The Journal of Fourier Analysis and Applications

Efficient reconstruction of functions on the sphere from scattered data

Jens Keiner, Stefan Kunis, and Daniel Potts

Communicated by Matthias Holschneider

ABSTRACT. Recently, fast and reliable algorithms for the evaluation of spherical harmonic expansions have been developed. The corresponding sampling problem is the computation of Fourier coefficients of a function from sampled values at scattered nodes. We consider a least squares approximation and an interpolation of the given data. Our main result is that the rate of convergence of the two proposed iterative schemes depends only on the mesh norm and the separation distance of the nodes. In conjunction with the nonequispaced FFT on the sphere, the reconstruction of N^2 Fourier coefficients from M reasonably distributed samples is shown to take $\mathcal{O}(N^2 \log^2 N + M)$ floating point operations. Numerical results support our theoretical findings.

1. Introduction

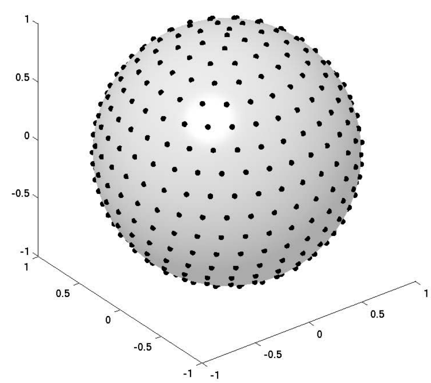
Motivated by the fact that most data collected over the surface of the earth is available at scattered nodes only, the least squares approximation and interpolation of such data has attracted much attention, see e.g. [4, 8, 2]. The most prominent approaches rely on so-called zonal basis function methods [28] or on finite expansions into spherical harmonics [19, 23]. We focus on the latter, i.e., the use of spherical polynomials since these allow for the fast spherical Fourier transforms, see for example [15, 13] and the references therein.

If we consider the problem of reconstructing a spherical polynomial of degree $N \in \mathbb{N}_0$ from sample values, one might set up a linear system of

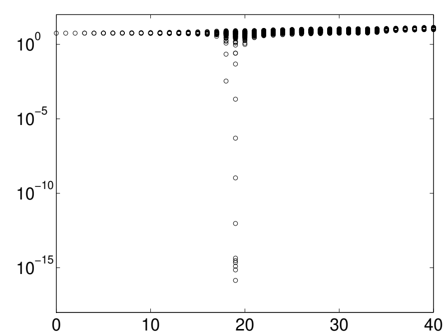
Math Subject Classifications. 65T50, 33C55, 65F10, 65T40.
Keywords and Phrases. approximation by spherical harmonics, scattered data interpolation, iterative methods, nonequispaced FFT on the sphere.

ISSN 1069-5869

equations with $M = (N+1)^2$ interpolation constraints which has to be solved for the unknown vector of Fourier coefficients $\hat{\boldsymbol{f}} \in \mathbb{C}^{(N+1)^2}$. If the nodes for interpolation are chosen such that the interpolation problem has always a unique solution, the sampling set is called a fundamental system. As can be seen in Figure 1(b), also, in a geometric sense, well distributed nodes on the sphere can lead to an ill conditioned square spherical Fourier matrix. Hence, we relax the condition that the number of equations M coincides with the number of unknowns $(N+1)^2$. Considering the overdetermined case $M > (N+1)^2$ or the underdetermined case $M < (N+1)^2$ leads to far better distributed singular values of the system matrix as seen in Figure 1(b).



(a) Generalised spiral nodes [25].



(b) Singular value distribution.

Figure 1 Distribution of the singular values of the spherical Fourier matrix $\boldsymbol{Y} \in \mathbb{C}^{M\times(N+1)^2}$ with respect to the polynomial degrees $N = 0, \dots, 40$ for $M = 400$ generalised spiral nodes.

Our main result is that for given sampling nodes the polynomial degree N can either be chosen small enough with respect to the inverse mesh norm or large enough with respect to the inverse separation distance of the sampling set to ensure a well conditioned spherical Fourier matrix. In both cases, the derived conditions are optimal within a reasonable constant.

In the first part, we consider the overdetermined case $M > (N+1)^2$, that is the least squares approximation by spherical polynomials to given data. The approximation by (univariate) trigonometric polynomials has been proven to be stable if the polynomial degree is less than the inverse of the mesh norm of the sampling set in [10]. Subsequently, Feichtinger, Gröchenig and Strohmer [5] developed the celebrated adaptive weights, conjugate gradient, Toeplitz method (ACT) for the fast iterative solution of the least squares problem. Moreover, so-called spherical L^p-Marcinkiewicz-Zygmund inequalities due to Mhaskar, Narcowich, Ward [18] and Filbir, Themistoclakis [7] yield stable least squares approximation for a quasi-uniform subset of a dense sampling set. We generalise the idea of adaptive weights to the sphere and prove the interesting L^2-Marcinkiewicz-Zygmund inequality for dense sampling sets in Theorem 1. The corresponding eigenvalue estimate and thus the rate of convergence of our fast iterative solver is given in Corollary 1.

On the other hand, we focus on the underdetermined case $M < (N+1)^2$, i.e., the interpolation of given data. Here, the (univariate) trigonometric interpolation has been considered in [16]. Applying the smoothness-decay principle for trigonometric polynomials from [16], we construct strongly localised zonal polynomials with strictly positive Fourier-Legendre coefficients in Lemma 7. In conjunction with a refined version of the packing argument from [20] in Lemma 5, we prove stable interpolation for well separated sampling sets in Theorem 2. The corresponding eigenvalue estimate and thus the rate of convergence of our fast iterative solver is given in Corollary 2.

The outline of this paper is as follows: Section 2 starts by introducing the necessary notation including the spaces of spherical harmonics. We consider the least squares approximation in Section 3. Section 4 discusses the optimal interpolation problem. Before giving our conclusions, Section 5 presents some numerical experiments.

2. Prerequisites

Let $\mathbb{S}^2 := \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\|_2 = 1\}$ be the two-dimensional unit sphere embedded in $\mathbb{R}^3$. A point $\boldsymbol{\xi} \in \mathbb{S}^2$ is identified in spherical coordinates by $\boldsymbol{\xi} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)^\top$, where the angles $(\vartheta, \varphi) \in [0,\pi] \times [-\pi,\pi)$ are the co-latitude and longitude of that point. The *geodesic distance* of $\boldsymbol{\xi}$, $\boldsymbol{\eta} \in \mathbb{S}^2$ is given by

$$\operatorname{dist}(\boldsymbol{\xi}, \boldsymbol{\eta}) := \arccos(\boldsymbol{\eta} \cdot \boldsymbol{\xi}).$$

Obviously, for $\boldsymbol{\eta}$, $\boldsymbol{\xi} \in \mathbb{S}^2$ the Euclidean distance fulfils $\|\boldsymbol{\eta} - \boldsymbol{\xi}\|_2^2 = 2 - 2\boldsymbol{\eta} \cdot \boldsymbol{\xi}$ and a node $\boldsymbol{\xi} \sim (\vartheta, \varphi)$ has geodesic distance $\arccos(\boldsymbol{e}_3 \cdot \boldsymbol{\xi}) = \vartheta$ to the north pole $\boldsymbol{e}_3 = (0,0,1)^\top$.

We measure the "nonuniformity" of a *sampling set* $\mathcal{X} := \{\boldsymbol{\xi}_j \in \mathbb{S}^2 : j = 0, \ldots, M-1\}$, $M \in \mathbb{N}$, by the mesh norm $\delta_{\mathcal{X}}$ and the separation distance $q_{\mathcal{X}}$, defined by

$$\begin{aligned}
\delta_{\mathcal{X}} &:= 2 \max_{\boldsymbol{\xi} \in \mathbb{S}^2} \min_{j=0,\ldots,M-1} \operatorname{dist}(\boldsymbol{\xi}_j, \boldsymbol{\xi}), \\
q_{\mathcal{X}} &:= \min_{0 \le j < l < M} \operatorname{dist}(\boldsymbol{\xi}_j, \boldsymbol{\xi}_l).
\end{aligned}$$

The sampling set $\mathcal{X}$ is called

- δ-dense for some $0 < \delta \le 2\pi$, if $\delta_{\mathcal{X}} \le \delta$, and
- q-separated for some $0 < q \le 2\pi$, if $q_{\mathcal{X}} \ge q$.

A sequence $(\mathcal{X}_M)_{M\in\mathbb{N}}$ with $|\mathcal{X}_M| = M$ is called quasi uniform, if $\delta_{\mathcal{X}_M} \le Cq_{\mathcal{X}_M}$ with a constant C independent of M.

In analogy to the complex exponentials $\mathrm{e}^{\mathrm{i}kx}$ on the circle, the spherical harmonics Y_k^n are the key to Fourier analysis on the sphere. The *Legendre polynomials* $P_k : [-1,1] \to \mathbb{R}$, $k \in \mathbb{N}_0$, are given by their Rodrigues formula

$$P_k\left(t\right) := \frac{1}{2^k k!}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\left(t^2-1\right)^k .$$

They are normalised by $P_k(1) = 1$ and obey the orthogonality relation $\int_{-1}^{1} P_k(t)P_l(t)\,\mathrm{d}t = 2\delta_{k,l}/(2k+1)$. The closely related *associated Legendre functions* $P_k^n : [-1,1] \to \mathbb{R}$, k, $n \in \mathbb{N}_0$, $n \le k$, are defined by

$$P_k^n\left(t\right) := \sqrt{\frac{(k-n)!}{(k+n)!}}\left(1-t^2\right)^{\frac{n}{2}}\frac{\mathrm{d}^n}{\mathrm{d}t^n}P_k\left(t\right) .$$

The *spherical harmonics* $Y_k^n : \mathbb{S}^2 \to \mathbb{C}$ of degree $k \in \mathbb{N}_0$ and order $n \in \mathbb{Z}$, $|n| \le k$, are the functions

$$Y_k^n\left(\boldsymbol{\xi}\right) = Y_k^n\left(\vartheta,\,\varphi\right) := \sqrt{\frac{2k+1}{4\pi}}P_k^{|n|}(\cos\vartheta)\mathrm{e}^{\mathrm{i}n\varphi} .$$

These functions form an orthonormal basis of $L^2(\mathbb{S}^2)$, i.e.,

$$\int_{\mathbb{S}^2} Y_k^n\left(\boldsymbol{\xi}\right)\overline{Y_l^m\left(\boldsymbol{\xi}\right)}\,\mathrm{d}\mu(\boldsymbol{\xi}) = \int_0^{2\pi}\int_0^{\pi} Y_k^n\left(\vartheta,\,\varphi\right)\overline{Y_l^m\left(\vartheta,\,\varphi\right)}\sin\vartheta\,\mathrm{d}\vartheta\mathrm{d}\varphi = \delta_{k,l}\delta_{n,m} ,$$

moreover the addition theorem

$$\sum_{n=-k}^{k}\overline{Y_k^n\left(\boldsymbol{\eta}\right)}Y_k^n\left(\boldsymbol{\xi}\right) = \frac{2k+1}{4\pi}P_k\left(\boldsymbol{\eta}\cdot\boldsymbol{\xi}\right) \tag{2.1}$$

is valid. We say that f is a *spherical polynomial of degree* N if $f = \sum_{k=0}^{N}\sum_{n=-k}^{k}\hat{f}_k^n Y_k^n$. The space of spherical polynomials of degree at most N is denoted by $\Pi_N(\mathbb{S}^2)$ and has dimension $(N+1)^2$. The *spherical Fourier matrix* is given by

$$\boldsymbol{Y} := \left(Y_k^n\left(\boldsymbol{\xi}_j\right)\right)_{j=0,\ldots,M-1;k=0,\ldots,N,|n|\le k} \in \mathbb{C}^{M\times(N+1)^2}.$$

The inverse spherical Fourier transform is the construction of a spherical polynomial f of predetermined degree N from given data points $(\boldsymbol{\xi}_j, y_j) \in \mathbb{S}^2 \times \mathbb{C}$, $j = 0,\ldots,M-1$, such that the approximate identity

$$f(\boldsymbol{\xi}_j) \approx y_j \tag{2.2}$$

is fulfilled. Thus, it is equivalent to solve the linear system of equations $\boldsymbol{Y}\hat{\boldsymbol{f}} \approx \boldsymbol{y} = (y_j)_{j=0,\dots,M-1} \in \mathbb{C}^M$ for the sought vector of Fourier coefficients $\hat{\boldsymbol{f}} = (\hat{f}_k^n)_{k=0,\dots,N,|n|\le k} \in \mathbb{C}^{(N+1)^2}$ of the approximating polynomial. In a sense, it is the inverse problem to the matrix vector multiplication $\boldsymbol{f} = \boldsymbol{Y}\hat{\boldsymbol{f}}$ which corresponds to evaluate a spherical polynomial on the sampling set. An efficient realisation of this matrix vector multiplication is known as fast spherical Fourier transform at arbitrary nodes, and has been proposed in [15, 13].

In what follows, we study sufficient conditions for the matrix $\boldsymbol{Y}$ to have full rank, as well as condition numbers for the overdetermined case $M > (N+1)^2$ and the underdetermined case $M < (N+1)^2$, respectively.

3. Least squares approximation

For $M > (N+1)^2$, the linear system (2.2) is over-determined. Hence, in general the given data $\boldsymbol{y} \in \mathbb{C}^M$ can only be approximated up to some residual $\boldsymbol{r} := \boldsymbol{y} - \boldsymbol{Y}\hat{\boldsymbol{f}}$. In order to compensate for clusters in the sampling set $\mathcal{X}$, it is also useful to incorporate weights $w_j > 0$ into our problem, i.e., to consider the weighted least squares problem

$$\|\boldsymbol{y} - \boldsymbol{Y}\hat{\boldsymbol{f}}\|_{\boldsymbol{W}}^2 = \sum_{j=0}^{M-1} w_j |y_j - f(\boldsymbol{\xi}_j)|^2 \xrightarrow{\hat{\boldsymbol{f}}} \min, \tag{3.1}$$

where $\boldsymbol{W} := \operatorname{diag}(w_j)_{j=0,\dots,M-1} \in \mathbb{R}^{M\times M}$.

Lemma 1. *The least squares problem* (3.1) *is equivalent to the normal equation of first kind*

$$\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{Y}\hat{\boldsymbol{f}} = \boldsymbol{Y}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{y}. \tag{3.2}$$

Proof. The assertion is due to [1, Thm. 1.1.2] for the matrix $\boldsymbol{W}^{1/2}\boldsymbol{Y}$. □

3.1 Voronoi partition of δ-dense nodes

In this subsection, we briefly survey Voronoi diagrams on the sphere $\mathbb{S}^2$ and define the corresponding weights w_j, $j = 0,\dots,M-1$, which prove usefull in (3.1). For every node $\boldsymbol{\xi}_j$ in the sampling set $\mathcal{X} \subset \mathbb{S}^2$, the set of all points $\boldsymbol{\xi} \in \mathbb{S}^2$ for which $\boldsymbol{\xi}_j$ is closest of all nodes in $\mathcal{X}$, with respect to the geodesic distance, is called the *Voronoi cell* of $\boldsymbol{\xi}_j$. The collection of all Voronoi cells tessellates the whole sphere and is denoted the *Voronoi partition* corresponding to the set $\mathcal{X}$. More formally, we have the following definition.

Definition 1. Let $\mathcal{X} = \{\boldsymbol{\xi}_j \in \mathbb{S}^2 : j = 0,\dots,M-1\}$ be a sampling set. Then, the Voronoi partition $\mathcal{R} := \{\Omega_j \subset \mathbb{S}^2 : j = 0,\dots,M-1\}$, the corresponding Voronoi cells Ω_j and their parameterisations $\tilde{\Omega}_j$, their characteristic functions χ_j, their weights w_j, and the partition norm $\|\mathcal{R}\|$ are given by

$$\begin{aligned}
\Omega_j &:= \left\{ \boldsymbol{\xi} \in \mathbb{S}^2 : \operatorname{dist}\left(\boldsymbol{\xi}, \boldsymbol{\xi}_j\right) \leq \min_{l=0,\ldots,M-1} \operatorname{dist}\left(\boldsymbol{\xi}, \boldsymbol{\xi}_l\right) \right\}, \\
\tilde{\Omega}_j &:= \left\{ (\vartheta, \varphi)^\top \in [0, \pi] \times [-\pi, \pi) : (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)^\top \in \Omega_j \right\}, \\
\chi_j\left(\boldsymbol{\xi}\right) &:= \begin{cases} 1 & \text{if } \boldsymbol{\xi} \in \Omega_j, \\ 0 & \text{otherwise}, \end{cases} \\
w_j &:= \int_{\mathbb{S}^2} \chi_j(\boldsymbol{\xi}) \mathrm{d}\mu\left(\boldsymbol{\xi}\right) = \int_{\Omega_j} \mathrm{d}\mu\left(\boldsymbol{\xi}\right) = \int_{\tilde{\Omega}_j} \sin\vartheta \mathrm{d}\vartheta \mathrm{d}\varphi, \\
\|\mathcal{R}\| &:= \max_{j=0,\ldots,M-1} \max_{\boldsymbol{\xi},\boldsymbol{\eta} \in \Omega_j} \operatorname{dist}\left(\boldsymbol{\xi}, \boldsymbol{\eta}\right).
\end{aligned}$$

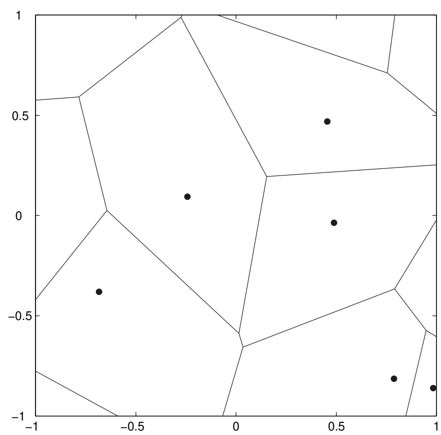
Obviously, the partition norm of the Voronoi partition obeys

$$\delta_{\mathcal{X}}/2 \leq \|\mathcal{R}\| \leq \delta_{\mathcal{X}}. \tag{3.3}$$

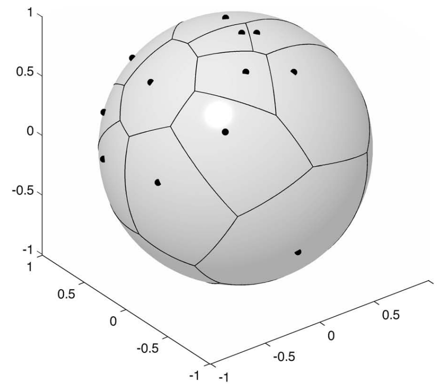
Remark 1. Figure 3.1(a) shows the most common example of Voronoi cells for points in the plane. Restricting the Euclidean Voronoi partition of $\mathbb{R}^3$ for a set on points $\mathcal{X}$ on the sphere to the sphere itself, yields exactly the Voronoi partition of $\mathbb{S}^2$ as defined above. An example is given in Figure 3.1(b). However, from a computational point of view, the three-dimensional procedure is not favourable since algorithms for computing the three-dimensional Voronoi partition can have quadratic complexity in the number of nodes M. In [24], Renka describes an algorithm which for M points on the sphere $\mathbb{S}^2$ computes all Voronoi cells with $\mathcal{O}\left(M \log M\right)$ arithmetic operations. For our computations, we used a C translation of Renka's original Fortran 77 library to compute the Voronoi cells Ω_j and weights w_j.

3.2 Spherical Marcinkiewicz-Zygmund inequalities

In order to estimate the convergence rate of our iterative solver we use a spherical Marcinkiewicz-Zygmund (MZ) inequality for L^2. In general, MZ-inequalities provide an equivalence between the continuous and discrete L^p-norms of polynomials and were first proven on the sphere by Mhaskar, Narcowich and Ward [18]. Recently, Filbir and Themistoclakis [7] used a particular de la Vallée Poussin kernel [6] in order to give a more concrete estimate in the L^1 case. Both versions of the L^p MZ-inequality were formulated for a quasi uniform subset of a dense sampling set. The existence of such a "compatible decomposition" was proven in [18]. However, its construction is somewhat lengthy and, more severe, part of the sampling information is discarded during the process.



(a) Voronoi cells for points in the plane.



(b) Voronoi cells for points on the sphere.

Figure 2 Voronoi cells in the plane and on the sphere.

We follow the lines of [18, 7] but improve previous estimates for the two dimensional unit sphere $\mathbb{S}^2$. In particular, we use the Voronoi partition of the sphere with respect to an arbitrary given sampling set.

The following technical Lemma 2 on a spherical de la Vallée Poussin kernel is due to [7]. In Lemma 3, we use the techniques of [18, 7] to estimate an L^1 variation between the kernel and its sampled version. Subsequently, we present in Theorem 1 an explicit MZ-inequality for L^2 which relies only on the degree N of the spherical polynomial and the mesh norm δ of the sampling set.

Lemma 2. *(Section 3 in [7])*
Let $N \in \mathbb{N}$ and define

$$n := \begin{cases} 3N/2 & N \text{ even}, \\ (3N-1)/2 & \text{otherwise}, \end{cases} \qquad m := \begin{cases} N/2 & N \text{ even}, \\ (N-1)/2 & \text{otherwise}. \end{cases}$$

Moreover, let $v_N : [-1,1] \to \mathbb{R}$ be given by

$$v_N(t) := \frac{4}{(m+1)^2} D_n(t) D_m(t) \quad \text{with} \quad D_N(t) := \sum_{k=0}^{N} \frac{2k+1}{2} P_k(t).$$

Then we have the following properties:

1. *((3.12) in [7]) v_N is a polynomial of degree at most $2N$.*
2. *(Proposition 3.4 in [7]) v_N is a reproducing kernel for spherical polynomials f of degree N, i.e., $f \in \Pi_N(\mathbb{S}^2)$ obeys*

$$f(\boldsymbol{\xi}) = \int_{\mathbb{S}^2} f(\boldsymbol{\eta}) \, v_N(\boldsymbol{\xi} \cdot \boldsymbol{\eta}) \, \mathrm{d}\mu(\boldsymbol{\eta}).$$

3. *((2.18) and (3.10) in [7] improving [7, Theorem 3.3] slightly)* v_N *is uniformly bounded by*

$$\|v_N\|_{L^\infty_{[-1,1]}} = \max_{t\in[-1,1]} v_N(t) \le \frac{9\left(N+\frac{2}{3}\right)}{16\pi} \le \frac{N^2}{2}.$$

4. *((3.13) in [7])* v_N *has bounded* L^1*-norm*

$$\sup_{N\in\mathbb{N}} \|v_N\|_{L^1_{[-1,1]}} = \sup_{N\in\mathbb{N}} \int_{-1}^{1} |v_N(t)|\,\mathrm{d}t \le \frac{3}{2\pi}\,.$$

Lemma 3. *Let* $N, M \in \mathbb{N}$, v_N *be the kernel in Lemma 2, and* $\mathcal{X} = \{\boldsymbol{\xi}_j \in \mathbb{S}^2 : j = 0,\dots,M-1\}$ *be a sampling set on the sphere with Voronoi partition* $\mathcal{R}$*, see Definition 1. Then we have for arbitrary fixed* $\boldsymbol{\eta} \in \mathbb{S}^2$ *and* $K := \left\lfloor \frac{\pi}{\|\mathcal{R}\|} \right\rfloor \ge 479N$ *the inequality*

$$\sum_{j=0}^{M-1} \int_{\Omega_j} \left|v_N(\boldsymbol{\xi}\cdot\boldsymbol{\eta}) - v_N\left(\boldsymbol{\xi}_j\cdot\boldsymbol{\eta}\right)\right| \mathrm{d}\mu(\boldsymbol{\xi}) \le 479\frac{N}{K}.$$

Proof. We use spherical coordinates with respect to the northpole $\boldsymbol{\eta}$ and set $\theta := \arccos(\boldsymbol{\xi}\cdot\boldsymbol{\eta})$ and $\rho_j := \arccos(\boldsymbol{\xi}_j\cdot\boldsymbol{\eta})$. Moreover, let $\theta_j^{\pm}$ denote the lowest and the highest values for the coordinate θ in the region $\tilde{\Omega}_j$, i.e.,

$$\theta_j^- := \min_{(\theta,\varphi)\in\tilde{\Omega}_j} \theta, \qquad \theta_j^+ := \max_{(\theta,\varphi)\in\tilde{\Omega}_j} \theta.$$

Of course, $\theta_j^- \le \rho_j, \theta \le \theta_j^+$ and hence we estimate for $j = 0,\dots,M-1$ the integral

$$\begin{aligned}
\int_{\Omega_j} \left|v_N(\boldsymbol{\xi}\cdot\boldsymbol{\eta}) - v_N\left(\boldsymbol{\xi}_j\cdot\boldsymbol{\eta}\right)\right| \mathrm{d}\mu(\boldsymbol{\xi}) &= \int_{\tilde{\Omega}_j} |v_N(\cos\theta) - v_N(\cos\rho_j)| \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi \\
&\le \int_{\tilde{\Omega}_j}\int_{\rho_j}^{\theta} \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right| \mathrm{d}t\ \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi \\
&\le \int_{\theta_j^-}^{\theta_j^+} \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right| \mathrm{d}t \int_{\tilde{\Omega}_j} \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi\,.
\end{aligned} \tag{3.4}$$

The next step is to cover the sphere with overlapping bands, see e.g. [18, Proposition 3.3], we define

$$\begin{aligned}
J_k &:= [(k-1)\pi/K, k\pi/K]\,, && k = 1,\dots K, \\
B_k &:= \left\{\boldsymbol{\xi} \in \mathbb{S}^2 : \theta \in J_k \cup J_{k+1}\right\}, && k = 1,\dots K-1, \\
\tilde{J}_k &:= J_k \cup J_{k+1}, && k = 1,\dots K-1.
\end{aligned}$$

Note that $\frac{\pi}{K} \geq \|\mathcal{R}\| \geq \theta_j^+ - \theta_j^-$ and thus if $\theta_j^- \in J_k$ then $[\theta_j^-, \theta_j^+] \subset \tilde{J}_k$ and $\Omega_j \subset B_k$. Due to $\cup_{k=1}^{K-1} B_k = \mathbb{S}^2$, we proceed with (3.4) by the following reordering of the sum over the nodes

$$\begin{aligned}
&\sum_{j=0}^{M-1} \int_{\theta_j^-}^{\theta_j^+} \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right| \mathrm{d}t \int_{\tilde{\Omega}_j} \sin\theta \,\mathrm{d}\theta\,\mathrm{d}\varphi \\
&\quad \leq \sum_{k=1}^{K-1} \sum_{\Omega_j \subset B_k} \int_{\theta_j^-}^{\theta_j^+} \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right| \mathrm{d}t \int_{\Omega_j} \mathrm{d}\mu(\boldsymbol{\xi}) \\
&\quad \leq \sum_{k=1}^{K-1} \int_{\tilde{J}_k} \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right| \mathrm{d}t \sum_{\Omega_j \subset B_k} \int_{\Omega_j} \mathrm{d}\mu(\boldsymbol{\xi}) \\
&\quad \leq \sum_{k=1}^{K-1} \int_{\tilde{J}_k} \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right| \mathrm{d}t \cdot 2\pi \int_{\tilde{J}_k} \sin\theta \mathrm{d}\theta,
\end{aligned} \tag{3.5}$$

where we have used that the Voronoi cells $\Omega_j \subset B_k$ share no interior points with each other. In order to avoid the introduction of the doubling weights as in [17, 18], we estimate the right hand side of (3.5) with a technique suggested in [7, Theorem 4.2]. We distinguish between bands B_k near the north/south pole and bands B_k well separated from the poles. We use the fact that away from the poles, i.e., for $t, \theta \in \tilde{J}_k$ we have

$$\sin\theta = \sin((\theta - t) + t) \leq \sin(\theta - t) + \sin(t) \leq 2\sin t, \quad k = 3, 4, \ldots, K-3\,.$$

Using the product rule

$$\left|\left(\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right)\sin t\right| \leq \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\sin t\right| + |v_N(\cos t)\cos(t)|\,,$$

we estimate

$$\begin{aligned}
2\pi \sum_{k=3}^{K-3} \int_{\tilde{J}_k} \int_{\tilde{J}_k} \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right| \sin\theta \,\mathrm{d}t\,\mathrm{d}\theta &\leq \frac{8\pi^2}{K} \sum_{k=3}^{K-3} \int_{\tilde{J}_k} \left|\frac{\mathrm{d}}{\mathrm{d}t} v_N(\cos t)\right| \sin t \,\mathrm{d}t \\
\leq \frac{16\pi^2}{K} \left(\int_0^\pi \left|\frac{\mathrm{d}}{\mathrm{d}t} [v_N(\cos t)\sin t]\right| \mathrm{d}t + \int_{\frac{2\pi}{K}}^{\frac{(K-2)\pi}{K}} |v_N(\cos t)\cos t| \,\mathrm{d}t \right).
\end{aligned} \tag{3.6}$$

By assumption, we have $\frac{2\pi}{K} \leq \frac{C}{N}$ for $C \geq 2\pi/479$. Using the fact $\frac{\cos t}{\sin t} \leq \frac{1}{t}$

for $t \in (0, \pi/2]$ and Lemma 2 yields for the last integral

$$\begin{aligned}
&\int_{\frac{2\pi}{K}}^{\frac{(K-2)\pi}{K}} |v_N(\cos t)\cos t\,|\,\mathrm{d}t \\
&\quad\le \left(\int_{\frac{2\pi}{K}}^{\frac{C}{N}} + \int_{\pi-\frac{C}{N}}^{\frac{(K-2)\pi}{K}}\right) |v_N(\cos t)\,|\,\mathrm{d}t + \int_{\frac{C}{N}}^{\pi-\frac{C}{N}} |v_N(\cos t)\cos t\,|\,\mathrm{d}t \\
&\quad\le \left(\int_{0}^{\frac{C}{N}} + \int_{\pi-\frac{C}{N}}^{\pi}\right) |v_N(\cos t)\,|\,\mathrm{d}t + \int_{\frac{C}{N}}^{\pi-\frac{C}{N}} \left|v_N(\cos t)\sin t\,\frac{\cos t}{\sin t}\,\right|\mathrm{d}t \\
&\quad\le \frac{2C}{N}\max_{t\in[0,\pi]}|v_N(\cos t)| + \frac{N}{C}\int_0^{\pi}|v_N(\cos t)\sin(t)|\,\mathrm{d}t \le \frac{4N}{\sqrt{2\pi}}
\end{aligned}$$

by choosing $C = \frac{1}{\sqrt{2\pi}}$. In conjunction with the Bernstein inequality for the first summand in (3.6), $N \ge 1$, and the fact that v_N has bounded L^1-norm, cf. Lemma 2, i.e.,

$$\begin{aligned}
\int_0^{\pi}\left|\frac{\mathrm{d}}{\mathrm{d}t}\left[v_N(\cos t)\sin t\right]\right|\mathrm{d}t &\le (2N+1)\int_0^{\pi}|v_N(\cos t)\sin t|\,\mathrm{d}t \\
&\le \frac{3(2N+1)}{2\pi} \le \frac{9N}{2\pi}
\end{aligned}$$

we conclude for the part away from the poles

$$2\pi\sum_{k=3}^{K-3}\int_{\tilde{J}_k}\int_{\tilde{J}_k}\left|\frac{\mathrm{d}}{\mathrm{d}t}v_N(\cos t)\right|\sin\theta\,\mathrm{d}t\,\mathrm{d}\theta \le \frac{72\pi N}{K} + \frac{64\pi^2 N}{\sqrt{2\pi}K} \le 478.2\frac{N}{K}\,. \tag{3.7}$$

Near the poles, i.e., for $\hat{I}_K := \{1, 2, K-2, K-1\}$, we proceed in (3.5) with $\sin\theta \le \theta$ and $\max_{t\in[0,\pi]}|\frac{\mathrm{d}}{\mathrm{d}t}v_N(\cos t)| \le 2N\|v_N\|_{L^\infty_{[-1,1]}} \le N^3$, cf. Lemma 2, by

$$\begin{aligned}
&2\pi\sum_{k\in\hat{I}_K}\int_{\tilde{J}_k}\left|\frac{\mathrm{d}}{\mathrm{d}t}v_N(\cos t)\right|\mathrm{d}t\int_{\tilde{J}_k}\sin\theta\,\mathrm{d}\theta \\
&\quad\le \frac{16\pi^2}{K}\max_{t\in[0,\pi]}\left|\frac{\mathrm{d}}{\mathrm{d}t}v_N(\cos t)\right|\cdot\max_{k=1,2}\int_{\tilde{J}_k}\theta\,\mathrm{d}\theta \le \frac{48\pi^4 N^3}{K^3} \le 0.03\frac{N}{K}\,,
\end{aligned}\tag{3.8}$$

where the last estimate follows from $N^2/K^2 \le 479^{-2}$. In summary, we estimate the left hand side of (3.5) by using (3.7) and (3.8) to obtain the assertion. □

Theorem 1. *Let a δ-dense sampling set $\mathcal{X} \subset \mathbb{S}^2$ of cardinality $M \in \mathbb{N}$ be given. Moreover let for $N \in \mathbb{N}$ with $154N\delta < 1$ and $\boldsymbol{W} = \operatorname{diag}(w_j)_{j=0,\dots,M-1}$, with Voronoi weights w_j, cf. Definition 1 be given. Then we have for arbitrary spherical polynomials $f \in \Pi_N(\mathbb{S}^2)$, for the vector $\boldsymbol{f} = \left(f(\boldsymbol{\xi}_j)\right)_{j=0,\dots,M-1}$ the weighted norm estimate*

$$(1 - 154N\delta)\, \|f\|_{L^2}^2 \leq \|\boldsymbol{f}\|_{\boldsymbol{W}}^2 \leq (1 + 154N\delta)\, \|f\|_{L^2}^2 .$$

Proof. We follow the lines of [18, 7] closely and use the Markov inequality

$$|f(\boldsymbol{\xi}) - f(\boldsymbol{\eta})| \leq N \operatorname{dist}(\boldsymbol{\xi}, \boldsymbol{\eta})\, \|f\|_{L^\infty}, \qquad \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{S}^2,$$

from [11]. Thus, for arbitrary $\boldsymbol{\xi} \in \mathbb{S}^2$ and its closest sampling node with index $j = \arg\min_{0 \leq l < M} \operatorname{dist}(\boldsymbol{\xi}, \boldsymbol{\xi}_l)$, we have

$$|f(\boldsymbol{\xi})| \leq \left|f(\boldsymbol{\xi}) - f\left(\boldsymbol{\xi}_j\right)\right| + \left|f\left(\boldsymbol{\xi}_j\right)\right| \leq \frac{N\delta}{2}\, \|f\|_{L^\infty} + \max_{0 \leq j < M} \left|f\left(\boldsymbol{\xi}_j\right)\right| .$$

Taking the maximum over $\boldsymbol{\xi} \in \mathbb{S}^2$ we conclude the L^∞-Marcinkiewicz-Zygmund inequality

$$(1 - \frac{N\delta}{2})\, \|f\|_{L^\infty} \leq \max_{0 \leq j < M} \left|f\left(\boldsymbol{\xi}_j\right)\right| \leq (1 + \frac{N\delta}{2})\, \|f\|_{L^\infty}, \tag{3.9}$$

where the right hand side is of course trivially fulfilled.

On the other hand, we have

$$\begin{aligned}
\left| \|f\|_{L^1} - \sum_{j=0}^{M-1} w_j \left|f\left(\boldsymbol{\xi}_j\right)\right| \right| &= \left| \|f\|_{L^1} - \left\| \sum_{j=0}^{M-1} f\left(\boldsymbol{\xi}_j\right) \chi_j \right\|_{L^1} \right| \\
&\leq \left\| f - \sum_{j=0}^{M-1} f\left(\boldsymbol{\xi}_j\right) \chi_j \right\|_{L^1} .
\end{aligned}$$

Using the reproducing property of v_N as first suggested in [7], one obtains

$$\begin{aligned}
&\left\| f - \sum_{j=0}^{M-1} f\left(\boldsymbol{\xi}_j\right) \chi_j \right\|_{L^1} = \sum_{j=0}^{M-1} \int_{\Omega_j} \left|f(\boldsymbol{\xi}) - f\left(\boldsymbol{\xi}_j\right)\right| \mathrm{d}\mu(\boldsymbol{\xi}) \\
&= \sum_{j=0}^{M-1} \int_{\Omega_j} \left| \int_{\mathbb{S}^2} \left(v_N(\boldsymbol{\xi} \cdot \boldsymbol{\eta}) - v_N\left(\boldsymbol{\xi}_j \cdot \boldsymbol{\eta}\right)\right) f(\boldsymbol{\eta})\, \mathrm{d}\mu(\boldsymbol{\eta}) \right| \mathrm{d}\mu(\boldsymbol{\xi}) \\
&\leq \sum_{j=0}^{M-1} \int_{\Omega_j} \int_{\mathbb{S}^2} \left|v_N(\boldsymbol{\xi} \cdot \boldsymbol{\eta}) - v_N\left(\boldsymbol{\xi}_j \cdot \boldsymbol{\eta}\right)\right| |f(\boldsymbol{\eta})|\, \mathrm{d}\mu(\boldsymbol{\eta})\, \mathrm{d}\mu(\boldsymbol{\xi}) \\
&\leq \|f\|_{L^1} \sup_{\boldsymbol{\eta} \in \mathbb{S}^2} \sum_{j=0}^{M-1} \int_{\Omega_j} \left|v_N(\boldsymbol{\xi} \cdot \boldsymbol{\eta}) - v_N\left(\boldsymbol{\xi}_j \cdot \boldsymbol{\eta}\right)\right| \mathrm{d}\mu(\boldsymbol{\xi}) .
\end{aligned}$$

In conjunction with Lemma 3 and the straightforward estimate

$$\frac{479N}{K} = \frac{479N}{\lfloor \pi / \|\mathcal{R}\| \rfloor} \leq \frac{479N \|\mathcal{R}\|}{\pi - \|\mathcal{R}\|} \leq \frac{480}{\pi} N\delta \leq 153 N\delta,$$

we obtain the explicit L^1-Marcinkiewicz-Zygmund inequality

$$(1 - 153N\delta) \|f\|_{L^1} \leq \sum_{j=0}^{M-1} w_j \left| f\left(\boldsymbol{\xi}_j\right) \right| \leq (1 + 153N\delta) \|f\|_{L^1}. \tag{3.10}$$

Now, let the sampling operator $\mathcal{S} : \Pi_N(\mathbb{S}^2) \to \Pi_N(\mathbb{S}^2)_{|\mathcal{X}}$, $f \mapsto (f(\boldsymbol{\xi}_j))_{0 \leq j < M}$, map each spherical polynomial f of degree at most N to its sample values $f(\boldsymbol{\xi}_j)$, $j = 0, \ldots, M-1$. We proceed by the Riesz-Thorin interpolation theorem [3, p. 32, Thm. 4.3] to bound the operator norm of $\mathcal{S}$, i.e., in our case this reads together with (3.9) and (3.10) as

$$\begin{aligned}
&\sup_{f \in \Pi_N(\mathbb{S}^2),\, f \neq 0} \sum_{j=0}^{M-1} w_j \left| f\left(\boldsymbol{\xi}_j\right) \right|^2 \Big/ \|f\|_{L^2}^2 \\
&\leq \sup_{f \in \Pi_N(\mathbb{S}^2),\, f \neq 0} \sum_{j=0}^{M-1} w_j \left| f\left(\boldsymbol{\xi}_j\right) \right| \Big/ \|f\|_{L^1} \times \sup_{f \in \Pi_N(\mathbb{S}^2),\, f \neq 0} \max_{0 \leq j < M} \left| f\left(\boldsymbol{\xi}_j\right) \right| \Big/ \|f\|_{L^\infty} \\
&\leq (1 + 153N\delta) \left(1 + \frac{1}{2} N\delta\right) \leq 1 + 154 N\delta.
\end{aligned}$$

Applying the Riesz-Thorin theorem to the inverse mapping $\mathcal{S}^{-1}$, we obtain

$$\sup_{f \in \Pi_N(\mathbb{S}^2),\, f \neq 0} \sum_{j=0}^{M-1} w_j \left| f\left(\boldsymbol{\xi}_j\right) \right|^2 \Big/ \|f\|_{L^2}^2 \geq 1 - 154 N\delta,$$

which in turn gives the assertion. Note, that this improves [18, Thm. 3.1] and [7, Thm. 4.2] by using the interpolation argument more directly. □

3.3 Eigenvalue estimate

In this subsection, we summarise the above results in order to answer the question under which condition the least squares problems (3.1) or the equivalent normal equations (3.2) are stable.

Corollary 1. *Let a δ-dense sampling set $\mathcal{X} \subset \mathbb{S}^2$ of cardinality $M \in \mathbb{N}$ be given and $\boldsymbol{W} = \operatorname{diag}((w_j)_{j=0,\ldots,M-1})$ denote the diagonal matrix with Voronoi weights w_j, cf. Definition 1. Then for $N \in \mathbb{N}$, $154N\delta < 1$, the matrix $\boldsymbol{Y}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{Y}$ in the normal equation of first kind* (3.2) *has bounded eigenvalues*

$$0 < 1 - 154N\delta \leq \lambda_{\min}\left(\boldsymbol{Y}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{Y}\right) \leq 1 \leq \lambda_{\max}\left(\boldsymbol{Y}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{Y}\right) \leq 1 + 154N\delta. \tag{3.11}$$

In particular, under the above conditions, the matrix $\boldsymbol{Y}$ has full (column) rank $(N+1)^2$. Moreover, the conjugate gradient method applied to (3.2) *converges linearly, i.e.,*

$$\|\boldsymbol{r}_l\|_{\boldsymbol{W}} \leq 2\,(154N\delta)^l\,\|\boldsymbol{r}_0\|_{\boldsymbol{W}} \tag{3.12}$$

with the initial residual $\boldsymbol{r}_0 := \boldsymbol{y}$ and the residual $\boldsymbol{r}_l := \boldsymbol{Y}\hat{\boldsymbol{f}}_l - \boldsymbol{y}$ of the l-th iterate $\hat{\boldsymbol{f}}_l$.

Proof. Due to $\boldsymbol{f} = \boldsymbol{Y}\hat{\boldsymbol{f}}$ and Parseval's identity $\|\hat{\boldsymbol{f}}\|_2 = \|f\|_{L^2}$, Theorem 1 indeed shows the eigenvalue estimate

$$1 - 154N\delta \leq \frac{\hat{\boldsymbol{f}}^{\mathsf{H}}\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{Y}\hat{\boldsymbol{f}}}{\hat{\boldsymbol{f}}^{\mathsf{H}}\boldsymbol{f}} \leq 1 + 154N\delta.$$

Thus, the matrix $\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{Y}$ possesses a bounded condition number

$$\operatorname{cond}(\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{Y}) \leq \frac{1+154N\delta}{1-154N\delta}.$$

Applying the standard estimate for the convergence of the conjugate gradient method, see e.g. [1, p. 289], yields the assertion (3.12). □

We solve problem (3.1) by a factorised variant of conjugate gradients (CGNR, N for "Normal equation" and R for "Residual minimisation"), where we use the nonequispaced fast spherical Fourier transform for fast matrix vector multiplications with $\boldsymbol{Y}$ and its adjoint $\boldsymbol{Y}^{\mathsf{H}}$, see [15, 13] for details. Note that for $154N\delta < 1$ a constant number of iterations suffices to decrease the residual to a certain fraction, i.e., the total number of floating point operations is bounded by the complexity $\mathcal{O}(N^2\log^2 N + M)$ of the fast spherical Fourier transform.

Remark 2. The condition on the mesh norm δ is optimal up to a constant. More formally, [26, Theorem 5.2] states the following necessary condition on the sampling set: If $\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{Y} = \boldsymbol{I}$ with $M = (N+1)^2$ and strictly positive weights $w_j > 0$, then $N\delta \lesssim 9.62$. In contrast, we gave the sufficient condition on the sampling set: If $N\delta \leq 1/154$, then the Voronoi weights $w_j > 0$ yield a well conditioned matrix $\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{Y}$.

4. Optimal interpolation

We define for given sample values $y_j \in \mathbb{C}$, $j = 0, \ldots, M-1$, and given weights $\hat{w}_k > 0$, $k = 0, \ldots, N$, the *optimal interpolation problem*

$$\min_{\hat{\boldsymbol{f}} \in \mathbb{C}^{(N+1)^2}} \sum_{k=0}^{N}\sum_{n=-k}^{k} \frac{\left|\hat{f}_k^n\right|^2}{\hat{w}_k} \quad \text{subject to} \quad \sum_{k=0}^{N}\sum_{n=-k}^{k} \hat{f}_k^n Y_k^n\left(\boldsymbol{\xi}_j\right) = y_j. \tag{4.1}$$

Lemma 4. *The optimal interpolation problem* (4.1) *is equivalent to the normal equations of second kind*

$$\boldsymbol{Y}\hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}}\tilde{\boldsymbol{f}} = \boldsymbol{y}, \qquad \hat{\boldsymbol{f}} = \hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}}\tilde{\boldsymbol{f}}, \tag{4.2}$$

where the weighting matrix is given by $\hat{\boldsymbol{W}} := \operatorname{diag}(\tilde{\boldsymbol{w}}) \in \mathbb{R}^{(N+1)^2\times(N+1)^2}$ *for the vector* $\tilde{\boldsymbol{w}} = (\tilde{w}_k^n)_{k=0,\ldots,N,|n|\le k}$ *with* $\tilde{w}_k^n = \hat{w}_k$, $k = 0,\ldots,N, |n| \le k$.

Moreover, with the help of the polynomial kernel $K_N : [-1,1] \to \mathbb{C}$ *and its associated matrix*

$$K_N(t) := \sum_{k=0}^{N} \frac{2k+1}{4\pi}\hat{w}_k P_k(t), \qquad \boldsymbol{K} := \left(K_N\left(\boldsymbol{\xi}_j \cdot \boldsymbol{\xi}_l\right)\right)_{j,l=0,\ldots,M-1} \tag{4.3}$$

we have $\boldsymbol{K} = \boldsymbol{Y}\hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}}$.

Proof. The first assertion is due to [1, Thm. 1.1.2] for the matrix $\boldsymbol{Y}\hat{\boldsymbol{W}}^{\mathbf{1/2}}$. The second assertion follows from the addition theorem (2.1). □

4.1 Ring partition of q-separated nodes

In this subsection, we define a partitioning of the sampling nodes into "rings" with increasing distance from the node $\boldsymbol{\xi}_0$, which is assumed without loss of generality to be the north pole $\boldsymbol{\xi}_0 = (0,0,1)^\top$.

Definition 2. For a separation distance $q \le \pi$, $0 \le m < \lfloor \pi q^{-1} \rfloor$, and the north pole $\boldsymbol{\xi}_0$ we define the sets

$$S_{q,m} := \left\{\boldsymbol{\xi} \in \mathbb{S}^2 : mq \le \arccos\left(\boldsymbol{\xi}_0 \cdot \boldsymbol{\xi}\right) < (m+1)\,q\right\}$$

and

$$S_{q,\lfloor \pi q^{-1}\rfloor} := \left\{\boldsymbol{\xi} \in \mathbb{S}^2 : \left\lfloor \pi q^{-1}\right\rfloor q \le \arccos\left(\boldsymbol{\xi}_0 \cdot \boldsymbol{\xi}\right) \le \pi\right\}.$$

Their restrictions to the set of sampling nodes is $S_{\mathcal{X},q,m} := S_{q,m} \cap \mathcal{X}$.The cardinality of these sets will be denoted by $|S_{\mathcal{X},q,m}|$.

We estimate the cardinality of $|S_{\mathcal{X},q,m}|$, i.e., we prove how many q-separated nodes can be placed in a certain distance to the north pole $\boldsymbol{\xi}_0$. In contrast to [20, Thm. 2.3], our estimate relies solely on the index m but no longer on the separation distance q. Figure 4.1 illustrates the basic idea behind the proof of Lemma 5.

Lemma 5. *Let a set of sampling nodes on* $\mathbb{S}^2$ *be* q*-separated, then the sets* $S_{\mathcal{X},q,m}$ *obey for* $m = 1,\ldots,\lfloor \pi q^{-1}\rfloor$ *the bound*

$$|S_{\mathcal{X},q,m}| \le 25m.$$

Proof. We use a packing argument from [20, Thm. 2.3] which states that for each node in $S_{\mathcal{X},q,m}$ the centred cap around it of colatitude $q/2$ is

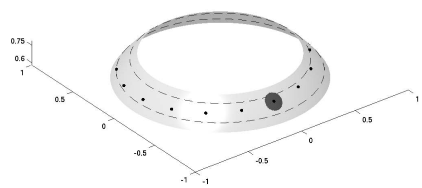


Figure 3 The set $S_{\mathcal{X},q,m}$, the ring $S_{q,m}$ (dashed), the larger ring $\tilde{S}_{q,m}$, and a spherical cap of colatitude $q/2$ centred at one node for generalised spiral nodes.

contained in the larger ring $\tilde{S}_{q,m} = S_{q,m-\frac{1}{2}} \cup S_{q,m+\frac{1}{2}}$ and has no interior points common with the cap of another node. Hence, we estimate for $m = 1, \ldots, \lfloor \pi q^{-1} \rfloor - 2$

$$|S_{\mathcal{X},q,m}| \leq \frac{\int_{\tilde{S}_{q,m}} \mathrm{d}\mu\left(\boldsymbol{\xi}\right)}{\int_{S_{q/2,0}} \mathrm{d}\mu\left(\boldsymbol{\xi}\right)} = \frac{\int_{\left(m-\frac{1}{2}\right)q}^{\left(m+\frac{3}{2}\right)q} \sin\theta \mathrm{d}\theta}{\int_0^{\frac{q}{2}} \sin\theta \mathrm{d}\theta} = \frac{\cos\left(\frac{(2m-1)q}{2}\right) - \cos\left(\frac{(2m+3)q}{2}\right)}{1-\cos\frac{q}{2}}.$$

Using an identity for the de la Vallée Poussin kernel, see e.g. [22, equation (3.4) and (3.5)], we calculate further

$$\begin{aligned}
|S_{\mathcal{X},q,m}| &\leq \frac{\cos\left((2m-1)\frac{q}{2}\right) - \cos\left((2m+3)\frac{q}{2}\right)}{1-\cos\frac{q}{2}} \\
&= \frac{\sin\left((2m+1)\frac{q}{2}\right)\sin\left(2\frac{q}{2}\right)}{\sin^2\frac{q}{4}} \\
&= 4 + 8\sum_{l=1}^{2m-1} \cos\frac{lq}{2} + 6\cos\frac{2mq}{2} + 4\cos\frac{(2m+1)q}{2} + 2\cos\frac{(2m+2)q}{2} \\
&\leq 8\left(2m+1\right).
\end{aligned}$$

In conjunction with a similar argument, starting from the south pole, see also the first estimate in [20, inequality (2.32)], i.e.,

$$|S_{\mathcal{X},q,\lfloor\pi q^{-1}\rfloor-1} \cup S_{\mathcal{X},q,\lfloor\pi q^{-1}\rfloor}| \leq \frac{1-\cos\left(5\frac{q}{2}\right)}{1-\cos\left(\frac{q}{2}\right)} \leq 25,$$

we obtain the assertion. □

4.2 Localised kernels

Localised kernels and well separated sampling sets yield stable interpolation for polynomials on the sphere. We construct localised polynomials on the sphere with the help of localised trigonometric polynomials. This construction was first suggested in [14]. We need the following lemma for the connection coefficients between the Chebyshev and Legendre polynomials.

Lemma 6. *For $k, l \in \mathbb{N}_0$ and the Chebyshev polynomials $T_l : [-1,1] \to \mathbb{R}$, $z \mapsto T_l(z) := \cos(l \arccos z)$ the matrix $\boldsymbol{P} := (P_{k,l})_{k,l\in\mathbb{N}_0}$ with entries*

$$P_{k,l} := \int_{-1}^{1} P_k\,(z)\, T_l\,(z)\, \mathrm{d}z \tag{4.4}$$

fulfils

$$\begin{aligned} P_{k,l} &= 2 \quad \text{if } k = l = 0, \\ P_{k,l} &> 0 \quad \text{if } k = l, \\ P_{k,l} &< 0 \quad \text{if } l > k \text{ and } (-1)^{k+l} = 1, \\ P_{k,l} &= 0 \quad \text{otherwise.} \end{aligned}$$

Moreover, the identity

$$\sum_{l=k}^{\infty} (2 - \delta_{l,0})\, P_{k,l} = 0 \tag{4.5}$$

is satisfied.

Proof. The matrix entries obey the explicit form, cf. [27, p. 99],

$$\begin{aligned} P_{2k,2l} &= \frac{-2 \prod_{s=0}^{k-1} \left((2l)^2 - (2s)^2\right)}{\prod_{s=0}^{k} \left((2l)^2 - (2s+1)^2\right)}, \\ P_{2k+1,2l+1} &= \frac{-2 \prod_{s=0}^{k-1} \left((2l+1)^2 - (2s+1)^2\right)}{\prod_{s=0}^{k} \left((2l+1)^2 - (2s+2)^2\right)} \end{aligned} \tag{4.6}$$

for $l \geq k \geq 0$. We comment on the zero entries of the matrix $\boldsymbol{P}$. For the lower triangular part, i.e., $l < k$, the k-th Legendre polynomial is orthogonal to all polynomials of degree at most l; furthermore, T_{2k}, P_{2k} and T_{2k+1}, P_{2k+1} are even and odd, respectively. Moreover, the only negative factor in the proposed identity appears for $s = k = l$ and thus the diagonal entries of $\boldsymbol{P}$ are positive, whereas the non-zeros of the upper triangular part are negative. The last assertion, i.e., the diagonal dominance of $\boldsymbol{P}$ and thus the strict diagonal dominance of its finite sections is due to

$$\sum_{l=0}^{\infty} P_{0,2l} = \sum_{l=0}^{\infty} \int_{-1}^{1} T_{2l}\,(x)\, \mathrm{d}x = \sum_{l=0}^{\infty} \frac{2}{1 - (2l)^2} = 1$$

and the following calculations, where we restrict to the even case (4.6). For notational convenience, let $p_{k,l} := P_{2k,2l}$, $k, l \in \mathbb{N}$. We prove for $k \in \mathbb{N}$ and by induction over $N \geq k$ that

$$\sum_{l=k}^{N-1+k} p_{k,l} = \frac{1}{2N-1} \prod_{r=0}^{2k-1} \frac{N+r}{N+r+\frac{1}{2}}. \tag{4.7}$$

For $N=1$ and $k=1$ equation (4.7) is fulfilled and due to

$$\begin{aligned}\frac{p_{k+1,k+1}}{p_{k,k}} &= \frac{\prod_{s=0}^{k}(k+1-s)(k+1+s)\cdot\prod_{s=0}^{k}\left(k-s-\frac{1}{2}\right)\left(k+s+\frac{1}{2}\right)}{\prod_{s=0}^{k+1}\left(k+1-s-\frac{1}{2}\right)\left(k+1+s+\frac{1}{2}\right)\cdot\prod_{s=0}^{k-1}(k-s)(k+s)}\\ &= \frac{(2k+1)(2k+2)}{\left(2k+\frac{3}{2}\right)\left(2k+\frac{5}{2}\right)}\end{aligned}$$

this extends to all $N=k$. We proceed for fixed k by

$$\begin{aligned}\sum_{l=k}^{N+k} p_{k,l} &= p_{k,N+k} + \sum_{l=k}^{N-1+k} p_{k,l}\\ &= -\frac{2\prod_{s=0}^{k-1}\left(4(N+k)^2-4s^2\right)}{\prod_{s=0}^{k}\left(4(N+k)^2-(2s+1)^2\right)} + \sum_{l=k}^{N-1+k} p_{k,l}\\ &= -\frac{\prod_{s=0}^{k-1}(N+k-s)(N+k+s)}{2\prod_{s=0}^{k}\left(N+k-s-\frac{1}{2}\right)\left(N+k+s+\frac{1}{2}\right)} + \sum_{l=k}^{N-1+k} p_{k,l}\end{aligned}$$

and apply the induction hypothesis (4.7), i.e.,

$$\begin{aligned}\sum_{l=k}^{N+k} p_{k,l} &= \frac{1}{2N-1}\left(\prod_{r=0}^{2k-1}\frac{N+r}{N+r+\frac{1}{2}}\right)\left(1-\frac{N+k}{N\left(N+2k+\frac{1}{2}\right)}\right)\\ &= \frac{N^2\left(N+2k+\frac{1}{2}\right)-N(N+k)}{(2N-1)\left(N+\frac{1}{2}\right)N\left(N+2k+\frac{1}{2}\right)}\left(\prod_{r=0}^{2k-2}\frac{N+1+r}{N+1+r+\frac{1}{2}}\right)\\ &= \frac{1}{2N+1}\prod_{r=0}^{2k-1}\frac{N+1+r}{N+1+r+\frac{1}{2}}.\end{aligned}$$

We conclude from (4.7) that the N-th partial sum of each even numbered row of $\boldsymbol{P}$ decays like $\mathcal{O}\left(N^{-1}\right)$. The same technique yields this property for all odd numbered rows and thus (4.5) follows. □

Definition 3. Let the normalised B-spline of order $\beta\in\mathbb{N}$ be defined by $g_\beta : [-\frac{1}{2},\frac{1}{2}]\to\mathbb{R}$, $g_\beta(z) := \beta N_\beta(\beta z+\frac{\beta}{2})$, with the cardinal B-spline given by

$$N_{\beta+1}(z) = \int_{z-1}^{z} N_\beta(\tau)\,\mathrm{d}\tau,\ \beta\in\mathbb{N},\qquad N_1(z) = \begin{cases}1 & 0<z<1,\\ 0 & \text{otherwise.}\end{cases}$$

Moreover, let for $N\in\mathbb{N}$ the *B-spline kernel* $B_{\beta,N} : [-1,1]\to\mathbb{R}$ be given by

$$B_{\beta,N}(t) := \frac{1}{\|g_\beta\|_{1,N}}\sum_{l=0}^{N}(2-\delta_{l,0})\,g_\beta\left(\frac{l}{2(N+1)}\right)T_l(t), \qquad (4.8)$$

where $\|\cdot\|_{1,N}$ denotes the discrete norm

$$\|g_\beta\|_{1,N} := \sum_{l=-N}^{N} g_\beta\left(\frac{l}{2\,(N+1)}\right).$$

Lemma 7. *The B-spline kernel $B_{\beta,N}$ obeys for $N \geq \beta - 1$ and $t \in [-1,1)$ the localisation property*

$$|B_{\beta,N}(t)| \leq c_\beta \,|(N+1)\arccos(t)|^{-\beta}, \qquad c_\beta := \frac{\left(2^\beta - 1\right)\zeta(\beta)\,\beta^\beta}{2^{\beta-1} - \zeta(\beta)\,\pi^{-\beta}}. \tag{4.9}$$

Moreover, it is normalised by $B_{\beta,N}(1) = 1$ and can be represented as

$$B_{\beta,N}(t) = \sum_{k=0}^{N} \frac{2k+1}{4\pi}\hat{w}_k P_k(t)$$

with **positive** *Fourier-Legendre coefficients*

$$\hat{w}_k = 2\pi \int_{-1}^{1} P_k(z)\, B_{\beta,N}(z)\,\mathrm{d}z, \qquad k = 0,\ldots,N.$$

Proof. First note that the B-splines are even functions and hence, the Chebyshev series in (4.8) can be rewritten with $t = \cos(2\pi x)$ as

$$\begin{aligned}\sum_{l=0}^{N}(2-\delta_{l,0})\,g_\beta\left(\tfrac{l}{2(N+1)}\right)T_l(t) &= \sum_{l=0}^{N}(2-\delta_{l,0})\,g_\beta\left(\tfrac{l}{2(N+1)}\right)\cos(2\pi l x)\\ &= \sum_{k=-N}^{N} g_\beta\left(\tfrac{k}{2(N+1)}\right)\mathrm{e}^{-2\pi \mathrm{i} k x}.\end{aligned}$$

Setting $x = 0$ yields the normalisation $B_{\beta,N}(1) = 1$.

Moreover, the smoothness-decay principle in [16, Lemma 3.2] states the localisation of a trigonometric polynomial with "smooth" Fourier coefficients. In particular, it has been proven that

$$\left|\sum_{k=-N}^{N} g_\beta\left(\frac{k}{2\,(N+1)}\right)\mathrm{e}^{-2\pi\mathrm{i}kx}\right| \leq \frac{\left(2^\beta-1\right)\zeta(\beta)\left|g_\beta^{(\beta-1)}\right|_V}{(4\,(N+1))^{\beta-1}\,|2\pi x|^\beta}, \tag{4.10}$$

$$\|g_\beta\|_{1,N} \geq 2\,(N+1)\left(1 - 2\zeta(\beta)\,(4\pi\beta)^{-\beta}\left|g_\beta^{(\beta-1)}\right|_V\right), \tag{4.11}$$

where $|g_\beta^{(\beta-1)}|_V$ denotes the total variation of the $(\beta-1)$-th derivative of g_β. Using $N_\beta'(z) = N_{\beta-1}(z) - N_{\beta-1}(z-1)$, we conclude

$$\left|g_\beta^{(\beta-1)}\right|_V = \beta^\beta\left|N_\beta^{(\beta-1)}\right|_V = \beta^\beta\left|\sum_{\tau=0}^{\beta-1}(-1)^\tau\binom{\beta-1}{\tau}N_1(\cdot-\tau)\right|_V = (2\beta)^\beta. \tag{4.12}$$

Combining (4.10)–(4.12) yields (4.9).

It remains to prove that every B-spline kernel $B_{\beta,N}$ possesses an expansion into Legendre polynomials with positive coefficients. We use the orthogonality of the Legendre polynomials and the linearity of the corresponding inner product and definition (4.4) and (4.8) to write the Fourier-Legendre coefficients as

$$\begin{aligned}\hat{w}_k &= 2\pi \int_{-1}^{1} P_k(z)\, B_{\beta,N}(z)\,\mathrm{d}z \\ &= \frac{2\pi}{\|g_\beta\|_{1,N}} \sum_{l=k}^{N} P_{k,l} \cdot (2-\delta_{l,0})\, g_\beta\left(\frac{l}{2(N+1)}\right). \qquad (4.13)\end{aligned}$$

Due to Lemma 6 and the fact that the weight function g_β is positive and nonincreasing for $0 \le z \le \frac{1}{2}$, we estimate

$$\hat{w}_k \ge \frac{2\pi g(0)}{\|g_\beta\|_{1,N}} \sum_{l=k}^{N} P_{k,l} \cdot (2-\delta_{l,0}) > 0,$$

where the positivity is due to the strict diagonal dominance of every finite section of the matrix $\boldsymbol{P}$ as shown in Lemma 6. □

Note, that in contrast to [20] the order β of the B-spline and the degree $N \in \mathbb{N}$ of the kernel $B_{\beta,N}$ are independent of each other. We shortly comment on two well known special cases. The case $\beta = 1$, i.e., the "top-hat function" $g_1(z) = 1$ for $|z| < \frac{1}{2}$ and $g_1(z) = 0$ elsewhere, leads to the well known *Dirichlet kernel* $B_{1,N}(\cos x) = \frac{1}{2N+1}\sum_{k=-N}^{N} \mathrm{e}^{+2\pi \mathrm{i} k x}$. Analogously, $\beta = 2$, i.e., the "hat function" $g_2(z) = 2 - 4|z|$ for $|z| \le \frac{1}{2}$ and $g_2(z) = 0$ elsewhere, leads to the *Fejér kernel*.

4.3 Eigenvalue estimate

In this subsection, we answer the question under which condition the optimal interpolation problems (4.1) or the equivalent normal equations (4.2) are stable.

Theorem 2. *Let a q-separated sampling set $\mathcal{X} \subset \mathbb{S}^2$ of cardinality $M \in \mathbb{N}$ and with $q \le \pi$ be given. Then for $N, \beta \in \mathbb{N}$, $N \ge \beta - 1 \ge 2$, the kernel matrix*

$$\boldsymbol{K} = (K_{j,l})_{j,l=0,\dots,M-1}, \qquad K_{j,l} = B_{\beta,N}(\boldsymbol{\xi}_j \cdot \boldsymbol{\xi}_l),$$

see also (4.3) *and Definition 3, has bounded eigenvalues*

$$|\lambda(\boldsymbol{K}) - 1| \le \frac{25 c_\beta \zeta(\beta - 1)}{((N+1)\, q)^\beta},$$

where c_β is given in (4.9).

Proof. We apply the Gershgorin circle theorem where we assume without loss of generality that $\boldsymbol{\xi}_0 = (0,0,1)^\top$. In conjunction with the partitioning in Definition 2, Lemma 5, and the localisation property in Lemma 7, we obtain

$$\begin{aligned}
|\lambda(\boldsymbol{K}) - 1| &\leq \sum_{j=1}^{M-1} \left| B_{\beta,N}\left(\boldsymbol{\xi}_0 \cdot \boldsymbol{\xi}_j\right)\right| \\
&\leq \sum_{m=1}^{\lfloor \pi q^{-1} \rfloor} |S_{\mathcal{X},q,m}| \max_{\boldsymbol{\xi} \in S_{q,m}} |B_{\beta,N}\left(\boldsymbol{\xi}_0 \cdot \boldsymbol{\xi}\right)| \\
&\leq \sum_{m=1}^{\lfloor \pi q^{-1} \rfloor} |S_{\mathcal{X},q,m}| \, c_\beta \left((N+1)\, mq\right)^{-\beta} \\
&\leq 25 c_\beta \zeta\left(\beta - 1\right) \left((N+1)\, q\right)^{-\beta} .
\end{aligned}$$

□

Corollary 2. *Let a q-separated sampling set $\mathcal{X} \subset \mathbb{S}^2$ of cardinality $M \in \mathbb{N}$ and with $q \leq \pi$ be given. Moreover, let $N \in \mathbb{N}$, $(N+1)q > 11.2$, and weights be given by the sampled cubic B-Spline, i.e.,*

$$\hat{w}_k = \frac{2\pi}{\|g_4\|_{1,N}} \sum_{l=k}^{N} P_{k,l} \cdot (2 - \delta_{l,0})\, g_4\left(\frac{l}{2\,(N+1)}\right) .$$

Then we have

$$0 < 1 - \left(\tfrac{11.2}{(N+1)q}\right)^4 \leq \lambda_{\min}(\boldsymbol{Y}\hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}}) \leq 1 \leq \lambda_{\max}(\boldsymbol{Y}\hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}}) \leq 1 + \left(\tfrac{11.2}{(N+1)q}\right)^4 . \tag{4.14}$$

In particular, under the above conditions, the matrix $\boldsymbol{Y}$ has full (row) rank M. Moreover, the conjugate gradient method applied to (4.2) *converges linearly, i.e.,*

$$\|\hat{\boldsymbol{e}}_l\|_{\hat{\boldsymbol{W}}^{-1}} \leq 2\left(\frac{11.2}{(N+1)q}\right)^{4l} \|\hat{\boldsymbol{e}}_0\|_{\hat{\boldsymbol{W}}^{-1}} \tag{4.15}$$

with the initial error $\hat{\boldsymbol{e}}_0 := \hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{K}^{-1}\boldsymbol{y}$ and the error $\hat{\boldsymbol{e}}_l := \hat{\boldsymbol{f}}_l - \hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{K}^{-1}\boldsymbol{y}$ of the l-th iterate $\hat{\boldsymbol{f}}_l$.

Proof. Setting $\beta = 4$ in Theorem 2 and using (4.9) yields $\sqrt[4]{25 c_4 \zeta(3)} \leq 11.2$ and finally (4.14). Applying the standard estimate for the convergence of the conjugate gradient method, see e.g. [1, p. 289], yields the assertion (4.15). □

We solve problem (4.2) by a factorised variant of conjugated gradients (CGNE, N for "Normal equation" and E for "Error minimisation"), where we use again the nonequispaced fast spherical Fourier transform for fast matrix vector multiplications with $\boldsymbol{Y}$ and its adjoint $\boldsymbol{Y}^{\mathsf{H}}$. Note that for $(N+1)q > 11.2$ a constant number of iterations suffices to decrease the error to a certain fraction, i.e., the total number of floating point operations is bounded by the complexity $\mathcal{O}(N^2 \log^2 N + M)$ of the fast spherical Fourier transform.

Remark 3. The condition on the separation distance q is optimal up to a constant. More formally, $M \in \mathbb{N}$ nodes can be distributed on the sphere such that their separation distance is approximately $3.8/\sqrt{M}$, see [25] for an introduction. Choosing the polynomial degree $N = \lfloor 3.8q^{-1} \rfloor - 1$ yields $M > (N+1)^2$ and thus $\operatorname{rank}(\boldsymbol{Y}) < M$. In contrast, we showed that $(N+1)q > 11.2$ suffices for full $\operatorname{rank}(\boldsymbol{Y}) = M$.

5. Numerical results

In this section, we present numerical examples in order to confirm our theoretical findings. We start in Example 1 and Example 2 with two MATLAB tests to shed some light on the typical behaviour of the least squares and interpolation problem.

Furthermore, the two conjugate gradient methods have been implemented in C and are tested with global atmospheric temperature data on a Intel Xeon TM 3.00 GHz CPU system with 12 GB main memory, SuSE-Linux (kernel 2.6.5-7.257-smp, gcc 3.3.3) in double precision arithmetic. We have used the libraries FFTW 3.0.1 [9] and NFFT 3.0 [12], the latter with default settings (pre-computed Kaiser–Bessel window function, cut-off parameter $m = 6$, oversampling factor $\sigma = 2$, and stabilisation threshold $\kappa = 1000$ for the NFFT on the sphere). Note that the NFFT 3.0 includes the nonequispaced fast spherical Fourier transforms from [15, 13].

Example 1. We consider the least squares problem (3.1) and the related condition number of $\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{Y} \in \mathbb{C}^{(N+1)^2 \times (N+1)^2}$ with respect to the mesh norm $\delta_{\mathcal{X}}$. The sampling set is initially given by the generalised spiral with $M_{\max} = 400$ nodes, see also Figure 1(a). By removing an increasing number of nodes, starting at the north pole, we introduce a larger and larger growing "hole" resulting in an increasing mesh norm $\delta_{\mathcal{X}}$, see Figure 4(a). In particular $\delta_{\mathcal{X}} = \pi/2$ corresponds to the case with no sampling nodes on the northern hemisphere. For each such sampling set, we consider the matrix $\boldsymbol{Y}$ of size $M \times (N+1)^2$ with $200 \leq M \leq 400$ nodes and for different polynomial degrees.

The same experiment has been run for randomly drawn nodes from the uniform distribution on the sphere in Figure 4(c).

As can be seen in Figures 4(b/d), the condition number increases strongly for $\delta_{\mathcal{X}} \to \pi/2$ – independent of the fact that the system $\boldsymbol{Y}\hat{\boldsymbol{f}} = \boldsymbol{y}$

is still largely overdetermined, i.e., $(N+1)^2 \ll M$. In contrast to this, the matrix $\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{Y}$ is well conditioned whenever the inverse mesh norm is bounded from below by the polynomial degree, hence $\delta_{\mathcal{X}}^{-1} > N$.

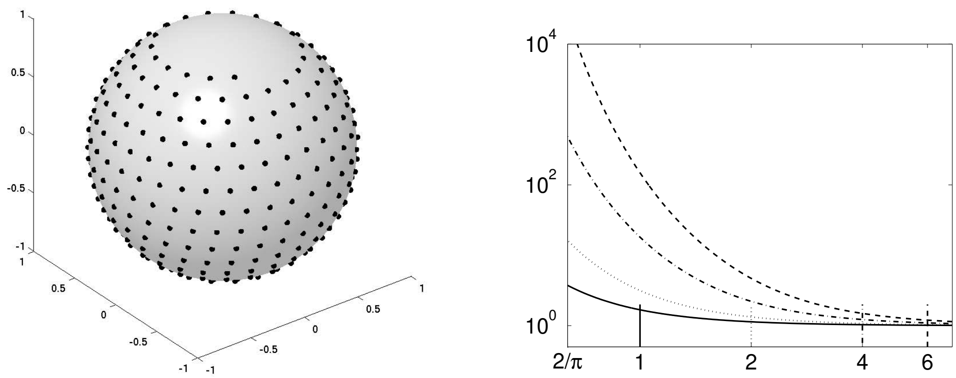
Example 2. We consider the optimal interpolation problem (4.1) and the eigenvalue distribution for different weights $\hat{w}_k$. Again, we use the generalised spiral with $M = 400$ nodes as sampling set. The weights $\hat{w}_k$, $k = 0, \ldots, N$, are computed by means of (4.13) from the samples of the β-th B-spline for $\beta = 1, 2, 3, 4$. Increasing the smoothness β of the Chebyshev coefficients in (4.8) yields a stronger localisation of the B-spline kernel and thus a more rapid off-diagonal decay in the matrix $\boldsymbol{K} = \boldsymbol{Y}\hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}} \in \mathbb{C}^{M \times M}$.

Besides a lower condition number, we see from Figure 3(b), that the eigenvalues $\lambda_0 \leq \ldots \leq \lambda_{399}$ cluster around one for larger values of β – which leads of course to much faster convergence of the conjugate gradient method. Also note that for a polynomial degree closer to some critical value $N \approx C q_{\mathcal{X}}$, strong localisation $\beta = 4$ does not pay off, cf. Figure 3(a). This behaviour probably results from an increased effective width of the main lobe of the B-spline kernel for larger β.

Example 3. In this example, we demonstrate both, the overdetermined and the underdetermined case for approximation with scattered data, corresponding to the least squares approximation and optimal interpolation cases, respectively. To this end, we chose a freely available data set from the NASA AMSU mission (see [21]) containing global atmospheric temperature data of the earth from 5 November 2006 measured by a satellite. Figure 7(a) shows the temperature data. Owing to the satellite's trajectory,

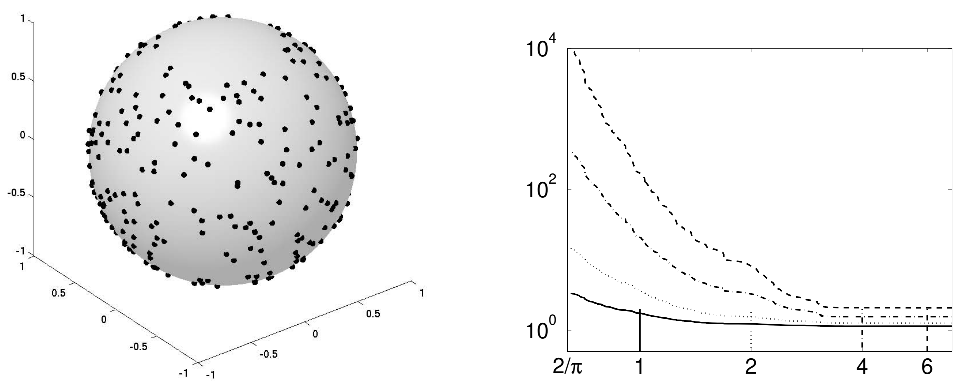
the set of data contains multiple, but in form and size similar, slits which are most wide at the equator. Within the slits, no data was acquired. The whole data comes already preprocessed and mapped onto a equiangular grid with 180 nodes in latitudinal direction and 360 nodes in longitudinal direction. From the theoretical amount of $180 \times 360 = 64,800$ nodes, only $53,996$ actually provide measured temperature values and are used for formulating the reconstruction problem. The rest lies in the uncovered regions.

In both cases, least squares approximation and optimal interpolation, we try to compute a reasonable continuation of the temperature distribution to these regions. In the first case, we iteratively compute a least squares approximation to the given data by solving the corresponding weighted normal equation of the first kind (3.2) using the CGNR algorithm. As weights, we choose the Voronoi weights w_j from Definition 1. Figure 7(b) shows the obtained approximation at degree $N = 128$ after 10 iterations. Compared to the original data, the approximation is smoother and does not fulfil any interpolation condition. Nevertheless, it might be taken as estimate for the data missing in the slits. Figure 7(c) shows the original data with values filled in from the approximation.



(a) Spiral nodes with gap at the north pole.

(b) Condition number of $\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{Y}$.

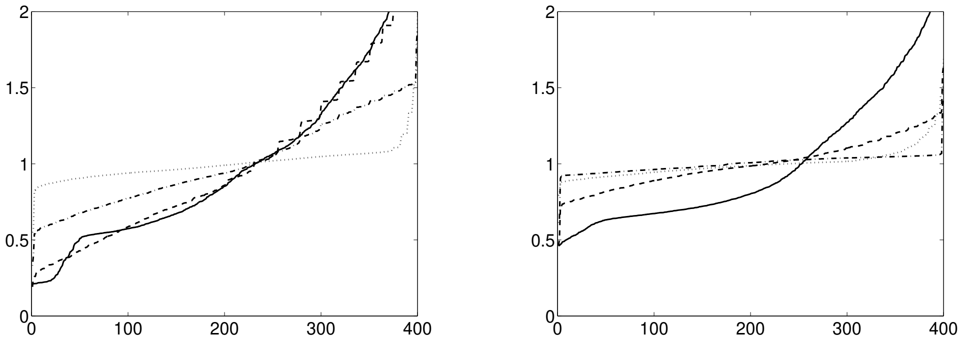


(c) Random nodes with gap at the north pole.

(d) Condition number of $\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{Y}$.

Figure 4 Condition number $\mathrm{cond}(\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{Y})$, $\boldsymbol{Y} \in \mathbb{C}^{M\times(N+1)^2}$, with respect to the approximate inverse mesh norm $\delta_{\mathcal{X}}^{-1} \in [2/\pi, \ldots, 7)$, i.e., $200 \leq M \leq 400$. The polynomial degree is $N = 1$ (solid), $N = 2$ (dotted), $N = 4$ (dash-dot), and $N = 6$ (dashed).

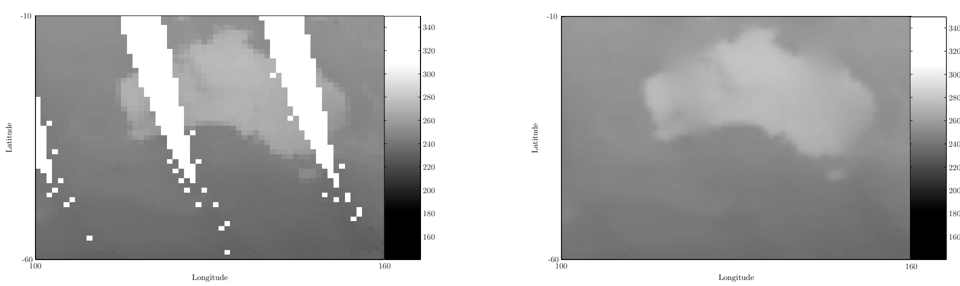
In the case of interpolation, we pick out a small portion of the temperature map in the region of Australia, shown in Figure 6(a), and choose a fairly high polynomial degree $N = 512$. In the corresponding normal equation of second kind (4.2), we use weights $\tilde{w}_k^n = (1 + k)^{-2}$. We did not choose B-spline weights, since the corresponding kernels are well localised and therefore not suited for smooth continuation of the interpolant to the missing regions. After 30 iterations using the CGNE algorithm, we evaluate the resulting interpolant on a refined grid with 4 fold resolution in each direc-



(a) Polynomial degree $N = 30$.

(b) Polynomial degree $N = 40$.

Figure 5 Eigenvalue distribution of the matrix $\boldsymbol{K} = \boldsymbol{Y}\hat{\boldsymbol{W}}\boldsymbol{Y}^{\mathsf{H}} \in \mathbb{C}^{M \times M}$ for weighting matrices obtained from the sampled B-spline of order $\beta = 1$ (solid), $\beta = 2$ (dotted), $\beta = 3$ (dash-dot), and $\beta = 4$ (dashed).



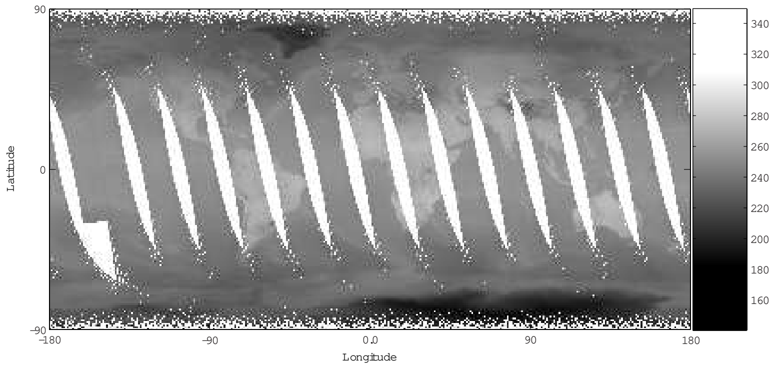
(a) Small portion of the temperature map in the region of Australia.

(b) Optimal interpolant with respect to the weights $\tilde{w}_k^n = (1+k)^{-2}$ at four fold resolution.

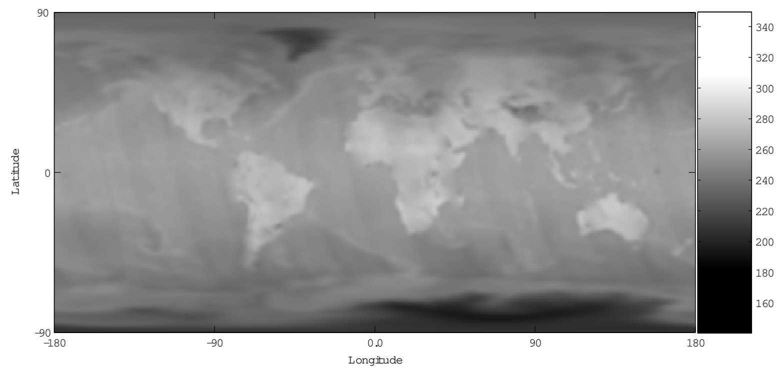
Figure 6 Optimal interpolation using a small portion of the temperature map.

tion, shown in Figure 6(b). Due to the weights $\tilde{w}_k^n$, smooth interpolants are favoured but the interpolation property enforces the conservation of sharp transitions at the edges of the continent. In the regions with missing data, however, the interpolant is smooth and does not totally recover sharp edges.

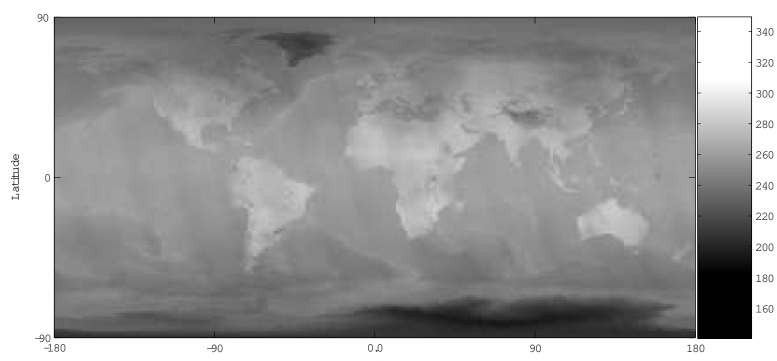
Remark 4. We shall mention here that the shown examples do not essentially require the use of nonequispaced spherical Fourier transforms. Owing to the equiangular distribution of the nodes, algorithms based on equispaced spherical Fourier transforms might also be employed to solve the problem. Nevertheless, this is due to the fact that the data has already been preprocessed. In most cases, it would be favourable to apply the described techniques directly to un-gridded sensor data for which we would not have to change a single line of code.



(a) The original atmospheric temperature data in Kelvin.



(b) Weighted least squares approximation at degree $N = 128$.



(c) Original data with filled-in values from the least squares approximation at degree $N = 128$.

Figure 7 Least squares approximation to the global temperature data.

6. Conclusions

We have shown that the spherical Fourier matrix $\boldsymbol{Y}$ has full rank if either

- for δ-dense sampling sets $\mathcal{X} \subset \mathbb{S}^2$ of cardinality $M \in \mathbb{N}$ we have

$$N < \frac{1}{154\,\delta},$$

- or for q-separated sampling sets $\mathcal{X} \subset \mathbb{S}^2$ of cardinality $M \in \mathbb{N}$ we have

$$N > \frac{11.2}{q} - 1.$$

In these cases, the condition numbers of the related normal equations are bounded, thus the total number of floating point operations is bounded by the complexity $\mathcal{O}(N^2 \log^2 N + M)$ of the fast spherical Fourier transform. Furthermore, our conditions for the least squares approximation as well as for the stable interpolation are best possible within a reasonable constant. Numerical results support our theoretical findings.

Acknowledgements

We highly appreciate numerous fruitful and enlightening discussions with Frank Filbir on Marcinkiewicz-Zygmund inequalities as proven in [7].

[1] Å. Björck. *Numerical Methods for Least Squares Problems.* SIAM, Philadelphia, 1996.
[2] J. P. Boyd. *Chebyshev and Fourier Spectral Methods.* Dover Press, New York, second edition, 2000.
[3] R. DeVore and G. Lorentz. *Constructive Approximation.* Springer–Verlag, Berlin, 1993.
[4] G. E. Fasshauer and L. L. Schumaker. Scattered data fitting on the sphere. In M. Dahlen, T. Lyche, and L. L. Schumaker, editors, *Mathematical Methods for Curves and Surfaces II*, pages 117 – 166, Nashville, 1998. Vanderbilt University Press.
[5] H. G. Feichtinger, K. Gröchenig, and T. Strohmer. Efficient numerical methods in non-uniform sampling theory. *Numer. Math.*, 69:423 – 440, 1995.
[6] F. Filbir and W. Themistoclakis. On the construction of de la Vallée Poussin means for orthogonal polynomials using convolution structures. *J. Comput. Anal. Appl.*, 6:297 – 312, 2004.
[7] F. Filbir and W. Themistoclakis. Polynomial approximation on the sphere using scattered data. Preprint 06-12, GSF München, 2006.
[8] W. Freeden, T. Gervens, and M. Schreiner. *Constructive Approximation on the Sphere.* Oxford University Press, Oxford, 1998.
[9] M. Frigo and S. G. Johnson. FFTW, C subroutine library. `http://www.fftw.org`.
[10] K. Gröchenig. Reconstruction algorithms in irregular sampling. *Math. Comput.*, 59:181 – 194, 1992.
[11] K. Jetter, J. Stöckler, and J. D. Ward. Errors estimates for scattered data interpolation on spheres. *Math. Comput.*, 68:733 – 747, 1999.
[12] J. Keiner, S. Kunis, and D. Potts. NFFT3.0, Softwarepackage, C subroutine library. `http://www.tu-chemnitz.de/~potts/nfft`, 2006.
[13] J. Keiner and D. Potts. Fast evaluation of quadrature formulae on the sphere. *Math. Comput.*, 2007. accepted.

[14] S. Kunis. Nonequispaced FFT, Generalisation and Inversion. Dissertation, Universität zu Lübeck, 2006.

[15] S. Kunis and D. Potts. Fast spherical Fourier algorithms. *J. Comput. Appl. Math.*, 161:75 – 98, 2003.

[16] S. Kunis and D. Potts. Stability results for scattered data interpolation by trigonometric polynomials. SIAM *J. Sci. Comput.*, 2007. accepted.

[17] G. Mastroianni and V. Totik. Weighted polynomial inequalities with doubling and A_∞ weights. *Constr. Approx.*, 16:37 – 71, 2000.

[18] H. N. Mhaskar, F. J. Narcowich, and J. D. Ward. Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature. *Math. Comput.*, 70:1113 – 1130, 2001. Corrigendum on the positivity of the quadrature weights in 71:453 – 454, 2002.

[19] C. Müller. *Spherical Harmonics*. Springer, Aachen, 1966.

[20] F. J. Narcowich, N. Sivakumar, and J. D. Ward. Stability results for scattered-data interpolation on euclidean spheres. *Adv. Comput. Math.*, 8:137 – 163, 1998.

[21] National Aeronautics and Space Administration. NASA AIRS Homepage. `http://disc.gsfc.nasa.gov/AIRS/index.shtml`.

[22] J. Prestin and K. Selig. On a constructive representation of an orthogonal trigonometric schauder basis for $C_{2\pi}$. In *Operator Theory: Advances and Applications*, pages 402 – 425. Birkhäuser, 2001.

[23] M. Reimer. *Multivariate Polynomial Approximation*. Birkhäuser Verlag, Basel, 2003.

[24] R. J. Renka. Algorithm 772: Stripack: Delaunay triangulation and voronoi diagram on the surface of a sphere. *ACM Trans. Math. Softw.*, 23:416 – 434, 1997. For accompanying software, see `http://www.acm.org/pubs/calgo`.

[25] E. B. Saff and A. B. J. Kuijlaars. Distributing many points on a sphere. *Math. Intelligencer*, 19:5 – 11, 1997.

[26] I. H. Sloan and R. S. Womersley. Extremal Systems of Points and Numerical Integration on the Sphere. *Adv. Comput. Math.*, 21:107 – 125, 2004.

[27] G. Szegő. *Orthogonal Polynomials*. Amer. Math. Soc., Providence, RI, 4th edition, 1975.

[28] H. Wendland. *Scattered Data Approximation*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge, 2005.

Received ... 2007

Universität zu Lübeck, Institut für Mathematik, Wallstrasse 40
23560 Lübeck
e-mail: keiner@math.uni-luebeck.de

Technische Universität Chemnitz, Fakultät für Mathematik, Reichenhainer Str. 39
09107 Chemnitz
e-mail: {kunis,potts}@mathematik.tu-chemnitz.de