REGULARIZED LEAST SQUARES APPROXIMATIONS ON THE SPHERE USING SPHERICAL DESIGNS

CONGPEI AN* XIAOJUN CHEN* IAN H. SLOAN†* AND ROBERT S. WOMERSLEY†

20 June 2011

Abstract. We consider polynomial approximation on the unit sphere $\mathbb{S}^2 = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 = 1\}$ by a class of regularized discrete least squares methods, with novel choices for the regularization operator and the point sets of the discretization. We allow different kinds of rotationally invariant regularization operators, including the zero operator (in which case the approximation includes interpolation, quasi-interpolation and hyperinterpolation); powers of the negative Laplace-Beltrami operator (which can be suitable when there are data errors); and regularization operator that yield filtered polynomial approximations. As node sets we use spherical t-designs, which are point sets on the sphere which when used as equal-weight quadrature rules integrate all spherical polynomials up to degree t exactly. More precisely, we use well conditioned spherical t-designs obtained in a previous paper by maximizing the determinants of the Gram matrices subject to the spherical design constraint. For $t \geq 2L$ and an approximating polynomial of degree L it turns out that there is no linear algebra problem to be solved, and the approximation in some cases recovers known polynomial approximation schemes, including interpolation, hyperinterpolation and filtered hyperinterpolation. For $t \in [L, 2L)$ the linear system needs to be solved numerically, Finally, we give numerical examples to illustrate the theoretical results, and show that well chosen regularization operator and well conditioned spherical t-designs can provide good polynomial approximation on the sphere, with or without the presence of data errors.

Key words. spherical polynomial, regularized least squares approximation, filtered approximation, rotationally invariant, spherical design, perturbation.

AMS subject classifications. 65D32, 65H10

1. Introduction.

In this paper, we consider a class of polynomial approximations on the unit sphere $\mathbb{S}^2 = \{\mathbf{x} = (x,y,z)^T \in \mathbb{R}^3 \ : \ x^2+y^2+z^2 = 1\}$ arising as minimizers of regularized discrete least squares problems of the form

$$\min_{p \in \mathbb{P}_L} \Big\{ \sum_{j=1}^{N} (p(\mathbf{x}_j) - f(\mathbf{x}_j))^2 + \lambda \sum_{j=1}^{N} (\mathcal{R}_L p(\mathbf{x}_j))^2 \Big\}, \tag{1.1}$$

where f is a given continuous function with values (possibly noisy) given at N points $\mathcal{X}_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^2$. Here $\mathbb{P}_L := \mathbb{P}_L(\mathbb{S}^2)$ is the linear space of spherical polynomials of degree $\leq L$, that is, the space of restrictions to $\mathbb{S}^2$ of polynomials of degree $\leq L$ in x, y and z, and $\mathcal{R}_L : \mathbb{P}_L \to \mathbb{P}_L$, the regularization operator, is a linear operator which can be chosen in different ways, and $\lambda > 0$ is a parameter. We shall assume always that the problem is well posed, which requires the number N to be at least $\dim(\mathbb{P}_L) = (L+1)^2$.

All approximations of the form (1.1) are special cases of the penalized least squares method, studied in a general context by [9]. In this paper, we will concentrate on aspects of penalized least squares that are special to polynomials on the sphere.

*Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China.

†School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia. E-mail addresses: andbach@163.com (C. An), maxjchen@polyu.edu.hk (X. Chen), i.sloan@unsw.edu.au (I. H. Sloan), r.womersley@unsw.edu.au (R. S. Womersley).

Many different approximations are included in the formulation (1.1), through the freedom to vary the point sets $\mathcal{X}_N$ and the regularization operator $\mathcal{R}_L$. We make the natural assumption that $\mathcal{R}_L$ is rotationally invariant [22, page 5], i.e, the form of $\mathcal{R}_L$ does not depend on the choice of the x, y, z axes. The simplest example is $\mathcal{R}_L = \mathbf{0}$, in which case the approximation is interpolation if $N = (L+1)^2$, or quasi-interpolation or hyperinterpolation (see below) if $N > (L+1)^2$. Another important example is $\mathcal{R}_L = -\Delta^*$, where Δ^* is the Laplace-Beltrami operator. This choice (or more generally a positive power of $-\Delta^*$) can yield a suitable smoothing approximation if there are errors in the data.

For choosing the point set $\mathcal{X}_N$, if as in many applications the point set is given by empirical data, then the only option is to selectively delete points so as to improve the distribution. If, on the other hand, the points may be freely chosen, then we shall see that there is merit in choosing $\mathcal{X}_N$ to be a "spherical t-design" for some appropriate value of t. A point set $\mathcal{X}_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^2$ is a spherical t-design if it satisfies

$$\frac{1}{N}\sum_{j=1}^{N} p\left(\mathbf{x}_j\right) = \frac{1}{4\pi}\int_{\mathbb{S}^2} p(\mathbf{x})\, d\omega(\mathbf{x}) \qquad \forall\, p \in \mathbb{P}_t, \tag{1.2}$$

where $d\omega(\mathbf{x})$ denotes area measure on the unit sphere. That is, $\mathcal{X}_N$ is a spherical t-design if a properly scaled equal-weight quadrature rule with nodes at the points of $\mathcal{X}_N$ integrates all (spherical) polynomials up to degree t exactly. For more details on spherical designs, see [10] and Section 2.2 below. In this paper we shall always assume that $\mathcal{X}_N$ is a spherical t-design, with $t \geq L$.

To reduce (1.1) to a linear system we choose a basis for $\mathbb{P}_L$. We take a basis of orthonormal spherical harmonics [16]:

$$\{\, Y_{\ell,k} \,:\, \ell = 0, 1, \ldots, L,\; k = 1, \ldots, 2\ell+1\}.$$

The spherical harmonics $Y_{\ell,k}$ with fixed ℓ form a basis for the $2\ell+1$-dimensional space $\mathbb{H}_\ell$ of homogeneous, harmonic polynomials of degree ℓ. The orthonormality is with respect to the L_2 inner product

$$(f, g)_{L_2} := \int_{\mathbb{S}^2} f(\mathbf{x}) g(\mathbf{x}) d\omega(\mathbf{x}), \tag{1.3}$$

which induces the norm $\|f\|_{L_2} := (f, f)_{L_2}^{\frac{1}{2}}$. Then for arbitrary $p \in \mathbb{P}_L$, there is a unique vector $\boldsymbol{\alpha} = (\alpha_{\ell,k}) \in \mathbb{R}^{(L+1)^2}$ such that

$$p(\mathbf{x}) = \sum_{\ell=0}^{L}\sum_{k=1}^{2\ell+1} \alpha_{\ell,k} Y_{\ell,k}(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{S}^2. \tag{1.4}$$

We can define the regularization operator $\mathcal{R}_L$ in its most general rotationally invariant form by its action on $p \in \mathbb{P}_L$:

$$\begin{aligned} \mathcal{R}_L p(\mathbf{x}) &= \sum_{\ell=0}^{L} \beta_\ell \sum_{k=1}^{2\ell+1} Y_{\ell,k}(\mathbf{x}) (Y_{\ell,k},\, p)_{L_2} \\ &= \sum_{\ell=0}^{L} \beta_\ell \int_{\mathbb{S}^2} \frac{(2\ell+1)}{4\pi} P_\ell(\mathbf{x}\cdot\mathbf{y}) p(\mathbf{y}) d\omega(\mathbf{y}), \end{aligned} \tag{1.5}$$

where $\beta_0, \beta_1, \ldots, \beta_L$ are at this point arbitrary non-negative numbers, which may depend on L. In the last step we used the addition theorem for spherical harmonics, see [16]:

$$\sum_{k=1}^{2\ell+1} Y_{\ell,k}(\mathbf{x})Y_{\ell,k}(\mathbf{y}) = \frac{2\ell+1}{4\pi} P_\ell\left(\mathbf{x}\cdot\mathbf{y}\right), \qquad \mathbf{x},\mathbf{y} \in \mathbb{S}^2, \tag{1.6}$$

with P_ℓ the Legendre polynomial of degree ℓ normalized to $P_\ell(1) = 1$.

Given a continuous function f defined on $\mathbb{S}^2$, let $\mathbf{f} := \mathbf{f}(\mathcal{X}_N)$ be the column vector

$$\mathbf{f} = [f(\mathbf{x}_1), \ldots, f(\mathbf{x}_N)]^T \in \mathbb{R}^N,$$

and let $\mathbf{Y}_L := \mathbf{Y}_L(\mathcal{X}_N) \in \mathbb{R}^{(L+1)^2\times N}$ be a matrix of spherical harmonics evaluated at the points of $\mathcal{X}_N$, with elements

$$Y_{\ell,k}(\mathbf{x}_j), \qquad \ell = 0, 1, \ldots, L, \quad k = 1, \ldots, 2\ell+1, \qquad j = 1, \ldots, N.$$

Substituting (1.4) into (1.1), the problem (1.1) reduces to the following discrete regularized least squares problem

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^{(L+1)^2}} \|\mathbf{Y}_L^T\boldsymbol{\alpha} - \mathbf{f}\|_2^2 + \lambda\|\mathbf{R}_L^T\boldsymbol{\alpha}\|_2^2, \qquad \lambda > 0, \tag{1.7}$$

where $\mathbf{R}_L := \mathbf{R}_L(\mathcal{X}_N) = \mathbf{B}_L\mathbf{Y}_L \in \mathbb{R}^{(L+1)^2\times N}$, with $\mathbf{B}_L$ a positive semi-definite diagonal matrix, defined by

$$\mathbf{B}_L := \operatorname{diag}(\beta_0, \underbrace{\beta_1, \beta_1, \beta_1}_{3}, \ldots, \underbrace{\beta_L, \ldots, \beta_L}_{2L+1}) \in \mathbb{R}^{(L+1)^2\times(L+1)^2}. \tag{1.8}$$

Thus the matrix $\mathbf{R}_L$ is determined by the elements of the diagonal matrix $\mathbf{B}_L$ and the choice of the points $\mathcal{X}_N$. The problem (1.7) is a convex unconstrained optimization problem. Its solution set coincides with the solution set of the system of linear equations

$$\mathbf{T}_L\boldsymbol{\alpha} = \mathbf{Y}_L\mathbf{f}, \tag{1.9}$$

where $\mathbf{T}_L := \mathbf{T}_L(\mathcal{X}_N)$ is given by

$$\mathbf{T}_L = (\mathbf{H}_L + \lambda\mathbf{B}_L\mathbf{H}_L\mathbf{B}_L) \in \mathbb{R}^{(L+1)^2\times(L+1)^2}, \tag{1.10}$$

$$\mathbf{H}_L := \mathbf{H}_L(\mathcal{X}_N) = \mathbf{Y}_L\mathbf{Y}_L^T \in \mathbb{R}^{(L+1)^2\times(L+1)^2}. \tag{1.11}$$

We shall always impose conditions on $\mathcal{X}_N$ that ensure that the matrix $\mathbf{H}_L$ is non-singular. In that case (since $\mathbf{B}_L\mathbf{H}_L\mathbf{B}_L$ is positive semi-definite) the solution of (1.9) is unique. We denote that solution by $\boldsymbol{\alpha} := \boldsymbol{\alpha}(L, \mathcal{X}_N, \mathbf{B}_L) \in \mathbb{R}^{(L+1)^2}$, and the corresponding polynomial approximation by

$$p_{L,N} = \sum_{\ell=0}^{L}\sum_{k=1}^{2\ell+1} \alpha_{\ell,k}Y_{\ell,k}. \tag{1.12}$$

As stated before, we assume that $\mathcal{X}_N$ is a spherical t-design for some $t \geq L$. It is useful to consider separately the cases $L \leq t < 2L$ and $t \geq 2L$, because in the first case important issues arise from the conditioning of the least squares problem (1.7), while in the second case, as we shall see in the following theorem, the matrix becomes diagonal and hence the linear algebra becomes trivial.

THEOREM 1.1. *Assume $f \in C(\mathbb{S}^2)$. Let $L \geq 0$ be given, and let $\mathcal{X}_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^2$ be a spherical t-design on $\mathbb{S}^2$ with $t \geq 2L$. Then*

$$\mathbf{H}_L = \mathbf{Y}_L \mathbf{Y}_L^T = \frac{N}{4\pi} \mathbf{I}_{(L+1)^2} \in \mathbb{R}^{(L+1)^2 \times (L+1)^2}, \tag{1.13}$$

while (1.9) *has a unique solution*

$$\alpha_{\ell,k} = \frac{4\pi}{N(1+\lambda\beta_\ell^2)} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) f(\mathbf{x}_j), \tag{1.14}$$

and the unique minimizer of (1.1) *is given by*

$$\begin{aligned} p_{L,N}(\mathbf{x}) &= \frac{4\pi}{N} \sum_{\ell=0}^{L} \sum_{k=1}^{2\ell+1} \frac{Y_{\ell,k}(\mathbf{x})}{1+\lambda\beta_\ell^2} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) f(\mathbf{x}_j) \\ &= \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j) f(\mathbf{x}_j). \end{aligned} \tag{1.15}$$

Proof. Under the conditions in the theorem $\mathbf{H}_L$ becomes diagonal, since by (1.11) and the definition (1.2) of a spherical t-design, for $t \geq 2L$ we have

$$(\mathbf{H}_L)_{\ell,k\,\ell',k'} = \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) Y_{\ell',k'}(\mathbf{x}_j) = \frac{N}{4\pi} \left(Y_{\ell,k},\, Y_{\ell',k'}\right)_{L_2} = \frac{N}{4\pi} \delta_{\ell\ell'} \delta_{kk'},$$

where $\ell, \ell' = 0, \ldots, L, \quad k = 1, \ldots, 2\ell+1, \quad k' = 1, \ldots, 2\ell'+1$. The middle equality holds because the product $Y_{\ell,k} Y_{\ell',k'} \in \mathbb{P}_{2L} \subset \mathbb{P}_t$ and that $\mathcal{X}_N$ is a spherical t-design. Thus (1.13) holds, and in turn

$$\mathbf{T}_L = \frac{N}{4\pi} \left(\mathbf{I}_{(L+1)^2} + \lambda \mathbf{B}_L^2\right). \tag{1.16}$$

Since $\mathbf{B}_L$ is diagonal with diagonal elements β_ℓ, the solution of (1.9) is given by (1.14) and the minimizer of (1.1) is therefore (1.15). ∎

Define the uniform norm of a continuous function f over the unit sphere $\mathbb{S}^2$ by

$$||f||_{C(\mathbb{S}^2)} := \sup_{x \in \mathbb{S}^2} |f(\mathbf{x})|. \tag{1.17}$$

In the limiting case $t \to \infty$ we obtain the following rather simple theorem.

THEOREM 1.2. *Let $f \in C(\mathbb{S}^2)$, and let $L \geq 0$ be given. Assume that the sets $\mathcal{X}_{N(t)}^{(t)} = \{\mathbf{x}_1^{(t)}, \ldots, \mathbf{x}_{N(t)}^{(t)}\}$ for $t = 1, 2, \ldots$ form a sequence of spherical t-designs with $t \geq L$. Then the unique minimizer $p_{L,N(t)} \in \mathbb{P}_L$ of (1.1) has the uniform limit p_L as $t \to \infty$, that is*

$$\lim_{t\to\infty} \|p_{L,N(t)} - p_L\|_{C(\mathbb{S}^2)} = 0, \tag{1.18}$$

where $p_L \in \mathbb{P}_L$ denotes the unique minimizer of the continuous regularized least squares problem

$$\min_{p\in\mathbb{P}_L} \left\{\|f-p\|_{L_2}^2 + \lambda\|\mathcal{R}_L\, p\|_{L_2}^2\right\}, \qquad \lambda > 0. \tag{1.19}$$

Proof. We have seen already that $p_{L,N}$ is uniquely determined when $t \geq 2L$, and that in this case $p_{L,N}$ is given explicitly by (1.15). It is easy to see that the minimizer of problem (1.19) is in a similar way given by

$$p_L(\mathbf{x}) = \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)4\pi} \int_{\mathbb{S}^2} P_\ell(\mathbf{x}\cdot\mathbf{y}) f(\mathbf{y}) d\omega(\mathbf{y}). \tag{1.20}$$

Since the sums over ℓ in (1.15) and (1.20) are finite, to prove the theorem it is sufficient to prove that for $0 \leq \ell \leq L$

$$\lim_{t\to\infty} \frac{1}{N} \sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j^{(t)}) f(\mathbf{x}_j^{(t)}) = \frac{1}{4\pi} \int_{\mathbb{S}^2} P_\ell(\mathbf{x}\cdot\mathbf{y}) f(\mathbf{y})\; d\omega(\mathbf{y}). \tag{1.21}$$

Noting that $P_\ell(\mathbf{x}\cdot\mathbf{y})f(\mathbf{y})$ is a continuous function of $\mathbf{y}$ for each fixed $\mathbf{x} \in \mathbb{S}^2$, the result now follows from the well known result that, for a positive weight quadrature rule with polynomial degree of accuracy t, the quadrature rule applied to a continuous function g converges to the integral of g as $t \to \infty$. For an explicit proof for the case of the sphere (and indeed for an error estimate) see [13, Th. 10] combined with Jackson's theorem (see (2.5)). ∎

In Section 2 we give necessary background information on polynomial spaces, spherical designs, and hyperinterpolation and its variants.

In Section 3 we discuss in detail several interesting choices of the regularization matrix $\mathbf{B}_L$:

1. The choice $\mathbf{B}_L = \mathbf{0}$, gives the unregularized problem. This choice includes (depending on N, L and t) the classical interpolation problem (when $N = (L+1)^2$), quasi-interpolation (when $N > (L+1)^2$), hyperinterpolation (when $t \geq 2L$ (see Subsection 2.3 below), and orthogonal projection in the limit $t \to \infty$ with fixed L.
2. Choices of $\mathbf{B}_{L-1}$ related to "filtered" polynomial approximation (see [24], [26]), in which the diagonal elements β_ℓ of $\mathbf{B}_{L-1}$ are chosen so that

$$\frac{1}{1+\lambda\beta_\ell^2} = h\left(\frac{\ell}{L}\right), \quad \ell = 0, \ldots, L-1, \tag{1.22}$$

   where $h(x)$ is a prescribed "filter" function on $\mathbb{R}^+$, vanishing for $x \geq 1$. The motivation is that, as we shall see, for $t \geq 2L$ such approximations can have excellent approximation properties in the uniform norm if h is well chosen. We shall see that they may also be good candidates when $t < 2L$. We shall call the approximation with β_ℓ satisfying (1.22) "filtered least squares".
3. Choices related to the Laplace-Beltrami operator Δ^*. (Laplace-Beltrami regularization operator).

In Section 4, we consider the condition number $\kappa(\mathbf{T}_L) = \kappa(\mathbf{T}_L(\mathcal{X}_N))$, where $\mathcal{X}_N$ is a spherical t-design with $t \geq L$. We see that the condition number generally becomes smaller (that is, the condition of the linear system improves) as t approaches $2L$. In the case $t = L$ and $N = (t+1)^2 = (L+1)^2$ the problem of conditioning the unregularized (interpolation) problem was already addressed by Chen and Womersley [5], through the concept of "extremal spherical designs". In the recent paper [7], the condition number was studied for $N \geq (t+1)^2$.

In Section 5, we discuss theoretical error bounds for various versions of the approximation.

In Section 6 we present numerical results of the approximation for both a smooth function and a nonsmooth function, using regularized least squares, with different choices for $\mathcal{R}_L$ and different spherical t-designs, and with and without data errors for both a smooth function and a nonsmooth function.

2. Background and notation.

2.1. Polynomial spaces on the unit sphere. For $\ell \geq 0$, let $\mathbb{H}_\ell := \mathbb{H}_\ell(\mathbb{S}^2)$ be the space of restrictions to $\mathbb{S}^2$ of the (real) homogeneous harmonic polynomials of degree $\ell \geq 0$. Its dimension is $\dim(\mathbb{H}_\ell) = 2\ell + 1$ [16]. Note that the rotationally invariant operator defined by (1.5) satisfies

$$\mathcal{R}_L p = \beta_\ell p, \qquad \text{for} \quad p \in \mathbb{H}_\ell, \quad \ell = 0, \ldots, L.$$

It is known that

$$\mathbb{P}_L = \sum_{\ell=0}^{L} \bigoplus \mathbb{H}_\ell,$$

and that the spaces $\mathbb{H}_\ell$ are mutually orthogonal with respect to the inner product (1.3): if $p \in \mathbb{H}_\ell$ and $p' \in \mathbb{H}_{\ell'}$ with $\ell \neq \ell'$, then $(p, p')_{L_2} = 0$.

The set of spherical harmonics

$$\{Y_{\ell,k} \;:\; k = 1, \ldots, 2\ell+1,\; \ell = 0, 1, \ldots\}$$

is a complete orthonormal basis of $L_2(\mathbb{S}^2)$. It follows that an arbitrary $f \in L_2(\mathbb{S}^2)$ can be represented in the L_2 sense by its Fourier (or Laplace) series [11] with respect to the spherical harmonics:

$$f = \sum_{\ell=0}^{\infty} \sum_{k=1}^{2\ell+1} \widehat{f}_{\ell,k} Y_{\ell,k}, \tag{2.1}$$

with the Fourier coefficients given by

$$\widehat{f}_{\ell,k} := (f, Y_{\ell,k})_{L_2} = \int_{\mathbb{S}^2} f(\mathbf{x}) Y_{\ell,k}(\mathbf{x}) \, d\omega(\mathbf{x}). \tag{2.2}$$

The orthogonal projection operator $\mathcal{P}_L : L_2(\mathbb{S}^2) \to \mathbb{P}_L$ onto $\mathbb{P}_L$ can be represented by

$$\mathcal{P}_L f(\mathbf{x}) = \sum_{\ell=0}^{L} \sum_{k=1}^{2\ell+1} \widehat{f}_{\ell,k} Y_{\ell,k}(\mathbf{x}). \tag{2.3}$$

We follow Reimer [21] in saying that, for a given positive integer k, $f : \mathbb{S}^2 \to \mathbb{R}$ is k times differentiable if all restrictions of f to a great circle are k times differentiable; that is, if

$$f_{\mathbf{x},\mathbf{y}}(\alpha) := f(\mathbf{x}\cos\alpha + \mathbf{y}\sin\alpha), \qquad \alpha \in \mathbb{R},$$

is k times differentiable for all choices of $\mathbf{x},\ \mathbf{y} \in \mathbb{S}^2$ with $\mathbf{x} \perp \mathbf{y}$. If so we then define

$$\left|\left|f^{(k)}\right|\right|_{C(\mathbb{S}^2)} := \sup\left\{ |f^{(k)}_{\mathbf{x},\mathbf{y}}(\alpha)| :\ \alpha \in [0,\, 2\pi],\, \mathbf{x},\, \mathbf{y} \in \mathbb{S}^2,\, \mathbf{x} \perp \mathbf{y} \right\}, \tag{2.4}$$

and $C^k(\mathbb{S}^2)$ may be defined as the set of real valued function f on $\mathbb{S}^2$ such that $\left|\left|f^{(k)}\right|\right|_{C(\mathbb{S}^2)}$ is finite. For a function $f \in C^k(\mathbb{S}^2)$, we have Jackson's theorem for the sphere, see [17, Th. 3.3], a simple version of which is

$$E_L(f) := \inf_{p \in \mathbb{P}_L} ||f - p||_{C(\mathbb{S}^2)} \le c(f,k) L^{-k}. \tag{2.5}$$

The reproducing kernel $G_L :\ \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}$ of the space $\mathbb{P}_L$ is

$$G_L(\mathbf{x},\mathbf{y}) = g_L(\mathbf{x}\cdot\mathbf{y}) = \sum_{\ell=0}^{L} \sum_{k=1}^{2\ell+1} Y_{\ell,k}(\mathbf{x}) Y_{\ell,k}(\mathbf{y}) = \sum_{\ell=0}^{L} \frac{2\ell+1}{4\pi} P_\ell(\mathbf{x}\cdot\mathbf{y}), \tag{2.6}$$

where the last equality is due to the addition theorem (1.6). It has the three properties needed for a reproducing kernel:

$$G_L(\mathbf{x},\cdot) \in \mathbb{P}_L, \quad \mathbf{x} \in \mathbb{S}^2; \qquad G_L(\mathbf{x},\mathbf{y}) = G_L(\mathbf{y},\mathbf{x}), \quad \mathbf{x},\mathbf{y} \in \mathbb{S}^2;$$

$$(p, G_L(\mathbf{x},\cdot))_{L_2} = p(\mathbf{x}), \quad \mathbf{x} \in \mathbb{S}^2, \quad p \in \mathbb{P}_L.$$

The projection $\mathcal{P}_L$ can be written in terms of the reproducing kernel:

$$\begin{aligned}\mathcal{P}_L f(\mathbf{x}) &= (f, G_L(\mathbf{x},\cdot))_{L_2} = \int_{\mathbb{S}^2} f(\mathbf{y}) g_L(\mathbf{x}\cdot\mathbf{y}) d\omega(\mathbf{y}) \\ &= \sum_{\ell=0}^{L} \frac{2\ell+1}{4\pi} \int_{\mathbb{S}^2} f(\mathbf{y}) P_\ell(\mathbf{x}\cdot\mathbf{y}) d\omega(\mathbf{y}), \qquad f \in L_2(\mathbb{S}^2), \quad \mathbf{x} \in \mathbb{S}^2.\end{aligned} \tag{2.7}$$

2.2. Spherical designs. In this paper we are interested in spherical t-designs (see Section 1 for the definition) with the number of points N close to $\dim(\mathbb{P}_t) = (t+1)^2$. While there is no proof that such spherical designs exist for all t, the paper of Chen and Womersley [5], and then that of Chen, Frommer and Lang [6], show by interval analysis that there exist "extremal spherical t-designs" with $N = (t+1)^2$ for all values of t up to 100. These are spherical designs for which the determinant of the matrix $\mathbf{H}_L$ (now $\mathbf{Y}_L$ is square) is maximized over $\mathcal{X}_N$. Recently, "well conditioned spherical designs" were defined and constructed in [1]. These are spherical designs with $N \ge (t+1)^2$ obtained by maximizing the determinant of $\mathbf{H}_L$ subject to the spherical design constraint. They are designed to have good properties for both interpolation and numerical integration. For further information about spherical t-designs, see [2] and [10].

2.3. Hyperinterpolation and its variants. *Hyperinterpolation* was introduced by Sloan [23] in 1995. The hyperinterpolation operator $\mathcal{L}_L$ is defined by replacing Fourier integrals in the L_2-orthogonal projection onto the space $\mathbb{P}_L$, see (2.3), by a quadrature rule that integrates exactly all spherical polynomials of degree up to $2L$. It is known that (see [23, Lemma 6]) for $L \geq 3$ the number of quadrature points in hyperinterpolation must exceed the dimension of the polynomial space, thus hyperinterpolation is intrinsically different from interpolation. In this paper, for the quadrature rules needed for hyperinterpolation we allow only spherical designs.

Using a spherical t-design $\mathcal{X}_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^2$ for $t \geq 2L$, we define the semi-inner product $(\cdot,\cdot)_N$ of two continuous functions $f, g \in C(\mathbb{S}^2)$ by

$$(f,g)_N := \frac{4\pi}{N}\sum_{j=1}^{N} f(\mathbf{x}_j)g(\mathbf{x}_j), \qquad j = 1, \ldots, N. \tag{2.8}$$

It is clear that

$$(p,q)_N = (p,q)_{L_2} = \int_{\mathbb{S}^2} p(\mathbf{x})q(\mathbf{x})d\omega(\mathbf{x}), \qquad p, q \in \mathbb{P}_L,$$

because $pq \in \mathbb{P}_{2L}(\mathbb{S}^2)$ and $t \geq 2L$. We note that for $f \in C(\mathbb{S}^2)$, $(f,f)_N = 0$ implies $f(\mathbf{x}_j) = 0,\ \ j = 1, \ldots, N$, but does not imply $f \equiv 0$. Thus (2.8) corresponds only to a semi-norm $\|f\|_N := \sqrt{(f,f)_N}$ in $C(\mathbb{S}^2)$.

The hyperinterpolant of a function $f \in C(\mathbb{S}^2)$ is defined by

$$\mathcal{L}_L f(\mathbf{x}) = \sum_{\ell=0}^{L}\sum_{k=1}^{2\ell+1} (f, Y_{\ell,k})_N Y_{\ell,k}(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{S}^2. \tag{2.9}$$

With the aid of the reproducing kernel G_L on the unit sphere, see (2.6), we can write the hyperinterpolant as

$$\begin{aligned}\mathcal{L}_L f(\mathbf{x}) =& (f, G_L(\mathbf{x},\cdot))_N = \frac{4\pi}{N}\sum_{j=1}^{N} f(\mathbf{x}_j) g_L(\mathbf{x}\cdot\mathbf{x}_j)\\ =& \sum_{\ell=0}^{L}\frac{2\ell+1}{N}\sum_{j=1}^{N} f(\mathbf{x}_j)P_\ell(\mathbf{x}\cdot\mathbf{x}_j), \qquad \mathbf{x} \in \mathbb{S}^2,\end{aligned} \tag{2.10}$$

which is just the discrete version of the orthogonal projection $\mathcal{P}_L f$ given by (2.7). In particular $\mathcal{L}_L f \in \mathbb{P}_L$, and by exactness of the quadrature rule for polynomials of degree$\leq 2L$ and orthogonality of the spherical harmonics, we have, for $0 \leq \ell \leq L$,

$$\begin{aligned}(\mathcal{L}_L f,\, Y_{\ell,k})_N =& (\mathcal{L}_L f,\, Y_{\ell,k})_{L_2} = \sum_{\ell'=0}^{L}\sum_{k'=1}^{2\ell+1} (f, Y_{\ell',k'})_N \left(Y_{\ell',k'}(\mathbf{x}),\, Y_{\ell,k}(\mathbf{x})\right)_{L_2}\\ =& (f,\, Y_{\ell,k})_N,\end{aligned} \tag{2.11}$$

giving an equivalent definition of hyperinterpolation:

$$\mathcal{L}_L f \in \mathbb{P}_L, \quad (f - \mathcal{L}_L f, p)_N = 0 \qquad \forall p \in \mathbb{P}_L. \tag{2.12}$$

A subset of $\mathbb{S}^2$

$$C(\mathbf{x}_c, r) = \{\mathbf{x} \in \mathbb{S}^2 : \mathbf{x} \cdot \mathbf{x}_c \geq \cos r\}, \qquad \mathbf{x}_c \in \mathbb{S}^2, \ r > 0, \tag{2.13}$$

is called a *spherical cap* (see [22, p. 195]) with center $\mathbf{x}_c$ and radius r.

The Lebesgue constant of the operator $\mathcal{L}_L f$, defined by

$$||\mathcal{L}_L||_{C(\mathbb{S}^2)} := \sup_{f \in C(\mathbb{S}^d)\setminus\{\mathbf{0}\}} \frac{||\mathcal{L}_L f||_{C(\mathbb{S}^2)}}{||f||_{C(\mathbb{S}^2)}} \tag{2.14}$$

was shown by [25] to satisfy

$$c\sqrt{L+1} \leq ||\mathcal{L}_L||_{C(\mathbb{S}^2)} \leq c_1\sqrt{L+1}, \qquad L = 0, 1, \ldots, \tag{2.15}$$

for some positive constants c, c_1, provided that the point set $\mathcal{X}_N$ satisfies a regularity condition of the form

$$\sum_{\substack{j=1 \\ \mathbf{x}_j \in \mathcal{X}_N \bigcap C(\mathbf{x}, \frac{1}{2L})}}^{N} 1 \leq \frac{c_0}{N}, \qquad \mathbf{x} \in \mathbb{S}^2,$$

for some positive constant c_0.

Subsequently Reimer [20] showed that the regularity condition is satisfied automatically for the points of a positive-weight quadrature rule with polynomial degree of precision $2L$, and therefore for the points of a spherical t-design with $t \geq 2L$. Reimer in that paper also gave a new proof of (2.15) and extended the result to spheres of arbitrary dimension d. The original proof of (2.15) in [25] was extended to arbitrary dimensions d by [12].

Filtered hyperinterpolation first appeared in the forthcoming paper [26]. It can be considered as an example of a large class of *generalized hyperinterpolation* approximations defined by Reimer [21]. However, it does not belong to the subclass preferred by Reimer, of approximations based on positive kernels. It is known that positive kernels lead to convergence for all continuous functions, but it is also known from a result of Korovkin [15] that their best possible rate of convergence is L^{-2}.

In contrast, it follows from (2.20) below that filtered hyperinterpolation has a rate of convergence of order $O(L^{-k})$ for k arbitrarily large, provided $f \in C^k(\mathbb{S}^2)$.

In this method of filtered hyperinterpolation the kernel G_L in (2.10) is replaced by a "filtered" kernel

$$H_L(\mathbf{x}, \mathbf{y}) = H_L(\mathbf{x} \cdot \mathbf{y}) = \sum_{\ell=0}^{L-1} h\left(\frac{\ell}{L}\right) \frac{2\ell+1}{4\pi} P_\ell(\mathbf{x} \cdot \mathbf{y}), \tag{2.16}$$

with $h : \mathbb{R}^+ \to \mathbb{R}^+$ a function with at least $C^1(\mathbb{R}^+)$ smoothness, satisfying

$$h(x) = \begin{cases} 1, & x \in [0, 1/2], \\ 0, & x \in [1, \infty). \end{cases}$$

Thus the filtered hyperinterpolant $\mathcal{F}_L f \in \mathbb{P}_{L-1}$ is defined by

$$\begin{aligned} \mathcal{F}_L f(\mathbf{x}) &= (f, H_L(\mathbf{x},\cdot))_N = \frac{4\pi}{N}\sum_{j=1}^{N} f(\mathbf{x}_j) H_L(\mathbf{x},\mathbf{x}_j) \\ &= \sum_{\ell=0}^{L-1}(2\ell+1)h\left(\frac{\ell}{L}\right)\frac{1}{N}\sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j) f(\mathbf{x}_j), \end{aligned} \tag{2.17}$$

which according to [26] can be shown to satisfy

$$||\mathcal{F}_L||_{C(\mathbb{S}^2)} \leq c, \tag{2.18}$$

$$\mathcal{F}_L p = p \qquad \forall p \in \mathbb{P}_{\lfloor L/2 \rfloor}, \tag{2.19}$$

and hence

$$||\mathcal{F}_L f - f||_{C(\mathbb{S}^2)} \leq c E_{\lfloor L/2 \rfloor}(f), \tag{2.20}$$

where $\lfloor\cdot\rfloor$ denotes the floor function and c is a constant, provided that

$$\Delta^3 h\left(\frac{\ell}{L}\right) \leq c\frac{1}{L^2}, \tag{2.21}$$

where Δ is the forward difference operator.

In this paper, we define a new filter function,

$$h(x) = \begin{cases} 1, & x \in [0, 1/2], \\ \sin^2 \pi x, & x \in [1/2, 1], \\ 0, & x \in [1, \infty) \end{cases} \tag{2.22}$$

to replace the quadratic spline function in [26]. For this function it is easily verified by direct calculation that (2.21) holds. The function (2.22) is shown in Figure 2.1.

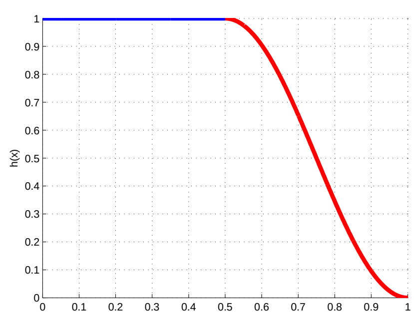


Fig. 2.1: The filter function $h(x)$

3. Choices of the regularization operators $\mathcal{R}_L$. According to (1.8), the regularization operator $\mathcal{R}_L$ is determined by the choice of the diagonal matrix $\mathbf{B}_L$ with diagonal elements β_ℓ. In the following subsections we present some interesting examples.

3.1. $\mathcal{R}_L = \mathbf{0}$. In the case that $\mathbf{B}_L$ is the zero matrix we obtain the classical least squares approximation, in which $p_{L,N}$ is the minimizer of

$$\min_{p\in\mathbb{P}_L} \|f - p\|_N. \tag{3.1}$$

It is known [23, Lemma 5] that for $t \geq 2L$ the minimizer is in this case the hyperinterpolant (2.9). If $L < t < 2L$ then the approximation is what is sometimes called quasi-interpolation. If $N = (L+1)^2$ then (regardless of the value of t) the approximation is polynomial interpolation.

3.2. Filtered least squares. The minimizer of the regularized least squares problem (1.9) can in some cases be considered as equivalent to a filtered polynomial approximation of the form in (2.17). Indeed, we have seen already, in Theorem 1.1, that for $t \geq 2L$ the minimizer of (1.9) is given by (1.15), which on setting $\lambda = 1$ coincides with the filtered polynomial approximation (2.17) if

$$\frac{1}{1+\beta_\ell^2} = h\left(\frac{\ell}{L}\right), \qquad \ell = 0, \ldots, L-1,$$

or correspondingly, if

$$\beta_\ell = \sqrt{\frac{1}{h(\ell/L)} - 1}, \qquad \ell = 0, \ldots, L-1. \tag{3.2}$$

With $\lambda = 1$ and the choice (3.2) the regularized least squares approximation thus coincides exactly with filtered hyperinterpolation [26] when $t \geq 2L$. But when $t < 2L$ the regularized least squares approximation with β_ℓ given by (3.2) and h by (2.22) is a new approximation, one not previously studied.

For filtered least squares the approximating polynomial is of degree at most $L-1$, thus we may replace L in (1.7) – (1.12) by $L-1$.

3.3. Laplace-Beltrami regularization operator. In this subsection we obtain choices of $\mathcal{R}_L$ related to the Laplace-Beltrami operator Δ^* [16, pp. 38-39] on $\mathbb{S}^2$, which is the angular part of the Laplace operator in three dimensions

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Defining the spherical polar coordinate system $(\theta, \varphi), 0 \leq \theta \leq \pi,\ 0 \leq \varphi < 2\pi$ in terms of the Cartesian coordinates x, y, z by

$$x = \sin\theta\,\cos\varphi, \qquad y = \sin\theta\,\sin\varphi, \qquad z = \cos\theta,$$

the Laplace-Beltrami operator as a differential operator is

$$\Delta^* := \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}.$$

The spherical harmonics have an intrinsic characterization as the eigenfunctions of the Laplace-Beltrami operator Δ^*, that is

$$\Delta^* Y_{\ell,k}(\mathbf{x}) = -\ell(\ell+1)Y_{\ell,k}(\mathbf{x}). \tag{3.3}$$

It follows that $-\Delta^*$ is a semi-positive operator, and for any $s > 0$ we may define $(-\Delta^*)^s$ by

$$(-\Delta^*)^s Y_{\ell,k}(\mathbf{x}) = [\ell(\ell+1)]^s Y_{\ell,k}(\mathbf{x}). \tag{3.4}$$

The corresponding matrix $\mathbf{B}_L$ is then

$$\mathbf{B}_L = \operatorname{diag}\left(0^s, 2^s, 2^s, 2^s, \ldots, \underbrace{[L(L+1)]^s, \ldots, [L(L+1)]^s}_{2L+1 \text{ times}}\right) \in \mathbb{R}^{(L+1)^2\times(L+1)^2}. \tag{3.5}$$

4. Condition number of regularized least squares approximation. In this section we study a perturbation bound for the regularized least squares problem. For convenience we denote $d_L = (L+1)^2$. For a symmetric matrix $\mathbf{M} \in \mathbb{R}^{d_L\times d_L}$, let $\lambda_1(\mathbf{M})$ and $\lambda_{d_L}(\mathbf{M})$ denote the largest and smallest eigenvalues of $\mathbf{M}$, and let $\kappa(\mathbf{M}) = \lambda_1(\mathbf{M})/\lambda_{d_L}(\mathbf{M})$ denote the condition number of $\mathbf{M}$ in the Euclidean norm. In this paper, since the diagonal elements of matrix $\mathbf{B}_L$ are in a non-decreasing order in the three choices of regularization operator, we have $\lambda_1(\mathbf{B}_L) = \beta_L$ and $\lambda_{d_L}(\mathbf{B}_L) = \beta_0$. All results in this section can be easily extended to a general diagonal matrix $\mathbf{B}_L$.

Let

$$\mathbf{A} = \begin{bmatrix} \mathbf{Y}_L^T \\ \sqrt{\lambda}\mathbf{R}_L^T \end{bmatrix} \in \mathbb{R}^{2N\times d_L}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{f} \\ \mathbf{O} \end{bmatrix} \in \mathbb{R}^{2N},$$

where $\mathbf{O}$ is an $N\times 1$ zero vector. Then the problem (1.7) can be written as

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^{d_L}} ||\mathbf{A}\boldsymbol{\alpha} - \mathbf{b}||_2^2, \tag{4.1}$$

which is equivalent to (1.9) since $\mathbf{T}_L = \mathbf{A}^T\mathbf{A}$ and $\mathbf{Y}_L\mathbf{f} = \mathbf{A}^T\mathbf{b}$.

THEOREM 4.1. *Let matrices $\mathbf{H}_L$, $\mathbf{T}_L$ and $\mathbf{B}_L$ be defined as in* (1.11),(1.10) *and* (1.8), *respectively, where $\mathcal{X}_N$ is a spherical t-design. Let $\mathbf{f}^\delta$ denote a perturbation of $\mathbf{f}$. Then for $t \geq L$, we have*

$$\kappa(\mathbf{T}_L) := \frac{\lambda_1(\mathbf{T}_L)}{\lambda_{d_L}(\mathbf{T}_L)} \leq \kappa(\mathbf{H}_L)\frac{1+\lambda\beta_L^2}{1+\lambda\beta_0^2} \tag{4.2}$$

and

$$\frac{||\boldsymbol{\alpha} - \boldsymbol{\alpha}^\delta||_2}{||\boldsymbol{\alpha}||_2} \leq \frac{||\mathbf{f} - \mathbf{f}^\delta||_2}{||\mathbf{f}||_2}\sqrt{\kappa(\mathbf{H}_L)\frac{1+\lambda\beta_L^2}{1+\lambda\beta_0^2}}. \tag{4.3}$$

Proof. Firstly we obtain the bound on the condition number of $\mathbf{T}_L$. From the eigenvalue inequalities for the product of symmetric matrices [14, page 224], we have

$$\lambda_1(\mathbf{B}_L\mathbf{H}_L\mathbf{B}_L) \leq \lambda_1(\mathbf{B}_L^2)\lambda_1(\mathbf{H}_L) = \beta_L^2\lambda_1(\mathbf{H}_L), \tag{4.4a}$$
$$\lambda_{d_L}(\mathbf{B}_L\mathbf{H}_L\mathbf{B}_L) \geq \lambda_{d_L}(\mathbf{B}_L^2)\lambda_{d_L}(\mathbf{H}_L) = \beta_0^2\lambda_{d_L}(\mathbf{H}_L). \tag{4.4b}$$

Combining (1.10), (4.4a), (4.4b) and Weyl's inequalities [14, page 181], we obtain

$$\lambda_1(\mathbf{T}_L) \le \lambda_1(\mathbf{H}_L) + \lambda_1(\mathbf{B}_L\mathbf{H}_L\mathbf{B}_L)\lambda \le (1+\lambda\beta_L^2)\lambda_1(\mathbf{H}_L),$$
$$\lambda_{d_L}(\mathbf{T}_L) \ge \lambda_{d_L}(\mathbf{H}_L) + \lambda_{d_L}(\mathbf{B}_L\mathbf{H}_L\mathbf{B}_L)\lambda \ge (1+\lambda\beta_0^2)\lambda_{d_L}(\mathbf{H}_L).$$

Therefore

$$\kappa(\mathbf{T}_L) = \frac{\lambda_1(\mathbf{T}_L)}{\lambda_{d_L}(\mathbf{T}_L)} \le \kappa(\mathbf{H}_L)\frac{1+\lambda\beta_L^2}{1+\lambda\beta_0^2}.$$

By applying the standard least squares perturbation bound (see [4, Theorem 1.4.6]) to (4.1) we find

$$\frac{\left|\left|\boldsymbol{\alpha}-\boldsymbol{\alpha}^\delta\right|\right|_2}{||\boldsymbol{\alpha}||_2} \le \frac{\left|\left|\mathbf{b}-\mathbf{b}^\delta\right|\right|_2}{||\mathbf{b}||_2}\kappa(\mathbf{A}) = \frac{\left|\left|\mathbf{f}-\mathbf{f}^\delta\right|\right|_2}{||\mathbf{f}||_2}\kappa(\mathbf{A}). \tag{4.5}$$

Then by using the singular value decomposition [4, page 28] and $\mathbf{A}^T\mathbf{A} = \mathbf{T}_L$, it can be verified that

$$\kappa(\mathbf{A}) = ||\mathbf{A}||_2\left|\left|\mathbf{A}^\dagger\right|\right|_2 = \sqrt{||\mathbf{A}^T\mathbf{A}||_2\,||(\mathbf{A}^T\mathbf{A})^{-1}||_2} = \sqrt{\kappa(\mathbf{T}_L)}, \tag{4.6}$$

with $\mathbf{A}^\dagger$ denoting the generalized inverse of $\mathbf{A}$. Now substituting (4.6) into (4.5) and using (4.2) we derive the perturbation bound (4.3). □

Remark 4.1. *For the case of filtered least squares, L in Theorem 4.1 should be replaced by $L-1$.*

Remark 4.2. *The estimate* (4.2) *is sharp since when $t \ge 2L$, the matrix $\mathbf{H}_L$ is a scalar multiple the identity matrix (see Theorem 1.1), and from* (1.16) *we find*

$$\kappa(\mathbf{T}_L) = \frac{1+\lambda\beta_L^2}{1+\lambda\beta_0^2}, \qquad t \ge 2L.$$

We discuss $\kappa(\mathbf{T}_L)$ for the three choices of the regularization operator.

1. If $\mathbf{B}_L = \mathbf{0}$, then $\kappa(\mathbf{T}_L) = \kappa(\mathbf{H}_L)$. When $t \ge 2L$, $\kappa(\mathbf{T}_L) = 1$. For $L \le t < 2L$, our well conditioned spherical t-designs [1] provide good condition numbers, see [7, Fig. 4.5].
2. For filtered least squares approximation, we consider the condition number of $\mathbf{T}_{L-1}$. From Subsection 3.2 and Theorem 4.1, we have $\beta_0 = 0$ and

   $$\kappa(\mathbf{T}_{L-1}) \le \kappa(\mathbf{H}_{L-1})(1+\lambda\beta_{L-1}^2),$$

   with equality for $t \ge 2L$. Fig. 4.1 shows for $L = 30$ and $\lambda = 1$ that $\kappa(\mathbf{T}_{29})$ generally goes down as t increases for $t = 30, 31, \ldots, 60$, with well conditioned spherical t-designs. When $t \ge 58$, $\kappa(\mathbf{T}_{29})$ is constant at $1 + \beta_{29}^2 = 1/\sin^2(\pi 29/30) \approx 91.5231$.
3. For the Laplace-Beltrami regularization operator, from (3.5) we have $\beta_0 = 1$, $\beta_L = (L(L+1))^{2s}$, and

   $$\kappa(\mathbf{T}_L) \le \kappa(\mathbf{H}_L)(1+\lambda(L(L+1))^{2s}),$$

   which monotonically increases as the parameter λ increases.

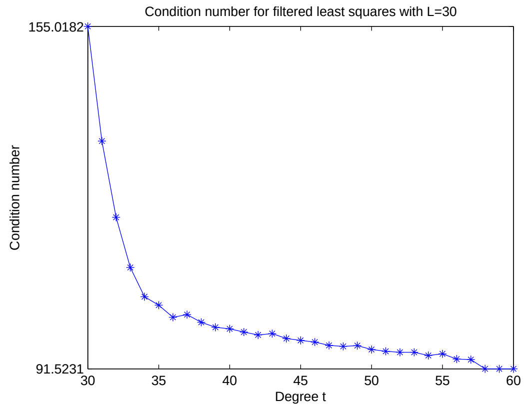


Fig. 4.1: Condition numbers of $\mathbf{T}_{29}$ for filtered polynomial approximation

The condition number of $\mathbf{T}_L$ can be large if the diagonal elements β_ℓ of $\mathbf{B}_L$ are large. For example, if the regularization operator is $\mathcal{R}_L = (-\Delta^*)^s,\ s > 0$, then

$$\kappa(\mathbf{T}_L) = 1 + \lambda(L(L+1))^{2s}, \qquad t \geq 2L, \tag{4.7}$$

which can be very large when s is large. We present an example of the condition number of the matrix $\mathbf{T}_L$. Using well conditioned spherical designs from [1], we calculate the condition number of the matrix $\mathbf{T}_L$ for $L = 4,\ \lambda = 1$ and the Laplace-Beltrami regularization operator with various values of $s > 0$. In Fig 4.2, it is seen that the condition number of the matrix $\mathbf{T}_4$ goes down as the degree t increases. As stated in (4.7), for $t \geq 2L = 8$ the condition number is a constant. However, with $s = 2$ and $t \geq 8$, the constant condition number is very large, even though the size of the matrix $\mathbf{T}_4$ is just 25×25. This phenomenon indicates the necessity of preconditioning the linear system (1.9).

5. Quality of approximation. In this section we study theoretically the approximation error.

In general we can write the solution of the regularized least squares problem (1.1) as

$$p_{L,N} = \mathcal{U}_L\, f \in \mathbb{P}_L, \tag{5.1}$$

where $\mathcal{U}_L := \mathcal{U}_L\,(\mathcal{X}_N, \boldsymbol{\beta})$ is a linear operator, and $\boldsymbol{\beta}$ stands for the values $\{\beta_0, \ldots, \beta_L\}$. For $t \geq 2L$ it is given explicitly by Theorem 1.1,

$$\mathcal{U}_L\, f(\mathbf{x}) = \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j) f(\mathbf{x}_j), \tag{5.2}$$

while for $L \leq t < 2L$ the construction of $\mathcal{U}_L$ involves inversion of a matrix or the solution of the linear system (1.9).

If f is replaced by a perturbed function f^δ and $p_{L,N}$ is correspondingly replaced by $p^\delta_{L,N}$, then it is clear that

$$\left|\left|p^\delta_{L,N} - p_{L,N}\right|\right|_{C(\mathbb{S}^2)} \leq ||\mathcal{U}_L\,||_{C(\mathbb{S}^2)} \left|\left|f^\delta - f\right|\right|_{C(\mathbb{S}^2)}, \tag{5.3}$$

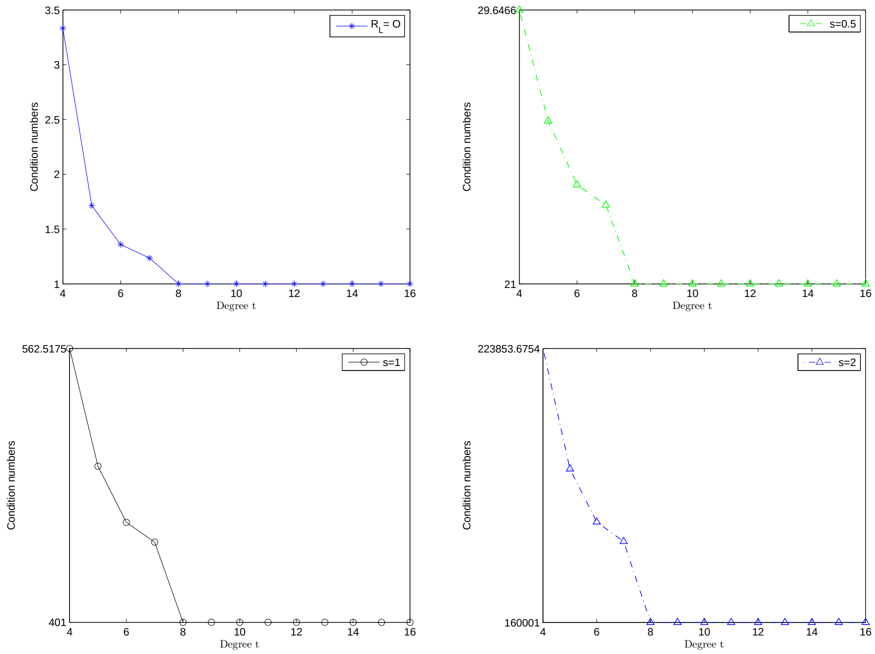


Fig. 4.2: Condition numbers of $\mathbf{T}_4$ for Laplace-Beltrami regularization operator with various values s

where $||\mathcal{U}_L||_{C(\mathbb{S}^2)}$ is the Lebesgue constant defined by replacing $\mathcal{L}_L$ by $\mathcal{U}_L$ in (2.14). Thus one use of the Lebesgue constant is to measure the sensitivity of the approximation to errors in the data. In some cases (see Subsections 5.1 and 5.2 below) the Lebesgue constant is also helpful in bounding the approximation error.

The following simple consequence of (5.2) will be useful. In Proposition 5.1 by $\boldsymbol{\beta}' \geq \boldsymbol{\beta}$ we mean $\beta'_\ell \geq \beta_\ell$ for $\ell = 0, \ldots, L$.

Proposition 5.1. *Let $\mathcal{U}_L(\mathcal{X}_N, \boldsymbol{\beta})$ be defined by (5.1), with $\mathcal{X}_N$ a spherical t-design. Assume $t \geq 2L$. Then the Lebesgue constant of $\mathcal{U}_L(\mathcal{X}_N, \boldsymbol{\beta})$ is given by*

$$||\mathcal{U}_L(\mathcal{X}_N, \boldsymbol{\beta})||_{C(\mathbb{S}^2)} = \max_{\mathbf{x}\in\mathbb{S}^2} \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} |P_\ell(\mathbf{x}\cdot\mathbf{x}_j)|. \tag{5.4}$$

If $\boldsymbol{\beta}' \geq \boldsymbol{\beta}$ then

$$||\mathcal{U}_L(\mathcal{X}_N, \boldsymbol{\beta})||_{C(\mathbb{S}^2)} \geq ||\mathcal{U}_L(\mathcal{X}_N, \boldsymbol{\beta}')||_{C(\mathbb{S}^2)}. \tag{5.5}$$

Proof. Since $t \geq 2L$, from the expression (5.2) for $\mathcal{U}_L f$, we have

$$|\mathcal{U}_L f(\mathbf{x})| \leq \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} |P_\ell(\mathbf{x}\cdot\mathbf{x}_j)|\, ||f||_{C(\mathbb{S}^2)},$$

and hence

$$||\mathcal{U}_L\, f||_{C(\mathbb{S}^2)} \le \max_{\mathbf{x}\in\mathbb{S}^2} \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} |P_\ell(\mathbf{x}\cdot\mathbf{x}_j)|\, ||f||_{C(\mathbb{S}^2)}\,. \tag{5.6}$$

Let $\mathbf{x}_0 \in \mathbb{S}^2$ achieve the maximum in (5.6), i.e.

$$\sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} |P_\ell(\mathbf{x}_0\cdot\mathbf{x}_j)| = \max_{\mathbf{x}\in\mathbb{S}^2} \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} |P_\ell(\mathbf{x}\cdot\mathbf{x}_j)|.$$

Then define $f^* \in C(\mathbb{S}^2)$, such that $||f^*||_{C(\mathbb{S}^2)} = 1$ and

$$f^*(\mathbf{x}_j) = \operatorname{sign} P_\ell(\mathbf{x}_0\cdot\mathbf{x}_j), \qquad j = 1,\dots,N.$$

By (5.2) we have

$$\mathcal{U}_L\, f^*(\mathbf{x}) = \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j)\operatorname{sign} P_\ell(\mathbf{x}_0\cdot\mathbf{x}_j),$$

and hence, setting $\mathbf{x} = \mathbf{x}_0$, we obtain

$$\mathcal{U}_L\, f^*(\mathbf{x}_0) = \sum_{\ell=0}^{L} \frac{2\ell+1}{(1+\lambda\beta_\ell^2)N} \sum_{j=1}^{N} |P_\ell(\mathbf{x}_0\cdot\mathbf{x}_j)|.$$

Thus the inequality in (5.6) becomes an equality for $f = f^*$, proving (5.4). The inequality (5.5) follows from (5.4). □

5.1. The case $\mathcal{R}_L = 0$. In this case $\beta_\ell = 0$ for all ℓ, and the approximation $p_{L,N}$ is exact if $f \in \mathbb{P}_L$; that is to say

$$\mathcal{U}_L\, p = p \qquad \text{for} \quad p \in \mathbb{P}_L.$$

Hence for $p \in \mathbb{P}_L$

$$||\mathcal{U}_L\ f - f||_{C(\mathbb{S}^2)} = ||\mathcal{U}_L\,(f-p) - (f-p)||_{C(\mathbb{S}^2)}\,,$$

and by making an appropriate choice of $p \in \mathbb{P}_L$

$$||\mathcal{U}_L\, f - f||_{C(\mathbb{S}^2)} \le (||\mathcal{U}_L\,||_{C(\mathbb{S}^2)} + 1)E_L(f), \tag{5.7}$$

where as in (2.5) $E_L(f)$ is the error of best uniform approximation of f by a polynomial of degree at most L.

For the case $t = L$ and $N = (L+1)^2$, where $\mathcal{U}_L$ is the polynomial interpolant, it seems that little is known theoretically about the Lebesgue constant, see [31], beyond a lower bound of the form $||\mathcal{U}_L\,||_{C(\mathbb{S}^2)} \ge c\sqrt{L}$, but there is convincing numerical evidence see [1], that $||\mathcal{U}_L\,||_{C(\mathbb{S}^2)} \le c_1 L$ for a sequence $\mathcal{X}_N$ of so-called well conditioned spherical L-designs (c and c_1 are some positive constants).

For $t \ge 2L$ the approximation $\mathcal{U}_L\, f$ is equivalent to hyperinterpolation $\mathcal{L}_L f$. In this case we have noted already, see (2.15), that $||\mathcal{L}_L||_{C(\mathbb{S}^2)}$ is of exact order $\sqrt{L+1}$. For intermediate values of t, that is satisfying $L < t < 2L$, it seems that nothing is known about the Lebesgue constant.

5.2. Filtered regularization operator. With h given by (2.22) and β_ℓ by (3.2) we have

$$\beta_\ell = 0 \quad \text{for} \quad 0 \le \ell \le \lfloor L/2 \rfloor.$$

From this we see that

$$\mathcal{R}_L p = 0 \quad \text{for} \quad p \in \mathbb{P}_{\lfloor L/2 \rfloor}.$$

In turn it follows from (1.1) that in this case

$$\mathcal{U}_L\, p = p \quad \text{for} \quad p \in \mathbb{P}_{\lfloor L/2 \rfloor},$$

and hence, by a similar argument to that used to prove (5.7),

$$||\mathcal{U}_L\, f - f||_{C(\mathbb{S}^2)} \le (||\mathcal{U}_L\,||_{C(\mathbb{S}^2)} + 1) E_{\lfloor L/2 \rfloor}(f). \tag{5.8}$$

For $t \ge 2L$ we know already (see (2.18)) that

$$||\mathcal{U}_L\,||_{C(\mathbb{S}^2)} = ||\mathcal{F}_L||_{C(\mathbb{S}^2)} \le c,$$

in which case both stability and convergence are assured. For $L \le t < 2L$ it seems that nothing is known about the Lebesgue constant.

5.3. Laplace-Beltrami regularization operator. If $\mathcal{R}_L = (-\Delta^*)^s$ with $s > 0$, or correspondingly $\beta_\ell = (\ell(\ell+1))^s$, by Proposition 5.1 the Lebesgue constant for $t \ge 2L$ is bounded by

$$||\mathcal{U}_L\,||_{C(\mathbb{S}^2)} \le \sum_{\ell=0}^{L} \frac{2\ell+1}{1+\lambda(\ell(\ell+1))^{2s}} < \sum_{\ell=0}^{\infty} \frac{2\ell+1}{1+\lambda(\ell(\ell+1))^{2s}}, \tag{5.9}$$

which is finite for $s > 1/2$. Thus for $t \ge 2L$ the Lebesgue constant is bounded independently of L when $s > 1/2$, with a bound that decreases monotonically with increasing s.

Note that for the Laplace-Beltrami regularization operator a knowledge of the Lebesgue constant does not give any useful information about the error, because the approximation in this case does not reproduce polynomials other than the constants.

The following theorem asserts that for $t \ge 2L$ and $L \to \infty$ the approximation with the Laplace-Beltrami regularization operator $(-\Delta^*)^s$ with $s > 1/2$ converges uniformly, not to f but rather to the "s-smoothed" solution f_s,

$$\begin{aligned} f_s(\mathbf{x}) :&= \sum_{\ell=0}^{\infty} \frac{1}{1+\lambda(\ell(\ell+1))^{2s}} \sum_{k=1}^{2\ell+1} \widehat{f}_{\ell,k} Y_{\ell,k}(\mathbf{x}) \\ &= \sum_{\ell=0}^{\infty} \frac{2\ell+1}{1+\lambda(\ell(\ell+1))^{2s}} \frac{1}{4\pi} \int_{\mathbb{S}^2} P_\ell(\mathbf{x}\cdot\mathbf{y}) f(\mathbf{y}) d\omega(\mathbf{y}), \end{aligned}$$

where the last equality uses (2.2) and the addition therorem (1.6).

Theorem 5.2. *Assume that the regularization operator is $\mathcal{R}_L = (-\Delta^*)^s$ with $s > 1/2$. Assume $t = t(L) \ge 2L$ as $L \to \infty$. Then with $p_{L,N} = p_{L,N(t)}$ as in Theorem 1.2, we have*

$$\lim_{L\to\infty} ||p_{L,N} - f_s||_{C(\mathbb{S}^2)} = 0.$$

Proof. From (5.2) we have

$$p_{L,N}(\mathbf{x}) = \mathcal{U}_L\, f(\mathbf{x}) = \sum_{\ell=0}^{L} \frac{2\ell+1}{1+\lambda(\ell(\ell+1))^{2s}} \frac{1}{N} \sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j) f(\mathbf{x}_j).$$

For fixed ℓ we have (since $L \to \infty$ implies $t \to \infty$, and since $P_\ell(\mathbf{x}\cdot\mathbf{y})f(\mathbf{y})$ is continuous in $\mathbf{y}$)

$$\frac{1}{N}\sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j) f(\mathbf{x}_j) \to \frac{1}{4\pi}\int_{\mathbb{S}^2} P_\ell(\mathbf{x}\cdot\mathbf{y}) f(\mathbf{y}) d\omega(\mathbf{y}).$$

Moreover,

$$\left| \frac{1}{N}\sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j) f(\mathbf{x}_j) \right| \le \frac{1}{N}\sum_{j=1}^{N} |f(\mathbf{x}_j)| \le ||f||_{C(\mathbb{S}^2)}\,,$$

and hence

$$\left| \sum_{\ell=0}^{L} \frac{2\ell+1}{1+\lambda(\ell(\ell+1))^{2s}} \frac{1}{N} \sum_{j=1}^{N} P_\ell(\mathbf{x}\cdot\mathbf{x}_j) f(\mathbf{x}_j) \right| \le \sum_{\ell=0}^{\infty} \frac{2\ell+1}{1+\lambda(\ell(\ell+1))^{2s}} ||f||_{C(\mathbb{S}^2)}\,,$$

which is finite because $s > 1/2$. The desired result is now an immediate consequence of Tannery's theorem [8, page 207]. □

5.4. Residuals for Laplace-Beltrami regularization operator. In this subsection we show that for $t \ge 2L$ the residual $A(\boldsymbol{\alpha}) := \sum_{j=1}^{N}(p_{L,N}(\mathbf{x}_j) - f(\mathbf{x}_j))^2$ will increase as the order s of the Laplace-Beltrami regularization operator increases.

Let $s > 0$, and let

$$\begin{aligned} \rho(s,\boldsymbol{\alpha}) :&= \|\mathbf{Y}_L^T\boldsymbol{\alpha} - \mathbf{f}\|_2^2 + \lambda\|\mathbf{R}_L^{(s)T}\boldsymbol{\alpha}\|_2^2 \\ &= A(\boldsymbol{\alpha}) + E(s,\boldsymbol{\alpha}), \end{aligned} \tag{5.10}$$

where $E(s,\boldsymbol{\alpha}) = \lambda\|\mathbf{R}_L^{(s)T}\boldsymbol{\alpha}\|_2^2$, and $\mathbf{R}_L^{(s)}$ is the Laplace-Beltrami regularization operator of order s, i.e.

$$\mathbf{R}_L^{(s)} = \mathbf{B}_L^{(s)}\mathbf{Y}_L,$$

with $(\mathbf{B}_L^{(s)})_{\ell,k,\ell',k'} = \delta_{\ell\ell'}\delta_{kk'}(\ell(\ell+1))^s$ for $\ell,\ell' = 0,\ldots,L,\quad k = 1,\ldots,2\ell+1,\quad k' = 1,\ldots,2\ell'+1$. For a given s, let $\boldsymbol{\alpha}^* = \boldsymbol{\alpha}^*(s)$ be the minimizer of $\rho(s,\boldsymbol{\alpha})$, i.e.,

$$\rho(s,\boldsymbol{\alpha}^*) \le \rho(s,\boldsymbol{\alpha}) \qquad \forall\, \boldsymbol{\alpha} \in \mathbb{R}^{(L+1)^2}. \tag{5.11}$$

In [9] it is shown that $A(\boldsymbol{\alpha}^*)$ is monotonic increasing with respect to increasing λ. In this section we use a similar argument to show (for $t \ge 2L$ only) that $A(\boldsymbol{\alpha}^*)$ is similarly monotonic increasing with respect to increasing order s.

Theorem 5.3. *Let $\mathcal{X}_N$ be a spherical t-design on $\mathbb{S}^2$, and for $s > 0$ let $\boldsymbol{\alpha}^*(s)$ and $A(\boldsymbol{\alpha})$ be defined as in* (5.10) *and* (5.11)*. Assume $t \ge 2L$. Then $A(\boldsymbol{\alpha}^*(s))$ is strictly increasing in s.*

Proof. Let s, $\widetilde{s}$ be given, with $0 < s < \widetilde{s}$. Temporarily we write

$$\boldsymbol{\alpha}^*(s) = \boldsymbol{\alpha}^*, \qquad \boldsymbol{\alpha}^*(\widetilde{s}) = \widetilde{\boldsymbol{\alpha}}^*.$$

Then the minimization property (5.11) for s gives

$$A(\boldsymbol{\alpha}^*) + E(s, \boldsymbol{\alpha}^*) \leq A(\widetilde{\boldsymbol{\alpha}}^*) + E(s, \widetilde{\boldsymbol{\alpha}}^*). \tag{5.12}$$

From (5.12) we have

$$A(\boldsymbol{\alpha}^*) - A(\widetilde{\boldsymbol{\alpha}}^*) \leq E(s, \widetilde{\boldsymbol{\alpha}}^*) - E(s, \boldsymbol{\alpha}^*). \tag{5.13}$$

Thus to show the desired result it is sufficient to show

$$E(s, \widetilde{\boldsymbol{\alpha}}^*) - E(s, \boldsymbol{\alpha}^*) < 0. \tag{5.14}$$

Now we use the known form of $E(s, \boldsymbol{\alpha})$,

$$E(s, \boldsymbol{\alpha}) = \lambda \boldsymbol{\alpha}^T \mathbf{B}_L^{(s)} \mathbf{Y}_L \mathbf{Y}_L^T \mathbf{B}_L^{(s)} \boldsymbol{\alpha}.$$

On specializing to $t \geq 2L$ we obtain using Theorem 1.1

$$E(s, \boldsymbol{\alpha}) = \frac{N}{4\pi} \lambda \boldsymbol{\alpha}^T \mathbf{B}_L^{(s)^2} \boldsymbol{\alpha},$$

and hence

$$\begin{aligned} E(s, \widetilde{\boldsymbol{\alpha}}^*) - E(s, \boldsymbol{\alpha}^*) &= \frac{N}{4\pi} \lambda \left(\widetilde{\boldsymbol{\alpha}}^{*T} \mathbf{B}_L^{(s)^2} \widetilde{\boldsymbol{\alpha}}^* - \boldsymbol{\alpha}^{*T} \mathbf{B}_L^{(s)^2} \boldsymbol{\alpha}^* \right) \\ &= \frac{N}{4\pi} \lambda \sum_{\ell=1}^{L} (\ell(\ell+1))^{2s} \sum_{k=1}^{2\ell+1} (\widetilde{\boldsymbol{\alpha}}_{\ell,k}^{*2} - \boldsymbol{\alpha}_{\ell,k}^{*2}). \end{aligned}$$

Now from (1.14) we have

$$\begin{aligned} \boldsymbol{\alpha}_{\ell,k}^* &= \frac{4\pi}{N(1 + \lambda(\ell(\ell+1))^{2s})} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) f(\mathbf{x}_j), \\ \widetilde{\boldsymbol{\alpha}}_{\ell,k}^* &= \frac{4\pi}{N(1 + \lambda(\ell(\ell+1))^{2\widetilde{s}})} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) f(\mathbf{x}_j). \end{aligned}$$

We observe that $|\widetilde{\boldsymbol{\alpha}}_{\ell,k}^*| < |\boldsymbol{\alpha}_{\ell,k}^*|$, from which it follows that

$$E(s, \widetilde{\boldsymbol{\alpha}}^*) - E(s, \boldsymbol{\alpha}^*) < 0,$$

so that the sufficient condition (5.14) is satisfied, completing the proof. ∎

6. Numerical results. In this section we present numerical results to illustrate the theoretical results derived in the previous sections, and show that well chosen regularization operator and well conditioned spherical t-designs can provide good polynomial approximation on the sphere for both exact data and contaminated data.

We choose two test functions for our numerical experiments. The first function is the Franke function as modified by Renka [19],

$$\begin{aligned}f_1(x,y,z) = {} & 0.75\exp(-(9x-2)^2/4-(9y-2)^2/4-(9z-2)^2/4)\\ & +0.75\exp(-(9x+1)^2/49-(9y+1)/10-(9z+1)/10)\\ & +0.5\exp(-(9x-7)^2/4-(9y-3)^2/4-(9z-5)^2/4)\\ & -0.2\exp(-(9x-4)^2-(9y-7)^2-(9z-5)^2), \quad (x,y,z)\in\mathbb{S}^2,\end{aligned}$$

which is in $C^\infty(\mathbb{S}^2)$. The second function is the sum of the Franke function f_1 and a function f_{cap} [29] with support on a spherical cap $C(\mathbf{x}_c,\, r)$, see (2.13), so

$$f_2 = f_1 + f_{\text{cap}}, \tag{6.1}$$

where

$$f_{\text{cap}}(\mathbf{x}) = \begin{cases} \rho\cos\left(\dfrac{\pi\arccos(\mathbf{x}_c\cdot\mathbf{x})}{2r}\right), & \mathbf{x}\in C(\mathbf{x}_c,\, r),\\ 0, & \text{otherwise},\end{cases}$$

and ρ is a positive number. This function is continuous on $\mathbb{S}^2$ but not differentiable on the boundary of the spherical cap $C(\mathbf{x}_c,\, r)$. In our numerical results $\mathbf{x}_c = (-\frac{1}{2},-\frac{1}{2},\frac{1}{\sqrt{2}})^T$, $\rho = 2$ and $r = \frac{1}{2}$, which is illustrated in Fig. 6.4 (a).

We use well conditioned spherical t-designs with $\mathcal{X}_N$, $t = 1,\ldots,60$, [1] and $N = (t+1)^2$.

For $t \geq 2L$, by Theorem 1.1 $\mathbf{T}_L$ is a diagonal matrix, and the solution of the system of linear equations (1.9) has the explicit form

$$\boldsymbol{\alpha} = \frac{4\pi}{N}\left[\mathbf{I}_{(L+1)^2} + \lambda\mathbf{B}_L^2\right]^{-1}\mathbf{Y}_L\mathbf{f}(\mathcal{X}_N), \qquad t\geq 2L.$$

For $L \leq t < 2L$, the coefficient matrix $\mathbf{T}_L$ is a symmetric positive definite $(L+1)^2$ by $(L+1)^2$ matrix. However it is not sparse. For $1 \leq L \leq 60$ (so the largest dimension is $61^2 = 3721$), the linear system can be efficiently solved using the Cholesky factorization [4, page 44]. Given $\boldsymbol{\alpha}$, the approximating polynomial has the form (1.4).

The uniform error of the approximation is estimated by

$$\|f - p_{L,N}\|_{C(\mathbb{S}^2)} \approx \max_{\mathbf{x}_i\in\mathcal{X}} |f(\mathbf{x}_i) - p_{L,N}(\mathbf{x}_i)|, \tag{6.2}$$

where $\mathcal{X}$ is a finite but large set of well distributed points over the sphere. In particular, for approximations to f_1, $\mathcal{X}$ is chosen as a set of 10^6 generalized spiral points [18, 3]. For estimating the approximation error for f_2, $\mathcal{X}$ is the union of the generalized spiral points and 1200 points around the boundary of the cap.

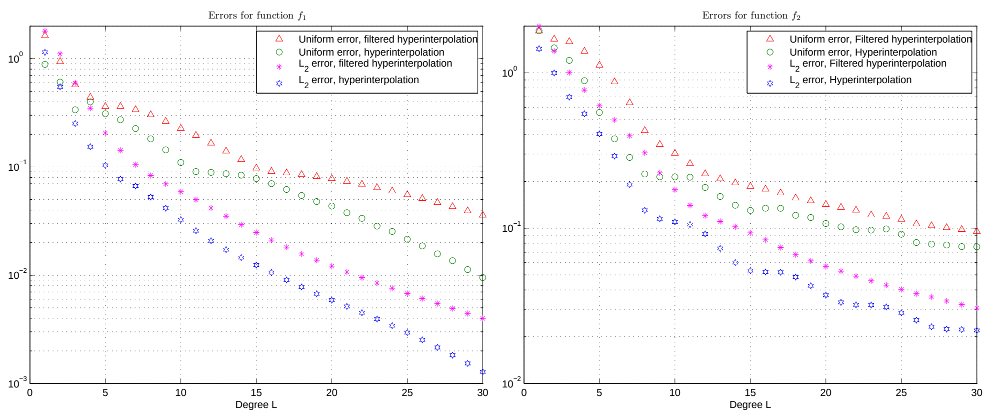
The L_2-norm of the approximation error is estimated by

$$\begin{aligned}||f - p_{L,N}||_{L_2} &:= \left(\int_{\mathbb{S}^2} |f - p_{L,N}(\mathbf{x})|^2 d\omega(\mathbf{x})\right)^{1/2}\\ &\approx \left(\frac{4\pi}{m}\sum_{j=1}^{m} |f - p_{L,N}(\mathbf{x}_j)|^2\right)^{1/2}.\end{aligned} \tag{6.3}$$

The set $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ can be the nodes of the spherical 100-design obtained in [6] (so $m = 101^2$) or generalized spiral points with $m = 10^6$ (which are approximate spherical designs [3]).

In the following, we consider two cases: filtered and zero regularization operator for exact data, and Laplace-Beltrami regularization operator for contaminated data.

6.1. Filtered and zero regularization operator for exact data. In this subsection we report numerical results to compare the filtered regularization operator with the hyperinterpolation. For a given L, we consider $L \le t \le 2L$ and set $N = (t+1)^2$. By Theorem 1.1, both the filtered regularization operator and the hyperinterpolation approximations have closed forms (1.15) with β_ℓ given by (3.2) and $\lambda = 1$ and $\lambda = 0$, respectively.



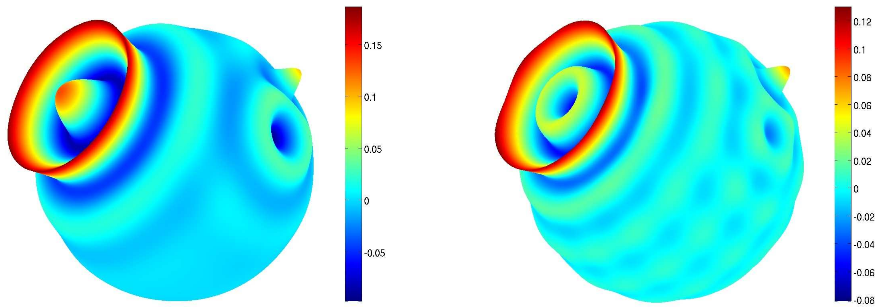
(a) Errors for f_1 (b) Errors for f_2

Fig. 6.1: Uniform and L_2 errors for hyperinterpolation and filtered hyperinterpolation with $t = 2L$, $N = (t+1)^2$ and $L = 1, \ldots, 30$

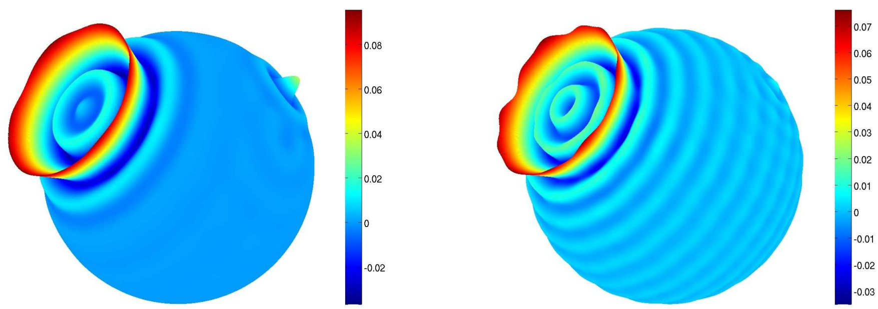
Fig. 6.1 (a) and (b) report the uniform error and the L_2 error of the approximations for the Franke functions f_1 and f_2 with $t = 2L$. Fig. 6.1 shows that the hyperinterpolation approximation has smaller uniform errors and L_2 errors than filtered approximation at every L. This reflects the error bounds (5.7) and (5.8), which show that the error of hyperinterpolation approximation is bounded by $cE_L(f)$, while the error of filtered approximation is bounded by $cE_{\lfloor L/2 \rfloor}(f)$, where c is a positive constant. Note that from definition (2.5), $E_L(f) \le E_{\lfloor L/2 \rfloor}(f)$.

Fig. 6.2 shows the errors $p_{L,N} - f_2$ for $L = 15, 30$, $t = 2L$ and $N = (t+1)^2$, for the filtered hyperinterpolation ($\lambda = 1$) in Fig. 6.2 (a) and (c) and for hyperinterpolation ($\lambda = 0$) in Fig. 6.2 (b) and (c). This clearly shows that for larger L, the uniform error is attained at a point around the boundary of the spherical cap, due to the non-differentiability of the function f_2 at the boundary. The uniform error for filtered hyperinterpolation is slightly larger, but more localized.

Fig. 6.3 shows the errors for both test functions f_1 and f_2 when solving a least squares problem with $t < 2L$, so the coefficients are given by the linear system (1.9). It is notable that in Fig. 6.3 (a) the errors change very little as t varies from L to



(a) $\mathcal{F}_L f_2 - f_2, L = 15$ (b) $\mathcal{L}_L f_2 - f_2, L = 15$



(c) $\mathcal{F}_L f_2 - f_2, L = 30$ (d) $\mathcal{L}_L f_2 - f_2, L = 30$

Fig. 6.2: Filtered and hyperinterpolation errors for $L = 15, 30$, $t = 2L$, $N = (2L+1)^2$

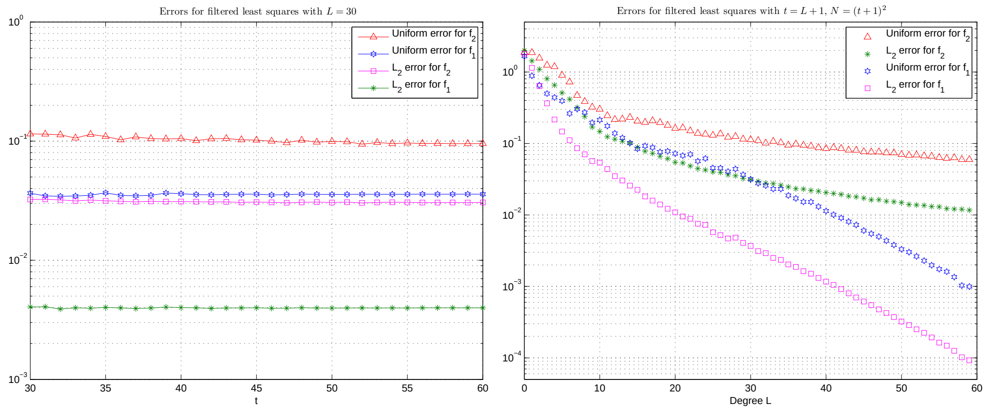
$2L$. As $N = (t+1)^2$ sample points are used, this means that fewer sample points are required, without significant loss of accuracy, if we are prepared to solve a linear system. Fig. 6.3 (b) shows the errors for f_1 and f_2 when $t = L+1$, $N = (t+1)^2$ and L varies from 1 to 60. Solving the linear system for the least squares problem allows us to use $t = L+1$ and $N = (t+1)^2$ sample points, and hence increase the degree of the approximating polynomial. As discussed in Section 4 the condition number of the linear system improves as t increases from L to $2L$.

6.2. Laplace-Beltrami regularization operator for contaminated data. In this subsection we report numerical results for reconstructing the nonsmooth function f_2 when the data has been contaminated with a high level of noise. We use Laplace-Beltrami regularization operator with $s = 1$ and different values of λ, so $\boldsymbol{\alpha}$ is given by the solution of (1.9) with $\mathbf{B}_L$ defined in (1.8).

Fig. 6.4 (a) illustrates the function f_2, while Fig. 6.4 (b) shows the contaminated function

$$f_2^\delta(\mathbf{x}) = f_2(\mathbf{x}) + \delta(\mathbf{x}),$$

where for each $\mathbf{x}$, $\delta(\mathbf{x})$ is a sample of a normal random variable with mean 0 and standard deviation $\sigma = 0.5$. Fig. 6.4 uses $N = 3721$ with $\mathbf{x}_i, i = 1, \ldots, N$ the



(a) Errors for $L = 30$, $L \leq t \leq 2L$

(b) Errors for $t = L + 1$, $1 \leq L \leq 60$

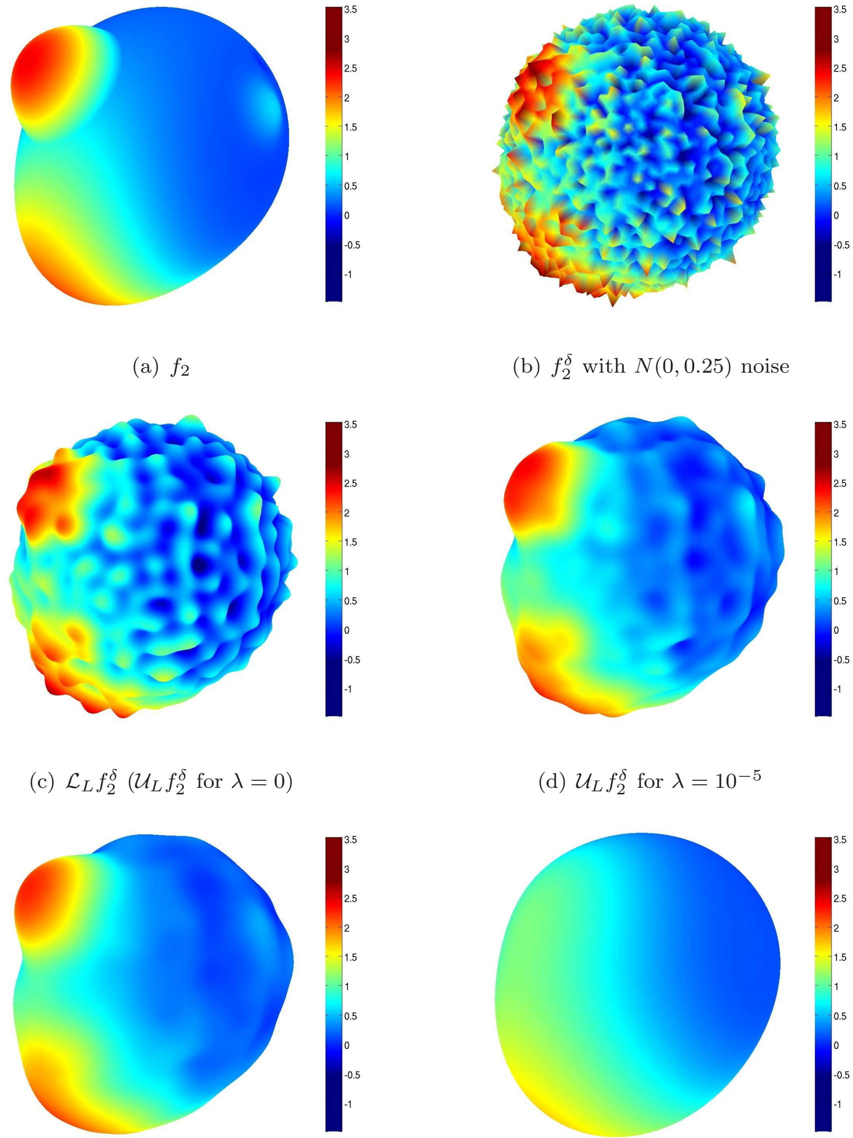
Fig. 6.3: Uniform and L_2 errors for filtered least squares

nodes of a well-conditioned spherical t-design with $t = 60$. The subplots (c) to (f) show the approximation for different values of λ when using the Laplace-Beltrami regularization operator with $L = 30$, $t = 2L$ and $N = (t + 1)^2 = 3721$. Fig. 6.4 (c) shows the approximation without using a regularization operator ($\lambda = 0$), so this is the hyperinterpolation approximation.

Fig. 6.4 shows that the least squares approximation with Laplace-Beltrami regularization operator is effective in recovering the underlying function from highly contaminated data. However, the choice of the regularization parameter λ is critical. Choosing too large a value of λ (for example $\lambda = 10^{-2}$ as in Fig 6.4 (f)), forces the approximation to be a low order polynomial, almost completely missing features such as the cap. How to automatically choose a good value of λ is a challenging problem, which we do not address here.

Fig. 6.5 reports the uniform and L_2 errors for recovering the function f_2 from contaminated data with various choices of regularization parameter λ and different strategies for choosing t in relation to L. Fig. 6.5 (a) shows the effect of varying t, $L \leq t \leq 2L$, where $L = 30$ and t is the degree of the spherical t-design where the (noisy) function values are evaluated. Apart from the least squares approximation with no regularization ($\lambda = 0$), varying t does not have a large influence on the quality of the approximation. This implies that it is possible to use the least squares approximation with $t < 2L$ without significant loss of accuracy. Fig. 6.5 (b) shows the effect of varying L while keeping $t = L + 1$ and solving the least squares problem. The choice $t = L+1$ uses $(L+2)^2$ function values in contrast to $t = 2L$ which requires $(2L + 1)^2$ function values.

7. Final Remarks. In this paper, we propose regularized discrete least squares methods using spherical t-designs with rotationally invariant regularization operators for continuous functions over $\mathbb{S}^2$. In particular, we investigate three kinds of regularization operator: zero regularization operator, filtered regularization operator, and Laplace-Beltrami regularization operator. For $t \geq 2L$, we give an explicit form of the unique solution of the regularized discrete least squares problem. For $L \leq t < 2L$,



(a) f_2

(b) f_2^δ with $N(0, 0.25)$ noise

(c) $\mathcal{L}_L f_2^\delta$ ($\mathcal{U}_L f_2^\delta$ for $\lambda = 0$)

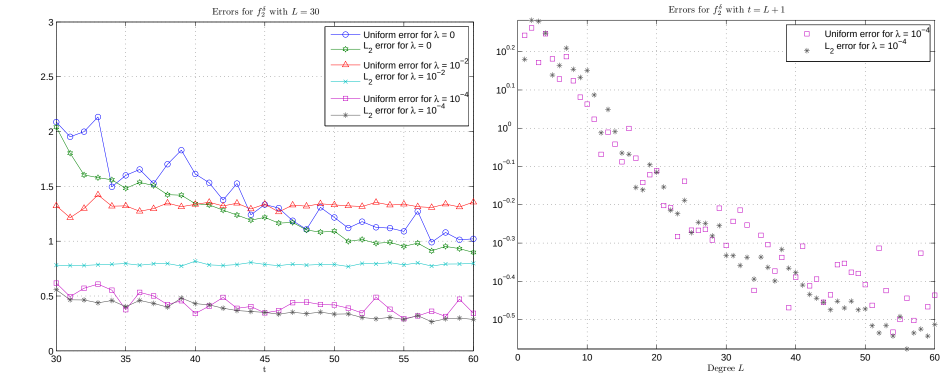
(d) $\mathcal{U}_L f_2^\delta$ for $\lambda = 10^{-5}$

(e) $\mathcal{U}_L f_2^\delta$ for $\lambda = 10^{-4}$

(f) $\mathcal{U}_L f_2^\delta$ for $\lambda = 10^{-2}$

Fig. 6.4: Laplace-Beltrami regularization operator with $L = 30, t = 2L, N = (L+1)^2$ to recover f_2 from contaminated data

we need to solve a system of linear equations to obtain the approximation polynomial. The condition number for this system is dependent on the choice of the set of points $\mathcal{X}_N$, the regularization operator $\mathbf{B}_L$ and the values of λ, t, L, N. We give an upper bound for the condition number. We prove that the solution of the regularized discrete least squares problem converges to the regularized continuous least squares



(a) $L = 30,\ L \leq t \leq 2L$ (b) $t = L + 1,\ 1 \leq L \leq 60$

Fig. 6.5: Errors with Laplace-Beltrami regularization operator for least squares approximation

problem as $t \to \infty$. Moreover, we derive error bounds in the uniform norm and the L_2 norm for the discrete approximation.

Numerical results show that the proposed regularized least squares approximations are effective not only for approximating smooth functions in $C^\infty(\mathbb{S}^2)$ but also for functions in $C^0(\mathbb{S}^2)$. Moreover, the method with a Laplace-Beltrami regularization Operator can be effective in recovering the original nonsmooth function from contaminated data.

Acknowledgements. The support of the Research Grants Council of Hong Kong and the Australian Research Council is gratefully acknowledged. The authors acknowledge the support of the Hong Kong Polytechnic University.

REFERENCES

[1] C. An, X, Chen, I. H. Sloan and R. S. Womersley, *Well conditioned spherical designs for integration and interpolation on the two-sphere*, SIAM J. Numer. Anal., 48 (2010), pp. 2135-2157.

[2] E. Bannai and E. Bannai, *A survey on spherical designs and algebraic combinatorics on spheres*, European J. Combin., 30 (2009), pp. 1392–1425.

[3] R. Bauer, *Distribution of points on a sphere with application to star catalogs*, J. Guidance, Control, and Dynamics, 23 (2000), pp. 130–137.

[4] A. Björck, *Numerical methods for least squares problems*, SIAM Press, Philadelphia, 1996.

[5] X. Chen and R. S. Womersley, *Existence of solutions to systems of undetermined equations and spherical designs*, SIAM J. Numer. Anal., 44 (2006), pp. 2326–2341.

[6] X. Chen, A. Frommer and B. Lang, *Computational existence proof for spherical t-designs*, Numer. Math., 117 (2011), pp. 289-205

[7] X. Chen, R. S. Womersley and J. Ye, *Minimizing the condition number of a Gram matrix* SIAM J. Optim. 21 (2011), pp. 127-148.

[8] E. T. Copson, *An Introduction to the Theory of Function of a Complex Variable*, Oxford Press, 1972.

[9] M. von Golitschek and L. L. Schumaker, *Data fitting by penalized least squares.* Algorithms for approximation, II, pp. 210-227, Chapman and Hall, London, 1990.

[10] P. Delsarte, J. M. Goethals and J. J. Seidel, *Spherical codes and designs*, Geom. Dedicata, 6 (1977), pp. 363–388.

[11] T. H. Gronwall, *On the degree of convergence of Laplace's series*, Trans. Amer. Math. Soc., 15 (1914), pp. 1–30.

[12] Q. T. Le Gia and I. H. Sloan *The uniform norm of hyperinterpolation on the unit sphere in an arbitrary number of dimensions*, Constr. Approx., 17 (2001), pp. 249-265.

[13] K. Hesse, I. H. Sloan and R. S. Womersley, *Numerical integration on the sphere, Handbook of Geomathematics*, edited by W. Freeden, Z. Nashed and T. Sonar, Springer-Verlag, Berlin, 2010.

[14] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge press, 1985.

[15] P. P. Korovkin, *Linear Operators and Approximation Theory*, Fitmatgiz, Moscow, 1959. English edition: Hindustan Publishing Co., Delhi, 1960.

[16] C. Müller, *Spherical Harmonics*, vol. 17 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, New-York, 1966.

[17] D. L. Ragozin, *Constructive polynomial approximation on spheres and projective spaces.* Trans. Amer. Math. Soc., 162 (1971), pp. 157-170.

[18] E. A. Rakhmanov, E. B. Saff and Y. M. Zhou, *Minimal discrete energy on the sphere*, Math. Res. Letters, 1 (1994) pp 647–882.

[19] R. J. Renka, *Multivariate interpolation of large sets of scattered data*, ACM Trans. Math. Software, 14 (1988), pp. 139–148.

[20] M. Reimer, *Hyperinterpolation on the sphere at the minimum projection order*, J. Approx. Theory, 104 (2000), pp. 272-286.

[21] M. Reimer, *Generalized hyperinterpolation on the sphere and the Newman-Shapiro operators*, Constr. Approx., 18 (2002), pp. 183-203.

[22] M. Reimer, *Multivariate Polynomial Approximation*, Birkhäuser Verlag, Berlin, 2003.

[23] I. H. Sloan, *Polynomial interpolation and hyperinterpolation over general regions*, J. Approx. Theory, 83 (1995), pp. 238–254.

[24] I. H. Sloan, *Polynomial approximation on sphere–generalizing de la Vallée-Pousin*, Manuscript, 2011.

[25] I. H. Sloan and R. S. Womersley, *Constructive polynomial approximation on the sphere*, J. Approx. Theory, 103 (2000), pp. 91-98.

[26] I. H. Sloan and R. S. Womersley, *Filtered hyperinterpolation: a constructive polynomial approximation on the sphere*, manuscript, 2011.

[27] I. H. Sloan and R. S. Womersley, *A variational characterisation of spherical designs*, J. Approx. Theory, 159 (2009), pp. 308–318.

[28] G. Szegö, *Orthogonal Polynomials*, In: American Mathematical Society Colloquium Publications, 4th ed., Volume 23, American Mathematical Society, Providence, Rhode Island, 1975.

[29] D. L. Williamson, J. B. Brake, J. J. Hack, R. Jakob and P. N. Swarztrauber, *A standard test set for numerical approximations to the shallow water equations in spherical geometry*, J. Comput. Phys., 102 (1992), pp. 211-224.

[30] R. S. Womersley, *Interpolation and Cubature on the Sphere,* http://web.maths.unsw.edu.au/ rsw/Sphere.

[31] R. S. Womersley and I. H. Sloan, *How good can polynomial interpolation on the sphere be?*, Adv. Comput. Math., 14 (2001), pp. 195–226.