Scattered data approximation on the bi-sphere and application to texture analysis

Frank Filbir and Daniel Potts

Abstract. The paper deals with the approximation and optimal interpolation of functions defined on the bisphere $\mathbb{S}^2 \times \mathbb{S}^2$ from scattered data. We demonstrate how the least square approximation to the function can be computed in a stable and efficient manner. The analysis of this problem is based on Marcinkiewicz–Zygmund inequalities for scattered data which we present here for the bisphere. The complementary problem of optimal interpolation is also solved by using well-localized kernels for our setting. Finally, we discuss the application of the developed methods to problems of texture analysis in material science.

1 Introduction

A typical problem in science is the development of a theoretical model for a hidden process from observational data. More precisely, we are given a set of measurements $\mathcal{A} = \mathcal{X} \times \mathbb{C} = \{(x_j, f_j) : j = 1, \dots, M\}$, where we assume that the sampling nodes $\mathcal{X} = \{x_j : j = 1, \dots, M\}$ are a finite subset of a metric space $(\mathbb{X}, d)$. We suppose that there exists a function p which generated the observed data. This assumption obviously leads to the equations $p(x_j) = f_j,\ j = 1, \dots, M$, or, even more realistic, $p(x_j) \approx f_j$ since the data are usually corrupted with noise or measurement errors. The function p will then be considered as a model for the underlying process. Dealing with real world problems, we can hardly expect that the sampling nodes are equally spaced. Indeed, due to experimental constraints, the sampling nodes are frequently scattered points. Consequently, we have to deal with interpolation respectively approximation of functions from scattered data. Moreover, we have to make assumptions on the complexity of our model, i.e., we have to restrict p to a certain function class. Often

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

this function class is assumed to be a linear space of finite dimension spanned by an orthogonal basis $\phi_1, \ldots, \phi_N$. This means that we suppose p to be a linear combination of some basis functions,

$$p(x) = \sum_{k=1}^{N} a_k \, \phi_k(x).$$

At this stage we prefer to be vague on the exact type of basis functions. They could be, for instance, complex exponentials, orthogonal polynomials or even some other type of basis functions. The problem we are confronted with is an efficient and stable computation of the coefficients a_k using the (scattered) data $\mathcal{A}$.

In case we are interested in the approximation problem, we look for a solution of a weighted least square problem. This means we choose the dimension of the space $\Pi_N := \mathsf{span}\{\phi_k : k = 1, \ldots, N\}$ to be smaller than the number of sampling points M and determine p according to

$$\sum_{j=1}^{M} w_j |f_j - p(x_j)|^2 \xrightarrow{p} \min .$$

The weights are incorporated in order to compensate for possible clusters in the data set. This weighted least square problem is equivalent to determine the solution of the normal equations of the first kind

$$\boldsymbol{\Phi}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{\Phi} \boldsymbol{a} = \boldsymbol{\Phi}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{f}, \tag{1.1}$$

where $\boldsymbol{\Phi} = (\phi_k(x_j))_{j=1,\ldots,M;k=1,\ldots,N}$, $\boldsymbol{a} = (a_k)_{k=1,\ldots,N}$, $\boldsymbol{f} = (f_j)_{j=1,\ldots,M}$, and $\boldsymbol{W} = \mathrm{diag}(w_1, \ldots, w_M)$.

In order to come up with a stable procedure for the computation of the vector $\boldsymbol{a}$ from (1.1) it is necessary to ensure nonsingularity of $\boldsymbol{\Phi}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{\Phi}$ and, moreover, to have control on the condition number of this matrix. This can be achieved by applying the so-called Marcinkiewicz–Zygmund inequalities (MZ inequalities for short) which establish an equivalence between the L^r-norm of p and the weighted norm $\sum_{j=1}^{M} w_j |p(x_j)|^r$. In the most important case of L^2, the MZ inequality reads as

$$A \|p\|_2 \leq \sum_{j=1}^{M} w_j |p(x_j)|^2 \leq B \|p\|_2.$$

This inequality immediately leads to the following estimate on the condition number of $\boldsymbol{\Phi}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{\Phi}$:

$$\mathrm{cond}(\boldsymbol{\Phi}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{\Phi}) \leq \frac{B}{A}.$$

The constants A and B depend on the degree N, the mesh norm $\delta_{\mathcal{X}} = \max_{x \in \mathbb{X}} d(x, \mathcal{X})$ of the set $\mathcal{X}$, and on the system $\{\phi_k\}$.

If it is necessary to interpolate the data, one might first think of choosing the parameter N such that the dimension of Π_N is equal to M. For this choice, we get a linear system $\boldsymbol{\Phi}\, \boldsymbol{a} = \boldsymbol{f}$ with $\boldsymbol{\Phi} \in \mathbb{C}^{M \times M}$. Even if we assume $\boldsymbol{\Phi}$ to be nonsingular, it turns out that this procedure is usually ill-conditioned. To overcome this obstacle, we choose N to be greater than M and look for a solution of the optimization problem

$$\min_{p \in \Pi_N} \sum_{k=1}^{N} \frac{|a_k|^2}{\hat{w}_k} \quad \text{subject to} \quad p(x_j) = f_j,\ j = 1, \ldots, M.$$

Certainly, we are not completely free in choosing the weights $\hat{w}_k$ but they have to be related in a certain sense to properties of the set of sampling nodes. How this can be done becomes understandable by the following observations. First of all, we see that the optimization problem is equivalent to the normal equation of second kind, i.e.,

$$\boldsymbol{\Phi} \hat{\boldsymbol{W}} \boldsymbol{\Phi}^{\mathsf{H}} \tilde{\boldsymbol{f}} = \boldsymbol{f}, \qquad \hat{\boldsymbol{p}} = \hat{\boldsymbol{W}} \boldsymbol{\Phi}^{\mathsf{H}} \tilde{\boldsymbol{f}}, \tag{1.2}$$

where now $\hat{\boldsymbol{W}} = \operatorname{diag}(\hat{w}_1, \ldots, \hat{w}_N)$. Secondly, an easy computation shows that the matrix entries $(\boldsymbol{\Phi} \hat{\boldsymbol{W}} \boldsymbol{\Phi}^{\mathsf{H}})_{j,l}$ are given in the form $P(x_j, x_l)$, where

$$P(x, y) = \sum_{k=1}^{N} \hat{w}_k\, \phi_k(x)\, \overline{\phi_k(y)}.$$

Note that due to the assumption $\hat{w}_k > 0$, the kernel P is positive definite. Now it is easy to understand how to choose the weights $\hat{w}_k$ in order to assure that $\hat{\boldsymbol{p}}$ can be computed in a stable manner from (1.2). Indeed, with a judicious choice of $\hat{w}_k$'s, we can achieve an excellent localization of the kernel P around the diagonal, i.e., $P(x, x) = 1$ and $P(x, y)$ being small away from the diagonal (see Kunis (2006); Mhaskar and Prestin (2009); Filbir et al. (2008)). Since our goal is to ensure $P(x_j, x_l)$ to be sufficiently small for $j \neq l$, the localization of P, and therefore the $\hat{w}_k$'s, have to be chosen according to the separation distance $q_{\mathcal{X}} = \min_{j \neq l} d(x_j, x_l)$ of the set $\mathcal{X}$. Now we can expect to obtain reasonable estimates for the eigenvalues of $\boldsymbol{\Phi} \hat{\boldsymbol{W}} \boldsymbol{\Phi}^{\mathsf{H}}$ by using Gerschgorin's theorem.

As one might expect, we have to restrict ourselves to concrete settings in order to carry out the program described above. This restriction stems from several facts. Firstly, the MZ inequalities for scattered data with concrete bounds A and B are not available in general. Secondly, it is not easy to come up with well-localized kernels, and, thirdly, algorithms for the fast matrix-vector multiplication are not always available. The methods were successfully applied in several concrete situations like d-dimensional torus Gröchenig (1992); Bass and Gröchenig (2004); Kunis (2006) or the unit sphere Mhaskar et al. (2001); Filbir and Themistoclakis (2008); Keiner et al. (2007); Kunis (2009). Recently, some progress was made in order to establish the theoretical foundation in case of the rotation group $SO(3)$, see Erb and Filbir (2008); Schmid (2008); Gräf and Kunis (2008).

The motivation for the present paper stems from a problem arising in crystallography. In order to get information on the polycrystalline structure of a material, one applies X-ray tomography. The mathematical problem now consists in the inversion of the Radon transform on $SO(3)$. This is usually an ill-posed problem and has rarely been studied with mathematical rigor even though there are several ad hoc methods most of which originate in material science, cf. Bunge (1982); Matthies et al. (1987); Randle and Engler (2000); Kocks et al. (1998). Hielscher et al. (2008) presented a novel approach for this problem. Overall the problem reads as follows. We are given a sampling set $\mathcal{A} = \{(\boldsymbol{x}_j, \boldsymbol{y}_j, f_j) : j = 1, \ldots, M\} \subset \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{C}$, and we have to construct a suitable polynomial defined on $\mathbb{S}^2 \times \mathbb{S}^2$. The aim of the paper is to give a mathematical foundation to this problem in the sense described above in order to explain the numerical observations in Hielscher et al. (2008).

We have organized the paper in the following way. In Section 2 we collect all basic material on the harmonic analysis on $\mathbb{S}^2 \times \mathbb{S}^2$. Section 3 explains shortly the basics for fast computation on $\mathbb{S}^2 \times \mathbb{S}^2$. Subsequently, we focus in Section 4 on the approximation and interpolation of scattered data in our setting. More precisely, we are able to prove the MZ inequality in Theorem 4.5, and hence we are able to estimate the condition number for the normal equation in Corollary 4.6. Furthermore, we focus on the optimal interpolation problem in Theorem 4.10 and estimate the condition number for the normal equation in Corollary 4.11. Finally, Section 5 provides a brief account on the texture analysis problem.

2 Prerequisites

In this section we will collect some basic material as far as it is necessary to understand the paper. Let $\mathbb{S}^2 := \{\boldsymbol{x} \in \mathbb{R}^3 : |\boldsymbol{x}|_2 = 1\}$ be the two-dimensional unit sphere embedded in $\mathbb{R}^3$, where $|\cdot|_2$ denotes the Euclidean norm on $\mathbb{R}^3$. The surface measure on $\mathbb{S}^2$ is denoted by σ, and we assume it is normalized so that

$$\int_{\mathbb{S}^2} \mathrm{d}\sigma = 4\pi. \tag{2.1}$$

The spaces $L^r(\mathbb{S}^2) := L^r(\mathbb{S}^2, \sigma)$, $1 \leq r \leq \infty$, corresponding are defined in the usual manner. The inner product on the Hilbert space $L^2(\mathbb{S}^2)$ is given by

$$\langle f, g \rangle = \int_{\mathbb{S}^2} f(\boldsymbol{x})\, \overline{g(\boldsymbol{x})}\, \mathrm{d}\sigma(\boldsymbol{x}), \quad f, g \in L^2(\mathbb{S}^2). \tag{2.2}$$

Throughout the paper we adopt the convention to denote the vectors in boldface type and the scalars in $\mathbb{R}$ in simple type.

Using polar coordinates for a representation of the sphere $\mathbb{S}^2$, for a point $\boldsymbol{x} \in \mathbb{S}^2$, we have the coordinate relation $\boldsymbol{x} = (\sin\vartheta \sin\varphi, \sin\vartheta\cos\varphi, \cos\vartheta)^\top$, where $(\vartheta, \varphi) \in [0, \pi] \times [-\pi, \pi)$. According to this parametrization, the surface measure is given by

$$\mathrm{d}\sigma(\boldsymbol{x}) = \sin\vartheta \; \mathrm{d}\varphi\, \mathrm{d}\vartheta, \quad \varphi \in [-\pi, \pi),\ \vartheta \in [0, \pi], \tag{2.3}$$

or, equivalently, by setting $t = \cos\vartheta$

$$\mathrm{d}\sigma(\boldsymbol{x}) = -\mathrm{d}\varphi\,\mathrm{d}t \quad \varphi \in [-\pi, \pi)\,, t \in [-1, 1]. \tag{2.4}$$

For any function $\phi \in L^1([-1, 1])$ and any $\boldsymbol{y} \in \mathbb{S}^2$, we have

$$\int_{\mathbb{S}^2} \phi(\boldsymbol{x} \cdot \boldsymbol{y})\ \mathrm{d}\sigma(\boldsymbol{x}) = 2\pi \int_{-1}^{1} \phi(t)\ \mathrm{d}t. \tag{2.5}$$

Here and in what follows, $\boldsymbol{x} \cdot \boldsymbol{y}$ denotes the usual inner product on $\mathbb{R}^3$.

Let $n \geq 0$ be a fixed integer. The restriction of a harmonic homogeneous polynomial of degree N to the unit sphere $\mathbb{S}^2$ is called a spherical harmonic polynomial of degree N. By Π_N we denote the space of all spherical polynomials of degree *at most* N and $\mathcal{H}_n$ denotes the space of spherical polynomials of *precise* degree n. The spaces $\mathcal{H}_n$ are mutually orthogonal, and we obviously have $\Pi_N = \oplus_{n=0}^{N} \mathcal{H}_n$. Moreover, it is well known that

$$L^2(\mathbb{S}^2) = c\ell \oplus_{n=0}^{\infty} \mathcal{H}_n. \tag{2.6}$$

An orthonormal basis for the space $\mathcal{H}_n$ can be constructed in a concrete form using Legendre polynomials. Let $(R_n)_{n\in\mathbb{N}_0}$ be the sequence of Legendre polynomials defined by

$$R_n(x) = F\left(-n, n+1, 1; \frac{1-x}{2}\right), \tag{2.7}$$

where $F(a, b, c; x)$ is the Gaussian hypergeometric function. Note that $R_n(1) = 1$ and, moreover, $\int_{-1}^{1} R_n(t)\, R_m(t)\ \mathrm{d}t = 2\,(2n+1)^{-1}\delta_{n,m}$. The closely related Legendre functions R_n^k, $k, n \in \mathbb{N}_0$, $n \leq k$, are defined as

$$R_n^k(t) = \sqrt{\frac{(n-k)!}{(n+k)!}}(1-t^2)^{k/2}\ \frac{\mathrm{d}^k}{\mathrm{d}t^k} R_n(t), \quad k \leq n. \tag{2.8}$$

We now define $Y_{n,k} : \mathbb{S}^2 \to \mathbb{C}$ by

$$Y_{n,k}(\boldsymbol{x}) = Y_{n,k}(\vartheta, \varphi) = \sqrt{\frac{2n+1}{4\pi}}\, R_n^{|k|}(\cos\vartheta)\ \mathrm{e}^{\mathrm{i}k\varphi}, \quad n \in \mathbb{N}_0,\ |k| \leq n. \tag{2.9}$$

These functions are simply called *spherical harmonics*. The set $\{Y_{n,k} : k = -n, \ldots, n\}$ is an orthonormal basis of the space $\mathcal{H}_n$. This fact and (2.6) give

$$\int_{\mathbb{S}^2} Y_{n,k}(\boldsymbol{x})\overline{Y_{m,l}(\boldsymbol{x})}\ \mathrm{d}\sigma(\boldsymbol{x}) = \int_{-\pi}^{\pi} \int_{0}^{\pi} Y_{n,k}(\vartheta, \varphi))\overline{Y_{m,l}(\vartheta, \varphi)}\ \mathrm{d}\vartheta\,\mathrm{d}\varphi = \delta_{n,m}\delta_{k,l}. \tag{2.10}$$

Consequently we obtain $\dim \mathcal{H}_n = 2n + 1$ and $\dim \Pi_N = (N + 1)^2$.

The following addition formula is of fundamental importance. It relates the spherical harmonics and the Legendre polynomials in the following way

$$\sum_{k=-n}^{n} Y_{n,k}(\boldsymbol{x})\overline{Y_{n,k}(\boldsymbol{y})} = \frac{2n+1}{4\pi}\, R_n(\boldsymbol{x} \cdot \boldsymbol{y}). \tag{2.11}$$

We consider the product space $\mathbb{S}^2 \times \mathbb{S}^2$ with product measure $\sigma \times \sigma$. To keep the notation simple we will also often use the symbol τ for the product measure $\sigma \times \sigma$. The L^r-spaces are defined accordingly and we have $L^r(\mathbb{S}^2 \times \mathbb{S}^2) = L^r(\mathbb{S}^2) \otimes L^r(\mathbb{S}^2)$, where $\otimes$ denotes the tensor product. For the Hilbert space $L^2(\mathbb{S}^2 \times \mathbb{S}^2)$ we have the orthonormal basis

$$Y_{n_1,k_1} \otimes Y_{n_2,k_2}(\boldsymbol{x}, \boldsymbol{y}) = Y_{n_1,k_1}(\boldsymbol{x}) Y_{n_2,k_2}(\boldsymbol{y}), \quad n_1, n_2 \in \mathbb{N}_0,\ |k_1| \leq n_1,\ |k_2| \leq n_2. \tag{2.12}$$

To keep the notation simple we write $S_{\boldsymbol{n},\boldsymbol{k}}$ for $Y_{n_1,k_1} \otimes Y_{n_2,k_2}$ with $\boldsymbol{n} = (n_1, n_2)$ and $\boldsymbol{k} = (k_1, k_2)$. Throughout this paper the function $S_{\boldsymbol{n},\boldsymbol{k}}$ will be called *tensor spherical harmonics*. We obviously have the following orthogonality relation

$$\int_{\mathbb{S}^2 \times \mathbb{S}^2} S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}, \boldsymbol{y}) S_{\boldsymbol{m},\boldsymbol{l}}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\tau(\boldsymbol{x}, \boldsymbol{y}) = \delta_{n_1,n_2} \delta_{k_1,k_2} \delta_{m_1,m_2} \delta_{l_1,l_2}. \tag{2.13}$$

Throughout the paper we will follow the convention that boldface indices are elements of $\mathbb{N}^2$ resp. $\mathbb{Z}^2$. Corresponding to (2.13) the spaces

$$\mathcal{H}_{\boldsymbol{n}} := \mathcal{H}_{n_1} \otimes \mathcal{H}_{n_2} = \operatorname{span}\{S_{\boldsymbol{n},\boldsymbol{k}} : |k_i| \leq n_i,\ i = 1, 2\}$$

respectively

$$\Pi_{\boldsymbol{N}} := \Pi_{N_1} \otimes \Pi_{N_2} = \operatorname{span}\{S_{\boldsymbol{n},\boldsymbol{k}} : n_1 \leq N_1,\ n_2 \leq N_2,\ |k_i| \leq n_i,\ i = 1, 2\}$$

contain the tensor spherical polynomials of total degree $|\boldsymbol{n}| = n_1 + n_2$ respectively of total degree at most $|\boldsymbol{N}| = N_1 + N_2$.

For the dimension of these spaces we obtain $\dim \mathcal{H}_{\boldsymbol{n}} = (2n_1 + 1)(2n_2 + 1)$ resp. $\dim \Pi_{\boldsymbol{N}} = (N_1 + 1)^2 \, (N_2 + 1)^2$. To keep the notational effort in reasonable bounds we introduce the following abbreviations for index sets

$$\begin{aligned} I_{\boldsymbol{N}} &= \{(\boldsymbol{n}, \boldsymbol{k}) \in \mathbb{N}_0^2 \times \mathbb{Z}^2 : n_1 \leq N_1,\ n_2 \leq N_2,\ |k_1| \leq n_1,\ |k_2| \leq n_2\}, \\ I_\infty &= \{(\boldsymbol{n}, \boldsymbol{k}) \in \mathbb{N}_0^2 \times \mathbb{Z}^2 : |k_1| \leq n_1,\ |k_2| \leq n_2\}. \end{aligned} \tag{2.14}$$

For a function $f \in L^1(\mathbb{S}^2 \times \mathbb{S}^2)$ the Fourier coefficients w.r.t. the tensor spherical harmonics are given by

$$\hat{f}_{\boldsymbol{n},\boldsymbol{k}} = \int_{\mathbb{S}^2 \times \mathbb{S}^2} f(\boldsymbol{x}, \boldsymbol{y}) \, \overline{S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}, \boldsymbol{y})} \, \mathrm{d}\tau(\boldsymbol{x}, \boldsymbol{y}), \quad (\boldsymbol{n}, \boldsymbol{k}) \in I_\infty. \tag{2.15}$$

and hence

$$f \sim \sum_{(\boldsymbol{n},\boldsymbol{k}) \in I_\infty} \hat{f}_{\boldsymbol{n},\boldsymbol{k}} \, S_{\boldsymbol{n},\boldsymbol{k}}. \tag{2.16}$$

Clearly, for $p \in \Pi_{\boldsymbol{N}}$ we have $p(\boldsymbol{x}, \boldsymbol{y}) = \sum_{(\boldsymbol{n},\boldsymbol{k}) \in I_{\boldsymbol{N}}} \hat{p}_{\boldsymbol{n},\boldsymbol{k}} \, S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}, \boldsymbol{y})$. The projection of $f \in L^2(\mathbb{S}^2 \times \mathbb{S}^2)$ onto $\mathcal{H}_{\boldsymbol{n}}$ is given by

$$\mathcal{P}_{\boldsymbol{n}} f(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k_1=-n_1}^{n_1} \sum_{k_2=-n_2}^{n_2} \left[\int_{\mathbb{S}^2 \times \mathbb{S}^2} f(\boldsymbol{u}, \boldsymbol{v}) \, \overline{S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{u}, \boldsymbol{v})} \, \mathrm{d}\tau(\boldsymbol{u}, \boldsymbol{v}) \right] S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}, \boldsymbol{y}). \tag{2.17}$$

As an easy consequence of (2.11) we get

$$\sum_{k_1=-n_1}^{n_1} \sum_{k_2=-n_2}^{n_2} S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x},\boldsymbol{y})\overline{S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{u},\boldsymbol{v})} = \frac{(2n_1+1)(2n_2+1)}{16\pi^2} R_{n_1}(\boldsymbol{x}\cdot\boldsymbol{u}) R_{n_2}(\boldsymbol{y}\cdot\boldsymbol{v}) \quad (2.18)$$

which in turn allows us to write the projection operator (2.17) in the form

$$\mathcal{P}_{\boldsymbol{n}} f(\boldsymbol{x},\boldsymbol{y}) = \frac{(2n_1+1)(2n_2+1)}{16\pi^2} \int_{\mathbb{S}^2}\int_{\mathbb{S}^2} f(\boldsymbol{u},\boldsymbol{v})\, R_{n_1}(\boldsymbol{x}\cdot\boldsymbol{u})\, R_{n_2}(\boldsymbol{y}\cdot\boldsymbol{v})\; \mathrm{d}\sigma(\boldsymbol{u})\,\mathrm{d}\sigma(\boldsymbol{v}). \quad (2.19)$$

The projection of f onto $\Pi_{\boldsymbol{N}}$ is given by

$$\begin{aligned} \mathcal{S}_{\boldsymbol{N}} f(\boldsymbol{x},\boldsymbol{y}) &= \sum_{n_1=1}^{N_1}\sum_{n_2=1}^{N_2} \mathcal{P}_{\boldsymbol{n}} f(\boldsymbol{x},\boldsymbol{y}) \\ &= \frac{1}{4\pi^2}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2} f(\boldsymbol{u},\boldsymbol{v})\, K_{N_1}(\boldsymbol{x}\cdot\boldsymbol{u},1)\, K_{N_2}(\boldsymbol{y}\cdot\boldsymbol{v},1)\; \mathrm{d}\sigma(\boldsymbol{u})\,\mathrm{d}\sigma(\boldsymbol{v}), \end{aligned} \quad (2.20)$$

where

$$K_n(s,t) = \sum_{k=0}^{n} \|R_k\|_2^{-2}\, R_k(s)\, R_k(t) \quad (2.21)$$

is the Darboux kernel corresponding to the Legendre polynomials. In the following we will simply write $K_n(s)$ for $K_n(s,1)$. Using the formulas (4.5.3) and (4.1.1) in Szegő (1975), we have

$$K_n(s) = \frac{(n+1)^2}{2}\, R_n^{(1,0)}(s), \quad (2.22)$$

where $R_n^{(1,0)}$ are the Jacobi polynomials with respect to the weight $(1-s)$. From (2.22) we get that $K_n(1) = \frac{(n+1)^2}{2}$.

Since we are dealing with scattered data on $\mathbb{S}^2\times\mathbb{S}^2$, it is necessary to introduce a metric on the set $\mathbb{S}^2\times\mathbb{S}^2$. In order to attain this, we start with the canonical metric on the sphere $\mathbb{S}^2$ which is defined by the geodesic distance. More precisely, for $\boldsymbol{x},\boldsymbol{y}\in\mathbb{S}^2$, the geodesic distance between these points is given by $d_0(\boldsymbol{x},\boldsymbol{y}) = \arccos(\boldsymbol{x}\cdot\boldsymbol{y})$. The geodesic distance is related to the Euclidean distance of the points $\boldsymbol{x},\boldsymbol{y}$ by $|\boldsymbol{x}-\boldsymbol{y}|_2 = 2-2\cos(d_0(\boldsymbol{x},\boldsymbol{y}))$. A family of product metrics on $\mathbb{S}^2\times\mathbb{S}^2$ can be defined by

$$\begin{aligned} d^{(s)}((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v})) &= \Big(d_0(\boldsymbol{x},\boldsymbol{u})^s + d_0(\boldsymbol{y},\boldsymbol{v})^s\Big)^{1/s}, \quad 1\le s<\infty, \\ d^{(\infty)}((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v})) &= \max\{d_0(\boldsymbol{x},\boldsymbol{u}), d_0(\boldsymbol{y},\boldsymbol{v})\}. \end{aligned} \quad (2.23)$$

Note that $d^{(1)}$ is the Riemannian metric on the product manifold $\mathbb{S}^2\times\mathbb{S}^2$ and, obviously, $d^{(1)}((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v})) \le d^{(\infty)}((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v}))$.

Now let $\mathcal{X} = \{(\boldsymbol{x}_j,\boldsymbol{y}_j)\in\mathbb{S}^2\times\mathbb{S}^2 : j=1,\ldots,M\}$ be a finite subset of points. In order to measure how uniform the points of $\mathcal{X}$ are distributed on $\mathbb{S}^2\times\mathbb{S}^2$, we introduce the terms *mesh norm* $\delta_{\mathcal{X}}$ and *separation distance* $q_{\mathcal{X}}$ according to (2.23). They are defined respectively as

$$\begin{aligned}\delta_{\mathcal{X}}^{(s)} &:= \max_{(\boldsymbol{x},\boldsymbol{y})\in\mathbb{S}^2\times\mathbb{S}^2}\ \min_{j=1,\dots,M} d^{(s)}((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{x}_j,\boldsymbol{y}_j)),\\ q_{\mathcal{X}}^{(s)} &:= \min_{j\neq l} d^{(s)}((\boldsymbol{x}_j,\boldsymbol{y}_j),(\boldsymbol{x}_l,\boldsymbol{y}_l))\,.\end{aligned} \tag{2.24}$$

Since we will work here mostly with the case $s=\infty$, we will ignore the superscript and write simply $\delta_{\mathcal{X}}$ resp. $q_{\mathcal{X}}$ instead of $\delta_{\mathcal{X}}^{(\infty)}$ resp. $q_{\mathcal{X}}^{(\infty)}$.

There is a natural decomposition of $\mathbb{S}^2\times\mathbb{S}^2$ coming along with the set $\mathcal{X}$. This decomposition consists of a finite collection $\mathcal{R}$ of closed regions $R_k\in\mathbb{S}^2\times\mathbb{S}^2,\ k=1,\dots,M$, with no common interior point, covering $\mathbb{S}^2\times\mathbb{S}^2$, and with exactly one point from $\mathcal{X}$ in every single R_k. We will refer to such a decomposition as an *admissible decomposition* of $\mathbb{S}^2\times\mathbb{S}^2$ w.r.t. $\mathcal{X}$. The partition s-norm of an admissible decomposition of $\mathbb{S}^2\times\mathbb{S}^2$ is defined as

$$\|\mathcal{R}\|_s := \max_{R\in\mathcal{R}}\left[\max_{(\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v})\in R} d^{(s)}((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v}))\right] \tag{2.25}$$

In the case $s=\infty$, we simply write $\|\mathcal{R}\|$.

3 Discrete spherical tensor Fourier transform on $\mathbb{S}^2\times\mathbb{S}^2$

In the present section we describe briefly the mathematical background of the discrete tensor Fourier transform and its fast realization. We consider a tensor spherical polynomial $p\in\Pi_{\boldsymbol{N}}$, i.e.,

$$p(\boldsymbol{x},\boldsymbol{y}) = \sum_{(\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}} \hat{p}_{\boldsymbol{n},\boldsymbol{k}}\ S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x},\boldsymbol{y}). \tag{3.1}$$

and a finite set of points $\mathcal{X}=\{(\boldsymbol{x}_j,\boldsymbol{y}_j): j=1,\dots,M\}$. Assume that the Fourier coefficients $\hat{p}_{\boldsymbol{n},\boldsymbol{k}}\in\mathbb{C},\ (\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}$, are given. Then the *discrete Fourier transform* is defined by

$$p(\boldsymbol{x}_j,\boldsymbol{y}_j) = \sum_{(\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}} \hat{p}_{\boldsymbol{n},\boldsymbol{k}}\ S_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}_j,\boldsymbol{y}_j),\quad j=1,\dots,M. \tag{3.2}$$

If we define the *tensor spherical Fourier matrix* by

$$\begin{aligned}\boldsymbol{\mathcal{Y}} &:= \left(S_{\boldsymbol{n},\boldsymbol{k}}\left(\boldsymbol{x}_j,\boldsymbol{y}_j\right)\right)_{j=1,\dots,M;(\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}}\\ &= \left(Y_{n_1,k_1}\left(\boldsymbol{x}_j\right)Y_{n_2,k_2}\left(\boldsymbol{y}_j\right)\right)_{j=1,\dots,M;(\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}} \in \mathbb{C}^{M\times(N_1+1)^2(N_2+1)^2},\end{aligned}$$

and the vectors $\boldsymbol{p}=(p(\boldsymbol{x}_j,\boldsymbol{y}_j))_{j=1,\dots,M}$, $\hat{\boldsymbol{p}}=(\hat{p}_{\boldsymbol{n},\boldsymbol{k}_{(\boldsymbol{n},\boldsymbol{k})\in I_n}})$, then (3.2) can be written as

$$\boldsymbol{p} = \boldsymbol{\mathcal{Y}}\,\hat{\boldsymbol{p}}. \tag{3.3}$$

This establishes a linear map from $\mathbb{C}^{(N_1+1)^2(N_2+1)^2}$ to $\mathbb{C}^M$. Its adjoint map $\mathbb{C}^M \to \mathbb{C}^{(N_1+1)^2(N_2+1)^2}$ defined as

$$\tilde{\boldsymbol{c}} = \boldsymbol{\mathcal{Y}}^{\mathsf{H}} \boldsymbol{c}$$

is called *adjoint discrete tensor spherical Fourier transform.*

A naive implementation of the discrete spherical tensor Fourier transform and of its adjoint transform for $M \in \mathbb{N}$ arbitrary nodes and for polynomial degree $\boldsymbol{N} = (N_1, N_2) \in \mathbb{N}^2$ requires $\mathcal{O}(N_1^2 N_2^2 M)$ floating point operations. Fortunately, this can be done in a much more efficient way.

Let

$$g_k(t) := \sum_{n=k}^{N} \hat{f}_{n,k} R_n^{|k|}(t)$$

for k even and

$$g_k(t) := \frac{1}{\sqrt{1-t^2}} \sum_{n=k}^{N} \hat{f}_{n,k} R_n^{|k|}(t)$$

for k odd. Note that g_k is a polynomial of degree N or $N-1$, respectively. Now we perform a change of basis via fast polynomial transform (see Potts et al. (1998)) to obtain the representation

$$g_k(\cos\vartheta) = \sum_{n=-N}^{N} c_{n,k} \mathrm{e}^{\mathrm{i}k\vartheta}, \tag{3.4}$$

with coefficients $c_{n,k} \in \mathbb{C}$ and $t = \cos(\vartheta)$. This transform takes only $\mathcal{O}(N \log^2 N)$ arithmetical operations. Now we consider (3.2) again. Rearranging the summation and using (2.9) gives us

$$\begin{aligned} p(\boldsymbol{x}_j, \boldsymbol{y}_j) = & \sum_{k_1=-N_1}^{N_1} \sum_{n_1=|k_1|}^{N_1} \sum_{k_2=-N_2}^{N_2} \sum_{n_2=|k_2|}^{N_2} \hat{p}_{n_1,n_2,k_1,k_2} \\ & \times \sqrt{\frac{2n_1+1}{4\pi}} \sqrt{\frac{2n_2+1}{4\pi}} R_{n_1}^{|k_1|}(\cos\theta_j) \mathrm{e}^{\mathrm{i}k_1\phi_j} R_{n_2}^{|k_2|}(\cos\vartheta_j) \mathrm{e}^{\mathrm{i}k_2\varphi_j}, \end{aligned} \tag{3.5}$$

where $\boldsymbol{x}_j = (\sin\theta_j \sin\phi_j, \sin\theta_j \cos\phi_j, \cos\theta_j)^\top$, $\boldsymbol{y}_j = (\sin\vartheta_j \sin\varphi_j, \sin\vartheta_j \cos\varphi_j, \cos\vartheta_j)^\top$. After rearranging the summation again, we apply the change of basis as described above two times to arrive at

$$p(\boldsymbol{x}_j, \boldsymbol{y}_j) = \sum_{k_1=-N_1}^{N_1} \sum_{n_1=-N_1}^{N_1} \sum_{k_2=-N_2}^{N_2} \sum_{n_2=-N_2}^{N_2} c_{n_1,n_2,k_1,k_2} \mathrm{e}^{\mathrm{i}n_1\theta_j} \mathrm{e}^{\mathrm{i}k_1\phi_j} \mathrm{e}^{\mathrm{i}n_2\vartheta_j} \mathrm{e}^{\mathrm{i}k_2\varphi_j}. \tag{3.6}$$

An evaluation of the coefficients c_{n_1,n_2,k_1,k_2} from the coefficients $\hat{p}_{n_1,n_2,k_1,k_2}$ can be realized by the fast polynomial transform and takes only $\mathcal{O}(N_1^2 N_2^2 (\log^2 N_1 + \log^2 N_2))$ arithmetical operations. Finally, the evaluation of $p(\boldsymbol{x}_j, \boldsymbol{y}_j)$, $j = 1, \ldots, M$, can be realized by the NFFT algorithm with $\mathcal{O}(N_1^2 N_2^2 (\log N_1 + \log N_2) + \log^4(1/\varepsilon) M)$ arithmetical operations, where ε is a prescribed accuracy. We refer to this algorithm as the *nonequispaced fast spherical tensor Fourier transform*. A fast algorithm for the adjoint transform, i.e., for the fast matrix-vector multiplication for $\boldsymbol{\mathcal{Y}}^{\mathsf{H}}\boldsymbol{c}$ follows in a straightforward way as suggested in case of the sphere $\mathbb{S}^2$, see Keiner and Potts (2008).

Algorithm 3.1. Input*: $\hat{p}_{\boldsymbol{n},\boldsymbol{k}} \in \mathbb{C}$ Fourier coefficients, $\mathcal{X}$ grid*
1. Compute from the given values $\hat{p}_{\boldsymbol{n},\boldsymbol{k}} \in \mathbb{C}$, $(\boldsymbol{k},\boldsymbol{n}) \in I_{N_1,N_2}$, the coefficients $c_{n_1,n_2,k_2,k_2} \in \mathbb{C}$ in (3.6) with the fast polynomial transform in $\mathcal{O}(N_1^2 N_2^2(\log^2 N_1 + \log^2 N_2))$.
2. Compute, for $j = 1,\dots,M$, the values $p(\boldsymbol{x}_j,\boldsymbol{y}_j) \in \mathbb{C}$ by the NFFT in $\mathcal{O}(N_1^2 N_2^2(\log N_1 + \log N_2) + \log^4(1/\varepsilon)M)$.
Output*: $p(\boldsymbol{x}_j,\boldsymbol{y}_j) \in \mathbb{C}, j = 1,\dots,M$.*

4 Approximation and interpolation from scattered data

We are now going to consider the main problem of the algorithmic construction of a polynomial model for a given data set. This means that we try to find, for a given data set $\mathcal{A} = \{(\boldsymbol{x}_j,\boldsymbol{y}_j,f_j) \in \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{C} : j = 1,\dots,M\}$, a polynomial $p \in \Pi_{\boldsymbol{N}}$ which approximates $(p(\boldsymbol{x}_j,\boldsymbol{y}_j) \approx f_j)$ or interpolates $(p(\boldsymbol{x}_j,\boldsymbol{y}_j) = f_j)$ the data. This is obviously equivalent to the problem of computing the Fourier coefficients using the data set $\mathcal{A}$. Applying the notation of the previous section, the problem consists of finding a vector $\hat{\boldsymbol{p}}$ with $\boldsymbol{\mathcal{Y}}\hat{\boldsymbol{p}} = \boldsymbol{f}$, where $\boldsymbol{f} = (f_j)_{j=1,\dots,M}$. Clearly the solvability of this problem depends on the relation between the number of data points and the degree of the polynomial (which reflects the complexity of our model). In case $M > \dim \Pi_{\boldsymbol{N}} = (N_1+1)^2(N_2+1)^2$, the problem is overdetermined, and our aim is in general not achievable. The best we can do is to find an approximation, i.e., we are looking for a polynomial with $p(\boldsymbol{x}_j,\boldsymbol{y}_j) \approx f_j$ or, equivalently, $\boldsymbol{\mathcal{Y}}\hat{\boldsymbol{p}} \approx \boldsymbol{f}$. In the underdetermined case $M < \dim \Pi_{\boldsymbol{N}} = (N_1+1)^2(N_2+1)^2$, we will look for an interpolation as described above. In the following we investigate both cases and give conditions for $\boldsymbol{\mathcal{Y}}$ to have full rank. In both cases we study the condition numbers of the problem related to the normal equation.

4.1 Least square approximation

In the first part we concentrate on the overdetermined case of the problem, i.e., $M > (N_1+1)^2(N_2+1)^2$. In general the given data $\boldsymbol{f} \in \mathbb{C}^M$ can only be approximated up to some residual $\boldsymbol{r} := \boldsymbol{f} - \boldsymbol{\mathcal{Y}}\hat{\boldsymbol{p}}$. 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. This means that we consider the weighted least squares problem

$$\|\boldsymbol{f} - \boldsymbol{\mathcal{Y}}\hat{\boldsymbol{p}}\|_{\boldsymbol{W}}^2 = \sum_{j=1}^{M} w_j |f_j - p(\boldsymbol{x}_j,\boldsymbol{y}_j)|^2 \xrightarrow{\hat{\boldsymbol{p}}} \min, \tag{4.1}$$

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

Lemma 4.1. *The least squares problem* (4.1) *is equivalent to the normal equation of first kind*

$$\boldsymbol{\mathcal{Y}}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{\mathcal{Y}}\hat{\boldsymbol{p}} = \boldsymbol{\mathcal{Y}}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{f}. \tag{4.2}$$

This assertion is due to (Björck, 1996, Thm. 1.1.2) for the matrix $\boldsymbol{W}^{1/2}\boldsymbol{\mathcal{Y}}$.

The stability analysis of the problem is based on the MZ inequalities. Roughly speaking, MZ inequalities compare continuous r-norms with discrete r-norms of a function. In case of scattered sampling points on the unit sphere in $\mathbb{R}^d$, MZ inequalities were first established by Mhaskar, Narcowich, and Ward in Mhaskar et al. (2001). Based on this work, a closer inspection was presented in Filbir and Themistoclakis (2008) which leads to concrete bounds of the MZ constants. We will show in the following presentation how to derive the Marcinkiewicz–Zygmund inequalities in the tensor product case. Although the main ideas are similar to the classical case, the constants differ slightly. Moreover, we present the material to keep the paper self-contained.

A fundamental role in our considerations will be played by certain kernels. We will now give the construction principle for those kernels. In case of univariate polynomial approximation on the unit interval $[-1, 1]$, these kernels have been studied before in Filbir and Themistoclakis (2004, 2008). We will present the material here in a slightly modified form suitable for our purposes in connection with fast algorithms on $\mathbb{S}^2 \times \mathbb{S}^2$.

We start with a reproducing kernel of de la Vallée Poussin type in the univariate case (see Filbir and Themistoclakis (2008)). For $n, m \in \mathbb{N}_0$ with $m \leq n$ let

$$V_m^n(t) = 2\,\frac{K_n(t)\,K_m(t)}{K_m(1)}, \tag{4.3}$$

where $K_n(t) = K_n(t, 1)$ and $K_n(\cdot, \cdot)$ is the Darboux kernel as defined in (2.21). Now let $\boldsymbol{n}, \boldsymbol{m} \in \mathbb{N}_0^2,\ \boldsymbol{n} = (n_1, n_2), \boldsymbol{m} = (m_1, m_2)$, with $m_1 \leq n_1,\ m_2 \leq n_2$ and define a tensor version of (4.3) in the usual way by

$$V_{\boldsymbol{m}}^{\boldsymbol{n}}(s,t) = V_{m_1}^{n_1}(s)\,V_{m_2}^{n_2}(t) = 4\,\frac{K_{n_1}(s)\,K_{m_1}(s)}{K_{m_1}(1)}\,\frac{K_{n_2}(t)\,K_{m_2}(t)}{K_{m_2}(1)}.$$

The following estimates are now an easy consequence of the inequalities presented in Section 3.1 in Filbir and Themistoclakis (2008).

Theorem 4.2. *For all $\boldsymbol{n}, \boldsymbol{m} \in \mathbb{N}_0^2$, with $m_1 \leq n_1,\ m_2 \leq n_2$, we have*

(i) $\|V_{\boldsymbol{m}}^{\boldsymbol{n}}\|_1 \leq 4\,\frac{(n_1+1)(n_2+1)}{(m_1+1)(m_2+1)}\sqrt{\frac{m_1!\,m_2!}{n_1!\,n_2!}}$

(ii) $\max\limits_{|s|,|t|\leq 1} |V_{\boldsymbol{m}}^{\boldsymbol{n}}(s,t)| \leq (n_1+2)^2\,(n_2+2)^2$.

In order to construct an operator sequence for approximating functions defined on $\mathbb{S}^2 \times \mathbb{S}^2$ we make the following choices for the parameters $\boldsymbol{n}, \boldsymbol{m} \in \mathbb{N}_0^2$. For $N_i \in \mathbb{N},\ i =$

1, 2, let

$$n_i = \begin{cases} \frac{3N_i}{2} & \text{if } N_i \text{ even}, \\ \frac{3N_i-1}{2} & \text{if } N_i \text{ odd}, \end{cases} \qquad m_i = \begin{cases} \frac{N_i}{2} & \text{if } N_i \text{ even}, \\ \frac{N_i-1}{2} & \text{if } N_i \text{ odd}, \end{cases} \tag{4.4}$$

and define

$$v_{\boldsymbol{N}}(s,t) = v_{N_1}(s)\, v_{N_2}(t), \quad \boldsymbol{N} = (N_1, N_2), \tag{4.5}$$

where $v_{N_i}(x) = \frac{1}{4\pi} V_{m_i}^{n_i}(x)$, $i = 1, 2$. Clearly, v_{N_i} is a polynomial of degree $2N_i - \chi_{N_i}$ for $i = 1, 2$, where χ_{N_i} is 0 if N_i is even and 1 if N_i is odd. Thus, we get

$$v_{\boldsymbol{N}}(s,t) = \frac{1}{16\pi^2} \sum_{k=0}^{2N_1-\chi_{N_1}} \sum_{l=0}^{2N_2-\chi_{N_2}} a_{N_1,k} a_{N_2,l} \|R_{N_1}\|_2^{-2} \|R_{N_2}\|_2^{-2} R_{N_1}(s)\, R_{N_2}(t). \tag{4.6}$$

The coefficients $a_{N_1,k}$ resp. $a_{N_2,l}$ can be calculated explicitly, see Filbir and Themistoclakis (2008). From (Filbir and Themistoclakis, 2008, equations (3.13), (3.14)), we obtain

$$\|v_{\boldsymbol{N}}\|_1 = \|v_{N_1}\|_1 \, \|v_{N_2}\|_1 \leq \frac{9}{4\pi^2} \quad \text{for all } \boldsymbol{N}, \tag{4.7}$$

and

$$\max_{|s|,|t|\leq 1} |v_{\boldsymbol{N}}(s,t)| \leq \max_{|s|\leq 1} |v_{N_1}(s)| \max_{|t|\leq 1} |v_{N_2}(t)| \leq \frac{N_1^2\, N_2^2}{\pi^2} \leq \frac{\tilde{N}^4}{\pi^2}, \quad \text{for all } \tilde{N} \geq 4, \tag{4.8}$$

where $\tilde{N} = \max\{N_1, N_2\}$.

Corresponding to $v_{\boldsymbol{N}}$ we define the operator

$$\mathcal{V}_{\boldsymbol{N}} f(\boldsymbol{x}, \boldsymbol{y}) = \int_{\mathbb{S}^2\times\mathbb{S}^2} f(\boldsymbol{u}, \boldsymbol{v})\, v_{\boldsymbol{N}}(\boldsymbol{x}\cdot\boldsymbol{u}, \boldsymbol{y}\cdot\boldsymbol{v})\; \mathrm{d}\tau(\boldsymbol{u}, \boldsymbol{v}), \tag{4.9}$$

which is the tensor version of the generalized de la Vallée Poussin operator studied in Filbir and Themistoclakis (2008). The tensor de la Vallée Poussin operator has the following properties.

Theorem 4.3. (i) *For all $\boldsymbol{N} \in \mathbb{N}_0^2$ we have*

$$\mathcal{V}_{\boldsymbol{N}} p = p \quad \textit{for every } p \in \Pi_{\boldsymbol{N}}.$$

(ii) *For all $\boldsymbol{N} \in \mathbb{N}_0^2$ and any $1 \leq r \leq \infty$, we have*

$$\|\mathcal{V}_{\boldsymbol{N}} f\|_r \leq 9 \|f\|_r \quad \textit{for all } f \in L^r(\mathbb{S}^2 \times \mathbb{S}^2).$$

The proofs follow easily from (Filbir and Themistoclakis, 2008, Proposition 3.4 resp. Theorem 3.5.).

We now make use of the de la Vallée Poissin operator in order to establish the MZ inequalities for scattered points on $\mathbb{S}^2 \times \mathbb{S}^2$. Let again $\mathcal{X} = \{(\boldsymbol{x}_j, \boldsymbol{y}_j) \in \mathbb{S}^2 \times \mathbb{S}^2 : j =$

$1, \ldots, M\}$ be a set of sampling points with corresponding tensor decomposition $\mathcal{R}$. The appropriate discrete r-norms of a function $f : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{C}$ are defined as

$$\|f\|_{\mathcal{X},r} := \begin{cases} \left(\sum_{j=1}^{M} |f(\boldsymbol{x}_j, \boldsymbol{y}_j)|^r \, \tau(R_j) \right)^{1/r} & \text{if } 1 \le r < \infty, \\ \max_{j=1,\ldots,M} |f(\boldsymbol{x}_j, \boldsymbol{y}_j)| & \text{if } r = \infty. \end{cases} \tag{4.10}$$

The $\tau(R_j)$ is referred to as the weight of the region R_j and will be later on also denoted by w_j.

For the proof of the MZ inequalities, we need the following lemma which establishes a Markov-type inequality for polynomials from $\Pi_{\boldsymbol{N}}$.

Lemma 4.4. *Let $p \in \Pi_{\boldsymbol{N}}$. Then for $(\boldsymbol{x}, \boldsymbol{y}), (\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{S}^2 \times \mathbb{S}^2$ and $s = 1$ or $s = \infty$ we have*

$$|p(\boldsymbol{x}, \boldsymbol{y}) - p(\boldsymbol{u}, \boldsymbol{v})| \le (2 - \delta_{1,s}) \tilde{N} \, d^{(s)}((\boldsymbol{x}, \boldsymbol{y}), (\boldsymbol{u}, \boldsymbol{v})) \, \|p\|_{\mathbb{S}^2 \times \mathbb{S}^2, \infty}, \tag{4.11}$$

where $\delta_{1,s}$ denotes the delta function which is 1 at $s = 1$ and 0 otherwise.

Proof. The proof follows from the Markov inequality for spherical polynomials Q on $\mathbb{S}^2$, see Jetter et al. (1999). For $q \in \Pi_N$ and $\boldsymbol{x}, \boldsymbol{u} \in \mathbb{S}^2$ this inequality reads as follows

$$|q(\boldsymbol{x}) - q(\boldsymbol{u})| \le N \, d_0(\boldsymbol{x}, \boldsymbol{u}) \, \|q\|_{\mathbb{S}^2, \infty}. \tag{4.12}$$

Now let $p \in \Pi_{\boldsymbol{N}}$. Then for fixed $\boldsymbol{y} \in \mathbb{S}^2$, the polynomial $p_{\boldsymbol{y}}(\boldsymbol{x}) = p(\boldsymbol{x}, \boldsymbol{y})$ is a spherical polynomial of degree N_1, and the Markov inequality (4.12) applies to it. Clearly, the analogous statement holds for $p_{\boldsymbol{u}}(\boldsymbol{y}) = p(\boldsymbol{u}, \boldsymbol{y})$ for fixed $\boldsymbol{u} \in \mathbb{S}^2$ and N_2 instead of N_1. For $(\boldsymbol{x}, \boldsymbol{y}), (\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{S}^2 \times \mathbb{S}^2$, we obtain

$$\begin{aligned} |p(\boldsymbol{x}, \boldsymbol{y}) - p(\boldsymbol{u}, \boldsymbol{v})| &= |p(\boldsymbol{x}, \boldsymbol{y}) - p(\boldsymbol{u}, \boldsymbol{y})| + |p(\boldsymbol{u}, \boldsymbol{y}) - p(\boldsymbol{u}, \boldsymbol{v})| \\ &\le |p_{\boldsymbol{y}}(\boldsymbol{x}) - p_{\boldsymbol{y}}(\boldsymbol{u})| + |p_{\boldsymbol{u}}(\boldsymbol{y}) - p_{\boldsymbol{u}}(\boldsymbol{v})| \\ &\le N_1 \, d_0(\boldsymbol{x}, \boldsymbol{u}) \|p_{\boldsymbol{y}}\|_{\mathbb{S}^2, \infty} + N_2 \, d_0(\boldsymbol{y}, \boldsymbol{v}) \|p_{\boldsymbol{u}}\|_{\mathbb{S}^2, \infty} \\ &\le (N_1 \, d_0(\boldsymbol{x}, \boldsymbol{u}) + N_2 \, d_0(\boldsymbol{y}, \boldsymbol{v})) \|p\|_{\mathbb{S}^2 \times \mathbb{S}^2, \infty} \\ &\le (2 - \delta_{1,s}) \tilde{N} \; d^{(s)}((\boldsymbol{x}, \boldsymbol{y}), (\boldsymbol{u}, \boldsymbol{v})) \; \|p\|_{\mathbb{S}^2 \times \mathbb{S}^2, \infty}. \end{aligned}$$

■

We are now prepared to formulate the main result of this section.

Theorem 4.5. *Let $\boldsymbol{N} \in \mathbb{N}^2$, and let $\mathcal{R}$ be an admissible decomposition of $\mathbb{S}^2 \times \mathbb{S}^2$ according to a sampling set $\mathcal{X}$ such that*

$$\|\mathcal{R}\|_1 \leq \frac{\eta}{21\,\tilde{N}}, \tag{4.13}$$

where $\eta \in (0,1)$ is arbitrarily fixed. Then for $r = 1$ or $r = \infty$ and any $p \in \Pi_{\boldsymbol{N}}$, we have

$$(1-\eta)\,\|p\|_{\mathbb{S}^2\times\mathbb{S}^2,r} \leq \|p\|_{\mathcal{X},r} \leq (1+\eta)\,\|p\|_{\mathbb{S}^2\times\mathbb{S}^2,r}. \tag{4.14}$$

Proof. We start with the case $r = \infty$. For proving the left-hand side, let $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{S}^2\times\mathbb{S}^2$ an arbitrary fixed point and assume that $(\boldsymbol{u},\boldsymbol{v}) \in \mathcal{X}$ is a point with $d((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v})) = d((\boldsymbol{x},\boldsymbol{y}),\mathcal{X}) = \delta_{\mathcal{X}}$. We conclude by Lemma 4.4

$$\begin{aligned} |p(\boldsymbol{x},\boldsymbol{y})| &\leq |p(\boldsymbol{x},\boldsymbol{y}) - p(\boldsymbol{u},\boldsymbol{v})| + |p(\boldsymbol{u},\boldsymbol{v})| \\ &\leq 2\tilde{N}\, d\left((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v})\right)\ \|p\|_{\mathbb{S}^2\times\mathbb{S}^2,\infty} + \max_{j=1,\ldots,M} |p(\boldsymbol{x}_j,\boldsymbol{y}_j)|\,. \end{aligned}$$

Taking the maximum over $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{S}^2\times\mathbb{S}^2$, we obtain, by using the fact $d((\boldsymbol{x},\boldsymbol{y}),(\boldsymbol{u},\boldsymbol{v})) = d((\boldsymbol{x},\boldsymbol{y}),\mathcal{X}) \leq \|\mathcal{R}\|_1$, the L^∞-MZ inequality

$$\left(1 - 2\tilde{N}\|\mathcal{R}\|_1\right)\|p\|_{\mathbb{S}^2\times\mathbb{S}^2,\infty} \leq \max_{j=1,\ldots,M}\left|p(\boldsymbol{\xi}_j,\boldsymbol{\eta}_j)\right| \leq \left(1 + 2\tilde{N}\|\mathcal{R}\|_1\right)\ \|p\|_{\mathbb{S}^2\times\mathbb{S}^2,\infty}. \tag{4.15}$$

This, in combination with (4.13), provides inequality (4.14) for $r = \infty$.
For $r = 1$, we are going to show both inequalities of (4.14) simultaneously by proving

$$\left|\|p\|_{\mathbb{S}^2\times\mathbb{S}^2,1} - \|p\|_{\mathcal{X},1}\right| \leq \eta\,\|p\|_{\mathbb{S}^2\times\mathbb{S}^2,1}.$$

Using Theorem 4.3(i), we obtain

$$\begin{aligned}
&\left|\int_{\mathbb{S}^2\times\mathbb{S}^2} |p(\boldsymbol{x},\boldsymbol{y})|\ \mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y}) - \sum_{j=1}^{M} |p(\boldsymbol{x}_j,\boldsymbol{y}_j)|\,\tau(R_j)\right| \\
=\ &\left|\sum_{j=1}^{M}\int_{R_j} |p(\boldsymbol{x},\boldsymbol{y})| - |p(\boldsymbol{x}_j,\boldsymbol{y}_j)|\ \mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y})\right| \\
\leq\ &\sum_{j=1}^{M}\int_{R_j} |p(\boldsymbol{x},\boldsymbol{y}) - p(\boldsymbol{x}_j,\boldsymbol{y}_j)|\ \mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y}) \\
\leq\ &\sum_{j=1}^{M}\int_{R_j} |\mathcal{V}_{\boldsymbol{N}}p(\boldsymbol{x},\boldsymbol{y}) - \mathcal{V}_{\boldsymbol{N}}p(\boldsymbol{x}_j,\boldsymbol{y}_j)|\ \mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y}) \\
\leq\ &\sup_{(\boldsymbol{u},\boldsymbol{v})\in\mathbb{S}^2\times\mathbb{S}^2}\left(\sum_{j=1}^{M}\int_{R_j} |v_{\boldsymbol{N}}(\boldsymbol{x}\cdot\boldsymbol{u},\boldsymbol{y}\cdot\boldsymbol{v}) - v_{\boldsymbol{N}}(\boldsymbol{x}_j\cdot\boldsymbol{u},\boldsymbol{y}_j\cdot\boldsymbol{v})|\ \mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y})\right) \\
&\times\int_{\mathbb{S}^2\times\mathbb{S}^2} |p(\boldsymbol{u},\boldsymbol{v})|\ \mathrm{d}\tau(\boldsymbol{u},\boldsymbol{v}).
\end{aligned}$$

It suffices to show that

$$\sup_{(\boldsymbol{u},\boldsymbol{v})\in\mathbb{S}^2\times\mathbb{S}^2}\left(\sum_{j=1}^{M}\int_{R_j} |v_{\boldsymbol{N}}(\boldsymbol{x}\cdot\boldsymbol{u},\boldsymbol{y}\cdot\boldsymbol{v}) - v_{\boldsymbol{N}}(\boldsymbol{x}_j\cdot\boldsymbol{u},\boldsymbol{y}_j\cdot\boldsymbol{v})|\ \mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y})\right) \leq \eta. \qquad (4.16)$$

To this end, we use spherical coordinates for the spheres with respect to the point $(\boldsymbol{u},\boldsymbol{v})\in\mathbb{S}^2\times\mathbb{S}^2$. This leads to $\boldsymbol{x}\cdot\boldsymbol{u}=\cos\xi,\ \boldsymbol{y}\cdot\boldsymbol{v}=\cos\eta$. We obtain

$$\begin{aligned}
\Sigma :=\ &\sum_{j=1}^{M}\int_{R_j} |v_{\boldsymbol{N}}(\boldsymbol{x}\cdot\boldsymbol{u},\boldsymbol{y}\cdot\boldsymbol{v}) - v_{\boldsymbol{N}}(\boldsymbol{x}_j\cdot\boldsymbol{u},\boldsymbol{y}_j\cdot\boldsymbol{v})|\ \mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y}) \\
=\ &\sum_{j=1}^{M}\int_{R_j} |v_{N_1}(\cos\xi)\,v_{N_2}(\cos\eta) - v_{N_1}(\cos\xi_j)\,v_{N_2}(\cos\eta_j)|\ \mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y}) \\
=\ &\sum_{j=1}^{M}\int_{R_j}\left[|v_{N_2}(\cos\eta)|\ \int_{\xi_j}^{\xi}\left|\frac{\mathrm{d}}{\mathrm{d}t}v_{N_1}(\cos t)\right|\ \mathrm{d}t + |v_{N_1}(\cos\xi_j)|\ \int_{\eta_j}^{\eta}\left|\frac{\mathrm{d}}{\mathrm{d}t}v_{N_2}(\cos t)\right|\ \mathrm{d}t\right] \\
&\times\mathrm{d}\tau(\boldsymbol{x},\boldsymbol{y}).
\end{aligned}$$

For further estimation of the above expression, we introduce a partition of $\mathbb{S}^2\times\mathbb{S}^2$ in the following way. Let $L=\left\lfloor\frac{\pi}{\|\mathcal{R}\|_1}\right\rfloor$ and define $\mathcal{I}_\nu=[(\nu-1)\pi L^{-1},(\nu+1)\pi L^{-1}],\ \nu=1,\ldots,L-1$. For $\ell_1,\ell_2\in\{1,\ldots,L-1\}$ let

$$B_{\ell_1,\ell_2} := \left\{(\boldsymbol{x},\boldsymbol{y})\in\mathbb{S}^2\times\mathbb{S}^2 : (\xi,\eta)\in\mathcal{I}_{\ell_1}\times\mathcal{I}_{\ell_2}\right\}.$$

Let $B_\nu = \{x \in \mathbb{S}^2 : \xi \in \mathcal{I}_\nu\}$. We then have $B_{\ell_1,\ell_2} = B_{\ell_1} \times B_{\ell_2}$, $\bigcup_{\ell_1,\ell_2=1}^{L-1} B_{\ell_1,\ell_2} = \mathbb{S}^2 \times \mathbb{S}^2$, and $B_{\ell_1,\ell_2} \cap B_{\ell_1',\ell_2'} = \emptyset$ if $\ell_1 > \ell_1' + 2$ or $\ell_1 < \ell_1' - 2$, and $\ell_2 > \ell_2' + 2$ or $\ell_2 < \ell_2' - 2$. This provides a decomposition of $\mathbb{S}^2 \times \mathbb{S}^2$ into overlapping regions. Let us denote by ξ_j^-, η_j^- and ξ_j^+, η_j^+ the minimal resp. maximal value of the coordinates ξ resp. η of the region R_j. Since

$$\max\{\xi_j^+ - \xi_j^-, \eta_j^+ - \eta_j^-\} \leq \mathsf{diam}(R_j) \leq \|\mathcal{R}\|_1 \leq \pi L^{-1},$$

the region R_j is contained completely in at least one B_{ℓ_1,ℓ_2} with $\ell_1, \ell_2 \in \{1, \ldots, L-1\}$. Accordingly, we have

$$\begin{aligned}\Sigma \leq \; & \sum_{\ell_1=1}^{L-1} \sum_{\ell_2=1}^{L-1} \sum_{R_j \subset B_{\ell_1,\ell_2}} \int_{R_j} |v_{N_2}(\cos\eta)| \int_{\xi_j^-}^{\xi} \left| \frac{\mathrm{d}}{\mathrm{d}t} v_{N_1}(\cos t) \right| \mathrm{d}t \; \mathrm{d}\tau(\boldsymbol{x}, \boldsymbol{y}) \\ & + \sum_{\ell_1=1}^{L-1} \sum_{\ell_2=1}^{L-1} \sum_{R_j \subset B_{\ell_1,\ell_2}} \int_{R_j} |v_{N_1}(\cos\xi_j)| \int_{\eta_j^-}^{\eta} \left| \frac{\mathrm{d}}{\mathrm{d}t} v_{N_2}(\cos t) \right| \mathrm{d}t \; \mathrm{d}\tau(\boldsymbol{x}, \boldsymbol{y}) \,.\end{aligned}$$

We denote the first sum on the right-hand side of the above inequality by $\Sigma^{(1)}$ and the second sum by $\Sigma^{(2)}$. Now we have to distinguish between different cases. These cases are

$$\begin{aligned}&(1.1): \; \ell_1 = 1 \wedge \{\ell_2 = 1 \vee \ell_2 = L-1\}, && (1.2): \; \ell_1 = L-1 \wedge (\ell_2 = 1 \vee \ell_2 = L-1, \\ &(1.3): \; \ell_2 = 1 \wedge (\ell_1 = 1 \vee \ell_1 = L-1), && (1.4): \; \ell_2 = L-1 \wedge (\ell_1 = 1 \vee \ell_1 = L-1);\end{aligned}$$

$$\begin{aligned}&(2.1): \; \ell_1 = 1 \wedge \ell_2 \in \{2, \ldots, L-2\}, && (2.2): \; \ell_1 = L-1 \wedge \ell_2 \in \{2, \ldots, L-2\}, \\ &(2.3): \; \ell_2 = 1 \wedge \ell_1 \in \{2, \ldots, L-2\}, && (2.4): \; \ell_2 = L-1 \wedge \ell_1 \in \{2, \ldots, L-2\};\end{aligned}$$

$$(3): \; \ell_1 \in \{2, \ldots, L-2\} \wedge \ell_2 \in \{2, \ldots, L-2\}$$

A detailed analysis of all these cases for both sums $\Sigma^{(1)}$ and $\Sigma^{(2)}$ finally gives us the following result (see Filbir and Potts (2009) for details):

$$\Sigma \leq 8832 \widetilde{N}^3 \, \|\mathcal{R}\|^3 + 9216 \widetilde{N}^5 \, \|\mathcal{R}\|^5 \tag{4.17}$$

Assume that $\|\mathcal{R}\|_1 \leq \frac{\eta}{C\widetilde{N}}$. Then we obtain

$$\Sigma \leq \left(8832 \frac{1}{C^3} + 9216 \frac{1}{C^5} \right) \eta \tag{4.18}$$

A short computation shows that for $C = 21$, the first factor in (4.18) is less than 1. This shows the desired result. ∎

For our purposes, we need the MZ inequality for the case $r = 2$. Since the polynomial spaces $\Pi_{\boldsymbol{N}}$ are finite-dimensional, we cannot get the desired inequalities by applying the Riesz–Thorin interpolation theorem. Alternatively, we can observe that if we consider instead of p the polynomial $p^2 \in \Pi_{2\boldsymbol{N}}$, the MZ inequality holds for L^2 with $2\boldsymbol{N}$ instead of $\boldsymbol{N}$. This leads to a constant which is two times the constant in Theorem 4.5. Applying this idea, we have the following result for the case $r = 2$.

Corollary 4.6. *Let $\mathcal{X} = \{(\boldsymbol{x}_j, \boldsymbol{y}_j) : j = 1, \ldots, M\} \subset \mathbb{S}^2 \times \mathbb{S}^2$ be a sampling set with corresponding tensor decomposition $\mathcal{R}$. Let $\boldsymbol{W} = \operatorname{diag}(w_j)_{j=1,\ldots,M}$ be the diagonal matrix of the corresponding tensor weights. If $\tilde{N} = \max\{N_1, N_2\}$ satisfies*

$$42\tilde{N}\|\mathcal{R}\|_1 \leq 1, \tag{4.19}$$

then for every polynomial $p \in \Pi_{\boldsymbol{N}}$, $\boldsymbol{N} = (N_1, N_2)$, the weighted norm estimate

$$\left(1 - 42\tilde{N}\|\mathcal{R}\|_1\right) \|p\|^2_{\mathbb{S}^2\times\mathbb{S}^2,2} \leq \sum_{j=1}^{M} w_j |p(\boldsymbol{x}_j, \boldsymbol{y}_j)|^2 \leq \left(1 + 42\tilde{N}\|\mathcal{R}\|_1\right) \|p\|_{\mathbb{S}^2\times\mathbb{S}^2,2} \tag{4.20}$$

holds.

Moreover, for the condition number of the matrix $\boldsymbol{\mathcal{Y}}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{\mathcal{Y}}$, we have

$$cond(\boldsymbol{\mathcal{Y}}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{\mathcal{Y}}) \leq \frac{1 + 42\tilde{N}\|\mathcal{R}\|_1}{1 - 42\tilde{N}\|\mathcal{R}\|_1}. \tag{4.21}$$

Proof. According to the remark above, we get

$$1 - 42\tilde{N}\|\mathcal{R}\|_1 \leq \|\mathcal{S}_{\mathcal{X}} p\|^2_{\mathbb{S}^2\times\mathbb{S}^2,2} \leq 1 + 42\tilde{N}\|\mathcal{R}\|_1 \,.$$

Due to $\boldsymbol{p} = \boldsymbol{\mathcal{Y}}\hat{\boldsymbol{p}}$ and Parseval's identity $\|\hat{\boldsymbol{p}}\|_2 = \|p\|_{\mathbb{S}^2\times\mathbb{S}^2,2}$, inequality (4.20) is equivalent to

$$1 - 42\tilde{N}\|\mathcal{R}\|_1 \leq \frac{\hat{\boldsymbol{p}}^{\mathsf{H}}\boldsymbol{\mathcal{Y}}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{\mathcal{Y}}\hat{\boldsymbol{p}}}{\hat{\boldsymbol{p}}^{\mathsf{H}}\hat{\boldsymbol{p}}} \leq 1 + 42\tilde{N}\|\mathcal{R}\|_1,$$

from which (4.21) follows. ∎

4.2 Optimal interpolation

In this section, we focus on the underdetermined case $M < (N_1 + 1)^2(N_2 + 1)^2$. As described at the beginning of Section 4, we will interpolate the data $\mathcal{A} = \{(\boldsymbol{x}_j, \boldsymbol{y}_j, f_j) \in \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{C} : j = 1, \ldots, M\}$. Instead of only asking for a polynomial $p(\boldsymbol{x}, \boldsymbol{y}) = \sum_{(\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}} \hat{p}_{\boldsymbol{n},\boldsymbol{k}} \boldsymbol{S}_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}, \boldsymbol{y})$ which satisfies $p(\boldsymbol{x}_j, \boldsymbol{y}_j) = f_j,\ j = 1, \ldots, M$, we will try to find an optimal interpolating polynomial, i.e., we try to find a polynomial which is a solution to

$$\min_{p\in\Pi_{\boldsymbol{N}}} \sum_{(\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}} \frac{|\hat{p}_{\boldsymbol{n},\boldsymbol{k}}|^2}{\hat{w}_{n_1}\hat{w}_{n_2}} \quad \text{subject to} \quad p(\boldsymbol{x}_j, \boldsymbol{y}_j) = f_j,\ j = 1, \ldots, M, \tag{4.22}$$

where $\hat{w}_{n_1}, \hat{w}_{n_2}, n_1 = 0, \ldots, N_1, n_2 = 0, \ldots, N_2$, are given positive real numbers.

Whereas for the least square problem the mesh norm was the critical parameter, this role is now played by the separation distance q of the set $\mathcal{X} = \{(\boldsymbol{x}_j, \boldsymbol{y}_j) : j = 1, \ldots, M\}$. The weights $\hat{w}_{n_1}, \hat{w}_{n_2}$ have to be chosen according to this parameter. Problem (4.22) can be equivalently restated in the following matrix form (see, (Björck, 1996, Thm. 1.1.2)).

Proposition 4.7. *The optimal interpolation problem* (4.22) *is equivalent to the normal equations of second kind*

$$\boldsymbol{\mathcal{Y}}\hat{\boldsymbol{W}}\boldsymbol{\mathcal{Y}}^{\mathsf{H}}\tilde{\boldsymbol{f}} = \boldsymbol{f}, \qquad \hat{\boldsymbol{p}} = \hat{\boldsymbol{W}}\boldsymbol{\mathcal{Y}}^{\mathsf{H}}\tilde{\boldsymbol{f}}, \tag{4.23}$$

where the weighting matrix is given by $\hat{\boldsymbol{W}} := \operatorname{diag}(\hat{\boldsymbol{w}}) \in \mathbb{R}^{(N_1+1)^2(N_2+1)^2\times(N_1+1)^2(N_2+1)^2}$ *for the vector* $\hat{\boldsymbol{w}} = (\hat{w}_{\boldsymbol{n},\boldsymbol{k}})_{(\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}}$ *with* $\hat{w}_{n_1,n_2,k_1,k_2} = \hat{w}_{n_1}\hat{w}_{n_2}$, $(\boldsymbol{n},\boldsymbol{k}) \in I_{\boldsymbol{N}}$.

It is the crucial observation that the matrix entry $(\boldsymbol{\mathcal{Y}}\hat{\boldsymbol{W}}\boldsymbol{\mathcal{Y}}^{\mathsf{H}})_{j;l}$ can be expressed as a product of certain polynomial kernels. Indeed, we have used the addition formula (2.18) and obtain

$$\begin{aligned}
\left(\boldsymbol{\mathcal{Y}}\hat{\boldsymbol{W}}\boldsymbol{\mathcal{Y}}^{\mathsf{H}}\right)_{j;l} &= \sum_{(\boldsymbol{n},\boldsymbol{k})\in I_{\boldsymbol{N}}} \hat{w}_{n_1}\hat{w}_{n_2}\boldsymbol{S}_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}_j,\boldsymbol{y}_j)\overline{\boldsymbol{S}_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}_l,\boldsymbol{y}_l)} \\
&= \sum_{n_1=0}^{N_1}\sum_{k_1=-n_1}^{n_1}\sum_{n_2=0}^{N_2}\sum_{k_2=-n_2}^{n_2} \hat{w}_{n_1}\hat{w}_{n_2}Y_{n_1,k_1}(\boldsymbol{x}_j)\overline{Y_{n_1,k_1}(\boldsymbol{x}_l)} \\
&\quad \times Y_{n_2,k_2}(\boldsymbol{y}_j)\overline{Y_{n_2,k_2}(\boldsymbol{y}_l)} \\
&= \sum_{n_1=0}^{N_1}\sum_{n_2=0}^{N_2} \frac{2n_1+1}{4\pi}\frac{2n_2+1}{4\pi}\hat{w}_{n_1}\hat{w}_{n_2}R_{n_1}(\boldsymbol{x}_j\cdot\boldsymbol{x}_l)R_{n_2}(\boldsymbol{y}_j\cdot\boldsymbol{y}_l) \\
&= P_{N_1}(\boldsymbol{x}_j\cdot\boldsymbol{x}_l)\; P_{N_2}(\boldsymbol{y}_j\cdot\boldsymbol{y}_l)\,,
\end{aligned}$$

where

$$P_{N_1}(t) := \sum_{n_1=0}^{N_1}\frac{2n_1+1}{4\pi}\hat{w}_{n_1}R_{n_1}(t)\,, \quad P_{N_2}(t) := \sum_{n_2=0}^{N_2}\frac{2n_2+1}{4\pi}\hat{w}_{n_2}R_{n_2}(t)\,. \tag{4.24}$$

Corresponding to this observation we introduce the shorter notation $\boldsymbol{P}$ for the matrix $\boldsymbol{\mathcal{Y}}\hat{\boldsymbol{W}}\boldsymbol{\mathcal{Y}}^{\mathsf{H}}$. Note that the positivity assumption on the coefficients $\hat{w}_{n_1}, \hat{w}_{n_2}$ ensures the positive definiteness of P_{N_1} resp. P_{N_2}. For convenience let us further assume that the kernels (4.24) are normalized, i.e.,

$$P_{N_1}(1) = 1, \quad P_{N_2}(1) = 1.$$

This normalization obviously does not change the type of the problem. The aim is now to give an estimate on the eigenvalues of the matrix $\boldsymbol{P}$. To this end we use the Gerschgorin circle theorem. Consider $(\boldsymbol{x}_1,\boldsymbol{y}_1) \in \mathcal{X}$ and let us assume without loss of generality that $(\boldsymbol{x}_1,\boldsymbol{y}_1) = \left((0,0,1)^\top,(0,0,1)^\top\right)$. Due to the structure of the matrix $\boldsymbol{P}$, for the eigenvalues $\lambda(\boldsymbol{P})$, we obtain the estimate

$$|\lambda(\boldsymbol{P}) - 1| \le \sum_{j=2}^{M}\boldsymbol{P}_{j,1} = \sum_{j=2}^{M} P_{N_1}(\boldsymbol{x}_1\cdot\boldsymbol{x}_j)\, P_{N_2}(\boldsymbol{y}_1\cdot\boldsymbol{y}_j).$$

Our aim is to make this estimate more precise and finally to get from this a criterion for the matrix $\boldsymbol{\mathcal{Y}}$ to have full row rank. In order to do so, we first have to examine the set $\mathcal{X}$ more precisely and, secondly, to construct well-localized kernels.

We define a partition of the sampling nodes $\mathcal{X}$ into "rings" with increasing distance from the node $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in \mathbb{S}^2 \times \mathbb{S}^2$. Let the separation distance of $\mathcal{X}$ satisfy $q \leq \pi$. For $0 \leq m < \lfloor \pi q^{-1} \rfloor$, let

$$S_{q,m} := \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{S}^2 \times \mathbb{S}^2 : mq \leq d\left((\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}, \boldsymbol{y})\right) < (m+1)\, q \right\}$$

and

$$S_{q,\lfloor \pi q^{-1} \rfloor} := \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{S}^2 \times \mathbb{S}^2 : \left\lfloor \pi q^{-1} \right\rfloor q \leq d\left((\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}, \boldsymbol{y})\right) \leq \pi \right\}.$$

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

The next proposition provides an estimate for the cardinality of the set $S_{\mathcal{X},q,m}$, i.e., we give an upper bound for the number of q-separated sampling points which can be placed within a certain distance to $(\boldsymbol{x}_1, \boldsymbol{y}_1)$. In contrast to (Narcowich et al., 1998, Thm. 2.3), our estimate relies solely on the index m but no longer on the separation distance q. We follow the lines in (Keiner et al., 2007, Lemma 4.3).

Proposition 4.8. *Let $\mathcal{X}$ be a set of q-separated sampling nodes with $q \leq \pi/3$. Then for $m = 1, \ldots, \lfloor \pi q^{-1} \rfloor$, we have*

$$|S_{\mathcal{X},q,m}| \leq 128m^3 + 192m^2 + 224m + 80\,.$$

Proof. We use a packing argument from (Narcowich et al., 1998, Thm. 2.3), which states that for each node in $S_{\mathcal{X},q,m}$, the centered cap of colatitude $q/2$ around it is 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, for $m = 1, \ldots, \lfloor \pi q^{-1} \rfloor - 2$, we estimate

$$
\begin{aligned}
|S_{\mathcal{X},q,m}| &\leq \frac{\displaystyle\int_{\tilde{S}_{q,m}} \mathrm{d}\tau\,(\boldsymbol{x},\boldsymbol{y})}{\displaystyle\int_{S_{q/2,0}} \mathrm{d}\tau\,(\boldsymbol{x},\boldsymbol{y})} \\
&= \frac{\displaystyle\int_{\tilde{S}_{(m+\frac{3}{2})q,0}} \mathrm{d}\tau\,(\boldsymbol{x},\boldsymbol{y}) - \int_{\tilde{S}_{(m-\frac{1}{2})q,0}} \mathrm{d}\tau\,(\boldsymbol{x},\boldsymbol{y})}{\displaystyle\int_{S_{q/2,0}} \mathrm{d}\tau\,(\boldsymbol{x},\boldsymbol{y})} \\
&= \frac{\displaystyle\int_0^{(m+\frac{3}{2})q} \sin\theta \mathrm{d}\theta \int_0^{(m+\frac{3}{2})q} \sin\vartheta \mathrm{d}\vartheta - \int_0^{(m-\frac{1}{2})q} \sin\theta \mathrm{d}\theta \int_0^{(m-\frac{1}{2})q} \sin\vartheta \mathrm{d}\vartheta}{\displaystyle\int_0^{\frac{q}{2}} \sin\theta \mathrm{d}\theta \int_0^{\frac{q}{2}} \sin\vartheta \mathrm{d}\vartheta} \\
&= \frac{\left(1-\cos\left(\frac{(2m+3)q}{2}\right)\right)^2 - \left(1-\cos\left(\frac{(2m-1)q}{2}\right)\right)^2}{\left(1-\cos\frac{q}{2}\right)^2} \\
&= \frac{\cos\left((2m-1)\frac{q}{2}\right) - \cos\left((2m+3)\frac{q}{2}\right)}{1-\cos\frac{q}{2}} \\
&\quad \times \left(\frac{1-\cos\left((2m+3)\frac{q}{2}\right)}{1-\cos\frac{q}{2}} + \frac{1-\cos\left((2m-1)\frac{q}{2}\right)}{1-\cos\frac{q}{2}}\right)
\end{aligned}
\tag{4.25}
$$

Using an identity for the trigonometric de la Vallée Poussin kernel, see, e.g. (Prestin and Selig, 2001, equation (3.4) and (3.5)), we estimate the first term in (4.25) as in (Keiner et al., 2007, Lemma 4.3) by

$$
\begin{aligned}
&\frac{\cos\left((2m-1)\frac{q}{2}\right) - \cos\left((2m+3)\frac{q}{2}\right)}{1-\cos\frac{q}{2}} = \frac{\sin\left((2m+1)\frac{q}{2}\right)\sin\left(2\frac{q}{2}\right)}{\sin^2\frac{q}{4}} \\
&= 4 + 8\sum_{l=1}^{2m-1} \cos\tfrac{lq}{2} + 6\cos\tfrac{2mq}{2} + 4\cos\tfrac{(2m+1)q}{2} + 2\cos\tfrac{(2m+2)q}{2} \\
&\leq 8\,(2m+1)\,.
\end{aligned}
\tag{4.26}
$$

We estimate the second terms in (4.25) by using the Dirichlet kernel

$$
\begin{aligned}
\frac{1-\cos\left((2m+3)\frac{q}{2}\right)}{1-\cos\frac{q}{2}} &= \left(\frac{\sin\left((2m+3)\frac{q}{4}\right)}{\sin\frac{q}{4}}\right)^2 = \left(1 + 2\sum_{k=1}^{m+1}\cos\left(k\frac{q}{2}\right)\right)^2 \\
&\leq \left(1+2(m+1)\right)^2 = 4m^2 + 12m + 9
\end{aligned}
\tag{4.27}
$$

and, similarly,

$$\frac{1-\cos\left((2m-1)\frac{q}{2}\right)}{1-\cos\frac{q}{2}} = \left(\frac{\sin\left((2m-1)\frac{q}{4}\right)}{\sin\frac{q}{4}}\right)^2 = \left(1+2\sum_{k=1}^{m-1}\cos\left(k\frac{q}{2}\right)\right)^2$$
$$\leq (1+2(m-1))^2 = 4m^2-4m+1\,. \tag{4.28}$$

Combining (4.25) – (4.28), we obtain the assertion. ■

We now give a construction of well-localized kernels. Our approach is based on B-spline kernels and is strongly related to the construction given by Kunis (Kunis, 2006, Definition 2.15). Recall that for $\beta \in \mathbb{N}$ the cardinal B-spline $N_{\beta+1}$ is defined recursively by

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

The normalized B-spline of order $\beta \in \mathbb{N}$ is defined as $g_\beta : [-\frac{1}{2}, \frac{1}{2}] \to \mathbb{R}$, $g_\beta(z) := \beta N_\beta(\beta z + \frac{\beta}{2})$. By using sample values of these functions we obtain a well-localized kernel as follows. For $\beta, N \in \mathbb{N}$ let $B_{\beta,N} : [-1,1] \to \mathbb{R}$ be given by

$$B_{\beta,N}(t) := \frac{1}{\|g_\beta\|_{1,N}} \sum_{l=0}^{N} (2-\delta_{l,0})\, g_\beta\left(\frac{l}{2(N+1)}\right)\cos(l \arccos t), \tag{4.29}$$

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

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

The kernel $B_{\beta,N}$ is called the B-spline kernel. The next proposition shows the localization property of the B-spline kernel. Based on (Kunis, 2006, Theorem 2.35), a proof was presented in (Keiner et al., 2007, Lemma 4.6).

Proposition 4.9. *The B-spline kernel $B_{\beta,N}$ obeys for $N \geq \beta - 1$ and $t \in [-1,1)$ the localization property*

$$|B_{\beta,N}(t)| \leq c_\beta\, |(N+1)\arccos(t)|^{-\beta}, \qquad c_\beta := \frac{\left(2^\beta-1\right)\zeta(\beta)\,\beta^\beta}{2^{\beta-1}-\zeta(\beta)\,\pi^{-\beta}}. \tag{4.30}$$

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

$$B_{\beta,N}(t) = \sum_{n=0}^{N} \frac{2n+1}{4\pi}\hat{w}_n R_n(t)$$

with positive Fourier–Legendre coefficients

$$\hat{w}_n = 2\pi \int_{-1}^{1} R_n(t)\, B_{\beta,N}(t)\,\mathrm{d}t, \qquad n = 0, \ldots, N.$$

We now turn to our main result, i.e., the estimate for the eigenvalues of the matrix $\boldsymbol{P}$. To get this estimate, we make use of the localization of the B-spline kernel (4.30) and the cardinality estimate of Proposition 4.8. Finally we will obtain an answer to our initial question on the stability of the optimal interpolation problem (4.22).

Theorem 4.10. *Let $\mathcal{X} = \{(\boldsymbol{x}_j, \boldsymbol{y}_j) : j = 1, \ldots, M\}$ be a sampling set of $\mathbb{S}^2 \times \mathbb{S}^2$ with separation distance q. For $N_1, N_2, \beta \in \mathbb{N}$, $N_1, N_2 \geq \beta \geq 5$ let $\boldsymbol{P} = (P_{j,l})_{j;l=1,\ldots,M}$ with $P_{j,l} = B_{\beta,N_1}(\boldsymbol{x}_j \cdot \boldsymbol{x}_l) B_{\beta,N_2}(\boldsymbol{y}_j \cdot \boldsymbol{y}_l)$. Then, for every eigenvalue $\lambda(\boldsymbol{P})$ of $\boldsymbol{P}$, we have*

$$|\lambda(\boldsymbol{P}) - 1| < \frac{c_\beta \left(128\,\zeta\,(\beta-3) + 192\,\zeta\,(\beta-2) + 224\,\zeta\,(\beta-1) + 80\,\zeta\,(\beta)\right)}{(N'+1)^\beta\, q^\beta}, \tag{4.31}$$

where c_β is given in (4.30) and $N' = \min\{N_1, N_2\}$.

Proof. We apply the Gershgorin circle theorem where we assume without loss of generality that $\boldsymbol{x}_1 = (0,0,1)^\top$ and $\boldsymbol{y}_1 = (0,0,1)^\top$. Using Proposition 4.8, the localization property in Proposition 4.9, and $\max_{t\in[-1,1]} |B_{\beta,N}(t)| \leq 1$, we obtain the estimate

$$\begin{aligned}
|\lambda(\boldsymbol{P}) - 1| &\leq \sum_{j=2}^{M} \left| B_{\beta,N_1}\left(\boldsymbol{x}_1 \cdot \boldsymbol{x}_j\right) B_{\beta,N_2}\left(\boldsymbol{y}_1 \cdot \boldsymbol{y}_j\right) \right| \\
&< \sum_{m=1}^{\lfloor \pi q^{-1} \rfloor} |S_{\mathcal{X},q,m}| \max_{(\boldsymbol{x},\boldsymbol{y}) \in S_{q,m}} |B_{\beta,N_1}\left(\boldsymbol{x}_1 \cdot \boldsymbol{x}\right)\, B_{\beta,N_2}\left(\boldsymbol{y}_1 \cdot \boldsymbol{y}\right)| \\
&< \sum_{m=1}^{\lfloor \pi q^{-1} \rfloor} |S_{\mathcal{X},q,m}| \frac{c_\beta}{\left((N'+1)\, mq\right)^\beta} \\
&< \frac{c_\beta}{(N'+1)^\beta q^\beta} \sum_{m=1}^{\lfloor \pi q^{-1} \rfloor} \frac{128m^3 + 192m^2 + 224m + 80}{m^\beta} \\
&< \frac{c_\beta \left(128\,\zeta\,(\beta-3) + 192\,\zeta\,(\beta-2) + 224\,\zeta\,(\beta-1) + 80\,\zeta\,(\beta)\right)}{(N'+1)^\beta q^\beta}
\end{aligned}$$

■

Although estimate (4.31) looks complicated at first glance, it allows us to give concrete conditions that ensure the stability of our optimal interpolation problem. More precisely, we have

Corollary 4.11. *Under the conditions of Theorem 4.10 and by choosing $\beta = 5$ in (4.31), for the eigenvalues of the matrix $\boldsymbol{P} = \boldsymbol{\mathcal{Y}} \hat{\boldsymbol{W}} \boldsymbol{\mathcal{Y}}^{\mathsf{H}}$, we obtain the estimate*

$$|\lambda(\boldsymbol{P}) - 1| < \left(\frac{22}{(N'+1)q} \right)^5 .$$

Moreover, let $(N'+1)q > 22$. Then the matrix $\boldsymbol{\mathcal{Y}}$ has full row rank M, and the conjugate gradient method applied to (4.23) converges linearly, i.e.,

$$\|\hat{\boldsymbol{e}}_l\|_{\hat{\boldsymbol{W}}^{-1}} \leq 2 \left(\frac{22}{(N'+1)q} \right)^{5l} \|\hat{\boldsymbol{e}}_0\|_{\hat{\boldsymbol{W}}^{-1}} \tag{4.32}$$

with the initial error $\hat{\boldsymbol{e}}_0 := \hat{\boldsymbol{W}}\boldsymbol{\mathcal{Y}}^{\mathrm{H}}\boldsymbol{P}^{-1}\boldsymbol{f}$ *and the error* $\hat{\boldsymbol{e}}_l := \hat{\boldsymbol{f}}_l - \hat{\boldsymbol{W}}\boldsymbol{\mathcal{Y}}^{\mathrm{H}}\boldsymbol{P}^{-1}\boldsymbol{f}$.

Proof. Applying the standard estimate for the convergence of the conjugate gradient method, see, e.g., (Björck, 1996, p. 289), yields assertion (4.32). ■

5 An Application to Texture Analysis

As pointed out in the introduction, the paper is motivated by an application. The Radon transform on $SO(3)$ is of central importance for the analysis of crystallographic preferred orientations, the technical term of which is texture analysis, cf. Wenk (1985). It establishes a relationship between the so-called *orientation density function* (ODF) $f\colon SO(3)/\mathcal{G} \to \mathbb{R}$, which models the distribution of crystal orientations within a polycrystalline specimen, and the so-called *pole density function* (PDF) $P\colon \mathbb{S}^2/\mathcal{G} \times \mathbb{S}^2 \to \mathbb{R}$, which models the distribution of crystallographic lattice planes within the specimen. Here $\mathcal{G} \subset SO(3)$ denotes a finite subgroup of $SO(3)$ which represents the crystal symmetries. In terms of the Radon transform $\mathcal{R}$, the relationship between the ODF $f \in L^2(SO(3)/\mathcal{G})$ and the PDF $P \in L^2(\mathbb{S}^2/\mathcal{G} \times \mathbb{S}^2)$ of a specimen reads as

$$P(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}\big(\mathcal{R}f(\boldsymbol{x}, \boldsymbol{y}) + \mathcal{R}f(-\boldsymbol{x}, \boldsymbol{y})\big). \tag{5.1}$$

PDFs can be experimentally sampled by diffraction techniques like X-ray, neutron, or synchrotron diffraction, whereas ODFs cannot directly be measured by these techniques. A central problem in texture analysis is the estimation of the ODF of a specimen, given its measured PDF, cf. Bunge (1982); Schaeben and v.d. Boogaart (2003); v.d. Boogaart et al. (2007); Hielscher (2007). Of particular importance are the lower-order Fourier coefficients of the ODF since they characterize the macroscopic properties of the specimen, e.g., the second-order Fourier coefficients characterize thermal expansion, optical refraction index, and electrical conductivity, whereas the fourth-order Fourier coefficients characterize the elastic properties of the specimen, cf. (Bunge, 1982, sec. 13).

The inversion of the Radon transform $\mathcal{R}$ on the rotation group $SO(3)$ is an ill-posed inverse problem. Hielscher et al. (2008) present a novel approach to the numerical inversion of the Radon transform on $SO(3)$. Based on a Fourier slice theorem, the discrete inverse Radon transform of a function sampled on the product space $\mathbb{S}^2 \times \mathbb{S}^2$ is determined as the solution of a minimization problem, which is iteratively solved using fast Fourier techniques for $\mathbb{S}^2$ and $SO(3)$. In this application the unknown function f is a spherical tensor polynomial of the form

$$f = \sum_{n=0}^{N} \sum_{k_1=-n}^{n} \sum_{k_2=-n}^{n} \hat{f}_{n,k_1,k_2} Y_{n,k_1} Y_{n,k_2}\,. \tag{5.2}$$

instead of (3.1). Clearly, $f \in \Pi_{\boldsymbol{N}}(\mathbb{S}^2 \times \mathbb{S}^2)$ with $\boldsymbol{N} = (N, N)$, and Theorem 4.5 and Theorem 4.10 hold for the special f in (5.2). Due to the special structure of the sampling set $\mathcal{X}$ in the application (see (Hielscher et al., 2008, Definition 3.3)), the authors developed a fast matrix times vector multiplication for the forward transform (see (Hielscher et al., 2008, Theorem 3.5)). In the iterative reconstruction this faster method was used instead of Algorithm 3.1. Indeed, it was observed that the optimal interpolation problem (4.22) is stable solvable with the CGNE method if $N^2 > M$, see (Hielscher et al., 2008, Figure 4.2 b)) and compare with Corollary 4.11. On the other hand, the authors could solve the approximation problem (4.1) if the sampling set was δ dense, i.e., if enough pole figures are available, see (Hielscher et al., 2008, Figure 4.2 a)) and compare with Corollary 4.6. Note that after solving the above approximation problems, we are able to reconstruct the crystallographic orientation density function by a fast Fourier transform on $SO(3)$ thanks to the Fourier slice theorem, see (Hielscher et al., 2008, Theorem 2.7).

6 Conclusions

In this paper we solved the approximation and optimal interpolation of functions defined on the bisphere $\mathbb{S}^2 \times \mathbb{S}^2$ from scattered data. We demonstrate that the least square approximation to the function can be computed in a stable and efficient manner. The analysis of this problem is based on Marcinkiewicz–Zygmund inequalities for scattered data which we present here for the bisphere. The complementary problem of optimal interpolation is solved by using well-localized kernels for our setting. Based on this theoretical findings, we discussed the application and developed methods to problems from texture analysis in material science. We are able to reconstruct the crystallographic orientation density function by a fast Fourier transform on $SO(3)$ thanks to the Fourier slice theorem, see (Hielscher et al., 2008, Theorem 2.7).

Acknowledgment

The authors gratefully acknowledge support by German Research Foundation within the project PO 711/9-1 and FI 883/3-1. Moreover, we would like to thank the referees for their valuable suggestions.

References

Bass RF, Gröchenig K (2004). Random sampling of multivariate trigonometric polynomials. SIAM J Math Anal 36: 773–795.

Björck Å (1996). SIAM, Philadelphia, PA, USA

vd Boogaart KG, Hielscher R, Prestin J, Schaeben H (2007). Kernel-based methods for inversion of the radon transform on SO(3) and their applications to texture analysis. J Comput Appl Math 199:122–140

Bunge HJ (1982). Texture Analysis in Material Science. Butterworths

Erb W, Filbir F (2008). Approximation by positive definite functions on compact groups. Numer Funct Anal Optim 29:1082–1107

Filbir F, Mhaskar HN, Prestin J (2008). On a filter for exponentially localized kernels based on Jacobi polynomials. J Approx Theory 160:256–280

Filbir F, Potts D (2009). A proof for the Marcinkiewicz–Zygmund inequalities on the bi-sphere, Technical Report. Chemnitz University of Technology, Faculty of Mathematics, http://www.tu-chemnitz.de/~potts/publicationen

Filbir F, Themistoclakis W (2004). On the construction of de la Vallée Poussin means for orthogonal polynomials using convolution structures. J Comput Anal Appl 6:297–312

Filbir F, Themistoclakis W (2008). Polynomial approximation on the sphere using scattered data. Math Nachr 281:650–668

Gräf M, Kunis S (2008). Stability results for scattered data interpolation on the rotation group. Electron Trans Numer Anal 31:30–39

Gröchenig K (1992). Reconstruction algorithms in irregular sampling. Math Comput 59:181–194

Hielscher R (2007). The Radon Transform on the Rotation Group – Inversion and Application to Texture Analysis Dissertation, Department of Geology, Technical University Bergakademie Freiberg

Hielscher R, Potts D, Prestin J, Schaeben H, Schmalz M (2008). The Radon transform on SO(3): A Fourier slice theorem and numerical inversion. Inverse Problems 24:025011

Jetter K, Stöckler J, Ward JD (1999). Errors estimates for scattered data interpolation on spheres. Math Comput 68:733–747

Keiner J, Kunis S, Potts D (2007). Efficient reconstruction of functions on the sphere from scattered data. J Fourier Anal Appl 13:435–458

Keiner J, Potts D (2008). Fast evaluation of quadrature formulae on the sphere. Math Comput 77:397–419

Kocks UF, Tomé CN, Wenk HR, Mecking H (1998). Texture and Anisotropy. Cambridge University Press

Kunis S (2006). Nonequispaced FFT - Generalisation and Inversion Dissertation, Institut für Mathematik, Universität zu Lübeck, http://www.tu-chemnitz.de/~skunis/

Kunis S (2009). A note on stability results for scattered data interpolation on Euclidean spheres. Adv Comput Math 30:303–314

Matthies S, Vinel G, Helmig K (1987). Volume 1. Akademie-Verlag Berlin

Mhaskar HN, Narcowich FJ, Ward JD (2001). Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature. Math Comput 70:1113–1130 corrigendum on the positivity of the quadrature weights. in 71:453-454, 2002

Mhaskar HN, Prestin J (2009). Polynomial operators for spectral approximation of piecewise analytic functions. Appl Comput Harmon Anal 26:121–142

Narcowich FJ, Sivakumar N, Ward JD (1998). Stability results for scattered-data interpolation on Euclidean spheres. Adv Comput Math 8:137–163

Potts D, Steidl G, Tasche M (1998). Fast algorithms for discrete polynomial transforms. Math Comput 67:1577–1590

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

Randle V, Engler O (2000). Introduction to Texture Analysis Macrotexture: Microtexture and Orientation Mapping. Gordon and Breach Science Publishers

Schaeben H, vd Boogaart KG (2003). Spherical harmonics in texture analysis. Tectonophysics 370:253–268

Schmid D (2008). Marcinkiewicz–Zygmund inequalities and polynomial approximation from scattered data on $SO(3)$. Numer Funct Anal Optim 29:855–882

Szegő G (1975) Orthogonal Polynomials Amer. Math. Soc., Providence, RI, USA, 4th edition

Wenk HR (1985). Preferred Orientation in Deformed Metals and Rocks: An introduction to modern Texture Analysis. Academic Press, Orlando