Vertical Slices of the Sphere (WT)

Ralf Hielscher Michael Quellmalz

May 22, 2015

We present a novel algorithm for the inversion of the vertical slice transform, i.e. the transform that associates to a function on the two-dimensional unit sphere all integrals along circles that are parallel to one fixed direction. Our approach makes use of the singular value decomposition and resembles the mollifier approach by applying numerical integration with a reconstruction kernel via a quadrature rule. Considering the inversion problem as a statistical inverse problem, we find a family of asymptotically optimal mollifiers that minimize the maximum risk of the mean integrated error for functions within a Sobolev ball. By using fast spherical Fourier transforms and the fast Legendre transform, our algorithm can be implemented with almost linear complexity. In numerical experiments, we compare our algorithm with other approaches and illustrate our theoretical findings.

Math Subject Classifications. 65T40, 45Q05, 65N21, 44A12, 62G05.

Keywords and Phrases. Radon transform, statistical inverse problems, minimax risk, asymptotic bounds, fast algorithms.

1. Introduction

The problem of reconstructing a function from its integrals along certain lower-dimensional submanifolds has been studied since the early twentieth century. It is associated with the terms reconstructive integral geometry [27] and geometric tomography [10]. In 1913, Funk [8] described the problem of reconstructing a function on the two-sphere knowing its mean values along all great circles of the sphere. The computation of these mean values is known as the Funk–Radon transform or simply the Funk transform or spherical Radon transform.

What Funk did for great circles can be generalized to other classes of circles on the unit sphere $\mathbb{S}^2$. One can consider the integration along all circles with a fixed radius, which is known as the translation operator or the spherical section transform, cf. [34, 33, 6]. Another example is the spherical slice transform (see [1], [15, II.1.C]), which computes the means along all circles that contain a fixed point of the sphere.

In this paper, we look at vertical slices of the sphere, i.e., circles that are the intersections of the sphere with planes parallel to the ξ_3–axis. Let us denote with $\boldsymbol{e}_\sigma = (\cos\sigma, \sin\sigma, 0)^\top$, $\sigma \in [0, 2\pi)$, the point on the equator of the sphere with the latitude σ. For a continuous function $f\colon \mathbb{S}^2 \to \mathbb{C}$, we define

$$\mathcal{T} f\,(\sigma, t) = \frac{1}{2\pi\sqrt{1-t^2}} \int_{\boldsymbol{\xi}\cdot\boldsymbol{e}_\sigma = t} f\,(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}, \quad (\sigma, t) \in \mathbb{X} = \mathbb{T} \times (-1, 1),$$

Technische Universität Chemnitz, Faculty of Mathematics, 09107 Chemnitz, Germany, mails ralf.hielscher@mathematik.tu-chemnitz.de, michael.quellmalz@mathematik.tu-chemnitz.de

where the integration is carried out with respect to the arc length. For $t = \pm 1$, we set $\mathcal{T}f(\sigma, \pm 1) = f(\pm \boldsymbol{e}_\sigma)$. We will call $\mathcal{T}$ the vertical slice transform on the sphere.

The vertical slice transform first arose in an article by Gindikin et. al. [12] from 1994 that included an explicit inversion formula of $\mathcal{T}$, which is numerically instable. It reappeared in 2010 as the circular mean transform in a problem related to photoacoustic tomography, when Zangerl and Scherzer [44] described an algorithm for inverting the vertical slice transform using a connection to the circular Radon transform in the plane.

In this paper, we consider the reconstruction of f from the point of view of statistical inverse problems, cf. [4]. More precisely, we consider discrete data

$$g(\boldsymbol{x}_m) = \mathcal{T}f(\boldsymbol{x}_m) + \varepsilon(\boldsymbol{x}_m), \quad m = 1, \ldots, M,$$

at sampling points $\boldsymbol{x}_m \in \mathbb{X}$, $m = 1, \ldots, M$, which is perturbed by an uncorrelated white noise random field ε; and aim at optimal estimators $\mathcal{E}f$ of f given the data $g(\boldsymbol{x}_m)$. As measure of optimality, we use the mean integrated squared error (MISE)

$$\operatorname{MISE}(\mathcal{E}, f) = \mathbb{E}\,\|f - \mathcal{E}f\|^2_{L^2(\mathbb{S}^2)}.$$

Assuming that the function f is in the Sobolev ball

$$\mathscr{F}(s, S) = \left\{ f \in H^s_{\mathrm{e}}(\mathbb{S}^2) : \|f\|_{H^s(\mathbb{S}^2)} \leq S \right\},$$

where $s, S > 0$ and $H^s_{\mathrm{e}}(\mathbb{S}^2)$ is the Sobolev space of degree s restricted to functions that are even in the third component, cf. Section 2, we look at the so-called minimax risk, cf. [40, 4],

$$\inf_{\mathcal{E}} \sup_{f \in \mathscr{F}(s,S)} \mathbb{E}\,\|f - \mathcal{E}f\|^2_{L^2(\mathbb{S}^2)}. \tag{1.1}$$

The minimax risk of spherical deconvolution has been examined in [21, 18]. In this paper we follow the ideas of [18] and restrict ourselves to a specific class of estimators which resemble the mollifier approach in inverse problems, cf. [23]. Let $\omega_m \in \mathbb{R}$, $m = 1, \ldots, M$ be some quadrature weights for the sampling points $\boldsymbol{x}_m \in \mathbb{X}$ and

$$\mathcal{L}_N g = \sum_{n=0}^{N} \sum_{k=-n}^{n} \left(\sum_{m=1}^{M} w_m g(\boldsymbol{x}_m) \overline{B^k_n(\boldsymbol{x}_m)} \right) B^k_n, \quad g \in C(\mathbb{X}),$$

the corresponding hyperinterpolation operator [36] of degree N for the orthonormal basis system $B^k_n(\sigma, t) = \sqrt{\frac{2n+1}{4\pi}} \mathrm{e}^{\mathrm{i}k\sigma} P_n(t)$, $(\sigma, t) \in \mathbb{X}$. Then we define for a mollifier $\psi \colon [-1, 1] \to \mathbb{R}$ the estimator

$$\mathcal{E}_{N,\psi} g = \psi \star \left(\mathcal{T}^\dagger \mathcal{L}_N g \right) m$$

where $\star$ denotes the spherical convolution and $\mathcal{T}^\dagger$ is the generalized inverse of $\mathcal{T} \colon L^2(\mathbb{S}^2) \to L^2(\mathbb{X})$. Our aim is to find optimal mollifiers ψ^* for which the infimum in (1.1), restricted to the class of estimators $\mathcal{E}_{N,\psi}$, is attained asymptotically, as the number of sampling points goes to infinity.

Outline. This paper is structured as follows. After collecting some basic facts about spherical harmonics in Section 2, we introduce in Section 3 the vertical slice transform $\mathcal{T}$ and give in Theorem 3.3 its singular value decomposition as well as the asymptotic behavior of the singular values. Section 4 introduces the hyperinterpolation operator and discusses the relationship to the underlying quadrature rule. The main result of the paper is formulated in Section 5 in Theorem 5.4, which gives an asymptotically optimal mollifier for the vertical slice transform $\mathcal{T}$ in dependency of the

number of sampling points, the noise level and the Sobolev ball $\mathscr{F}(s, S)$. The proof of Theorem 5.4 has been postponed to Section 6, where we derive lower and upper bounds for the bias as well for the variance part of the MISE. Section 7 is devoted to numerical tests illustrating our theoretical findings. First of all we give a fast algorithm for our optimal estimator. This algorithm is based on the fast spherical Fourier transform [22] in conjunction with a fast Legendre transform and has the numerical complexity $\mathcal{O}(M \log^2 M)$. In the subsequent section, we compare our algorithms with two algorithms for the inversion of the planar Radon transform. This is possible since, by orthogonal projection, the vertical slice transform of a function f can be expressed as the planar Radon transform of the projected function multiplied with the weight $(1 - |\boldsymbol{x}|^2)^{-1/2}$. The main drawback of this approach is that the weight increases to infinity at the boundary of the disc. This becomes clearly visible when inverting the planar Radon transform with a backprojection-type algorithm [19] and a little less visible when an algorithm based on orthogonal polynomial expansion on the unit disc, cf. [42], is used. The reason for this improvement is that the latter algorithm uses, similarly as our algorithm, a sampling grid that becomes more dense close to the boundary.

2. Preliminaries and notation

In this section we are going to summarize some basic facts on harmonic analysis on the sphere as it can be found, e.g., in [6]. We denote with $\mathbb{Z}$ the set of integers and with $\mathbb{N}_0$ the nonnegative integers. We define the two-dimensional sphere $\mathbb{S}^2 = \{\boldsymbol{\xi} \in \mathbb{R}^3 \mid |\boldsymbol{\xi}| = 1\}$ as the set of unit vectors $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)^\top$ in the three-dimensional Euclidean space and make use of its parametrization in terms of the polar angles

$$\boldsymbol{\xi}(\theta, \rho) = (\cos\rho \sin\theta, \sin\rho \sin\theta, \cos\theta)^\top, \quad \theta \in [0, \pi],\ \rho \in [0, 2\pi).$$

Let $f\colon \mathbb{S}^2 \to \mathbb{C}$ be some measurable function. With respect to polar angles, the surface measure $\mathrm{d}\boldsymbol{\xi}$ on the sphere reads

$$\int_{\mathbb{S}^2} f(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} = \int_0^\pi \int_0^{2\pi} f(\boldsymbol{\xi}(\theta, \rho))\, \mathrm{d}\rho \sin\theta\, \mathrm{d}\theta,$$

The Hilbert space $L^2(\mathbb{S}^2)$ is the space of all measurable functions $f\colon \mathbb{S}^2 \to \mathbb{C}$, whose norm $\|f\|_{L^2} = \sqrt{\langle f, f\rangle}$ is finite, where $\langle f, g\rangle = \int_{\mathbb{S}^2} f(\boldsymbol{\xi})\overline{g(\boldsymbol{\xi})}\, \mathrm{d}\boldsymbol{\xi}$ denotes the usual L^2–inner product.

Spherical harmonics. We define the associated Legendre polynomials

$$P_n^k(t) = \frac{(-1)^k}{2^k k!} \left(1 - t^2\right)^{k/2} \frac{\mathrm{d}^{n+k}}{\mathrm{d}t^{n+k}} \left(t^2 - 1\right)^k, \quad t \in [-1, 1],$$

for all

$$(n, k) \in I := \{(n, k) \in \mathbb{N}_0 \times \mathbb{Z} \mid |k| \le n\}.$$

They satisfy the three-term recurrence relation [13, 8.735.2]

$$\sqrt{1 - t^2} P_n^{k+1}(t) = (n - k)\, t P_n^k(t) - (n + k)\, P_{n-1}^k(t), \quad t \in [-1, 1]. \tag{2.1}$$

The Legendre polynomials $P_n = P_n^0$ of degree $n \in \mathbb{N}$ form a system of orthogonal polynomials in $L^2([-1, 1])$ and satisfy the three-term recurrence relation

$$P_n(t) = \frac{2n - 1}{n} t P_{n-1}(t) - \frac{n - 1}{n} P_{n-2}(t), \quad t \in [-1, 1], \tag{2.2}$$

for $n \ge 1$ with the initialization $P_0(t) \equiv 1$ and $P_{-1}(t) \equiv 0$.

An orthonormal basis in the Hilbert space $L^2(\mathbb{S}^2)$ of square integrable functions on the sphere is formed by the so-called spherical harmonics

$$Y_n^k(\boldsymbol{\xi}(\theta,\rho)) = \sqrt{\frac{2n+1}{4\pi}\frac{(n-k)!}{(n+k)!}}\, P_n^k(\cos\theta)\mathrm{e}^{\mathrm{i}k\rho}, \quad (n,k)\in I. \tag{2.3}$$

Accordingly, any function $f \in L^2(\mathbb{S}^2)$ can be expressed by its Fourier series

$$f = \sum_{(n,k)\in I} \hat{f}(n,k) Y_n^k$$

with the Fourier coefficients

$$\hat{f}(n,k) = \int_{\mathbb{S}^2} f(\boldsymbol{\xi})\overline{Y_n^k(\boldsymbol{\xi})}\,\mathrm{d}\boldsymbol{\xi}, \quad (n,k)\in I,$$

and it satisfies Parseval's equality

$$\|f\|^2_{L^2(\mathbb{S}^2)} = \sum_{(n,k)\in I} \left|\hat{f}(n,k)\right|^2. \tag{2.4}$$

Sobolev spaces. A function $p\colon \mathbb{S}^2 \to \mathbb{C}$ that has a finite representation

$$p = \sum_{(n,k)\in I_N} \hat{p}(n,k) Y_n^k, \quad I_N = \{(n,k)\in I \mid n \le N\}$$

with respect to spherical harmonics is called a spherical polynomial of degree $N \in \mathbb{N}_0$ if $\hat{p}(N,k) \neq 0$ for some k. For such a spherical polynomial p of degree up to N and some $s \in \mathbb{R}$, we introduce the Sobolev norm

$$\|p\|^2_{H^s(\mathbb{S}^2)} = \sum_{(n,k)\in I_N} (n+\tfrac{1}{2})^{2s} \left|\hat{p}(n,k)\right|^2. \tag{2.5}$$

As usual, the Sobolev spaces $H^s(\mathbb{S}^2)$ are defined as the completion of the space of all spherical polynomials with respect to the Sobolev norm $\|\cdot\|_{H^s(\mathbb{S}^2)}$, cf. [7]. If $s > 1$, the Sobolev space $H^s(\mathbb{S}^2)$ is embedded in the space of continuous functions $C(\mathbb{S}^2)$. We define $H^s_{\mathrm{e}}(\mathbb{S}^2)$ as the subspace of $H^s(\mathbb{S}^2)$ that contains only functions that are even in the third coordinate. This definition is equal to saying that $H^s_{\mathrm{e}}(\mathbb{S}^2)$ is the subspace of functions whose Fourier coefficients vanish outside the set $I^{\mathrm{e}} = \{(n,k)\in I \mid n+k \text{ even}\}$.

Asymptotic analysis. We define some symbols comparing the asymptotic growth. Let $(a_n)_{n\in\mathbb{N}_0}$ and $(b_n)_{n\in\mathbb{N}_0}$ be two sequences of real numbers. We say $a_n \sim b_n$ for $n\to\infty$ if there exist constants $c_1, c_2 \in (0,\infty)$ and $n_0 \in \mathbb{N}_0$ such that $c_1 b_n \le a_n \le c_2 b_n$ for any $n \ge n_0$. We say $a_n \simeq b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$. Furthermore, we write $a_n \lesssim b_n$ if $\limsup_{n\to\infty} a_n/b_n \le 1$ and analogously $a_n \gtrsim b_n$ if $\liminf_{n\to\infty} a_n/b_n \ge 1$.

3. The vertical slice transform on the sphere

In the following definition we denote by $\mathbb{T} = [0, 2\pi)$ the 2π–periodic torus.

Definition 3.1. For $f \in C(\mathbb{S}^2)$, we define the vertical slice transform

$$\mathcal{T}f(\sigma,t) = \begin{cases} \dfrac{1}{2\pi\sqrt{1-t^2}} \displaystyle\int_{\boldsymbol{\xi}(\frac{\pi}{2},\sigma)\cdot\boldsymbol{\eta}=t} f(\boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\eta}, & \sigma\in\mathbb{T},\ t\in(-1,1) \\ f\left(\pm\boldsymbol{\xi}\left(\frac{\pi}{2},\sigma\right)\right), & \sigma\in\mathbb{T},\ t=\pm1, \end{cases} \tag{3.1}$$

where $\cdot$ denotes the scalar product. For simplicity, we will write $\boldsymbol{x} = (\sigma,t) \in \mathbb{X} = \mathbb{T}\times[-1,1]$.

The vector $\boldsymbol{\xi}(\pi/2, \sigma) = (\cos\sigma, \sin\sigma, 0)^\top$ is located at the equator of the sphere and has latitude σ. The vertical slice transform $\mathcal{T}$ computes the average values along all circles that are parallel to the ξ_3–axis. Since all these circles are symmetric with respect to the ξ_1–ξ_2–plane, $\mathcal{T}f$ vanishes for functions f that are odd in the third coordinate, i.e.

$$f(\xi_1, \xi_2, \xi_3) = -f(\xi_1, \xi_2, -\xi_3), \quad \boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3) \in \mathbb{S}^2.$$

Another symmetry property of the vertical slice transform is given by

$$\mathcal{T}f(\sigma, t) = \mathcal{T}f(\sigma + \pi, -t), \quad (\sigma, t) \in \mathbb{X},$$

where equality is to be understood modulo 2π in the σ variable. Hence, we can define an equivalence relation $\sim$ on $\mathbb{X}$ by saying $(\sigma, t) \sim (\sigma + \pi, -t)$. Then the quotient space of $\mathbb{X}$ by this equivalence relation is isomorphic to the Möbius strip.

An explicit inversion formula for operator $\mathcal{T}$ was shown in [12]. It is based on the inversion formula [15] of the Radon transform on the plane applied to the function

$$\tilde{f}\colon \mathbb{R}^2 \to \mathbb{R}, \quad \tilde{f}(\xi_1, \xi_2) = \begin{cases} \frac{f\left(\xi_1, \xi_2, \sqrt{1-\xi_1^2-\xi_2^2}\right)}{\sqrt{1-\xi_1^2-\xi_2^2,}}, & \text{for } \xi_1^2 + \xi_2^2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Theorem 3.2 (Gindikin et. al. [12])**.** *Let $f\colon \mathbb{S}^2 \to \mathbb{C}$ be even in the third coordinate. Then*

$$f(\boldsymbol{\xi}) = \frac{-\sqrt{1-\xi_1^2-\xi_2^2}}{4\pi} \int_{-\infty}^{\infty} \frac{1}{q} \int_0^{2\pi} \frac{\partial}{\partial q} \mathcal{T}f(\sigma, \xi_1\cos\sigma + \xi_2 \sin\sigma + q)\, d\sigma\, dq, \quad \boldsymbol{\xi} \in \mathbb{S}^2, \tag{3.2}$$

where the integral with respect to q has to be understood in the sense of Cauchy principal value and where $\mathcal{T}f(\sigma, t) = 0$ for $|t| > 1$.

The inversion formula (3.2) is unfavorable for numeric computations. The singular value decomposition of $\mathcal{T}$, which is given in the following theorem, promises more numerical stability.

Theorem 3.3 (Singular value decomposition)**.** *The vertical slice transform $\mathcal{T}$, as defined in* (3.1)*, extends continuously to a compact linear operator $\mathcal{T} : L^2\left(\mathbb{S}^2\right) \to L^2\left(\mathbb{X}\right)$ with the singular value decomposition*

$$\mathcal{T}Y_n^k = \lambda_n^k B_n^k, \quad (n, k) \in I, \tag{3.3}$$

consisting of the spherical harmonics Y_n^k from (2.3)*, the basis functions*

$$B_n^k\colon \mathbb{X} \to \mathbb{C},\ (\sigma, t) \mapsto \sqrt{\frac{2n+1}{4\pi}} \mathrm{e}^{\mathrm{i}k\sigma} P_n(t), \tag{3.4}$$

and the singular values

$$\lambda_n^k = \begin{cases} (-1)^{\frac{n+k}{2}} \frac{(n+k-1)!!}{(n-k)!!} \sqrt{\frac{(n-k)!}{(n+k)!}} = (-1)^{\frac{n+k}{2}} \frac{\sqrt{(n+k)!(n-k)!}}{2^n (\frac{n+k}{2})!(\frac{n-k}{2})!}, & (n, k) \in I^{\mathrm{e}} \\ 0, & (n, k) \in I \backslash I^{\mathrm{e}} \end{cases} \tag{3.5}$$

where $I^{\mathrm{e}} := \{(n, k) \in I \mid n + k \text{ even}\}$ and !! denotes the double factorial with $(-1)!! = 1$. Furthermore, we have for $(n, k) \in I^{\mathrm{e}}$

$$2^{1/2}(\pi n)^{-1/2} \simeq \left|\lambda_n^0\right| \leq \left|\lambda_n^k\right| \leq \left|\lambda_n^n\right| \simeq (\pi n)^{-1/4}, \quad n \to \infty. \tag{3.6}$$

Proof. Let $f \in C(\mathbb{S}^2)$, $(n,k) \in I$ and $(\sigma,t) \in \mathbb{X}$. The following generalization of the Funk–Hecke formula was proven in [3, Section 4.2],

$$\frac{1}{2\pi\sqrt{1-t^2}} \int_{\boldsymbol{\xi}\cdot\boldsymbol{\eta}=t} Y_n^k(\boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\eta} = P_n(t) Y_n^k(\boldsymbol{\xi}), \quad \boldsymbol{\xi} \in \mathbb{S}^2,\ t \in (-1,1). \tag{3.7}$$

Hence, the vertical slice transform applied to a spherical harmonic Y_n^k gives

$$\mathcal{T}Y_n^k(\sigma,t) = P_n(t)\, Y_n^k\left(\boldsymbol{\xi}\left(\frac{\pi}{2},\sigma\right)\right) = P_n(t)\sqrt{\frac{2n+1}{4\pi}}\sqrt{\frac{(n-k)!}{(n+k)!}}\mathrm{e}^{\mathrm{i}k\sigma} P_n^k(0).$$

In the next step, we compute the associated Legendre polynomials P_n^k at zero. According to the recurrence relation (2.2) for the Legendre polynomials, we have

$$P_n^0(0) = P_n(0) = \frac{-(n-1)P_{n-2}(0)}{n} = \frac{1+(-1)^n}{2}(-1)^{n/2}\frac{(n-1)!!}{n!!}$$

for $n \geq 1$ as well as $P_0(0) = 1$. With the recurrence relation (2.1) of the associated Legendre polynomials, we obtain for $n,k \geq 1$,

$$P_n^k(0) = (-1)^k \frac{(n+k-1)!!}{(n-k-1)!!} P_{n-k}^0(0) = (-1)^{(n+k)/2}\frac{(n+k-1)!!}{(n-k)!!}\frac{1+(-1)^{n-k}}{2}.$$

For $k<0$, we have by the symmetry relation

$$P_n^k(0) = (-1)^k \frac{(n+k)!}{(n-k)!} P_n^{-k}(0) = (-1)^{(n+k)/2}\frac{(n+k-1)!!}{(n-k)!!}\frac{1+(-1)^{n-k}}{2}$$

Hence, we have for all $(n,k) \in I^{\mathrm{e}}$

$$\mathcal{T}Y_n^k(\sigma,t) = (-1)^{(n+k)/2}\frac{(n+k-1)!!}{(n-k)!!}\frac{1+(-1)^{n-k}}{2}\sqrt{\frac{2n+1}{4\pi}}\sqrt{\frac{(n-k)!}{(n+k)!}} P_n(t)\,\mathrm{e}^{\mathrm{i}k\sigma},$$

which proves (3.3) and the left part of (3.5). The right side of (3.5) follows by replacing the double factorials with factorials using the well-known equalities $(2m)!! = 2^m m!$ and $(2m-1)!! = (2m)!/(2^m m!)$ for $m \in \mathbb{N}_0$.

Next, we prove (3.6). For $(n,k) \in I^{\mathrm{e}}$ with $0 \leq k \leq n-2$, we obtain the recurrence relation

$$\left|\lambda_n^{k+2}\right| = \left(\frac{n+k+1}{n+k+2}\frac{n-k}{n-k-1}\right)^{1/2}\left|\lambda_n^k\right| \geq \left|\lambda_n^k\right|.$$

Since $\lambda_n^k = (-1)^n\lambda_n^{-k}$, we observe that $\left|\lambda_n^0\right| \leq \left|\lambda_n^k\right| \leq |\lambda_n^n|$ for all $(n,k) \in I^{\mathrm{e}}$. The application of Stirling's formula $n! \simeq \sqrt{2\pi n}\, n^n \mathrm{e}^{-n}$ yields for $n \to \infty$

$$\lambda_n^0 = (-1)^{n/2}\frac{n!}{2^n\left(\frac{n}{2}!\right)^2} \simeq \frac{\sqrt{2\pi n}\, n^n \mathrm{e}^{-n}}{2^n \pi n n^n 2^{-n}\mathrm{e}^{-n}} = (-1)^{n/2}\sqrt{\frac{2}{\pi n}}$$

and

$$\lambda_n^n = (-1)^n\frac{\sqrt{(2n)!}}{2^n n!} \simeq \frac{\sqrt[4]{4\pi n}\, 2^n n^n \mathrm{e}^{-n}}{2^n\sqrt{2\pi n}\, n^n \mathrm{e}^{-n}} = (-1)^n\frac{1}{\sqrt[4]{\pi n}}.$$ ∎

The functions $\{B_n^k \mid n \in \mathbb{N}_0,\ k \in \mathbb{Z}\}$ from (3.4) form an orthonormal basis in $L^2(\mathbb{X})$. The fact that the singular values λ_n^k vanish for all $(n,k) \in I \backslash I^{\mathrm{e}}$ shows that the nullspace of $\mathcal{T}$ consists of all functions that are odd with respect to the third coordinate. Hence, the vertical slice transform $\mathcal{T}\colon L^2_{\mathrm{e}}(\mathbb{S}^2) \to L^2(\mathbb{X})$ is injective.

4. Quadrature and hyperinterpolation

For the set $\mathbb{X} = \mathbb{T} \times [-1, 1]$, we use a quadrature rule

$$Qg = \sum_{m=1}^{M} w_m g(\boldsymbol{x}_m), \quad g \in C(\mathbb{X}),$$

with quadrature nodes $\boldsymbol{x}_m \in \mathbb{X}$ and the respective weights w_m, $m = 1, \ldots, M$. For the orthonormal system of functions B_n^k from (3.4) in the space $L^2(\mathbb{X})$, we define an analogue to the trigonometric interpolation operator, namely

$$\mathcal{L}_N g := \sum_{(n,k)\in I_N^{\mathrm{e}}} \left(\sum_{m=1}^{M} w_m g(\boldsymbol{x}_m) \, \overline{B_n^k(\boldsymbol{x}_m)} \right) B_n^k, \quad g \in C(\mathbb{X}), \tag{4.1}$$

where $I_N^{\mathrm{e}} = I^{\mathrm{e}} \cap I_N = \{(n,k) \in \mathbb{N}_0 \times \mathbb{Z} \mid n \le N,\ |k| \le n,\ n+k \text{ even}\}$. In [36], the operator $\mathcal{L}_N$ is called hyperinterpolation of degree $N \in \mathbb{N}_0$. We call $\mathcal{L}_N$ exact if it is a projection operator, i.e., if $\mathcal{L}_N \mathcal{L}_N g = \mathcal{L}_N g$ for all $g \in C(\mathbb{X})$. The exactness of $\mathcal{L}_N$ implies that $Qg = \int g$ for all g in the span of $\{B_n^k \mid (n,k) \in I_N^{\mathrm{e}}\}$. Conversely, if $Qg = \int g$ for all g in the span of $\{B_n^k B_{n'}^{k'} \mid (n,k), (n',k') \in I_N^{\mathrm{e}}\}$, then the hyperinterpolation $\mathcal{L}_N$ is exact. We are going to impose the following condition on the hyperinterpolation $\mathcal{L}_N$.

Definition 4.1. Let $M(N)$, $N \in \mathbb{N}_0$, be a sequence of integers and $\mathcal{L}_N$, $N \in \mathbb{N}_0$, be a sequence of $M(N)$–point hyperinterpolations on $\mathbb{X}$ as defined in (4.1) with positive weights $w_m^N > 0$ and the nodes $\boldsymbol{x}_m^N \in \mathbb{X}$, $m = 1, \ldots, M(N)$. We call this sequence applicable if the following three conditions hold: $\mathcal{L}_N$ is exact for all $N \in \mathbb{N}_0$, $M(N) \in \mathcal{O}\left(N^2\right)$ for $N \to \infty$ and there exist constants $\gamma_1, \gamma_2 \in (0, \infty)$ independent of N such that for every sufficiently large N

$$\gamma_1 \frac{4\pi}{M(N)} \le \sum_{m=1}^{M(N)} \left| w_m^N B_n^k\left(\boldsymbol{x}_m^N\right) \right|^2 \le \gamma_2 \frac{4\pi}{M(N)}, \quad (n,k) \in I_N. \tag{4.2}$$

In the following theorem, we give an example of an applicable sequence of hyperinterpolations. These hyperinterpolations of degree N, $N \in \mathbb{N}_0$, make use of the tensor product of the $(2N+1)$–point equidistant quadrature on the torus $\mathbb{T}$ with the $(2N+1)$–point Fejér quadrature [11] on the unit interval $[-1, 1]$.

Theorem 4.2. *We define the Fejér hyperinterpolation of degree $N \in \mathbb{N}_0$ for $g \in C(\mathbb{X})$*

$$\mathcal{L}_N g := \sum_{(n,k)\in I_N^{\mathrm{e}}} \frac{1}{2N+1} \sum_{\ell,j=1}^{2N+1} \omega_j^{2N+1} g\left(\frac{2\pi\ell}{2N+1}, \cos\theta_j^{2N+1}\right) \overline{B_n^k\left(\frac{2\pi\ell}{2N+1}, \cos\theta_j^{2N+1}\right)} B_n^k, \tag{4.3}$$

where

$$\omega_j^{2N+1} = \frac{4}{2N+2} \sin\left(\theta_j^{2N+1}\right) \sum_{r=1}^{N} \frac{\sin\left((2r-1)\,\theta_j^{2N+1}\right)}{2r-1}, \qquad \theta_j^{2N+1} = \frac{j\pi}{2N+2} \tag{4.4}$$

for $j = 1, \ldots, 2N+1$. Then the so-defined sequence of hyperinterpolations $\mathcal{L}_N$ is applicable.

Proof. The proof is included in Section A of the appendix. ∎

5. Reconstruction

5.1. The mollifier method

We want to reconstruct a function f given its vertical slice transform $g = \mathcal{T}f$. The singular value decomposition in Theorem 3.3 shows that the nullspace of $\mathcal{T}$ contains exactly the functions that are odd in the third coordinate. Hence only the even part of f can be reconstructed. In the following, we assume that f is even in its third coordinate. Like it is the case for many other problems of this type, the inversion of the vertical slice transform is an ill-posed problem. We consider the discrete noisy data

$$g_m^\varepsilon = \mathcal{T}f(\boldsymbol{x}_m) + \varepsilon_m, \quad m = 1, \ldots, M$$

at data points $\boldsymbol{x}_m \in \mathbb{X},\ m = 1, \ldots, M$. We assume that the data error ε is a white noise field with $\varepsilon_m = \varepsilon(\boldsymbol{x}_m)$, $m = 1, \ldots, M$, which is

- unbiased, i.e. $\mathbb{E}(\varepsilon_m) = 0$,
- uncorrelated, i.e. $\mathbb{E}(\varepsilon_m \overline{\varepsilon_l}) = 0$ for $m \neq l$,
- and has variance $\delta^2 = \mathbb{E}(\varepsilon_m^2)$, where $\mathbb{E}$ denotes the expected value.

To overcome the ill-posedness of the inverse problem, we use the mollifier method, cf. [23]. We consider a polynomial $\psi\colon [-1, 1] \to \mathbb{R}$ of degree N and aim at computing the convolution

$$\psi \star f(\boldsymbol{\xi}) = \int_{\mathbb{S}^2} f(\boldsymbol{\eta})\, \psi(\boldsymbol{\eta} \cdot \boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\eta}, \quad \boldsymbol{\xi} \in \mathbb{S}^2. \tag{5.1}$$

For a sufficiently good choice of the so-called mollifier ψ, the convolution $\psi \star f$ is an approximation of the desired function f. The Funk–Hecke formula [9, 14] states that

$$\psi \star f = \sum_{(n,k) \in I^\mathrm{e}} \hat{\psi}(n) \left(\int_{\mathbb{X}} f(\boldsymbol{\xi}) \overline{Y_n^k(\boldsymbol{\xi})}\, \mathrm{d}\boldsymbol{\xi} \right) Y_n^k \tag{5.2}$$

with the Legendre coefficients

$$\hat{\psi}(n) = 2\pi \int_{-1}^{1} \psi(t) P_n(t)\, \mathrm{d}t, \quad n \in \mathbb{N}_0.$$

For $g = \mathcal{T}f$, the singular value decomposition from Theorem 3.3 yields

$$\psi \star f = \psi \star \mathcal{T}^\dagger g = \sum_{(n,k) \in I^\mathrm{e}} \frac{\hat{\psi}(n,k)}{\lambda_n^k} \left(\int_{\mathbb{X}} g(\boldsymbol{x}) \overline{B_n^k(\boldsymbol{x})}\, \mathrm{d}\boldsymbol{x} \right) Y_n^k.$$

Replacing the integral in the last equation with the quadrature rule from Section 4, we define for any $g \in C(\mathbb{S}^2)$ the estimator

$$\mathcal{E}_{N,\psi} g := \psi \star \mathcal{T}^\dagger \mathcal{L}_N g = \sum_{(n,k) \in I^\mathrm{e}} \frac{\hat{\psi}(n)}{\lambda_n^k} \sum_{m=1}^{M} w_m g(\boldsymbol{x}_m) \overline{B_n^k(\boldsymbol{x}_m)}\, Y_n^k. \tag{5.3}$$

5.2. Asymptotically optimal mollifiers

As a measure for the accuracy of the estimator $\mathcal{E}_{N,\psi}$, we use the mean integrated squared error (MISE) defined by

$$\operatorname{MISE}\left(\mathcal{E}_{N,\psi}, f\right) := \mathbb{E}\left\|\mathcal{E}_{N,\psi}(\mathcal{T}f + \varepsilon) - f\right\|^2_{L^2(\mathbb{S}^2)}.$$

We want to bound the MISE over some Sobolev balls

$$\mathscr{F}(s, S) := \left\{ f \in H^s_{\mathrm{e}}\left(\mathbb{S}^2\right) : \|f\|_{H^s(\mathbb{S}^2)} \leq S \right\}$$

where $s > 1$ and $S > 0$ are some constants and $H^s_{\mathrm{e}}(\mathbb{S}^2)$ is the Sobolev space of functions that are even in the third coordinate, cf. (**??**). The condition $s > 1$ ensures that f and therefore $\mathcal{T}f$ is continuous and hence the estimator $\mathcal{E}_{N,\psi}(\mathcal{T}f)$, which requires point evaluations, is well-defined. We are interested in the maximum risk

$$\sup_{f \in \mathscr{F}(s,S)} \operatorname{MISE}\left(\mathcal{E}_{N,\psi}, f\right). \tag{5.4}$$

We call a mollifier ψ^* optimal for degree N, if it minimizes (5.4) amongst all mollifiers. If a mollifier is optimal, it achieves the minimax error

$$\operatorname{MiniMax}(N) = \inf_{\psi} \sup_{f \in \mathscr{F}(s,S)} \operatorname{MISE}\left(\mathcal{E}_{N,\psi}, f\right).$$

We call a sequence of mollifiers ψ^*_N asymptotically optimal, if the maximum risk (5.4) for the mollifier ψ^*_N is asymptotically equal to the minimax rate for $N \to \infty$, i.e.

$$\sup_{f \in \mathscr{F}(s,S)} \operatorname{MISE}\left(\mathcal{E}_{N,\psi^*_N}, f\right) \simeq \inf_{\psi} \sup_{f \in \mathscr{F}(s,S)} \operatorname{MISE}\left(\mathcal{E}_{N,\psi}, f\right).$$

A class of asymptotically optimal mollifiers is given in the following definition.

Definition 5.1. For positive numbers s and L, we define the mollifiers $\psi^s_L \colon [-1, 1] \to \mathbb{R}$ as polynomials of degree $\lfloor L \rfloor$ by

$$\psi^s_L = \sum_{n=0}^{L} \frac{2n+1}{4\pi} \left(1 - \left(\frac{n + \frac{1}{2}}{L + \frac{1}{2}}\right)^s\right) P_n. \tag{5.5}$$

Remark 5.2. The mollifiers from Definition 5.2 can be seen as a generalization of some well-known mollifiers. They lie between the Fejér kernel, whose Legendre coefficients decrease linearly from 0 to L like those of ψ^1_L, and the Dirichlet kernel, which is the limit of ψ^s_L for $s \to \infty$. Furthermore, these mollifiers are similar to the CuP (cubic polynomial) scaling functions, whose Legendre coefficients are given by $\hat{\psi}(n) = (1 - n/L)^2(1 + 2n/L)$ for $n \leq L$ and 0 elsewhere, see [25]. The Fourier coefficients of the CuP functions are smoother at the upper end of the bandwidth $n = L$ compared to the mollifier (5.5), see Figure 1. □

The following theorem shows that the family of mollifiers (5.5) is asymptotically optimal for the class $\mathscr{F}(s, S)$.

Theorem 5.3. *Let $s > (1 + \sqrt{10})/2$, $S > 0$, and $\delta > 0$. Let the sequence $\mathcal{L}_N$, $N \in \mathbb{N}_0$, of $M(N)$–point hyperinterpolations be applicable in the sense of Definition 4.1. Furthermore, let the data error ε satisfy the conditions from Section 5.1 and have the standard deviation δ. Then there exist parameters $L(N)$ such that the sequence of mollifiers $\psi^s_{L(N)}$, $N \in \mathbb{N}_0$, is asymptotically optimal for*

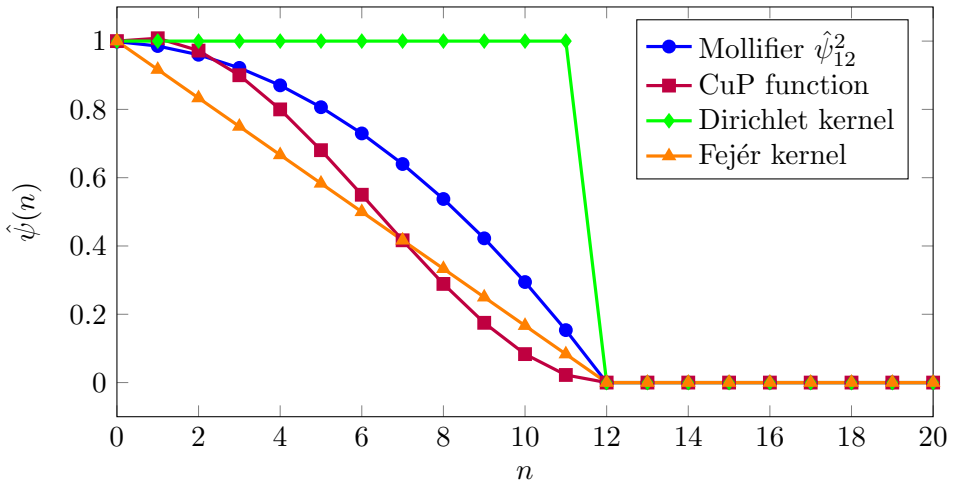


Figure 1: Legendre coefficients of the different mollifiers mentioned in Remark 5.3

the estimator $\mathcal{E}_{N,\psi}$ for the inversion of the vertical slice transform $\mathcal{T}$ of functions f belonging to the class $\mathscr{F}(s,S)$. In particular, we have for $N \to \infty$

$$\sup_{f\in\mathscr{F}(s,S)} \operatorname{MISE}\left(f, \mathcal{E}_{N,\psi^s_{L(N)}}\right) \simeq \inf_{\psi} \sup_{f\in\mathscr{F}(s,S)} \operatorname{MISE}\left(f, \mathcal{E}_{N,\psi}\right). \tag{5.6}$$

Furthermore, the minimax risk of the MISE is asymptotically for $N \to \infty$ bounded by

$$\begin{aligned}&\left(\gamma_1\delta^2 d_s \left(\frac{2sS^2}{3\delta^2 d_s}\right)^{\frac{3}{2s+3}} + S^2\left(\frac{2sS^2}{3\delta^2 d_s}\right)^{\frac{-2s}{2s+3}}\right) M^{\frac{-2s}{2s+3}} \lesssim \inf_{\psi} \sup_{f\in\mathscr{F}(s,S)} \operatorname{MISE}\left(\mathcal{E}_{N,\psi}, f\right)\\ &\lesssim \left(\gamma_2\delta^2 d_s \left(\frac{2sS^2}{3\delta^2 d_s}\right)^{\frac{3}{2s+3}} + S^2\left(\frac{2sS^2}{3\delta^2 d_s}\right)^{\frac{-2s}{2s+3}}\right) M^{\frac{-2s}{2s+3}},\end{aligned} \tag{5.7}$$

where

$$d_s = \frac{\pi^3 s^2}{3\,(2s+3)\,(s+3)}. \tag{5.8}$$

Remark 5.4. The mollifier approach (5.1) can be generalized to Fourier multiplications, where the coefficients $\hat{\psi}(n)$ in the Funk–Hecke formula (5.2) are replaced by coefficients that depend not only on n also on k. However, Theorem 5.4 would not change for this approach since the shape of the optimal mollifier ψ_N^s depends only on the Sobolev space $H^s(\mathbb{S}^2)$, which we will see in Lemma 6.7. More precisely, in the definition of the Sobolev norm in (**??**), the Fourier coefficients of p are multiplied by a factor independent of k. □

6. Proof of Theorem 5.4

In following, we will prepare the proof of Theorem 5.4, which marks the very end of this section.

Theorem 6.1. *The MISE can be decomposed into a bias and a variance term*

$$\mathbb{E}\,\|f - \mathcal{E}_{N,\psi}(\mathcal{T}f+\varepsilon)\|^2_{L^2(\mathbb{S}^2)} = \|f - \mathcal{E}_{N,\psi}\mathcal{T}f\|^2_{L^2(\mathbb{S}^2)} + \mathbb{E}\,\|\mathcal{E}_{N,\psi}\varepsilon\|^2_{L^2(\mathbb{S}^2)}. \tag{6.1}$$

Proof. Since the estimator $\mathcal{E}_{N,\psi}$ is linear, we have

$$\mathbb{E}\,\|f - \mathcal{E}_{N,\psi}(\mathcal{T}f+\varepsilon)\|^2_{L^2(\mathbb{S}^2)} = \mathbb{E}\,\|f - \mathcal{E}_{N,\psi}\mathcal{T}f\|^2_{L^2(\mathbb{S}^2)} + \mathbb{E}\,\|\mathcal{E}_{N,\psi}\varepsilon\|^2_{L^2(\mathbb{S}^2)} - 2\mathbb{E}\,\langle f - \mathcal{E}_{N,\psi}\mathcal{T}f, \mathcal{E}_{N,\psi}\varepsilon\rangle,$$

where the last summand vanishes because $\mathbb{E}(\mathcal{E}_{N,\psi}\varepsilon) = 0$. ■

The decomposition (6.1) is well-known, see [40, Section 1.2.1]. In the following two subsections, we derive bounds for the variance and bias error. In Section 6.3, the proof of the optimality of the mollifiers (5.5) follows eventually.

6.1. Variance error

Proposition 6.2. Let $f \in H^s_{\mathrm{e}}(\mathbb{S}^2)$ for $s > 1$, $N \in \mathbb{N}$, $\psi \in L^2([-1,1])$, and the hyperinterpolation $\mathcal{L}_N$ be applicable in the sense of Definition 4.1. If N is sufficiently large, the variance term of (6.1) can be estimated by

$$\gamma_1 \frac{4\pi}{M}\delta^2 \left\|\hat{\psi}\right\|_\lambda^2 \leq \mathbb{E}\left\|\mathcal{E}_{N,\psi}\varepsilon\right\|^2_{L^2(\mathbb{S}^2)} \leq \gamma_2 \frac{4\pi}{M}\delta^2 \left\|\hat{\psi}\right\|_\lambda^2, \tag{6.2}$$

where the constants γ_1, γ_2 are from Definition 4.1 and

$$\left\|\hat{\psi}\right\|_\lambda^2 := \sum_{(n,k)\in I_N^{\mathrm{e}}} \left|\frac{\hat{\psi}(n)}{\lambda_n^k}\right|^2.$$

Proof. By Parseval's equality (2.4), (5.2) and the uncorrelatedness of the noise ε from Section 5.1,

$$\begin{aligned}
\mathbb{E}\left\|\mathcal{E}_{N,\psi}(\varepsilon)\right\|^2_{L^2(\mathbb{S}^2)} &= \sum_{(n,k)\in I_N^{\mathrm{e}}} \mathbb{E}\left|\frac{\hat{\psi}(n)}{\lambda_n^k}\sum_{m=1}^M w_m \varepsilon_m \overline{B_n^k(\boldsymbol{x}_m)}\right|^2 \\
&= \sum_{(n,k)\in I_N^{\mathrm{e}}} \left|\frac{\hat{\psi}(n)}{\lambda_n^k}\right|^2 \sum_{m,l=1}^M w_m w_l \overline{B_n^k(\boldsymbol{x}_m)} B_n^k(\boldsymbol{x}_l) \mathbb{E}\varepsilon_m \overline{\varepsilon_l} \\
&= \delta^2 \sum_{(n,k)\in I_N^{\mathrm{e}}} \left|\frac{\hat{\psi}(n)}{\lambda_n^k}\right|^2 \sum_{m=1}^M \left|w_m B_n^k(\boldsymbol{x}_m)\right|^2.
\end{aligned}$$

The applicability condition (4.2) implies (6.2). ∎

Lemma 6.3. *Let ψ_L^s be the mollifier defined in* (5.5)*. Then for $L \to \infty$*

$$\left\|\hat{\psi}_L^s\right\|_\lambda^2 = d_s L^3 + \mathcal{O}(L^2), \tag{6.3}$$

where d_s is given in (5.8)*.*

Proof. Let $(n,k) \in I^{\mathrm{e}}$. At first we calculate an asymptotic expression of the singular values λ_n^k. The following version of Stirling's formula was proven in [32]. For $n = 1, 2, \ldots$

$$n! = \sqrt{2\pi} n^{n+1/2} \mathrm{e}^{-n} \mathrm{e}^{r(n)},$$

where $0 < r(n) < 1/(12n)$. The application of this formula to the singular values yields

$$\begin{aligned}
\left|\lambda_n^k\right|^2 =& \frac{2\pi\,(n+k)^{n+k+1/2}\,\mathrm{e}^{-(n+k)}\mathrm{e}^{r(n+k)} \cdot (n-k)^{n-k+1/2}\,\mathrm{e}^{-(n-k)}\mathrm{e}^{r(n-k)}}{2^{2n}(2\pi)^2\left(\frac{n+k}{2}\right)^{n+k+1}\mathrm{e}^{-(n+k)}\mathrm{e}^{2r((n+k)/2)} \cdot \left(\frac{n-k}{2}\right)^{n-k+1}\mathrm{e}^{-(n-k)}\mathrm{e}^{2r((n-k)/2)}} \\
=& \frac{2}{\pi}(n+k)^{-1/2}(n-k)^{-1/2}\mathrm{e}^{R(n,k)}
\end{aligned} \tag{6.4}$$

for $|k| < n$ with the error term

$$R(n,k) = r(n+k) + r(n-k) - 2r\left(\frac{n+k}{2}\right) - 2r\left(\frac{n-k}{2}\right),$$

which is for $k \geq 0$ bounded by

$$|R(n,k)| < \frac{2}{3(n-k)}.$$

In the next step, we insert the approximation (6.4) and the Taylor series of the exponential function in order to compute the sum

$$\sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \frac{1}{|\lambda_n^k|^2} = \frac{\pi}{2} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \sqrt{n^2-k^2} \mathrm{e}^{-R(n,k)} = \frac{\pi}{2} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \sqrt{n^2-k^2} \Big(1 - R(n,k) + \dots\Big). \tag{6.5}$$

The first summand of the right hand side of (6.5) gives

$$\frac{\pi}{2} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \sqrt{n^2-k^2} = \frac{\pi}{4} \int_{-n}^{n} \sqrt{n^2-k^2}\,\mathrm{d}k + \mathcal{O}(n) = \frac{\pi^2}{8} n^2 + \mathcal{O}(n).$$

For the second summand of the right side of (6.5), we see that

$$\begin{aligned}
\frac{\pi}{2} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \sqrt{n^2-k^2}\,|R(n,k)| \leq & \pi \sum_{\substack{k=0 \\ 2|(n+k)}}^{n-1} \frac{\sqrt{n+k}\sqrt{n-k}}{n-k} + \mathcal{O}(n) \\
\leq & \pi \int_0^n \sqrt{\frac{n+k}{n-k}}\,\mathrm{d}k + \mathcal{O}(n) = \frac{1}{2} n\pi(2+\pi) + \mathcal{O}(n).
\end{aligned}$$

Since the other summands are even smaller than the second, these parts can also be bounded by $\mathcal{O}(n)$. Now we can insert the Fourier coefficients of the mollifier (5.5) and see that for $L \to \infty$

$$\begin{aligned}
\|\hat{\psi}_L^s\|_\lambda^2 = \sum_{n=0}^{L} \left|\hat{\psi}_L^s(n)\right|^2 \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \frac{1}{|\lambda_n^k|^2} &= \frac{\pi^2}{8} \sum_{n=0}^{L} \left(1 - \left(\frac{n+1/2}{L+1/2}\right)^s\right)^2 \left(n^2 + \mathcal{O}(n)\right) \\
&= \frac{\pi^2}{8} \int_0^L n^2 \left(1 - \left(\frac{n+1/2}{L+1/2}\right)^s\right)^2 \mathrm{d}n + \mathcal{O}\left(L^2\right) \\
&= \frac{\pi^2 s^2}{12\,(2s+3)\,(s+3)} L^3 + \mathcal{O}\left(L^2\right),
\end{aligned}$$

which proves (6.3). ■

6.2. Bias error

With the triangle inequality, we split up the bias error into a smoothing and an aliasing part,

$$\left\| f - \psi \star \mathcal{T}^\dagger \mathcal{L}_N \mathcal{T} f \right\|_{L^2(\mathbb{S}^2)} + \left\| \psi \star \left(f - \mathcal{T}^\dagger \mathcal{L}_N \mathcal{T} f \right) \right\|_{L^2(\mathbb{S}^2)}.$$

6.2.1. Smoothing error

If $f \in H^s\left(\mathbb{S}^2\right)$ for $s \geq 0$, then Parseval's equality (2.4) and (5.2) imply that the smoothing error is bounded by

$$\|f - \psi \star f\|_{L^2(\mathbb{S}^2)} \leq \sup_{n \in \mathbb{N}_0} \left(\frac{\left|1 - \hat{\psi}(n)\right|}{\left(n + \frac{1}{2}\right)^s} \right) \|f\|_{H^s(\mathbb{S}^2)}.$$

In particular, for the mollifier ψ_L^s from (5.5) and a function $f \in \mathscr{F}(s,S)$, we have the smoothing error

$$\|f - \psi_L^s \star f\|_{L^2(\mathbb{S}^2)} \leq \frac{S}{\left(L + \frac{1}{2}\right)^s}. \tag{6.6}$$

6.2.2. Aliasing error

To derive an upper bound for the aliasing error $\|\psi \star (f - \mathcal{T}^\dagger \mathcal{L}_N \mathcal{T} f)\|_{L^2(\mathbb{S}^2)}$, we define the Sobolev space $H^s_{\mathcal{T}}(\mathbb{X}) := \mathcal{T} H^s(\mathbb{S}^2)$ with the corresponding norm for $g = \mathcal{T} f$

$$\|g\|_{H^s_{\mathcal{T}}(\mathbb{X})} := \|f\|_{H^s(\mathbb{S}^2)} = \sum_{(n,k)\in I^{\mathrm{e}}} \left| \hat{g}(n,k) \left(n + \frac{1}{2}\right)^s \frac{1}{\lambda_n^k} \right|^2 . \tag{6.7}$$

Lemma 6.4. *Let $g \in C(\mathbb{X})$, $N \in \mathbb{N}$, and the hyperinterpolation $\mathcal{L}_N$ from (4.1) be exact. Denote with $\mathcal{P}_N$ the orthogonal projection onto the range of $\mathcal{L}_N$. Then*

$$\|\mathcal{L}_N\|_{C(\mathbb{X})\to L^2(\mathbb{X})} = \sqrt{4\pi}. \tag{6.8}$$

The previous lemma was shown in [36]. The constant 4π equals the surface area of the manifold $\mathbb{X}$. The following lemma and the theorem thereafter are based on [16], where they were proven similarly for the Sobolev spaces $H^s(\mathbb{S}^2)$ on the two-sphere.

Lemma 6.5. *For $s > 1$, the Sobolev space $H^s_{\mathcal{T}}(\mathbb{X})$ can be embedded continuously into the space of continuous functions $C(\mathbb{X})$. Let $N \in \mathbb{N}$, then there exists a constant $c > 0$ independent of N such that for $g \in \operatorname{span}\{B_n^k \mid (n,k) \in I^{\mathrm{e}} \backslash I_N^{\mathrm{e}}\}$,*

$$\|g\|_{C(\mathbb{X})} \le c \left(N + \frac{1}{2}\right)^{1-s} \|g\|_{H^s_{\mathcal{T}}(\mathbb{X})} . \tag{6.9}$$

and

$$\|g\|_{H^0_{\mathcal{T}}(\mathbb{X})} \le \left(N + \frac{1}{2}\right)^{-s} \|g\|_{H^s_{\mathcal{T}}(\mathbb{X})} . \tag{6.10}$$

Proof. The embedding property follows from the Sobolev embedding theorem, which says that $H^s(\mathbb{S}^2)$ is embedded in the continuous functions for $s > 1$, and the fact that $\mathcal{T}$ maps continuous functions to continuous functions. We have $I^{\mathrm{e}} \backslash I_N^{\mathrm{e}} = \{(n,k) \mid n \in \mathbb{N}_0,\ k = -n, \dots, n,\ n+k \text{ even}\}$. From the definition of the spherical harmonics in (2.3), we observe that

$$\sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left|\lambda_n^k\right|^2 = \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \frac{(n-k)!}{(n+k)!} \left|P_n^k(0)\right|^2 \le \frac{4\pi}{2n+1} \sum_{k=-n}^{n} \left|Y_n^k(0, \frac{\pi}{2})\right|^2 = 1,$$

where the last equality follows directly from the addition formula for spherical harmonics, cf. [6, (1.2.8)]. We denote with $\mathcal{P}_N$ the L^2-orthogonal projection onto the range of $\mathcal{L}_N$. Let $g \in H^s_{\mathcal{T}}(\mathbb{X})$. Then, by the Cauchy–Schwartz inequality and the definition of B_n^k, we have for all $(\sigma, t) \in \mathbb{X}$

$$\begin{aligned}
|[g - \mathcal{P}_N g](\sigma,t)| &= \left| \sum_{(n,k)\in I^{\mathrm{e}}\backslash I_N^{\mathrm{e}}} \hat{g}(n,k)\, B_n^k(\sigma,t) \right| \\
&= \sum_{(n,k)\in I^{\mathrm{e}}\backslash I_N^{\mathrm{e}}} \left| \hat{g}(n,k) \frac{\left(n+\frac{1}{2}\right)^s}{\lambda_n^k} \right| \left| \sqrt{\frac{2n+1}{4\pi}} \mathrm{e}^{\mathrm{i}k\sigma} P_n(t) \lambda_n^k \right| \left(n + \frac{1}{2}\right)^{-s} \\
&\le \|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{X})} \sqrt{\sum_{n=N+1}^{\infty} \left(n + \frac{1}{2}\right)^{1-2s}} .
\end{aligned}$$

Since $s > 1$, the last sum converges and there exists a constant c_1 independent of N such that (6.9) holds. Analogously, (6.10) follows by the calculation

$$\|g - \mathcal{P}_N g\|^2_{H^0_{\mathcal{T}}(\mathbb{X})} = \sum_{(n,k)\in I^{\mathrm{e}}\backslash I_N^{\mathrm{e}}} |\hat{g}(n,k)|^2 \frac{(n+\frac{1}{2})^{2s}}{(n+\frac{1}{2})^{2s}} \le \left(N + \frac{1}{2}\right)^{-2s} \|g - \mathcal{P}_N g\|^2_{H^s_{\mathcal{T}}(\mathbb{X})} . \qquad \blacksquare$$

Now we are able to prove the following upper bound for the aliasing error.

Theorem 6.6. *Let $s > 1$, $f \in H^s_{\mathrm{e}}(\mathbb{S}^2)$ and $g = \mathcal{T} f$. Furthermore, let for every $N \in \mathbb{N}$ the hyperinterpolation $\mathcal{L}_N$ be exact. If the mollifier ψ satisfies $\left|\hat{\psi}(n)\right| \leq 1$ for all $n \in \mathbb{N}_0$, then there exists a constant $c > 0$ that is independent of f and N so that*

$$\left\| \psi \star \mathcal{T}^\dagger \left(g - \mathcal{L}_N g \right) \right\|_{L^2(\mathbb{S}^2)} \leq c \left(N + \frac{1}{2} \right)^{3/2-s} \|f\|_{H^s(\mathbb{S}^2)} . \tag{6.11}$$

Proof. We first note that, because $s > 1$, g is continuous and hence $\mathcal{L}_N g$ is well-defined. We will show the estimate

$$\left\| \mathcal{T}^\dagger \left(\mathcal{L}_N g - g \right) \right\|_{L^2(\mathbb{S}^2)} \leq c \left(N + \frac{1}{2} \right)^{3/2-s} \|f\|_{H^s(\mathbb{S}^2)} ,$$

from which the claimed formula (6.11) follows by the assumptions on ψ and (5.2). Let $\mathcal{P}_N$ denote the orthogonal projection onto the range of $\mathcal{L}_N$. Since $\mathcal{L}_N$ is a projector, we have $\mathcal{L}_N \mathcal{P}_N = \mathcal{P}_N$ and therefore

$$\begin{aligned} \left\| \mathcal{T}^\dagger \left(\mathcal{L}_N g - g \right) \right\|_{L^2(\mathbb{S}^2)} &= \|\mathcal{L}_N g - g\|_{H^0_{\mathcal{T}}(\mathbb{X})} \\ &= \|\mathcal{L}_N \left(g - \mathcal{P}_N g \right) - \left(g - \mathcal{P}_N g \right)\|_{H^0_{\mathcal{T}}(\mathbb{X})} \\ &\leq \|\mathcal{L}_N \left(g - \mathcal{P}_N g \right)\|_{H^0_{\mathcal{T}}(\mathbb{X})} + \|g - \mathcal{P}_N g\|_{H^0_{\mathcal{T}}(\mathbb{X})} . \end{aligned} \tag{6.12}$$

Now we examine both summands of (6.12). For the second summand, we obtain by (6.10) from the previous lemma

$$\|g - \mathcal{P}_N g\|_{H^0_{\mathcal{T}}(\mathbb{X})} \leq \left(N + \frac{1}{2} \right)^{-s} \|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{X})} . \tag{6.13}$$

For the first summand of (6.12), we have by (3.6)

$$\begin{aligned} \|\mathcal{L}_N \left(g - \mathcal{P}_N g \right)\|_{H^0_{\mathcal{T}}(\mathbb{X})} &= \sum_{(n,k) \in I^{\mathrm{e}}} \left| \frac{[\mathcal{L}_N (g - \mathcal{P}_N g)]^\wedge (n,k)}{\lambda_n^k} \right|^2 \\ &\leq c \left(N + \frac{1}{2} \right)^{1/2} \|\mathcal{L}_N (g - \mathcal{P}_N g)\|_{L^2(\mathbb{X})} . \end{aligned} \tag{6.14}$$

By Lemma 6.4,

$$\|\mathcal{L}_N \left(g - \mathcal{P}_N g \right)\|_{L^2(\mathbb{X})} \leq \sqrt{4\pi} \, \|g - \mathcal{P}_N g\|_{C(\mathbb{X})} . \tag{6.15}$$

By (6.9), there exists a constant c_1 independent of g and N such that

$$\|g - \mathcal{P}_N g\|_{C(\mathbb{X})} \leq c_1 \left(N + \frac{1}{2} \right)^{1-s} \|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{X})} . \tag{6.16}$$

Combining (6.12), (6.13), (6.14), (6.15), and (6.16) yields

$$\left\| \mathcal{T}^\dagger \left(\mathcal{L}_N g - g \right) \right\|_{L^2(\mathbb{S}^2)} \leq c \left(N + \frac{1}{2} \right)^{3/2-s} \|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{X})}$$

where c is some positive constant. Since $\mathcal{P}_N$ is a projection, $\|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{X})} \leq \|g\|_{H^s_{\mathcal{T}}(\mathbb{X})}$, which finally proves the theorem. ∎

6.3. Proof of the optimality

Lemma 6.7. *Let $\psi \in L^2(\mathbb{S}^2)$, $s > 1$ and $S > 0$. Furthermore, let the estimator $\mathcal{E}_{N,\psi}$ of the function $f \in H^s(\mathbb{S}^2)$ and the noise ε be as in Section 5.1, and the hyperinterpolation $\mathcal{L}_N$ be exact. We define L^* as the largest number $L \in \mathbb{R}$ such that $\hat{\psi}_L^s(n) \leq \left|\hat{\psi}(n)\right|$ for all $n \in \mathbb{N}_0$. Then, for every sufficiently large N, the maximum risk (5.4) is bounded from below by*

$$\sup_{f \in \mathscr{F}(s,S)} \mathbb{E}\,\|f - \mathcal{E}_{N,\psi}(\mathcal{T}f + \varepsilon)\|_{L^2(\mathbb{S}^2)}^2 \geq \mathbb{E}\,\|\mathcal{E}_{N,\psi_{L^*}^s}\varepsilon\|_{L^2(\mathbb{S}^2)}^2 + \frac{S^2}{\left(L^* + \frac{1}{2}\right)^{2s}}. \tag{6.17}$$

Proof. We choose an integer $n^* \in \mathbb{N}_0$ for which $\hat{\psi}_{L^*}^s(n^*) = \left|\hat{\psi}(n^*)\right|$. Depending on wheather n^* is even or odd, we define the function f^* to be a multiple of a spherical harmonic $Y_{n^*}^0$ or $Y_{n^*}^1$, respectively, such that $\|f^*\|_{H^s(\mathbb{S}^2)} = S$. Hence we have $f^* \in \mathscr{F}(s,S)$. By (6.1),

$$\begin{aligned}\sup_{f \in \mathscr{F}(s,S)} \mathbb{E}\,\|f - \mathcal{E}_{N,\psi}(\mathcal{T}f + \varepsilon)\|_{L^2(\mathbb{S}^2)}^2 &\geq \mathbb{E}\,\|f^* - \mathcal{E}_{N,\psi}(\mathcal{T}f^* + \varepsilon)\|_{L^2(S^2)}^2 \\ &= \|f^* - \mathcal{E}_{N,\psi}\mathcal{T}f^*\|_{L^2(\mathbb{S}^2)}^2 + \mathbb{E}\,\|\mathcal{E}_{N,\psi}\varepsilon\|_{L^2(\mathbb{S}^2)}^2.\end{aligned}$$

These two summands are the bias and variance error with respect to f^*. If N is larger than n^*, the aliasing error $\left\|\psi \star \mathcal{T}^\dagger\left(\mathcal{T}f^* - \mathcal{L}_N\mathcal{T}f^*\right)\right\|_{L^2(\mathbb{S}^2)}$ vanishes since f^* is a polynomial of degree n^*. Using (5.2) and the definition of $\psi_{L^*}^s$, the smoothing error reads

$$\begin{aligned}\|f^* - \mathcal{E}_{N,\psi}(\mathcal{T}f^*)\|_{L^2(\mathbb{S}^2)}^2 &= \|f^* - \psi \star f^*\|_{L^2(\mathbb{S}^2)}^2 \\ &= \sum_{(n,k) \in I_N^{\mathrm{e}}} \left|\left(1 - \hat{\psi}(n)\right)\hat{f}^*(n,k)\right|^2 \\ &= \left(1 - \hat{\psi}_{L^*}^s(n^*)\right)^2 \frac{S^2}{\left(n^* + \frac{1}{2}\right)^{2s}} = \|f^* - \psi_{L^*}^s \star f^*\|_{L^2(\mathbb{S}^2)}^2.\end{aligned}$$

By the definition of L^*, we have $\left|\hat{\psi}(n)\right| \geq \hat{\psi}_{L^*}^s(n)$ for all $n \in \mathbb{N}_0$. Therefore, Proposition 6.2 implies that the variance error for the mollifier $\psi_{L^*}^s$ is not smaller than for ψ. The claimed formula (6.17) follows with Proposition 6.2. ∎

Now we can finally prove the main theorem. The idea behind this proof is, to calculate the parameter L such that the variance and smoothing error are about equal and to show that the aliasing error is asymptotically smaller if the smoothness parameter s is sufficiently large.

Proof of Theorem 5.4. We note that if $N \to \infty$, also the number of points $M = M(N)$ must go to infinity since $\mathcal{L}_N$ is assumed to be exact. If N and thus M is sufficiently large, then, by Lemma 6.7, there exists an $L > 0$ such that

$$\inf_{\psi \in L^2[-1,1]} \sup_{f \in \mathscr{F}(s,S)} \mathbb{E}\,\|f - \mathcal{E}_{N,\psi}f\|_{L^2(\mathbb{S}^2)}^2 \geq \mathbb{E}\,\|\mathcal{E}_{N,\psi_L^s}\varepsilon\|_{L^2(\mathbb{S}^2)}^2 + \frac{S^2}{\left(L + \frac{1}{2}\right)^{2s}} \tag{6.18}$$

$$\gtrsim \gamma_1 \frac{\delta^2 d_s L^3}{M} + \frac{S^2}{\left(L + \frac{1}{2}\right)^{2s}}, \tag{6.19}$$

where we have plugged in (6.3) and (6.2), and the constant d_s is defined in (5.8). We want to minimize the right side of (6.18) with respect to L and denote the optimal argument with $L(M)$.

In order to minimize (6.19), we compute the zeros of its derivative with respect to L,

$$(-2s)\,S^2 L^{-2s-1} + 3\gamma_1 \frac{\delta^2 d_s}{M} L^2 = 0 \iff 3\gamma_1 \frac{3\delta^2 d_s}{M} L^{2s+3} = 2sS^2.$$

Hence, (6.19) is asymptotically minimized for $M \to \infty$ by the choice

$$L = L_1(M) := \left(\frac{2sS^2}{3\gamma_1\delta^2 d_s} M\right)^{\frac{1}{2s+3}}. \tag{6.20}$$

Plugging this value for $L_1(M)$ for L into (6.19) yields the following asymptotic lower bound of the maximum risk for $M \to \infty$,

$$\mathbb{E}\,\|\mathcal{E}_{N,\psi_L^s}\varepsilon\|^2_{L^2(\mathbb{S}^2)} + \frac{S^2}{\left(L+\frac{1}{2}\right)^{2s}} \gtrsim \left(\gamma_1\delta^2 d_s \left(\frac{2sS^2}{3\gamma_1\delta^2 d_s}\right)^{\frac{3}{2s+3}} + S^2\left(\frac{2sS^2}{3\gamma_1\delta^2 d_s}\right)^{\frac{-2s}{2s+3}}\right) M^{\frac{-2s}{2s+3}},$$

which gives the first inequality of (5.7). Here we also see that, for the choice $L_1(M)$, both terms of the sum in the right side of (6.19) decrease of the order $M^{\frac{-2s}{2s+3}}$.

In the second part of the proof, we derive the upper bound of the maximum risk. By the decomposition (6.1) combined with the upper bounds from (6.3), (6.6) and (6.11), there exists a constant $c > 0$ such that for $f \in \mathscr{F}(s,S)$

$$\begin{aligned}\mathbb{E}\left\|f - \mathcal{E}_{N,\psi_L^s}(\mathcal{T}f+\varepsilon)\right\|^2_{L^2(\mathbb{S}^2)} &\le \left(\frac{S}{\left(L+\frac{1}{2}\right)^s} + \left\|\psi_L^s \star \mathcal{T}^\dagger(\mathcal{T}f - \mathcal{L}_N\mathcal{T}f)\right\|_{L^2(\mathbb{S}^2)}\right)^2 + \mathbb{E}\left\|\mathcal{E}_{N,\psi_L^s}\varepsilon\right\|^2_{L^2(\mathbb{S}^2)} \\ &\lesssim \left(\frac{S}{L^s} + c\left(N+\frac{1}{2}\right)^{3/2-s} S\right)^2 + \gamma_2\frac{\delta^2 d_s L^3}{M}.\end{aligned} \tag{6.21}$$

Now we plug in $L_1(M)$ from the first part of the proof and we obtain

$$\begin{aligned}\mathbb{E}\left\|f - \mathcal{E}_{N,\psi_L^s} f\right\|^2_{L^2(\mathbb{S}^2)} \lesssim \left(S\left(\frac{2sS^2}{3\gamma_1\delta^2 d_s}\right)^{\frac{-s}{2s+3}} M^{\frac{-s}{2s+3}} + c\left(N+\frac{1}{2}\right)^{3/2-s} S\right)^2 \\ + \gamma_2\delta^2 d_s\left(\frac{2sS^2}{3\gamma_1\delta^2 d_s}\right)^{\frac{3}{2s+3}} M^{\frac{-2s}{2s+3}}.\end{aligned} \tag{6.22}$$

By the applicability, we have $M = \mathcal{O}\left(N^2\right)$. We have also assumed that $s > (1+\sqrt{10})/2$ which implies $3/2 - s < -s/(2s+3)$. Hence, only those terms in the sum in (6.21) that grow like $M^{-2s/(2s+3)}$ play a significant role as M goes to infinity. Consequently, the term $c(N+1/2)^{3/2-s}S$, which comes from the aliasing error, is asymptotically negligible for $M \to \infty$. This yields the upper bound

$$\sup_{f\in\mathscr{F}(s,S)} \mathbb{E}\left\|f - \mathcal{E}_{N,\psi^s_{L_1(M)}} f\right\|^2_{L^2(\mathbb{S}^2)} \cdot M^{\frac{2s}{2s+3}} \lesssim \gamma_2\delta^2 d_s\left(\frac{2sS^2}{3\gamma_1\delta^2 d_s}\right)^{\frac{3}{2s+3}} + S^2\left(\frac{2sS^2}{3\gamma_1\delta^2 d_s}\right)^{\frac{-2s}{2s+3}} \tag{6.23}$$

for N and thus $M(N) \to \infty$, which proves the second inequality of (5.7).

It is left to show the asymptotic optimality (5.6) of the family of mollifiers ψ_L^s to complete the proof. When we insert the optimal value $L(M)$ into both (6.18) and (6.20), those two bounds coincide except for the aliasing term, but since $L(M)$ grows of the same asymptotic order as $L_1(M)$, the aliasing part is again negligible. ■

7. Numerical tests

7.1. Spherical Fourier algorithm

The algorithm. The estimator $\mathcal{E}_{N,\psi}g$ can be computed numerically in a fast and stable way. To this end, we decompose (5.3) into a three-step process. Suppose we have given a quadrature rule on $\mathbb{X}$ with the nodes $\boldsymbol{x}_m \in \mathbb{X}$ and the weights w_m as well as function values $g(\boldsymbol{x}_m)$ for $m = 1, \ldots, M$.

i) Compute the Fourier coefficients of g:

$$\widehat{\mathcal{L}_N g}(n,k) := \sum_{m=1}^{M} w_m g(\boldsymbol{x}_m)\overline{B_n^k(\boldsymbol{x}_m)}, \quad (n,k) \in I_N^{\mathrm{e}}. \tag{7.1}$$

If the quadrature uses the tensor product structure like the one in Theorem 4.2, this computation can be accelerated yielding to a complexity of $\mathcal{O}(N^2 \log^2 N)$ operations, which we will show below.

ii) Do the inversion and regularization in the Fourier space:

$$\widehat{\mathcal{E}_{N,\psi} f}(n,k) := \frac{\hat{\psi}(n)}{\lambda_n^k} \widehat{\mathcal{L}_N g}(n,k), \quad (n,k) \in I_N^{\mathrm{e}},$$

which needs $\mathcal{O}(N^2)$ arithmetic operations.

iii) Compute the estimator at some nodes $\boldsymbol{\xi}_j \in \mathbb{S}^2$ from its Fourier coefficients:

$$\mathcal{E}_{N,\psi} f(\boldsymbol{\xi}_j) := \sum_{(n,k)\in I^{\mathrm{e}}} \widehat{\mathcal{E}_{N,\psi} f}(n,k) Y_n^k(\boldsymbol{\xi}_j), \quad j = 1, \dots, J.$$

This computation can also be done in $\mathcal{O}(N^2 \log^2 N)$ steps using the fast spherical Fourier transform from [22], provided that the number of evaluation points $J \in \mathcal{O}(N^2)$.

This algorithm has an overall numeric complexity of $\mathcal{O}(N^2 \log^2 N)$.

Efficient computation of the first step. Next, we are going to use a tensor product ansatz in step i). Suppose we have given a tensor product quadrature rule on $\mathbb{X}$ with the nodes $(\sigma_\ell, t_j) \in \mathbb{X}$ and the weights $w_{\ell,j}$ as well as function values $g(\sigma_\ell, t_j)$ for $\ell = 1, \dots, L$ and $j = 1, \dots, J$ and the mollifier coefficients $\hat{\psi}(n)$. For simplicity, we assume that $L = 2N+1$ and $J = N+1$. Furthermore, we are going to use equidistant nodes $\sigma_\ell = 2\pi\ell/L$ on the torus. Then equation (7.1) becomes for $n = 0, \dots, N$ and $k = -n, \dots, n$

$$\widehat{\mathcal{L}_N g}(n,k) := \frac{2\pi}{L}\sqrt{\frac{2n+1}{4\pi}} \sum_{j=1}^{J} \omega_j \left[\sum_{\ell=1}^{L} g\left(\frac{2\pi\ell}{L}, t_j\right) \mathrm{e}^{-\mathrm{i}k2\pi\ell/L}\right] P_n(t_j) \tag{7.2}$$

$$= \frac{2\pi}{L}\sqrt{\frac{2n+1}{4\pi}} \sum_{j=1}^{J} \omega_j A(k', j) P_n(t_j), \tag{7.3}$$

where $A(k', j)$ is the term in brackets from equation (7.2) and $k' := k + N \in \{0, \dots, 2N\}$. We first look at the computation of

$$A(k', j) = \sum_{l=0}^{L-1} g\left(\frac{2\pi\ell}{L}, t_j\right) \mathrm{e}^{-2\pi\mathrm{i}\ell k'/L} \mathrm{e}^{2\pi\mathrm{i}\ell N/L} = \mathrm{FFT}\left[\left(g\left(\frac{2\pi\ell}{L}, t_j\right)\right)_{\ell=0}^{L-1}\right](k') \mathrm{e}^{2\pi\mathrm{i}k'N/L},$$

where the fast Fourier transform FFT [5] of a vector $(z_\ell)_{\ell=0}^{L-1}$ is defined by

$$\mathrm{FFT}\left[(z_\ell)_{\ell=0}^{L-1}\right](k) = \sum_{l=0}^{L-1} z_\ell \mathrm{e}^{-2\pi\mathrm{i}\ell k/L}, \quad k = 0, \dots, L-1.$$

The FFT of length L has an arithmetic complexity of $\mathcal{O}(L \log L)$. We compute J FFTs of length L, which implies a complexity of $\mathcal{O}(JL \log L)$.

In equation (7.3), the sum over j is a discrete Legendre transform of length J, which is done $2N+1$ times (for every $k = -n, \dots, n$). Discrete Legendre transforms of length J can be computed in $\mathcal{O}(J \log^2 J)$ operations, cp. [29]. So the computation of (7.3) has a complexity of $\mathcal{O}(NJ \log^2 J)$.

7.2. Inversion via the planar Radon transform

The vertical slice transform on the sphere is closely related to the Radon transform in the plane which was first described in [30]. Before we show this connection, we need the following lemma, that gives a parametric formula for computing the vertical slice transform $\mathcal{T}$.

Lemma 7.1.

$$\mathcal{T}f\,(\sigma,t) = \frac{1}{2\pi}\int_0^{2\pi} f\left(t\begin{pmatrix}\cos\sigma\\ \sin\sigma\\ 0\end{pmatrix} + \sqrt{1-t^2}\begin{pmatrix}-\sin\sigma\cos\varphi\\ \cos\sigma\cos\varphi\\ \sin\varphi\end{pmatrix}\right)\mathrm{d}\varphi, \quad (\sigma,t)\in\mathbb{X}.$$

Proof. We first look at the case $\sigma = 0$. Then the integration domain of (3.1) is a circle perpendicular to the ξ_1 axis with center $(t,0,0)^\top$ and radius $\sqrt{1-t^2}$. Note that the definition of $\mathcal{T}$ is normalized with respect to the circumference of these circles. We have

$$\mathcal{T}f(0,t) = \frac{1}{2\pi}\int_0^{2\pi} f\left(t\begin{pmatrix}1\\ 0\\ 0\end{pmatrix} + \sqrt{1-t^2}\begin{pmatrix}0\\ \cos\varphi\\ \sin\varphi\end{pmatrix}\right)\mathrm{d}\varphi.$$

The claimed formula follows by rotation around the ξ_3 axis with the angle σ. ∎

The (planar) Radon transform on the unit disc $\mathbb{B}^2 = \{\boldsymbol{\xi}\in\mathbb{R}^2 \mid \xi_1^2+\xi_2^2<1\}$ computes the integrals of a function $\tilde{f}:\mathbb{R}^2\to\mathbb{C}$ along all line segments, (cf. [30]), i.e.

$$\mathcal{R}\tilde{f}(\sigma,t) = \int_{-\sqrt{1-t^2}}^{\sqrt{1-t^2}} \tilde{f}\left(t\begin{pmatrix}\cos\sigma\\ \sin\sigma\end{pmatrix} + \begin{pmatrix}-\sin\sigma\\ \cos\sigma\end{pmatrix}u\right)\mathrm{d}u, \quad (\sigma,t)\in\mathbb{T}\times(-1,1). \tag{7.4}$$

Theorem 7.2. *For a function $f\colon\mathbb{S}^2\to\mathbb{C}$ that is even in the third coordinate, denote with*

$$\tilde{f}:\mathbb{B}^2\to\mathbb{C},\ (\xi_1,\xi_2)\mapsto \frac{f\left((\xi_1,\xi_2,\sqrt{1-\xi_1^2-\xi_2^2})^\top\right)}{\pi\sqrt{1-\xi_1^2-\xi_2^2}},$$

its weighted orthogonal projection onto the ξ_1–ξ_2–plane. Then

$$\mathcal{T}f(\sigma,t) = \mathcal{R}\tilde{f}\,(\sigma,t)\,, \quad (\sigma,t)\in\mathbb{T}\times(-1,1)\,.$$

Remark. The function $\tilde{f}$ can be interpreted as follows: the function f defined on the sphere is projected orthogonally onto the ξ_1–ξ_2 plane, and then divided by the weight $\pi\sqrt{1-\xi_1^2-\xi_2^2}$. This projection is well-defined since f is even with respect to ξ_3.

Proof. According to Lemma 7.1 and because f is even in the third component, $\mathcal{T}f$ can be expressed in terms of $\tilde{f}$ by

$$\mathcal{T}f\,(\sigma,t) = \int_0^{\pi}\tilde{f}\left(t\begin{pmatrix}\cos\sigma\\ \sin\sigma\end{pmatrix} + \sqrt{1-t^2}\cos\varphi\begin{pmatrix}-\sin\sigma\\ \cos\sigma\end{pmatrix}\right)\sqrt{1-t^2-(1-t^2)\,(\cos\varphi)^2}\,\mathrm{d}\varphi$$

for all $(\sigma,t)\in\mathbb{T}\times(-1,1)$ since

$$\left\|t\begin{pmatrix}\cos\sigma\\ \sin\sigma\end{pmatrix} + \sqrt{1-t^2}\begin{pmatrix}-\cos\varphi\sin\sigma\\ \cos\varphi\cos\sigma\end{pmatrix}\right\|^2 = t^2 + \left(1-t^2\right)(\cos\varphi)^2\,.$$

Performing the substitution

$$u = \sqrt{1-t^2}\cos\varphi, \quad \mathrm{d}u = \sqrt{1-t^2}\sin\varphi\,\mathrm{d}\varphi = \sqrt{1-t^2}\sqrt{1-\frac{u^2}{1-t^2}}\,\mathrm{d}\varphi = \sqrt{1-t^2-u^2}\,\mathrm{d}\varphi$$

leads to

$$\mathcal{T}f(\sigma,t) = \frac{1}{2\pi}\int_{-\sqrt{1-t^2}}^{\sqrt{1-t^2}}\tilde{f}\left(t\begin{pmatrix}\cos\sigma\\ \sin\sigma\end{pmatrix} + u\begin{pmatrix}-\sin\sigma\\ \cos\sigma\end{pmatrix}\right)\mathrm{d}u.$$

∎

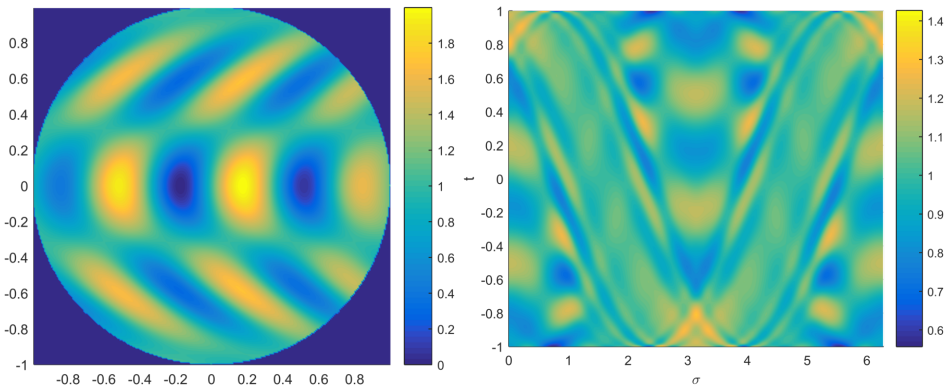


Figure 2: Test function f from (7.6) projected onto the ξ_1–ξ_2 plane (left) and its vertical slice transform $\mathcal{T}f$ (right)

This theorem gives a nice way of reconstructing a function $f \in C(\mathbb{S}^2)$ given its vertical slice transform $\mathcal{T}f$, namely

$$f(\boldsymbol{\xi}) = \pi\sqrt{1-\xi_1^2-\xi_2^2}\,\mathcal{R}^{-1}[\mathcal{T}f](\xi_1,\xi_2), \quad \boldsymbol{\xi} \in \{(\xi_1,\xi_2,\xi_3) \in \mathbb{S}^2 \mid \xi_3 \neq 0\}, \tag{7.5}$$

where we first compute the inverse Radon transform of $\mathcal{T}f$ and then multiply it by $\pi\sqrt{1-\xi_1^2-\xi_2^2}$. However, if the function f does not vanish on the equator, the respective function $\tilde{f}(\xi_1,\xi_2)$ grows infinitely near the boundary of the unit disc.

The inversion of the Radon transform has been treated by many authors, see e.g. [30, 35, 19, 26, 28, 43, 37]. For our comparison we focus on two algorithms: the standard filtered back projection algorithm, [19], and an algorithm based on orthogonal polynomial expansion on the unit disc (OPED), [42]. One important difference between both algorithms is that in the standard filtered back projection algorithm the Radon transform $\mathcal{R}\tilde{f}(\sigma,t)$ is sampled uniformly in t while the OPED algorithm requires a sampling at Chebyshev points. For the standard backprojection algorithm, we use the `iradon` routine from the Matlab imaging toolbox, and for the algorithm based on orthogonal polynomial expansion on the unit disc, we use the so-called fast OPED algorithm, [43]. The numerical complexity of both algorithms is $\mathcal{O}(N^3)$, provided we have sampled $\mathcal{R}\tilde{f}(\sigma,t)$ at $\mathcal{O}(N^2)$ points and we compute f at the same number of points.

7.3. Comparison of the algorithms

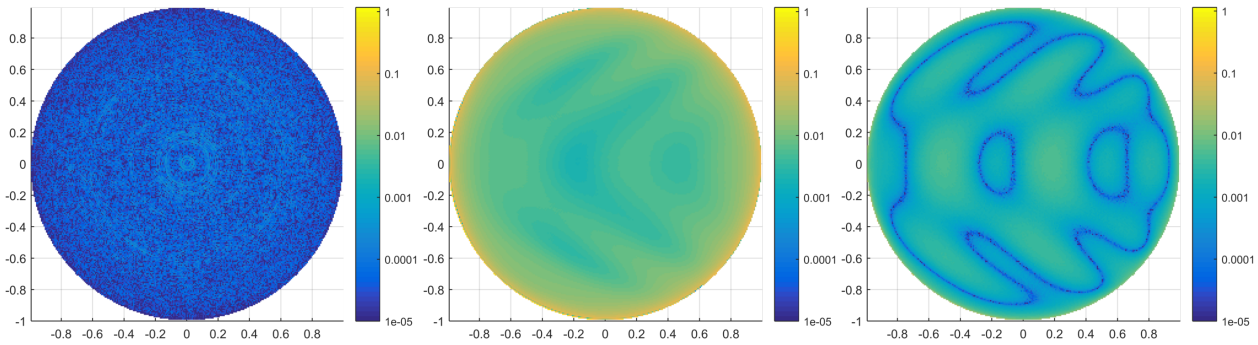
In this section we numeric results to compare our algorithm based on spherical harmonics with algorithms based on the planar Radon transform.

Smooth test function. We chose the test function

$$f(\boldsymbol{\xi}) = \sin(9(\xi_1-\xi_2^2))\exp(-\xi_1^4-\xi_2^2)\cos(5\xi_2)+1, \quad \boldsymbol{\xi} \in \mathbb{S}^2. \tag{7.6}$$

Since f is even with respect to ξ_3, it suffices to consider the projection of f onto the ξ_1–ξ_2 plane, which is depicted in Figure 2. Its vertical slice transform $\mathcal{T}f$ is computed by quadrature applied to (3.1). Although the above algorithms we want to compare could easily be inverted, we apply a simple quadrature-based method for the computation of the forward transform to avoid the so-called inverse crime [41].

At first, we apply the three algorithms to undisturbed data. The first two algorithms are applied to same number of $M = 513 \cdot 257 = 131841$ data points, while the OPED algorithm uses $513^2 = 263169$ ponts. However, while for the backprojection algorithm the grid is equally spaced; the grid



(a) NFSFT (MSE $5.48 \cdot 10^{-9}$) (b) Radon (MSE $1.96 \cdot 10^{-3}$) (c) OPED (MSE $5.58 \cdot 10^{-5}$)

Figure 3: Logarithmic plot of the error of the reconstruction with exact data for the different algorithms with the respective mean squared errors (MSE) in parentheses.

for the OPED algorithms as well for our algorithm based on spherical harmonics is denser close to the boundary $t = \pm 1$.

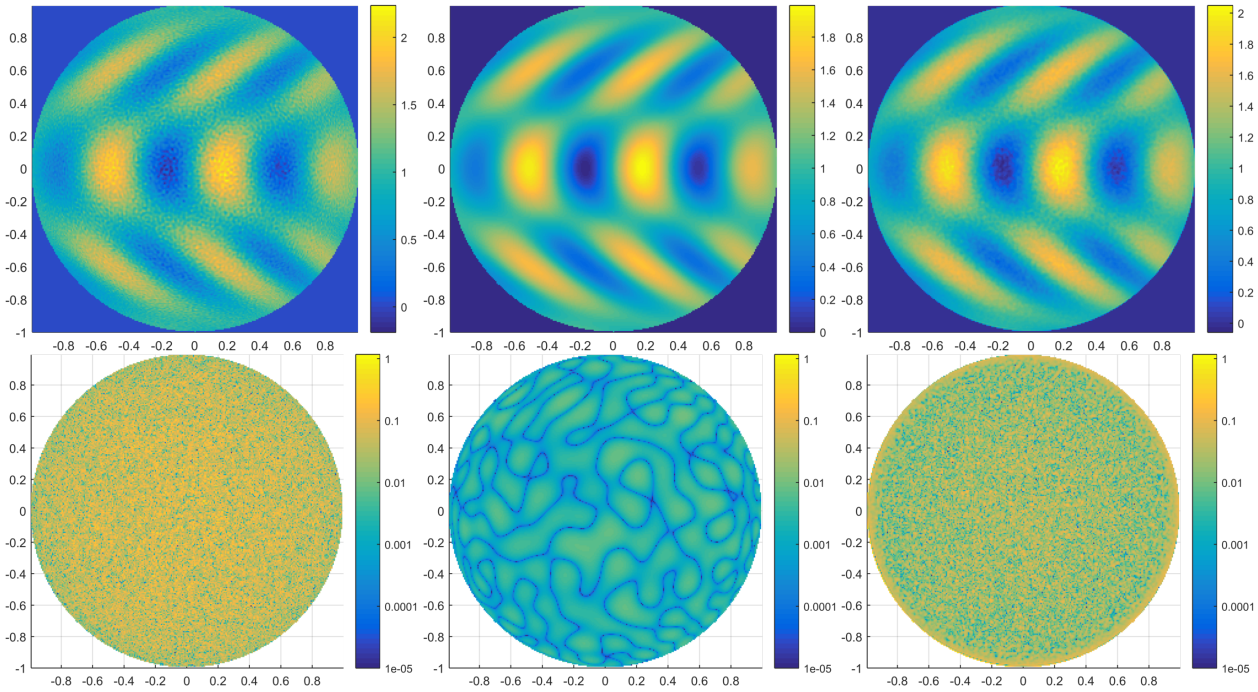
In Figure 3, the reconstruction error for the three methods is illustrated. One can clearly see the main disadvantage of the Radon algorithm, which is due to the function $\tilde{f}$ being unbounded near the boundary on the unit disc. Even though $\tilde{f}$ is multiplied by $\sqrt{1-\xi_1^2-\xi_2^2}$, the error of the reconstruction is still considerably larger near the boundary. The better error of the OPED reconstruction close to the boundary of the disc compare to back projection algorithm is due to the different sampling and the fact that the algorithm is specifically designed for functions supported on the disc. Nevertheless, the OPED reconstruction does not meet the accuracy of the algorithm based on spherical harmonics.

For our next test, we add some Gaussian noise with a standard deviation of 0.01 to the data. In Figure 4, one can see the reconstruction error for the spherical Fourier algorithm without any regularization, which is the same as using the Dirichlet kernel of degree $N = 256$ as mollifier. Figure 4b shows the error with regularization where the regularization parameter L is chosen optimally to reduce the L^2–error. The reconstruction with the Radon algorithm shown in Figure 4c looks slightly better then one with the unregularized Fourier algorithm, as long as we stay away from the boundary. For the inverse Radon transform, we used the Hann window, which already performs a regularization. The OPED reconstruction is not shown because it does not include a regularization.

To illustrate the minimax rate of the MISE from Theorem 5.4, we did the following. We chose the parameter $s = 2$ and computed the norm $S = \|f\|_{H^2(\mathbb{S}^2)} \approx 166$ numerically. Then, for different degrees N, we computed the parameter L as in (**??**) (with $\gamma_1 = \gamma_2 = 1$), reconstructied the function f using the Fourier algorithm as we did above with this value of L, and computed the integrated squared error $\|f - \mathcal{E}_{N,\psi_N^2}(\mathcal{T}f + \varepsilon)\|^2_{L^2(\mathbb{S}^2)}$. Here, we used the Gauss–Legendre nodes on the sphere. In order to get an estimate of the MISE, we repeated this procedure **20** times with a new instance of the error ε each time, computed the mean of the integrated squared error and compared it with the theoretical minimax rate (5.7). As one can see in Figure **??**, our achieved error almost coincides with the theoretical rate. Actually, the achieved error is slightly higher than the theoreical rate, which may be explained by tow facts: we set $\gamma_2 = 1$, where this value might be slightly bigger than one, and inaccuracies caused by rounding in the numerical computation occur.

todo

Non-smooth test function. For this test, we chose a function whose vertical slice transform is known analytically. In [39], explicit formulas for the Radon transform of some simple functions are given. If we take such a function $\tilde{f}(\xi_1, \xi_2)$ supported in the unit disc $\xi_1^2 + \xi_2^2 < 1$ in the plane, we



(a) NFSFT with Dirichlet kernel (MSE $6.97 \cdot 10^{-3}$)

(b) NFSFT with regularization (MSE $8.28 \cdot 10^{-6}$)

(c) Radon with Hann filter (MSE $2.97 \cdot 10^{-3}$)

Figure 4: Reconstruction from noisy data with standard deviation 0.01 in the top row and a logarithmic plot of respective errors $|f - \mathcal{E}f|$ in the bottom row

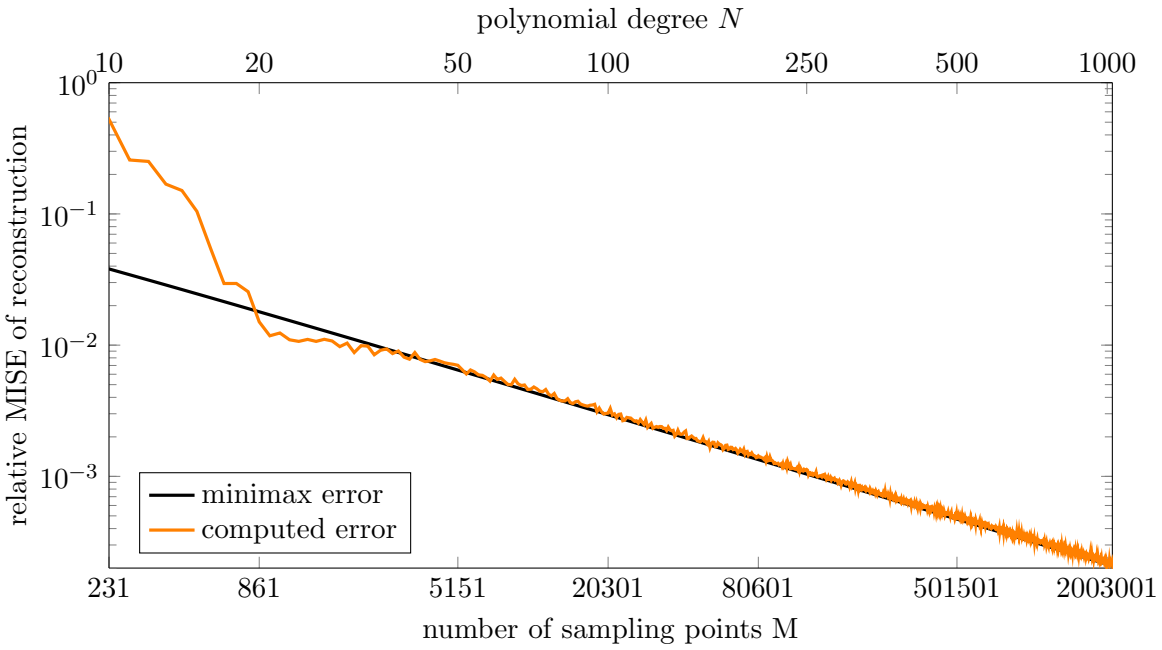


Figure 5: Log–log plot of the MISE of the Fourier reconstruction compared with the minimax rate

can use the construction from Theorem 7.2 to define a test function on the sphere by

$$f(\boldsymbol{\xi}) = \pi\sqrt{1-\xi_1^2-\xi_2^2}\,\tilde{f}(\xi_1,\xi_2), \quad \boldsymbol{\xi} \in \mathbb{S}^2,$$

and Theorem 7.2 gives us $\mathcal{T}f = \mathcal{R}\tilde{f}$. As such a function, we define

$$\tilde{h}(\xi_1,\xi_2) = \begin{cases} 1-\sqrt{\xi_1^2+\xi_2^2}, & \xi_1^2+\xi_2^2 < 1 \\ 0, & \text{otherwise,} \end{cases}$$

whose graph is a circular cone with height 1 and radius 1, centered at the origin. Since $\tilde{h}$ is radially symmetric with respect to the origin, its Radon transform $\mathcal{R}\tilde{h}(\sigma,t)$ is independent of σ. Hence, we can choose $\sigma = 0$ and see that $\mathcal{R}\tilde{h}(\sigma,t)$ is the integral along the line $\xi_1 = t$ and obtain

$$\mathcal{R}\tilde{h}(\sigma,t) = \int_{-\sqrt{1-t^2}}^{\sqrt{1-t^2}} 1-\sqrt{r^2+t^2}\,\mathrm{d}r = \sqrt{1-t^2} - t^2\,\mathrm{Arcsch}\left(\frac{x}{\sqrt{1-t^2}}\right), \quad (\sigma,t) \in \mathbb{X},$$

where Arcsch denotes the inverse hyperbolic cosecant function, see [2, Section 4.6]. To define the function $\tilde{f}$, we first scale h with the factors $a, b > 0$ in direction of ξ_1 and ξ_2, then rotate it through the angle α about the origin and shift it to $(\zeta_1,\zeta_2) \in \mathbb{R}^2$. We set

$$\tilde{f}(\xi_1,\xi_2) := \tilde{h}\left(\frac{(\xi_1-\zeta_1)\cos(\alpha)+(\xi_2-\zeta_2)\sin(\alpha)}{a}, \frac{-(\xi_1-\zeta_1)\sin(\alpha)+(\xi_2-\zeta_2)\cos(\alpha)}{b}\right). \quad (7.7)$$

Then, by [39, Appendix B], its Radon transform is given by

$$\mathcal{R}\tilde{f}(\sigma,t) = ab\mathcal{R}\tilde{h}\left(\arctan\left(\frac{b}{a\tan\sigma}\right), (t-\zeta_1\cos\sigma-\zeta_2\sin\sigma)\sqrt{a^2\cos(\sigma-\alpha)^2+b^2\sin(\sigma-\alpha)^2}\right).$$

We made a particular choice of $\tilde{f}_0$ as the linear combination of functions of the form (7.7), which is illustrated in Figure 5. Since the Radon transform is linear, we also have an explicit formula for its vertical slice transform $\mathcal{T}f = \mathcal{R}\tilde{f}$. The reconstructions of f with the three algorithms for exact data are plotted in Figure 6. The Radon algorithm yields much better results then before. The overshoot of Radon reconstruction around the boundary of the disc does not occur here because the test function f vanishes around the equator (or equivalently $\tilde{f}$ vanishes at the boundary of the unit disc). However, it still does not achieve the accuracy of the Fourier method. Moreover, the error of the Radon reconstruction is concentrated where f is not smooth, while this concentration is not as strong in the Fourier reconstruction. The accuracy of the OPEC algorithm is comparable to the Radon algorithm.

A. Proof of the applicability of the Fejér quadrature

This section is about the proof of Theorem 4.2. Due to the exactness of the Fejér quadrature and the equidistant quadrature on the torus and the unit interval, respectively, the hyperinterpolation $\mathcal{L}_N$ from (4.3) is also exact. Furthermore, we have $M(N) = (2N+1)^2 \in \mathcal{O}(N^2)$. So the first two conditions of Definition 4.1 are satisfied. Proving the last condition requires more work. Since $\left|B_n^k(\sigma,t)\right| = |P_n(t)|$, we observe that

$$\sum_{\ell,j=1}^{2N+1}\left|\frac{2\pi}{2N+1}\omega_j^{2N+1}B_n^k\left(\frac{2\pi\ell}{2N+1},\cos\theta_j^{2N+1}\right)\right|^2 = \frac{2\pi\left(n+\frac{1}{2}\right)}{2N+1}\sum_{j=1}^{2N+1}\left|\omega_j^{2N+1}P_n(\cos\theta_j^{2N+1})\right|^2 \quad (A.1)$$

In order to prove (4.2), we have to show that (A.1) is equal to $4\pi/(2N+1)^2$ up to some constants. We are going to do this with the help of the following two lemmas.

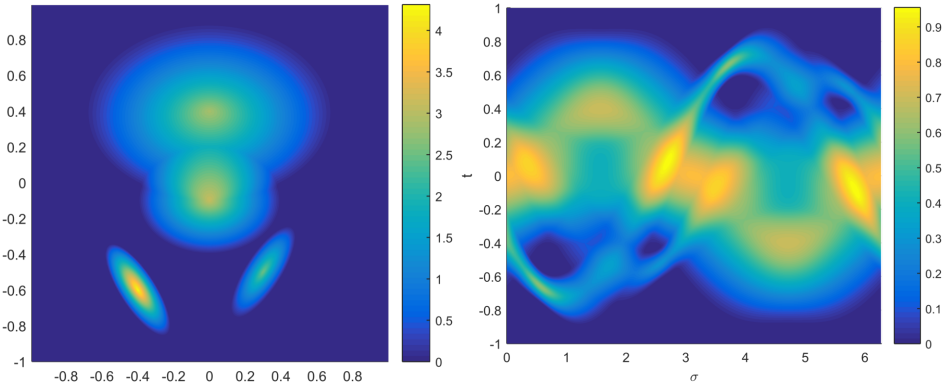
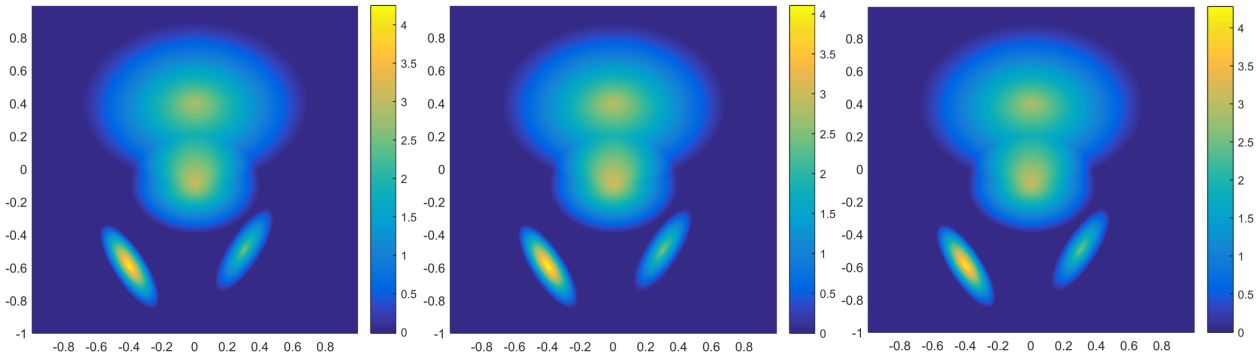
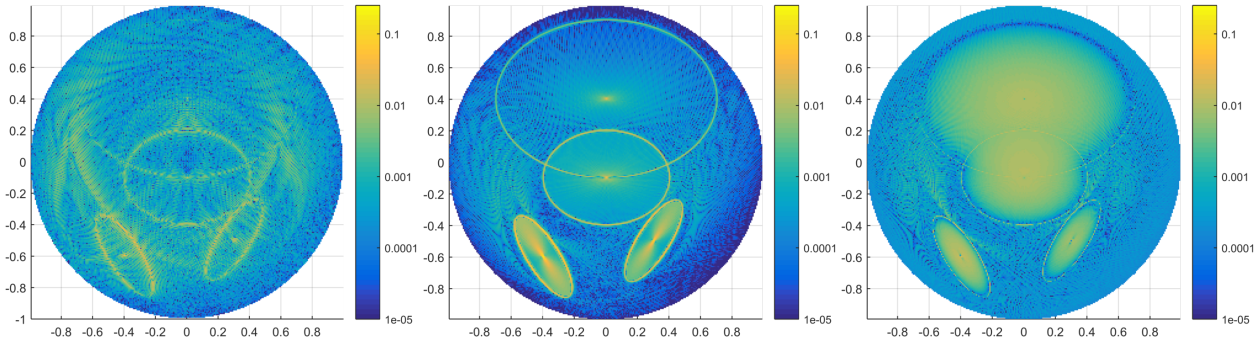


Figure 6: Non-smooth test function f_0 (left) and its vertical slice transform $\mathcal{T}f_0$ (right)



(a) Fourier reconstruction (b) Radon reconstruction (c) OPED reconstruction



(d) Error of Fourier reconstruction (MSE $3.93 \cdot 10^{-6}$) (e) Error of Radon reconstruction (MSE $2.29 \cdot 10^{-5}$) (f) Error of OPED reconstruction (MSE $1.17 \cdot 10^{-5}$)

Figure 7: Reconstruction of non-smooth test function from exact data

Lemma A.1. *For $N \in \mathbb{N}_0$ and $n \leq N$ we have*

$$\sum_{j=1}^{2N+1} \left|\omega_j^{2N+1} P_n\left(\cos\theta_j^{2N+1}\right)\right|^2 = \gamma_N^2 \frac{\pi}{2N+1} \int_{-1}^{1} \sqrt{1-t^2}\,|P_n(t)|^2 \,\mathrm{d}t,$$

where for all $N \in \mathbb{N}_0$

$$0.9028233 =: \gamma_1 < \gamma_N < \gamma_2 := 1.1789797. \tag{A.2}$$

Proof. We want to derive an approximation of the quadrature weights ω_j^{2N+1} from (4.4). To this end, we define for $N \in \mathbb{N}$ the function

$$S_N(\theta) = \sum_{r=1}^{N} \frac{\sin\left((2r-1)\,\theta\right)}{2r-1}, \quad \theta \in (-\pi, \pi)\,, \tag{A.3}$$

which is equal to the sum in (4.4) if $\theta = \theta_j^{2N+1}$. The function S_N is the N–th partial sum of the Fourier series of the Heaviside step function $\theta \mapsto \pi/4 \cdot \operatorname{sgn}(\theta)$, $\theta \in (-\pi, \pi)$, where sgn denotes the sign function, see [13, 1.442]. So $S_N(\theta)$ converges for $N \to \infty$ to the constant $\pi/4$ for all $\theta \in (0, \pi)$, but this convergence is not uniform in θ. In fact, $S_N(\theta)$ oscillates heavily around 0 and π, which is known as the Gibbs phenomenon. In [17], it was shown that $S_N(\theta_j^{2N+1}) = \gamma_{N,j} \cdot \pi/4$ with some constants $\gamma_{N,j}$ satisfying $\gamma_1 < \gamma_{N,j} < \gamma_2$ for all $j = 1, \ldots, 2N+1$, and $N \in \mathbb{N}_0$. Hence, we have

$$\omega_j^{2N+1} = \frac{\pi}{2N+2} \gamma_{N,j} \sin\theta_j^{2N+1}. \tag{A.4}$$

The Gauss–Chebyshev quadrature of the second kind (see [2]) uses the same nodes $\cos\theta_j^{2N+1}$ as the Fejér quadrature and the weights $\frac{\pi}{2N+2}(\sin\theta_j^{2N+1})^2$. Because the $(2N+1)$–point Gauss–Chebyshev quadrature of the second kind is exact of degree $4N+1$ for the integration with respect to the measure $\sqrt{1-t^2}\,\mathrm{d}t$, we have for $n \leq N$

$$\int_{-1}^{1} |P_n(t)|^2 \sqrt{1-t^2}\,\mathrm{d}t = \frac{\pi}{2N+2} \sum_{j=1}^{2N+1} \left(\sin\theta_j^{2N+1}\right)^2 \left|P_n\left(\cos\theta_j^{2N+1}\right)\right|^2,$$

where P_n is the Legendre polynomial of degree n. Considering this and (A.4) leads to

$$\begin{aligned}
\sum_{j=1}^{2N+1} \left|\omega_j^{2N+1} P_n\left(\cos\theta_j^{2N+1}\right)\right|^2 &= \left(\frac{\pi}{2N+2}\right)^2 \sum_{j=1}^{2N+1} \left(\sin\theta_j^{2N+1}\right)^2 \gamma_{N,j}^2 \left|P_n\left(\cos\theta_j^{2N+1}\right)\right|^2 \\
&= \gamma_N^2 \frac{\pi}{2N+2} \sum_{j=1}^{2N+1} \frac{\pi}{2N+2} \left(\sin\theta_j^{2N+1}\right)^2 \left|P_n\left(\cos\theta_j^{2N+1}\right)\right|^2 \\
&= \gamma_N^2 \frac{\pi}{2N+2} \int_{-1}^{1} \sqrt{1-t^2}\,|P_n(t)|^2 \,\mathrm{d}t,
\end{aligned}$$

where γ_N satisfies the inequality (A.2). ∎

Lemma A.2. *For $n \to \infty$, the Legendre polynomials P_n satisfy the equation*

$$\left(n + \frac{1}{2}\right) \int_{-1}^{1} |P_n(t)|^2 \sqrt{1-t^2}\,\mathrm{d}t = \frac{2}{\pi} + \mathcal{O}\left(n^{-1/2}\right). \tag{A.5}$$

Proof. This proof is based on the asymptotic approximation of the Legendre polynomials by Stieltjes' generalization of the Laplace–Heine formula, cf. [38, 8.21], which states that for $n \to \infty$

$$P_n(\cos\theta) = \sqrt{\frac{2}{\pi n}} \frac{\cos\left(\left(n+\frac{1}{2}\right)\theta - \frac{\pi}{4}\right)}{\sqrt{\sin\theta}} + \mathcal{O}\left((n\sin\theta)^{-3/2}\right), \quad 0 < \theta < \pi.$$

Thus, we have

$$\begin{aligned}
\int_{-1}^{1} |P_n(t)|^2 \sqrt{1-t^2}\,\mathrm{d}t &= \int_0^{\pi} \left(P_n(\cos\theta)\right)^2 (\sin\theta)^2\,\mathrm{d}\theta \\
&= \int_0^{\pi} \left(\frac{2}{\pi n} \frac{\left(\cos\left(\left(n+\frac{1}{2}\right)\theta - \frac{\pi}{4}\right)\right)^2}{\sin\theta} + \mathcal{O}\left((n\sin\theta)^{-3/2}\right) \right) (\sin\theta)^2\,\mathrm{d}\theta \\
&= \frac{2}{\pi n} \int_0^{\pi} \left(\cos\left(\left(n+\frac{1}{2}\right)\theta - \frac{\pi}{4} \right) \right)^2 \sin\theta\,\mathrm{d}\theta + \mathcal{O}\left(n^{-3/2}\right).
\end{aligned}$$

With the formula $(\cos x)^2 = (\cos 2x + 1)/2$, we observe that

$$\left(\cos\left(\left(n+\frac{1}{2}\right)\theta - \frac{\pi}{4} \right) \right)^2 = \frac{\cos\left((2n+1)\theta - \frac{\pi}{2}\right) + 1}{2} = \frac{\sin\left((2n+1)\theta\right) + 1}{2}.$$

Using this equation and the integral formula [13, 2.532.1], we obtain for $n \geq 1$

$$\begin{aligned}
\int_0^{\pi} \left(\cos\left(\left(n+\frac{1}{2}\right)\theta - \frac{\pi}{4} \right) \right)^2 \sin\theta\,\mathrm{d}\theta &= \frac{1}{2} \int_0^{\pi} \left(\sin\left((2n+1)\theta\right)\sin\theta + \sin\theta \right) \mathrm{d}\theta \\
&= \frac{1}{2} \left[\frac{\cos(2n\theta)}{4n} - \frac{\sin\left((2n+2)\theta\right)}{2(2n+2)} - \cos\theta \right]_0^{\pi} = 1,
\end{aligned}$$

from which the claimed formula (A.5) follows. ∎

References

[1] A. Abouelaz and R. Daher. Sur la transformation de Radon de la sphère S^d. *Bull. Soc. math. France*, 121(3):353 – 382, 1993.

[2] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions*. National Bureau of Standards, Washington, DC, USA, 1972.

[3] H. Berens, P. L. Butzer, and S. Pawelke. Limitierungsverfahren von Reihen mehrdimensionaler Kugelfunktionen und deren Saturationsverhalten. *Publ. Res. Inst. Math. Sci.*, 4(2):201 – 268, Mar. 1968.

[4] L. Cavalier. Nonparametric statistical inverse problems. *Inverse Problems*, 24(3):034004, June 2008.

[5] J. W. Cooley and J. W. Tukey. An algorithm for machine calculation of complex Fourier series. *Math. Comput.*, 19:297 – 301, 1965.

[6] F. Dai and Y. Xu. *Approximation theory and harmonic analysis on spheres and balls*. Springer Monographs in Mathematics. Springer, New York, 2013.

[7] W. Freeden, T. Gervens, and M. Schreiner. *Constructive Approximation on the Sphere*. Oxford University Press, Oxford, 1998.

[8] P. Funk. Über Flächen mit lauter geschlossenen geodätischen Linien. *Math. Ann.*, 74(2):278 – 300, June 1913.

[9] P. Funk. Beiträge zur Theorie der Kugelfunktionen. *Math. Ann.*, 77(1):136 – 152, 1915.

[10] R. J. Gardner. *Geometric tomography.* Number 58 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge; New York, second edition, 2006.

[11] W. Gautschi. Numerical quadrature in the presence of a singularity. *SIAM J. Numer. Anal.*, 4(3):357 – 362, Sept. 1967.

[12] S. Gindikin, J. Reeds, and L. Shepp. Spherical tomography and spherical integral geometry. In E. T. Quinto, M. Cheney, and P. Kuchment, editors, *Tomography, Impedance Imaging, and Integral Geometry*, volume 30 of *Lectures in Appl. Math*, pages 83 – 92. South Hadley, Massachusetts, 1994.

[13] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* Academic Press New York, seventh edition, 2007.

[14] E. Hecke. Über orthogonal-invariante Integralgleichungen. *Math. Ann.*, 78(1):398 – 404, 1917.

[15] S. Helgason. *The Radon Transform.* Birkhäuser, 2nd edition, 1999.

[16] K. Hesse and I. H. Sloan. Hyperinterpolation on the sphere. In N. K. Govil, H. N. Mhaskar, R. N. Mohapatra, Z. Nashed, and J. Szabados, editors, *Frontiers in Interpolation and Approximation*, Pure and Applied Mathematics. Taylor & Francis Books, Boca Raton, Florida, 2006.

[17] E. Hewitt and R. E. Hewitt. The Gibbs–Wilbraham phenomenon: An episode in Fourier analysis. *Arch. History Exact Sci.*, 21(2):129 – 160, 1979.

[18] R. Hielscher and M. Quellmalz. Optimal mollifiers for spherical deconvolution. *Preprint 2015-04, Faculty of Mathematics, Technische Universität Chemnitz*, 2015.

[19] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, New York, NY, USA, 1987.

[20] J. Keiner, S. Kunis, and D. Potts. NFFT 3.0, C subroutine library. `http://www.tu-chemnitz.de/~potts/nfft`.

[21] P. T. Kim and J. Koo. Optimal spherical deconvolution. *J. Multivariate Anal.*, 80:21 – 42, 2002.

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

[23] A. K. Louis and P. Maass. A mollifier method for linear operator equations of the first kind. *Inverse Problems*, 6(3):427–440, June 1990.

[24] A. K. Louis, M. Riplinger, M. Spiess, and E. Spodarev. Inversion algorithms for the spherical Radon and cosine transform. *Inverse Problems*, 27(3):035015, Mar. 2011.

[25] V. Michel. *Lectures on Constructive Approximation : Fourier, Spline, and Wavelet Methods on the Real Line, the Sphere, and the Ball.* Birkhäuser, New York, 2013.

[26] F. Natterer. *The Mathematics of Computerized Tomography.* John Wiley & Sons Ltd, 1986.

[27] V. P. Palamodov. *Reconstructive Integral Geometry*, volume 98 of *Monographs in Mathematics Series*. Birkhäuser, Basel, 2004.

[28] D. Potts and G. Steidl. A new linogram algorithm for computerized tomography. *IMA J. Numer. Anal.*, 21:769 – 782, 2001.

[29] D. Potts, G. Steidl, and M. Tasche. Fast algorithms for discrete polynomial transforms. *Math. Comput.*, 67:1577 – 1590, 1998.

[30] J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Ber. Verh. Sächs. Akad. Wiss. Leipzig. Math. Nat. Kl.*, 69:262 – 277, 1917.

[31] M. Riplinger and M. Spiess. Numerical inversion of the spherical Radon transform and the cosine transform using the approximate inverse with a special class of locally supported mollifiers. *J. Inverse Ill-Posed Probl.*, 22(4):497 – 536, Dec. 2013.

[32] H. Robbins. A remark on Stirling's formula. *Amer. Math. Monthly*, pages 26–29, 1955.

[33] B. Rubin. Generalized Minkowski-Funk transforms and small denominators on the sphere. *Fract. Calc. Appl. Anal.*, 3(2):177 – 203, 2000.

[34] W. Rudin. Uniqueness theory for Laplace series. *Trans. Amer. Math. Soc.*, 68(2):287 – 303, Mar. 1950.

[35] L. A. Shepp and J. B. Kruskal. Computerized tomography: The new medical X-ray technology. *Amer. Math. Monthly*, 85(6):420 – 439, June 1978.

[36] I. H. Sloan. Polynomial interpolation and hyperinterpolation over general regions. *J. Approx. Theory*, 83:238 – 254, 1995.

[37] M. Storath, A. W. abd J. Frikel, and M. Unser. Joint image reconstruction and segmentation using the potts model. *Inverse Problems*, 31:025003, 2015.

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

[39] P. Toft. *The Radon Transform - Theory and Implementation*. PhD thesis, Technical University of Denmark, 1996.

[40] A. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Verlag, Berlin, 2009.

[41] A. Wirgin. The inverse crime. *arXiv preprint math-ph/0401050*, 2004.

[42] Y. Xu. A new approach to the reconstruction of images from Radon projections. *Adv. in Appl. Math.*, 36(4):388 – 420, 2006.

[43] Y. Xu and O. Tischenko. Fast OPED algorithm for reconstruction of images from Radon data. *arXiv preprint math/0703617*, 2007.

[44] G. Zangerl and O. Scherzer. Exact reconstruction in photoacoustic tomography with circular integrating detectors II: Spherical geometry. *Math. Methods Appl. Sci.*, 33(15):1771 – 1782, Aug. 2010.