesn't really depend on $\mathbb{S}^2$, a nearly identical proof works for any of the kernels with exponentially decaying Lagrange functions considered in [11, 12]. Specifically, we have this: Theorem 5.3 holds for compact, two-point homogeneous spaces with polyharmonic kernels satisfying $\mathcal{L}_m \perp \Pi$ (cf. [12]) and for any compact, C^∞ Riemannian manifold, with the kernels being the Sobolev splines given in [11].

6. Truncating the Lagrange basis. We now discuss truncating the kernel expansion Lagrange function $\chi_\xi = \sum_{\zeta\Xi} A_{\xi,\zeta}k_m(\cdot,\zeta) + p_\xi \in S_m(\Xi)$, replacing it with an expansion of the form

$$\widetilde{\chi}_\xi = \sum_{\zeta\in\Upsilon(\xi)} \widetilde{A}_{\xi,\zeta}k_m(\cdot,\zeta) + p_\xi \in S_m(\Xi), \tag{6.1}$$

where $\Upsilon(\xi) \subset \Xi$ is a set of centers contained in a ball $B(\xi, r(h))$ centered at ξ, where $r(h)$ and the $\widetilde{A}_{\xi,\zeta}$'s will be determined by $A_{\xi,\zeta}$ with $\zeta \in \Upsilon(\xi)$. We also assume that $\xi \in \Upsilon(\xi)$. Finally, to avoid notational clutter, we will simply use Υ rather than $\Upsilon(\xi)$.

Our goal is to show that if χ_ξ satisfies the properties (5.5), (5.6), and (5.7), then we may take $r(h) = Kh|\log(h)|$ with $K = K(m) > 0$, while maintaining algebraic decay in h of the error $\|\widetilde{\chi}_\xi - \chi_\xi\|_\infty$. For this choice of $r(h)$, a simple volume estimate (given at the end of section 6.3) shows that the number of terms required for $\widetilde{\chi}_\xi$ is just $\mathcal{O}((\log N)^2) \ll N$, far fewer than the N needed for χ_ξ.

Simply truncating at a fixed radius $r(h) = Kh|\log(h)|$ is not suitable, however, because the truncated function $\widetilde{\chi}_\xi$ will no longer be in the space $S_m(\Xi)$ (and thus $\{\widetilde{\chi}_\xi\}$ will not act as a basis). To treat this, we must slightly realign coefficients to satisfy the moment conditions.

A remark before proceeding with the analysis: Finding $\widetilde{\chi}_\xi$ in the way described below requires knowing the expansion for χ_ξ and carrying out the truncation above. This is expensive, although it does have utility in terms of speeding up evaluations for interpolation when the same set of centers is to be used repeatedly. The main point is that we now know roughly how many basis elements are required to obtain a good approximation to χ_ξ. The question of producing a good basis efficiently is left to the next section.

6.1. Constraint conditions on the coefficients. We would like $\widetilde{\chi}_\xi$ to be in the space $S_m(\Xi)$, and so the $\widetilde{A}_{\xi,\zeta}$'s have to satisfy the constraints in the system (2.2):

$$\sum_{\zeta\in\Upsilon} \widetilde{A}_{\xi,\zeta}\overline{\phi_j}(\zeta) = 0,\ j \in \mathcal{J} := (1,\dots,m^2), \tag{6.2}$$

where $\{\phi_j\}_{j=1}^{m^2}$ is an orthonormal basis for Π_{m-1}. Since the original χ_ξ's are in $S_m(\Xi)$, the $A_{\xi,\zeta}$'s in their expansions satisfy the constraint equations in (2.2). Splitting these equations into sums over Υ and its complement in Ξ and manipulating the result, we see that

$$\sum_{\zeta\in\Upsilon} A_{\xi,\zeta}\overline{\phi_j}(\zeta) + \sigma_j, \quad \text{where } \sigma_j := \sum_{\zeta\notin\Upsilon} A_{\xi,\zeta}\overline{\phi_j}(\zeta),\ j \in \mathcal{J}. \tag{6.3}$$

The way that we will relate the two sets of coefficients is to define the vector $(\widetilde{A}_{\xi,\zeta})_{\zeta\in\Upsilon}$ to be the orthogonal projection of $(A_{\xi,\zeta})_{\zeta\in\Upsilon}$ onto the constraint space, which is the orthogonal complement of $\operatorname{span}\{\,\phi_j|_\Upsilon\,, j \le m^2\}$, in the usual inner product for $\ell_2(\Upsilon)$. The equations below then follow:

$$\begin{aligned} (\widetilde{A}_{\xi,\zeta})_{\zeta\in\Upsilon} - (A_{\xi,\zeta})_{\zeta\in\Upsilon} &= \sum_{j\in\mathcal{J}} \tau_j \phi_j|_\Upsilon \in \operatorname{span}\{(\phi_j(\zeta))_{\zeta\in\Upsilon}, j \le m^2\}, \\ \|(\widetilde{A}_{\xi,\zeta})_{\zeta\in\Upsilon} - (A_{\xi,\zeta})_{\zeta\in\Upsilon}\|^2_{\ell_2(\Upsilon)} &= \tau^* G_\Upsilon \tau, \quad [G_\Upsilon]_{k,j} := \sum_{\zeta\in\Upsilon} \overline{\phi_k}(\zeta)\phi_j(\zeta), \end{aligned} \tag{6.4}$$

where τ is a column vector having the τ_j's as entries. Let σ be a column vector with the σ_j's as entries. From the first equation above together with (6.2) and (6.3), τ and σ are related by $\sigma = G_\Upsilon\tau$. If we make the rather mild assumption that Υ is unisolvent for the space Π_{m-1}, then we can invert G_Υ: $\tau = G_\Upsilon^{-1}\sigma$, thereby obtaining $\tau^* G_\Upsilon \tau = \sigma^* G_\Upsilon^{-1}\sigma$. Using this in (6.4) and applying Schwarz's inequality, we obtain the following bound:

$$\left\| \sum_{\zeta\in\Upsilon} (\widetilde{A}_{\xi,\zeta} - A_{\xi,\zeta}) k_m(\cdot,\zeta) \right\|_\infty \le \sqrt{\#\Upsilon\, \|G_\Upsilon^{-1}\|_2}\, \|k_m\|_\infty \|\sigma\|_2, \tag{6.5}$$

which we will make use of to establish the estimates below.

PROPOSITION 6.1. *Assume that Υ is unisolvent for Π_{m-1} and that $\|G_\Upsilon^{-1}\|_2 = \mathcal{O}(|\log h|^{-2}\, h^{-2\mu})$ for some $\mu \ge 0$. If we take $r(h) = Kh|\log(h)|$, where K is chosen so that $J := K\nu - 2m - \mu > 0$, then for h sufficiently small,*

$$\|\widetilde{\chi}_\xi - \chi_\xi\|_\infty \le Ch^J, \tag{6.6}$$

$$|\widetilde{\chi}_\xi(x)| \le C\big(1 + \operatorname{dist}(x,\xi)/h\big)^{-J}. \tag{6.7}$$

Furthermore, when $J > 2$, the set $\{\widetilde{\chi}_\xi\}$ is L_p stable: there are $C_1, C_2 > 0$ for which

$$C_1 q^{2/p} \|\mathbf{a}\|_{\ell_p(\Xi)} \le \left\| \sum_{\xi\in\Xi} a_\xi \widetilde{\chi}_\xi \right\|_{L_p(\mathbb{S}^2)} \le C_2 q^{2/p} \|\mathbf{a}\|_{\ell_p(\Xi)}. \tag{6.8}$$

Proof. From (5.6) and $N \le 4\pi/\mathrm{vol}(B(\xi, q)) \le Cq^{-2}$, we have that

$$\sum_{\zeta\notin\Upsilon} |A_{\xi,\zeta}| = \mathcal{O}\big(Nq^{2-2m} \exp(-\nu r(h)/h)\big) \le Ch^{K\nu-2m}. \tag{6.9}$$

Applying it to the σ_j's defined in (6.3) results in $\|\sigma\|_2 \le Ch^{K\nu-2m}$. Using this in connection with (6.5), $\|G_\Upsilon^{-1}\|_2 = \mathcal{O}(|\log h|^{-2} h^{-2\mu})$, (6.9), and

$$\widetilde{\chi}_\xi - \chi_\xi = \sum_{\zeta\in\Upsilon} (\widetilde{A}_{\xi,\zeta} - A_{\xi,\zeta}) k_m(\cdot,\zeta) - \sum_{\zeta\notin\Upsilon} A_{\xi,\zeta} k_m(\cdot,\zeta)$$

yields (6.6). Next, from (5.5) we have

$$|\chi_\xi| \le C \exp\left(-\nu \frac{\mathrm{dist}(x,\xi)}{h}\right) \le C \exp\left(-K\nu \frac{\mathrm{dist}(x,\xi)}{Kh}\right) \le C\left(1 + \frac{\mathrm{dist}(x,\xi)}{Kh}\right)^{-K\nu}.$$

Combining this with (6.6), using $J = K\nu - 2m - \mu > 0$, and manipulating, we arrive at (6.7).

It remains to demonstrate the L_p stability of $(\tilde{\chi}_\xi)$ for $1 \le p \le \infty$. When $p = 1$, we consider a sequence $\mathbf{a} = (a_\xi)_{\xi\in\Xi} \in \ell_1(\Xi)$. Let $s := \sum a_\xi \chi_\xi$ and $\tilde{s} := \sum a_\xi \widetilde{\chi}_\xi$. From Hölder's inequality and (6.6), we have $\|\tilde{s} - s\|_{L_1(\mathbb{S}^2)} \le C\|\mathbf{a}\|_{\ell_1(\Xi)} h^J$ and

$$\|\tilde{s} - s\|_{L_\infty(\mathbb{S}^2)} \le C\|\mathbf{a}\|_{\ell_\infty(\Xi)} \underbrace{\sum_{\xi\in\Xi} |\widetilde{\chi}_\xi(x) - \chi_\xi(x)|}_{\le N \max_\xi \|\widetilde{\chi}_\xi - \chi_\xi\|_{L_\infty(\mathbb{S}^2)}} \le C\|\mathbf{a}\|_{\ell_\infty(\Xi)} h^J q^{-2}.$$

Interpolating between these two inequalities—i.e., interpolating the finite rank operator $\mathbf{a} \mapsto (s - \tilde{s})$—gives

$$\begin{aligned} \|s - \tilde{s}\|_{L_p(\mathbb{S}^2)} &\le Ch^J q^{-2(1-1/p)} \|\mathbf{a}\|_{\ell_p(\Xi)} \\ &\le Ch^{J-2} q^{2/p} \|\mathbf{a}\|_{\ell_p(\Xi)}. \quad (q^{-2} \sim h^{-2}). \end{aligned}$$

After some manipulation, this bound and (5.7) imply that

$$c_1 q^{2/p} \|\mathbf{a}\|_{\ell_p(\Xi)} (1 - Ch^{J-2}) \le \|\tilde{s}\|_{L_p(\mathbb{S}^2)} \le c_2 q^{2/p} \|a\|_{\ell_p(\Xi)} (1 + Ch^{J-2}).$$

Choosing h so that $Ch^{J-2} \le 1/2$ and letting $C_1 = c_1/2$ and $C_2 = 3c_2/2$, we obtain (6.8). ∎

Remark 6.2. When there are no constraint conditions on the coefficients, this result holds for any of the strictly positive definite kernels mentioned in Remark 5.4. In particular it holds for Sobolev splines on a compact C^∞ Riemannian manifold.

6.2. Norm of the inverse Gram matrix. We now demonstrate that the conditions on G_Υ^{-1} in Proposition 6.1 are automatically satisfied. We will state and prove the results below for caps on $\mathbb{S}^d$, rather than just $\mathbb{S}^2$. Also, it is more convenient to use with Π_L rather than Π_{m-1}, because m is notationally tied to the polyharmonic

kernels k_m as well as the spherical harmonics on $\mathbb{S}^2$. That said, we begin with the lemma below.

LEMMA 6.3. *Suppose that $S_r := B(\xi, r) \subset \mathbb{S}^d$ is a cap of fixed radius $r < \pi$, and that $\mathscr{C} \subset S_r$ is finite and has mesh norm $h_{\mathscr{C}} := h_{S_r,\mathscr{C}}$. In addition, let $L \geq 0$ be a fixed integer and take Π_L to be the space of all spherical harmonics of degree at most L. Then, there exists a constant $c_0 := c_0(d, L) > 0$ such that when $h_{\mathscr{C}} \leq c_0 r$ we have*

$$\sum_{\zeta\in\mathscr{C}} |\varphi(\zeta)|^2 \geq \mu(S_r)^{-1} \int_{S_r} |\varphi(x)|^2 d\mu(x) \ \forall \varphi \in \Pi_L. \tag{6.10}$$

Moreover, the set $\mathscr{C}$ is unisolvent for Π_L. Finally, for every basis for Π_L the corresponding Gram matrices $G_{\mathscr{C}}$ and G_{S_r}, relative to the inner products on $\ell^2(\mathscr{C})$ and S_r, respectively, satisfy

$$\|G_{\mathscr{C}}^{-1}\|_2 \leq \mu(S_r)\|G_{S_r}^{-1}\|_2. \tag{6.11}$$

Proof. Since $\varphi(x)$ and $\overline{\varphi}(x)$ are spherical harmonics in Π_L, their product is a spherical harmonic of degree at most $2L$. Thus, applying the nonnegative-weight quadrature formula in [17, Theorem 2.1] to spherical harmonics of order $2L$ yields

$$\sum_{\zeta\in\mathscr{C}} w_\zeta |\varphi(\zeta)|^2 = \int_{S_r} |\varphi(x)|^2 d\mu(x).$$

Since $0 \leq w_\zeta \leq \sum_{\zeta\in\mathscr{C}} w_\zeta = \mu(S_r)$, we have $\sum_{\zeta\in\mathscr{C}} w_\zeta |\varphi(\zeta)|^2 \leq \mu(S_r) \sum_{\zeta\in\mathcal{C}} |\varphi(\zeta)|^2$. The inequality (6.10) follows immediately from the quadrature formula. To prove that $\mathcal{C}$ is unisolvent, suppose that $\varphi \in \Pi_L$ vanishes on $\mathcal{C}$. By (6.10), we have that $\int_{S_r} |\varphi(x)|^2 d\mu(x) = 0$. Since φ is in Π_L, it is a polynomial in sines and cosines of the angles used in the standard parameterization of $\mathbb{S}^d$ with ξ being the "north" pole. As a consequence, it is continuous on S_r and, because $\int_{S_r} |\varphi(x)|^2 d\mu(x) = 0$, it is identically 0 on S_r. Finally, as a function of the angular variables in the complex plane, it is analytic, entire in fact, and can be expanded in a power series in these variables. The fact that it vanishes identically for *real* values of the angular variables is enough to show that the coefficients in the series are all zero. Hence, $\varphi \equiv 0$ on $\mathbb{S}^d$ and $\mathscr{C}$ is unisolvent for Π_L. To establish (6.11), note that (6.10) implies that $G_{\mathscr{C}} - \mu(S_r)^{-1} G_{S_r}$ is positive semidefinite. From the Courant–Fischer theorem, the lowest eigenvalue of $G_{\mathscr{C}}$ is greater than that of $\mu(S_r)^{-1} G_{S_r}$. This inequality then yields (6.11), since these eigenvalues are $\|G_{\mathscr{C}}^{-1}\|_2^{-1}$ and $\mu(S_r)\|G_{S_r}^{-1}\|_2^{-1}$, respectively. ∎

We now need to compute the Gram matrix for the *canonical basis* of Π_L. This basis is described in [26, Chapter IX, section 3.6] and consists of spherical harmonics. Let $\ell, k_1, \ldots, k_{d-1}$ be integers satisfying $\ell \geq k_1 \geq k_2 \geq \cdots \geq k_{d-1} \geq 0$, and take $K := (k_1, \ldots, \pm k_{d-1})$. A spherical harmonic of degree ℓ [26, p. 466]) will be denoted by $Y_K^\ell(\theta_1, \ldots, \theta_d)$. The angles are the usual ones from spherical coordinates in $\mathbb{R}^{d+1}$ (cf. [26, p. 435]). The basis for Π_L is then the set of all Y_K^ℓ, $0 \leq \ell \leq L$. The entries in the Gram matrix are $[G_{S_r}]_{(\ell,K),(\ell',K')} = \langle Y_K^\ell, Y_{K'}^{\ell'} \rangle_{S_r}$. Following the argument in [26, Chapter IX, section 3.6], one may show that

$$\langle Y_K^\ell, Y_{K'}^{\ell'} \rangle_{S_r} = B_{\ell,K} B_{\ell',K} \delta_{K,K'} \int_0^r C_{\ell-k_1}^{\frac{d-1}{2}+k_1}(\cos\theta) C_{\ell'-k_1}^{\frac{d-1}{2}+k_1}(\cos\theta) \sin^{2k_1+d-1}\theta \, d\theta, \tag{6.12}$$

where $C_n^s(t)$ is the Gegenbauer polynomial of degree n and type s, and B_{ℓ,k_1} is a normalization factor. In the case where $d = 2$ and $L = 1$, G_{S_r} is 4×4 and has six nonzero entries,

$$G_{(0,0),(0,0)} = \frac{1}{2}(1-\cos r),\ G_{(0,0),(1,0)} = G_{(1,0),(0,0)} = \frac{\sqrt{3}}{4}(1-\cos r)(1+\cos r),$$
$$G_{(1,0),(1,0)} = \frac{1}{2}(1-\cos r)(1+\cos r+\cos^2 r),\ G_{(1,\pm1),(1,\pm1)} = \frac{1}{4}(1-\cos r)^2(2+\cos r).$$

Since $\mu(S_r) = 2\pi(1-\cos r)$, the formulas for the entries above imply that $G_{S_r}/\mu(S_r)$ is a polynomial in $\cos r$. In fact, a straightforward calculation shows that the minimum eigenvalue of this matrix is $r^4/(256\pi)+\mathcal{O}(r^6)$. Lemma 6.3 then implies that $\|G_{\mathcal{C}}^{-1}\|_2 \le \mu(S_r)\|G_{S_r}^{-1}\|_2 = 256\pi r^{-4} + \mathcal{O}(r^{-2})$. A less precise, but similar result, holds in the general case.

LEMMA 6.4. *Under the assumptions of Lemma* 6.3, *for general $L \ge 0$, $d \ge 2$, and r sufficiently small, there is an integer $\iota = \iota(L,d) \ge L$ and a constant $C = C(L,d) > 0$ such that $\|G_{\mathcal{C}}^{-1}\|_2 \le Cr^{-2\iota}$. For $L = 1$ and $d = 2$, we may take $\iota = 2$.*

Proof. From the expression in (6.12) for the entries in G_{S_r}, we see that each of them is entire in r and has a zero of order d or greater at $r = 0$. In addition, $\mu(S_r)$ is also entire in r and has a zero of order d. It follows that the matrix $\widetilde{G}(r) = G_{S_r}/\mu(S_r)$ is entire, even in r, and for real r, it is real, self-adjoint, and positive semidefinite. (In fact, for d even, it is a polynomial in $\cos r$.) In addition, the 2×2 block in $\widetilde{G}(0)$ corresponding to $k_1 = 0$, $\ell = 0, 1$, is rank 1 and therefore has 0 as an eigenvalue; consequently, $\widetilde{G}(0)$ also has 0 as an eigenvalue—it's lowest, in fact. As Rellich [21, p. 91] shows, the eigenvalues of $\widetilde{G}(r)$ are analytic functions of r. For $r > 0$, these eigenvalues are proportional to those of the Gram matrix G_{S_r} and therefore must be positive. None of these eigenvalues are identically 0. In particular, the eigenvalues splitting off from the 0 eigenvalue of $\widetilde{G}(0)$ are not identically 0. As functions of r they thus have a zero of finite order at $r = 0$; the order is an even integer because $\widetilde{G}(r)$ is even in r. The smallest eigenvalue then behaves like $\lambda_{\min}(r) = r^{2\iota}(a_0 + \mathcal{O}(r^2))$, where $a_0 > 0$, $\iota > 0$ is an integer and r is sufficiently small. Furthermore, from (6.12) we see that the diagonal entry, with $\ell = \ell' = k_1 = L$, is $\mathcal{O}(r^{2L})$. Since this bounds the minimum eigenvalue from above, we must have $2\iota \ge 2L$, so $\iota \ge L$. The result then follows from Lemma 6.3 and the observation that $\mu(S_r)\|G_{S_r}^{-1}\|_2 = \lambda_{\min}^{-1}$. The calculation for $L = 1$ and $d = 2$ was done above. ∎

6.3. Local Lagrange bases. We now turn to the *local* Lagrange basis. Recall that the function $\check{\chi}_\xi \in S_m(\Xi)$, with the kernel representation

$$\check{\chi}_\xi = \sum_{\zeta\in\Upsilon} \check{A}_{\xi,\zeta} k_m(\cdot,\zeta) + \sum_{j=1}^{m^2} \check{b}_j \phi_j \in S_m(\Xi), \tag{6.13}$$

is a local Lagrange function centered at ξ if it satisfies $\check{\chi}_\xi|_\Upsilon = \mathbf{e}_\xi$, where $\mathbf{e}_\xi(\zeta) = \delta_{\xi,\zeta}$, that is, $\mathbf{e}_\xi$ is the vector $(1, 0, \ldots, 0)^T$. Since $\check{\chi}_\xi \in S_m(\Xi)$, the vector $\check{A}_\xi = (\check{A}_{\xi,\zeta})_{\zeta\in\Upsilon}$ is in the constraint space. This vector and the coefficients $\check{b}_j$ then satisfy $\check{\chi}_\xi|_\Upsilon = \mathbf{e}_\xi$. Of course, the (full) Lagrange function $\chi_\xi = \sum_{\zeta\in\Xi} A_{\xi,\zeta}\kappa(\cdot,\zeta) + \sum_{j=1}^{m^2} b_j\phi_j$ restricted to Υ also satisfies $\chi_\xi|_\Upsilon = \mathbf{e}_\xi$. Consequently, $\mathfrak{D}_\xi := \check{\chi}_\xi - \chi_\xi$ satisfies $\mathfrak{D}_\xi|_\Upsilon = \mathbf{0}$. We can rewrite this difference as $\mathfrak{D}_\xi = \check{\chi}_\xi - \widetilde{\chi}_\xi + \widetilde{\chi}_\xi - \chi_\xi$ (with $\widetilde{\chi}_\xi$ the truncated basis

function introduced in the last section). It follows that

$$\mathfrak{D}_\xi = \underbrace{\sum_{\zeta\in\Upsilon}(\underbrace{\check{A}_\zeta - \widetilde{A}_\zeta}_{\alpha_\zeta})\kappa(\cdot,\zeta) + \sum_{j=1}^{\#\mathcal{J}}(\underbrace{\check{b}_j - b_j}_{\beta_j})\phi_j + \widetilde{\chi}_\xi - \chi_\xi}_{\check{\chi}_\xi - \widetilde{\chi}_\xi}.$$

Evaluating this on Υ then gives the system $\mathrm{K}_\Upsilon\alpha + \Phi\beta + (\widetilde{\chi}_\xi - \chi_\xi)|_\Upsilon = \mathbf{0}$. By linearity, it is clear that $\Phi^*\alpha = \mathbf{0}$. Finally, letting $\mathbf{y} = (\chi_\xi - \widetilde{\chi}_\xi)|_\Upsilon$, we arrive at the system

$$\begin{pmatrix} \mathrm{K}_\Upsilon & \Phi \\ \Phi^* & \mathbf{0} \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}. \tag{6.14}$$

Proposition 5.2 applies to (6.14) with Ξ replaced by Υ; thus, noting that $\|\mathbf{y}\|_{\ell_\infty(\Upsilon)} \le \|\widetilde{\chi}_\xi - \chi_\xi\|_\infty$, and writing $\mathcal{J} = (1,\dots,m^2)$, we see that

$$\begin{aligned}
&\|\alpha\|_{\ell_2(\Upsilon)} \le \vartheta^{-1}\sqrt{\#\Upsilon}\|\widetilde{\chi}_\xi - \chi_\xi\|_\infty, \\
&\|\beta\|_{\ell_2(\mathcal{J})} \le 2\|k_m\|_\infty \|G_\Upsilon^{-1}\|^{1/2}\vartheta^{-1}(\#\Upsilon)^{3/2}\|\widetilde{\chi}_\xi - \chi_\xi\|_\infty.
\end{aligned}$$

From this we obtain the following inequalities:

$$\begin{aligned}
\|\check{\chi}_\xi - \widetilde{\chi}_\xi\|_\infty &\le \|k_m\|_\infty\sqrt{\#\Upsilon}\|\alpha\|_{\ell_2(\Upsilon)} + mC_m\|\beta\|_{\ell_2(\mathcal{J})},\ C_m = \max_{j\le m^2}\|\phi_j\|_\infty \\
&\le \|k_m\|_\infty \#\Upsilon\vartheta^{-1}\left(1 + 2mC_m\sqrt{\#\Upsilon\|G_\Upsilon^{-1}\|}\right)\|\widetilde{\chi}_\xi - \chi_\xi\|_\infty.
\end{aligned}$$

Moreover, using $\|\check{\chi}_\xi - \chi_\xi\|_\infty \le \|\check{\chi}_\xi - \widetilde{\chi}_\xi\|_\infty + \|\chi_\xi - \widetilde{\chi}_\xi\|_\infty$, we see that

$$\|\check{\chi}_\xi - \chi_\xi\|_\infty \le 2\|k_m\|_\infty \#\Upsilon\vartheta^{-1}\left(1 + 2mC_m\sqrt{\#\Upsilon\|G_\Upsilon^{-1}\|}\right)\|\widetilde{\chi}_\xi - \chi_\xi\|_\infty. \tag{6.15}$$

Finally, from Proposition 6.1, if $K > (2m+2\mu)/\nu$, it is easy to see that this holds:

$$\|\check{\chi}_\xi - \chi_\xi\|_\infty \le C\frac{|\log h|^2}{\vartheta}\left(1 + mh^{-\mu}\right)h^{K\nu-2m-\mu} \le C\frac{|\log h|^2}{\vartheta}h^{K\nu-2m-2\mu}. \tag{6.16}$$

To proceed further, we need to estimate ϑ. Such estimates are known for surface splines in the Euclidean case [20, section 6]. Simply repeating the proofs of [20, Corollary 2.2] and [20, Theorem 2.4] for a set of points in $\mathbb{R}^3$ restricted to $\mathbb{S}^2$ yields the desired estimate. For the collocation matrix associated with k_m and Ξ, we have

$$\vartheta \ge Cq^{2m-2}, \tag{6.17}$$

where C depends only on m.[4] Thus, for k_m, we have $\|\check{\chi}_\xi - \chi_\xi\|_\infty \le Ch^{K\nu-4m+2-2\mu}$, where the constant K has to be increased slightly to absorb $|\log h|^2$. With this in mind we have the following result, whose proof, being similar to Proposition 6.1, we omit.

[4] For any d-dimensional sphere or projective space and any conditionally positive definite polyharmonic kernel with associated polynomial operator $\mathcal{L}_m = Q(-\Delta)$, where Q is a polynomial of degree m, the coefficients in the expansion for k_m are given by $\widehat{k}(j) = Q(\lambda_j)^{-1},\ j \notin \mathcal{J}$. For large λ_j, all of these have the asymptotic behavior λ_j^{-m}, which is the same as that of the coefficients for the m-d thin-plate spline. This implies that the matrix $P^\perp \mathrm{K}_\Upsilon P^\perp$ in Proposition 5.2 (here, $\Xi \to \Upsilon$) will have a lowest eigenvalue value, that is, up to a constant multiple, dependent only on m and d. Consequently, the bound $\vartheta \ge Cq^{2m-d}$ holds for all k_m associated with $\mathcal{L}_m$ in dimension d.

THEOREM 6.5. *Let the notation and assumptions of Theorem* 5.3 *hold. Suppose that* $K > 0$ *is chosen so that* $K > \frac{4m-2+2\mu}{\nu}$ *and, for each* $\xi \in \Xi$, $\Upsilon(\xi) := \Xi \cap B(\xi, Kh|\log h|)$. *If* $\check{\chi}_\xi$ *is a local Lagrange function for* $\Upsilon(\xi)$ *centered at* ξ, *then set* $\{\check{\chi}_\xi\}_{\xi\in\Xi}$ *is a basis for* $S_m(\Xi)$. *Moreover, with* $J := K\nu - 4m + 2 - 2\mu$, *we have*

$$\|\check{\chi}_\xi - \chi_\xi\|_\infty \le C\ h^J, \tag{6.18}$$

$$|\check{\chi}_\xi(x)| \le C\big(1 + \mathrm{dist}(x,\xi)/h\big)^{-J}. \tag{6.19}$$

Furthermore, when $J > d$, *the set* $\{\check{\chi}_\xi\}$ *is* L_p *stable: there are* $C_1, C_2 > 0$ *for which*

$$C_1 q^{2/p}\|\mathbf{a}\|_{\ell_p(\Xi)} \le \left\| \sum_{\xi\in\Xi} a_\xi \check{\chi}_\xi \right\|_{L_p(\mathbb{S}^2)} \le C_2 q^{2/p}\|\mathbf{a}\|_{\ell_p(\Xi)}. \tag{6.20}$$

Quasi-interpolation. It follows that the operator

$$Q_\Xi f = \sum_{\xi\in\Xi} f(\xi)\check{\chi}_\xi$$

provides L_∞ convergence at the same asymptotic rate as interpolation I_Ξ. Indeed,

$$|I_\Xi f(x) - Q_\Xi f(x)| \le \sum_{\xi\in\Xi} |\check{\chi}_\xi(x) - \chi_\xi(x)||f(\xi)| \le Cq^{-2}\|f\|_\infty h^{K\nu-4m-2\mu} \le C\|f\|_\infty h^{2m}$$

provided that $K > \frac{6m+2\mu+2}{\nu}$. It is shown in [12, Corollary 5.9] that restricted surface spline interpolation exhibits $\|I_\Xi f - f\|_\infty \le Ch^\sigma$ for $f \in C^{2m}(\mathbb{S}^2)$ when $\sigma = 2m$ and for $f \in B^\sigma_{\infty,\infty}(\mathbb{S}^2)$ for $\sigma < 2m$. So Q_Ξ has the same rate of approximation (without needing to solve a large system of equations).

Constructing basis functions in terms of N. For a set of scattered points Ξ it is possible, in fact often desirable, to use N as the basic parameter instead of h. Therefore we wish to express the number of nearest neighbors needed as a function of the total cardinality N instead of those within a $Kh|\log h|$ neighborhood. Consider a cap $B(\alpha, r)$. A simple volume argument gives

$$\#(B(\alpha,r)\cap\Xi) \le \frac{36}{11}\left(\frac{r}{q}\right)^2. \tag{6.21}$$

Indeed, one arrives at this bound by first considering caps of radius q around each node in $B(\alpha, r)$. If q is small enough, say, $q < r$, then at least $1/3$ the volume of each cap will be contained in $B(\alpha, r)$. Using this and a Taylor expansion of the volume formula $2\pi(1 - \cos(q))$ leads to (6.21). Thus, the greatest number of points in a cap of radius $Kh|\log h|$ is $\frac{36}{11}\rho^2(K|\log(h)|)^2$. Also, it is not hard to show that $2h^{-2} \le N$, and hence it follows that the number of points is bounded by $\frac{36}{11}\rho^2(\frac{K}{2}\log(N))^2 = \frac{9}{11}(\rho K)^2(\log(N))^2$, and it suffices to take for Υ the nearest $\frac{9}{11}(\rho K)^2(\log(N))^2$ neighbors.

As a final remark, we point out that bound on $n := \#\Upsilon$ derived above is pessimistic. To see why, suppose that ρ is large, but only a few points are roughly $2q$ apart, with the other points being on the order of $2h$ apart. The number of points in the disk of radius $Kh|\log(h)|$ will still be $n \sim K^2(\log(N))^2$. Any dependence on ρ will be weak, and certainly not $\mathcal{O}(\rho^2)$. The important point to be made here is that, with weak ρ dependence, n depends chiefly on the kernel used and the number of points in Ξ. This is born out in the results of numerical experiments we conducted, where we used $n = 7(\log(N))^2$. See Table 7.1.

The constants ν and K. Before we turn to a discussion of preconditioning, we wish to comment on the constants ν and K above. These two constants come into play in a crucial way in many of our estimates.

The decay constant ν first comes up in the proof of Theorem 5.3. (Although we do not mention it in the theorem, the proof produces two different decay constants: ν_L and ν_C, the former for the Lagrange function and the latter for the coefficients.) The estimate for ν is, of course, a lower bound on the decay constant itself; it is independent of ρ, but weakly dependent on m. Because of the nature of such estimates, it is very likely that they are much lower than ν_{actual}. How ν_{actual} behaves as a function of ρ is an open question.

There is another open question concerning K. We know that it must be bounded below by $\frac{4m-2+2\mu}{\nu}$. Thus a better estimate on ν would produce a better lower bound on K. This in turn means using smaller caps and fewer points in constructing the local Lagrange interpolant—i.e., giving it a smaller "footprint." On the other hand, the larger we make K the better the approximation to χ_ξ we get. Since K can be made as large as we please, the question then becomes this: What is an *optimal* choice for K? Indeed, what does the term *optimal* mean here?

7. Preconditioning with local Lagrange functions. In this section we illustrate how the local Lagrange functions can also be used as an effective preconditioner for linear systems associated with interpolation using the standard restricted spline basis. Our focus is on the restricted surface spline k_2 (i.e., the restricted thin plate spline), for which the interpolant to $f|_\Xi$ in the standard basis takes the form

$$I_\Xi f = \sum_{\xi\in\Xi} a_\xi k_2(\cdot,\xi) + \sum_{j=1}^{4} c_j\phi_j(\cdot), \tag{7.1}$$

where ϕ_j are a basis for the spherical harmonics of degree ≤ 1. We note that this interpolant can also be written with respect to the local Lagrange basis for $S_2(\Xi)$ as

$$I_\Xi f = \sum_{\xi\in\Xi} \check{a}_\xi \check{\chi}_\xi(\cdot); \tag{7.2}$$

see section 3 for the details on constructing this basis.

Using the properties of the local Lagrange basis, we can write the linear system for determining the interpolation coefficients $\check{a}_\xi$ in (7.2) as

$$\begin{bmatrix} \mathrm{K}_\Xi & \Phi \end{bmatrix} \begin{bmatrix} \mathsf{A}_\Upsilon \\ \mathsf{C}_\Upsilon \end{bmatrix} \begin{bmatrix} \check{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \end{bmatrix}, \tag{7.3}$$

where $(\mathrm{K}_\Xi)_{i,j} = k_2(\xi_i,\xi_j)$, $i,j = 1,\dots,N$, and $\Phi_{i,j} = \phi_j(\xi_i)$, $i = 1,\dots,N$, $j = 1,\dots,4$. The matrix A_Υ is an N-by-N *sparse* matrix where each column contains $n = M(\log N)^2$ entries corresponding to the values of the interpolation coefficients $A_{\xi,\zeta}$ for the local Lagrange basis in (3.2). The matrix C_Υ is a 4-by-N matrix with each column containing the values of the interpolation coefficients $c_{\xi,j}$ in (3.2). With the linear system written in this way, one can view the matrix $[\mathsf{A}_\Upsilon\ \mathsf{C}_\Upsilon]^T$ as a *right*

preconditioner for the standard kernel interpolation matrix. Once $\check{\mathbf{a}}$ is determined from (7.3), we can then find the interpolation coefficients a_ξ and c_j in (7.1) from

$$[\mathbf{a}\ \mathbf{c}]^T = [\mathsf{A}_\Upsilon\ \mathsf{C}_\Upsilon]^T \check{\mathbf{a}}. \tag{7.4}$$

If the local Lagrange basis decays sufficiently fast then the linear system (7.3) should be "numerically nice" in the sense that the matrix $\mathrm{K}_\Xi \mathsf{A}_\Upsilon + \Phi \mathsf{C}_\Upsilon$ should have decaying elements from its diagonal and should be well conditioned. As discussed in the previous section, the decay is controlled by the number of nearest neighbors n used in constructing each local Lagrange function and that $n = M(\log N)^2$. In the experiments below, we found that choosing $n = 7\lceil(\log N)^2/(\log 10)^2)\rceil = 7\lceil(\log_{10} N)^2\rceil$ gave very good results over several decades of N.

To solve the preconditioned linear system (7.3) we will use the generalized minimum residual method (GMRES) [23]. This is a Krylov subspace method which is applicable to non-symmetric linear systems and only requires computing matrix-vector products. Each matrix-vector product involving the preconditioner matrix $[A_\Upsilon\ \mathsf{C}_\Upsilon]^T$ requires $\mathcal{O}(N(\log N)^2)$ operations, while each matrix-vector product involving $[\mathrm{K}_\Xi\ \Phi]$ requires $\mathcal{O}(N^2)$ operations. However, Keiner, Kunis, and Potts have shown that this latter product can be done in $\mathcal{O}(N \log N)$ operations using fast algorithms for spherical Fourier transforms [14]. As we are primarily interested in the exploring the effectiveness of the local Lagrange basis as a preconditioner, we have not used these fast algorithms in the results below. In a followup study, we will investigate these fast algorithms in combination with the preconditioner in much more detail.

For the first numerical experiment we test the preconditioner on two different families of node sets $\Xi \subset \mathbb{S}^2$ of increasing cardinality. The first is the icosahedral family of nodes, which are quasi-uniformly distributed over the sphere with a small mesh ratio (at least for the sizes we consider). These nodes are popular in computational geosciences (see, for example, [10, 24, 22, 16]) where interpolation between node sets is often required. The second family of nodes we use are the Hammersley nodes, which are formed by transformed van der Corput sequences and provide so-called low discrepancy sequences on the sphere (see [5] for more details). While these nodes are equidistributed over the sphere, they are not quasi-uniform, and the mesh ratio can be quite large. They are thus not covered by the above theory. However, we have included them to illustrate how the preconditioner may perform on less uniform data. Additionally, low discrepancy sequences like the Hammersley nodes are also popular for numerical tests of kernel methods [6].

Table 7.1 displays the number of GMRES iterations required to compute an approximate solution to the linear systems (7.3) for these two node families using various values of N and different tolerances. The values of the target function f in these tests were chosen from a random uniform distribution between $[-1, 1]$. For both node families, the results show that the number of iterations is small and stays relatively constant as N increases. Additionally, there are only minor increases in the number of iterations as the tolerances are made stricter. The results for the Hammersley family show that the number iterations are slightly higher than the icosahedral family, but they are still small. This is encouraging considering that the mesh ratios (displayed in the third column) are considerably larger for these nodes and that we used the same formula as the icosahedral family for determining nodes in each local Lagrange basis.

For the final numerical experiment, we use the preconditioner to interpolate a field taken from a numerical simulation on the icosahedral node sets to a regular latitude-longitude grid. As mentioned above, this is often necessary for purposes of comparing solutions from different computational models, plotting solutions, or coupling different

TABLE 7.1

Number of GMRES iterations required for computing an approximate solution to (7.3) *using the quasi-uniform icosahadedral and the low discrepancy Hammersley families of nodes on the sphere. Here N is the cardinality of the node set, $n = 7\lceil(\log_{10} N)^2\rceil$ corresponds to the number of nodes used to construct the local basis, and tol refers to the tolerance on the relative residual in the GMRES method. The mesh ratio ρ_X is also displayed for each node set for comparison purposes but is not needed in the computations. The right-hand side was set to random values uniformly distributed between $[-1, 1]$, and the initial guess for GMRES was set equal to the function values.*

			Number of GMRES iterations			
			$tol = 10^{-6}$	$tol = 10^{-8}$	$tol = 10^{-10}$	$tol = 10^{-12}$
N	n	ρ_X	Icosahedral nodes			
2562	84	1.650	7	8	9	10
10242	119	1.679	5	7	8	9
23042	140	1.688	6	7	8	9
40962	154	1.693	5	7	7	8
92162	175	1.688	6	8	9	10
163842	196	1.701	5	7	7	8
N	n	ρ_X	Hammersley nodes			
4000	91	24.56	8	10	11	12
8000	112	34.74	8	9	11	12
16000	126	49.13	7	9	10	11
32000	147	69.48	7	8	10	11
64000	168	98.26	7	9	10	12

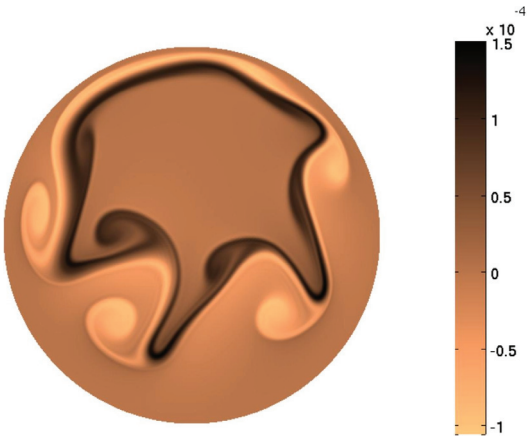

FIG. 7.1. *Interpolated relative vorticity from a numerical simulation of the shallow water wave equations on the $N = 163842$ icosahedral node sets. The original values for the relative vorticity come from* [9] *and have been interpolated to a regular 300×600 latitude-longitude based grid using the restricted kernel spline $k_2(x, \alpha) = (1 - x \cdot \alpha)\log(1 - x \cdot \alpha)$. The interpolation coefficients were computed using GMRES on the preconditioned system* (7.3).

models together. The data we use comes from [9] and represents the relative vorticity of a fluid described by the shallow water wave equations on the surface of a rotating sphere. The initial conditions for the model lead to the development of a highly nonlinear wave with rapid energy transfer from large to small scales, resulting in complex vortical dynamics. The numerical solution was computed on the $N = 163842$ node set and we interpolated it to a regular 300×600 latitude-longitude based grid. Figure 7.1 displays the resulting interpolated relative vorticity from the simulation at time $t = 6$ days. The figure clearly shows that the complex flow structure has been maintained after the interpolation. As in the numerical examples above, the

approximate solution to (7.3) with this data was obtained in seven iterations of the GMRES method using a tolerance of 10^{-8}.

REFERENCES

[1] R. Beatson and L. Greengard, *A short course on fast multipole methods*, in Wavelets, Multilevel Methods and Elliptic PDEs, Numer. Math. Sci. Comput., Oxford University Press, New York, 1997, pp. 1–37.

[2] R. K. Beatson, J. B. Cherrie, and C. T. Mouat, *Fast fitting of radial basis functions: Methods based on preconditioned GMRES iteration*, Adv. Comput. Math., 11 (1999), pp. 253–270.

[3] R. K. Beatson and M. J. D. Powell, *An iterative method for thin plate spline interpolation that employs approximations to Lagrange functions*, in Numerical Analysis 1993, Pitman Res. Notes Math. Ser. 303, Longman Scientific and Technical, Harlow, UK, 1994, pp. 17–39.

[4] R. K. Beatson, M. J. D. Powell, and A. M. Tan, *Fast evaluation of polyharmonic splines in three dimensions*, IMA J. Numer. Anal., 27 (2007), pp. 427–450.

[5] J. Cui and W. Freeden, *Equidistribution on the sphere*, SIAM J. Sci. Comput., 18 (1997), pp. 595–609.

[6] G. E. Fasshauer, *Meshfree Approximation Methods with MATLAB*, Interdiscip. Math. Sci. 6, World Scientific Publishing, Hackensack, NJ, 2007.

[7] A. C. Faul, G. Goodsell, and M. J. D. Powell, *A Krylov subspace algorithm for multiquadric interpolation in many dimensions*, IMA J. Numer. Anal., 25 (2005), pp. 1–24.

[8] A. C. Faul and M. J. D. Powell, *Proof of convergence of an iterative technique for thin plate spline interpolation in two dimensions*, Adv. Comput. Math., 11 (1999), pp. 183–192.

[9] N. Flyer, E. Lehto, S. Blaise, G. B. Wright, and A. St-Cyr, *A guide to RBF-generated finite differences for nonlinear transport: Shallow water simulations on a sphere*, J. Comput. Phys., 231 (2012), pp. 4078–4095.

[10] F. X. Giraldo, *Lagrange-Galerkin methods on spherical geodesic grids*, J. Comput. Phys., 136 (1997), pp. 197–213.

[11] T. Hangelbroek, F. J. Narcowich, and J. D. Ward, *Kernel approximation on manifolds* I: *Bounding the Lebesgue constant*, SIAM J. Math. Anal., 42 (2010), pp. 1732–1760.

[12] T. Hangelbroek, F. J. Narcowich, and J. D. Ward, *Polyharmonic and related kernels on manifolds: Interpolation and approximation*, Found. Comput. Math., 12 (2012), pp. 1–46.

[13] M. J. Johnson, *A symmetric collocation method with fast evaluation*, IMA J. Numer. Anal., 29 (2009), pp. 773–789.

[14] J. Keiner, S. Kunis, and D. Potts, *Fast summation of radial functions on the sphere*, Computing, 78 (2006), pp. 1–15.

[15] L. Ling and E. J. Kansa, *A least-squares preconditioner for radial basis functions collocation methods*, Adv. Comput. Math., 23 (2005), pp. 31–54.

[16] D. Majewski, D. Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen, and J. Baumgardner, *The operational global icosahedral-hexagonal gridpoint model GME: Description and high-resolution tests*, Mon. Wea. Rev., 130 (2002), pp. 319–338.

[17] H. N. Mhaskar, *Local quadrature formulas on the sphere*, J. Complexity, 20 (2004), pp. 753–772.

[18] 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.

[19] F. J. Narcowich, *Generalized Hermite interpolation and positive definite kernels on a Riemannian manifold*, J. Math. Anal. Appl., 190 (1995), pp. 165–193.

[20] F. J. Narcowich and J. D. Ward, *Norm estimates for the inverses of a general class of scattered-data radial-function interpolation matrices*, J. Approx. Theory, 69 (1992), pp. 84–109.

[21] F. Rellich, *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach Science Publishers, New York, 1969.

[22] T. D. Ringler, R. P. Heikes, and D. A. Randall, *Modeling the atmospheric general circulation using a spherical geodesic grid: A new class of dynamical cores*, Mon. Wea. Rev., 128 (2000), pp. 2471–2490.

[23] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Comput., 7 (1986), pp. 856–869.

[24] G. R. Stuhne and W. R. Peltier, *New icosahedral grid-point discretizations of the shallow water equations on the sphere*, J. Comput. Phys., 148 (1999), pp. 23–53.

[25] T. Tran, Q. T. Le Gia, I. H. Sloan, and E. P. Stephan, *Preconditioners for pseudodifferential equations on the sphere with radial basis functions*, Numer. Math., 115 (2010), pp. 141–163.
[26] N. J. Vilenkin, *Special Functions and the Theory of Group Representations*, Transl. Math. Monogr. 22, AMS, Providence, RI, 1968.
[27] R. Womersley, *Minimum Energy Points on the Sphere* $\mathbb{S}^2$, http://web.maths.unsw.edu.au/~rsw/Sphere/Energy/index.html (2003).