Journal of Multivariate Analysis 119 (2013) 119–143

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Kernel density estimation on the rotation group and its application to crystallographic texture analysis

Ralf Hielscher

TU Chemnitz, Fakultät für Mathematik, 09107 Chemnitz, Germany

ARTICLE INFO

Article history:
Received 6 October 2011
Available online 5 April 2013

AMS subject classifications:
62G07
2H99

Keywords:
Kernel density estimation
Minimax rates
Rotation group
Harmonic analysis
Crystallographic texture analysis

ABSTRACT

We are concerned with kernel density estimation on the rotation group SO(3). We prove asymptotically optimal convergence rates for the minimax risk of the mean integrated squared error for different function classes including bandlimited functions, functions with bounded Sobolev norm and functions with polynomially decaying Fourier coefficients and give optimal kernel functions. Furthermore, we consider kernel density estimation with nonnegative kernel functions and prove analogous saturation behavior as it is known for the Euclidean case, i.e., the optimal minimax rate does not improve for smoothness classes of functions which are more than two times differentiable. We also benchmark several families of kernel functions with respect to their capability for kernel density estimation. To make our finding applicable, we give a fast algorithm for the computation of the kernel density estimator for large sampling sets and illustrate our theoretical findings by numerical experiments. Finally, we apply our results to answer a long standing question in crystallographic texture analysis on the number of orientation measurements needed to estimate the underlying orientation density function up to a given accuracy.

1. Introduction

Kernel density estimation has been proven to be a powerful and flexible technique to estimate the underlying probability density function of a given random sample [40]. In this paper we are concerned with kernel density estimation on the rotation group SO(3). Our major motivation to consider this specific domain comes from crystallographic texture analysis, where kernel density estimation on the rotation group is used to determine the orientation density function (ODF) of a specimen from electron back scattering diffraction (EBSD) data [32,11]. The main open question in ODF estimation is: How many measurements are needed to achieve a given accuracy [3,6,39,27,10]. This question will be answered in Section 3.4.

Let λ be the Haar measure on SO(3) and let $X_1, \dots, X_N \in \mathrm{SO}(3)$ be a random sample corresponding to a square integrable, probability density function $f \in L^2(\mathrm{SO}(3))$ with respect to λ. Then we call any measurable function

$$\mathcal{E}_N : \bigotimes_{n=1}^{N} \mathrm{SO}(3) \to L^2(\mathrm{SO}(3))$$

that assigns to a random sample $X_1, \dots, X_N \in \mathrm{SO}(3)$ a function in $L^2(\mathrm{SO}(3))$ a *square integrable density estimator*. If we assume $\bigotimes_{n=1}^{N} \mathrm{SO}(3)$ to be endowed with the product measure corresponding to the density function f we write $\mathcal{E}_N^f$ and call it the *square integrable density estimator of* f. In our paper we are interested in the *mean integrated squared error* (MISE),

$$\mathrm{MISE}(\mathcal{E}_N^f) = \mathbb{E}\|f - \mathcal{E}_N^f\|_2^2 = \mathbb{E} \int_{\mathrm{SO}(3)} \left| f(x) - \mathcal{E}_N^f(x) \right|^2 \, \mathrm{d}\lambda(x), \tag{1}$$

E-mail addresses: ralf.hielscher@mathematik.tu-chemnitz.de, Ralf.Hielscher@mathematik.tu-chemnitz.de.

0047-259X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmva.2013.03.014

as a measure for the mean discrepancy between the probability density function f and the square integrable density estimator $\mathcal{E}_N^f$ of f. More specifically, we want to prove asymptotic lower and upper bounds for the so called minimax risk

$$\inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}} \mathrm{MISE}(\mathcal{E}_N^f), \tag{2}$$

where the density function f is known to be within a certain smoothness class $\mathcal{F} \subset L^2(\mathrm{SO}(3))$ and the infimum is taken over all square integrable density estimators $\mathcal{E}_N$ based on a random sample of size N.

Asymptotic upper and lower bounds of the minimax risk have been considered by many authors in the more general setting of d-dimensional compact groups. Upper bounds of rate $N^{-\frac{2s}{2s+d}}$ for the smoothness class of density functions with bounded Sobolev norm of order $s > d/2$ have been found in [14,21,13,31]. Lower bounds of the same rate have been found in [24] by following a general framework described in [35].

A specific class of density estimators are kernel density estimators. For a kernel function $\psi \in L^2(\mathrm{SO}(3))$ with $\int_{\mathrm{SO}(3)} \psi(x)\,\mathrm{d}\lambda(x) = 1$ the corresponding *kernel density estimator*, cf. [31], is defined by

$$f_\psi^*(x) = \frac{1}{N}\sum_{n=1}^{N} \psi(X_n^{-1}x) \quad x \in \mathrm{SO}(3). \tag{3}$$

Kernel density estimators with specific kernel functions ψ have been used in [14,21,13,31] to derive upper bounds for the minimax risk (2). The purpose of this paper is to find optimal kernel functions ψ and asymptotically optimal lower and upper bounds with explicit constants for the minimax risk (2) for different smoothness classes $\mathcal{F}$.

The findings of our paper are as follows. For the simple case that $\mathcal{F}$ is the class $\mathcal{F}_L$ of bandlimited functions of bandwidth L we show in Theorem 4 that the minimax risk restricted to kernel density estimators satisfies asymptotically

$$\lim_{N\to\infty} \inf_{\psi\in L^2(\mathrm{SO}(3))} \sup_{f\in\mathcal{F}_L} \mathrm{MISE}(f_\psi^*)\cdot N = \frac{(L+1)(2L+1)(2L+3)-3}{3},$$

with an asymptotically optimal kernel function being the Dirichlet kernel with bandwidth L. Since $\mathcal{F}_L$ is finite dimensional the density estimation problem is actually a parametric one.

For the smoothness class $\mathcal{F}_{s,S}^2$, cf. (A.18), of density functions with bounded Sobolev norm of order $s > 0$ we show in Theorem 5 that the minimax risk restricted to kernel density estimators satisfies asymptotically

$$\lim_{N\to\infty} \inf_{\psi\in L^2(\mathrm{SO}(3))} \sup_{f\in\mathcal{F}_{s,S}^2} \mathrm{MISE}(f_\psi^*)\cdot N^{\frac{2s}{2s+3}} = \left(\left(\frac{2s}{3}\right)^{-\frac{2s}{2s+3}} + \left(\frac{2s}{3}\right)^{\frac{3}{2s+3}}\right) C^{\frac{2s}{2s+3}} S^{\frac{6}{2s+3}}, \tag{4}$$

where $C = \left(\frac{4}{3} - \frac{8}{s+3} + \frac{4}{2s+3}\right)$. The corresponding asymptotically optimal kernel function is the Jackson type kernel (11). Surprisingly, the sharp constant in (4) differs from the Pinsker–Weyl bound reported in [22], indicating that there is still a gap in understanding the asymptotic equivalence of kernel density estimation and white noise experiments for the manifold case. For the Euclidean case results have been recently reported by Reiß [34].

Finally, we consider the smoothness class $\mathcal{F}_{s,S}^\infty$, cf. (A.19), of density functions with polynomially decaying Fourier coefficients with order $s > \frac{1}{2}$. In Theorem 6 we show that the minimax risk restricted to kernel density estimators satisfies asymptotically

$$\lim_{N\to\infty} \inf_{\psi\in L^2(\mathrm{SO}(3))} \sup_{f\in\mathcal{F}_{s,S}^\infty} \mathrm{MISE}(f_\psi^*)\cdot N^{\frac{2s-1}{2s+2}} = \frac{2^{\frac{s-2}{s+1}}\pi}{(s+1)\sin\frac{3\pi}{2s+2}} S^{\frac{2}{2s+2}}$$

and give an asymptotically optimal kernel function in (15).

Of practical importance is the case when the kernel function ψ is assumed to be nonnegative in order to guarantee a nonnegative kernel density estimator. In analogy to the Euclidean case we prove in Theorem 7 an upper bound for the MISE of two times weakly differentiable density functions $f \in \mathcal{F}_{2,S}^2$,

$$\mathrm{MISE}(f_\psi^*) \le \mu_2(\psi)^2 \|\tilde{\Delta} f\|_2^2 + N^{-1}\|\psi\|_2^2,$$

that depend on the second moment

$$\mu_2(\psi) = \frac{4}{3\pi}\int_0^1 (1-t^2)\psi(t)\sqrt{1-t^2}\,\mathrm{d}t$$

of the nonnegative kernel function ψ. In Theorem 8 we show that kernel density estimation with nonnegative kernel functions ψ saturates at the convergence rate $N^{-\frac{4}{7}}$, i.e., even for stricter smoothness classes $\mathcal{F}_{s,S}^2$, $s > 2$ the minimax rate does not improve. An example of a nonnegative kernel function ψ that attains this rate is the de la Vallée Poussin kernel (17).

Finally, we consider the minimax risk (2) without restricting the class of estimators to kernel density estimators. Following the framework presented in [35] we extend the lower bound given in [24] for the class of density functions with bounded Sobolev norm to the class of functions with polynomially decaying Fourier coefficients. More specifically, we

prove in Theorem 12 the existence of a constant C such that

$$\lim_{N\to\infty} \inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}_{s,S}^{\infty}} \mathrm{MISE}(\mathcal{E}_N^f)\cdot N^{\frac{2s-1}{2s+2}} > C.$$

It should be stressed that throughout this paper we assume that the smoothness class where the density belongs to is known before, i.e., we consider only non-adaptive methods. For adaptive methods the reader is referred to [35] and the references therein.

In order to make our finding applicable, we give in Section 3 a new fast algorithm for the numerical computation of the kernel density estimator as well as for its Fourier coefficients for large sampling sets. The algorithm is based on the nonequispaced fast Fourier transform [33] and fast summation [17] on the rotation group. Furthermore, we describe in Section 3.1 an algorithm for drawing a random sample from a distribution on the rotation group. These algorithms are applied to illustrate our theoretical findings by numerical simulations.

In Section 3.4 we generalize our results to the quotient $\mathrm{SO}(3)/S$ where $S \subset \mathrm{SO}(3)$ is a finite symmetry group. This allows us to apply our results and algorithms to the ODF estimation problem from individual orientation measurements in crystallographic texture analysis [32,11]. In particular, we answer the long standing question on the number of orientation measurements needed to estimate the underlying orientation density function up to a given accuracy [3].

All the algorithms described in this paper are freely available as part of the texture analysis toolbox MTEX [15]. In Section 3.4 the capabilities of this toolbox are demonstrated on a real world example.

Appendix A contains a tight representation of harmonic analysis on the rotation group including some results from approximation theory. Most notably, we derive in Lemmas 14 and 19 inequalities for the Fourier coefficients of the nonnegative function on the rotation group and prove in Theorem 18 for $f \in \mathcal{F}_{s,S}^2$ and $\psi \in L^2(\mathrm{SO}(3))$ a nonnegative, zonal function the estimate

$$\|f - f*\psi\|_2 \le \frac{1}{2}\big(1-\hat{\psi}(1)\big)S.$$

A similar result was already known for the spherical case [41]. Appendix B contains some proofs that have been skipped in the previous sections.

2. Lower and upper bounds for kernel density estimators

2.1. Basic properties of the MISE

Let throughout this section $f \in L^2(\mathrm{SO}(3))$ be a probability density function and $\psi \in L^2(\mathrm{SO}(3))$ a kernel function with $\int_{\mathrm{SO}(3)} \psi(x)\,\mathrm{d}\lambda(x) = 1$. We start with the following well known decomposition result of the MISE (1) of the kernel density estimator f_ψ^* into a bias term and a variance term, see e.g. [36], which actually holds true for the much more general setting of locally compact groups.

Lemma 1. *The MISE (1) of the kernel density estimator f_ψ^*, cf. (3), allows for the decomposition*

$$\mathrm{MISE}(f_\psi^*) = \|f - \mathbb{E}f_\psi^*\|_2^2 + \mathbb{E}\|f_\psi^* - \mathbb{E}f_\psi^*\|_2^2$$

into a bias term

$$\|f - \mathbb{E}f_\psi^*\|_2^2 = \|f - f*\psi\|_2^2$$

and a variance term

$$\mathbb{E}\|f_\psi^* - \mathbb{E}f_\psi^*\|_2^2 = \frac{1}{N}\Big(\|\psi\|_2^2 - \|f*\psi\|_2^2\Big).$$

In particular, we have with respect to L^2-convergence

$$\lim_{N\to\infty} f_\psi^* = f*\psi.$$

Proof. In order to keep the paper self contained we have included the proof in Appendix B. □

In the following we consider only zonal kernel functions, i.e., $\psi(x)$ depends only on the rotational angle $\omega(x)$ of $x \in \mathrm{SO}(3)$. Those functions can be seen as the group counterpart of radially symmetric functions in the Euclidean case. Some basic facts on zonal functions on the rotation group can be found in Appendix A.2. Most importantly, we will utilize that a zonal function $\psi \in L^2(\mathrm{SO}(3))$ can be expanded into a Chebyshev series

$$\psi(x) = \sum_{\ell=0}^{\infty} \hat{\psi}(\ell)(2\ell+1)\mathcal{U}_{2\ell}\left(\cos\frac{\omega(x)}{2}\right),$$

where $\mathcal{U}_{2\ell}$ denotes the Chebyshev polynomials of second kind and order 2ℓ and the convergence of the sum is meant in $L^2(\mathrm{SO}(3))$. We call the coefficients $\hat{\psi}(\ell)$ Chebyshev coefficients of ψ and because of the normalization of ψ to integral one we have $\hat{\psi}(0) = 1$.

Since the convolution of f with a zonal function ψ possesses a simple representation in the Fourier space, cf. (A.17), the decomposition of the MISE in Lemma 1 allows for a simple representation in the Fourier domain.

Lemma 2. *Let $\hat{f}(\ell,k,k')$, $\ell\in\mathbb{N}$, $k,k'=-\ell,\dots,\ell$ be the Fourier coefficients of a probability density function $f\in L^2(\mathrm{SO}(3))$ as defined in Appendix A.1, and let*

$$\hat{f}_\ell^2=\frac{1}{(2\ell+1)^2}\sum_{k,k'=-\ell}^{\ell}\left|\hat{f}(\ell,k,k')\right|^2,\quad \ell\in\mathbb{N}. \tag{5}$$

Then the MISE of the kernel density estimator f_ψ^* has the representation

$$\mathrm{MISE}(f_\psi^*)=\sum_{\ell=1}^{\infty}\left((2\ell+1)^2\hat{f}_\ell^2(1-\hat{\psi}(\ell))^2+\frac{(2\ell+1)^2}{N}\hat{\psi}(\ell)^2\left(1-\hat{f}_\ell^2\right)\right). \tag{6}$$

Proof. Representation (6) is a direct consequence of Parseval's identities (A.9), (A.16) and the convolution formula (A.17). □

Note that in the Fourier representation (6) of the MISE the sum starts from $\ell=1$. This is due to the fact that we assumed $\hat{\psi}(0)=1$ and $\hat{f}(0,0,0)=\hat{f}_0=1$.

2.2. Optimal kernel functions and rates of convergence

Our goal is to find optimal kernel functions that minimize the MISE (1) for certain smoothness classes $\mathcal{F}\subset L^2(\mathrm{SO}(3))$ and to investigate the asymptotic decay rates of the MISE. Let us first consider the trivial case that the class of functions $\mathcal{F}=\{f\}, f\in L^2(\mathrm{SO}(3))$ consists only of the true density function itself.

Theorem 3. *The MISE optimal kernel function for the class $\mathcal{F}=\{f\}$ is*

$$\psi_{f,N}(x)=\sum_{\ell=0}^{\infty}(2\ell+1)\frac{N\hat{f}_\ell^2}{(N-1)\hat{f}_\ell^2+1}\mathcal{U}_{2\ell}\left(\cos\frac{\omega(x)}{2}\right), \tag{7}$$

where the convergence of the sum is meant in $L^2(\mathrm{SO}(3))$. The corresponding MISE is

$$\mathrm{MISE}(f_{\psi_{f,N}}^*)=\sum_{\ell=1}^{\infty}(2\ell+1)^2\frac{\hat{f}_\ell^2(1-\hat{f}_\ell^2)}{(N-1)\hat{f}_\ell^2+1}. \tag{8}$$

In particular, we have $f_{\psi_{f,N}}^*\to f$ in $L^2(\mathrm{SO}(3))$ as $N\to\infty$ and the best possible rate of convergence is N^{-1} unless f is constant.

Proof. See Appendix B. □

A second simple case is that the class of functions $\mathcal{F}$ is the space of bandlimited probability density functions

$$\mathcal{F}_L=\left\{f\in L^2(\mathrm{SO}(3))\mid f\ge 0,\ \int f(x)\,\mathrm{d}\lambda(x)=1,\ \text{and}\ \hat{f}(\ell,k,k')=0\ \text{for all}\ \ell>L\right\}.$$

In this case the density estimation problem becomes a parametric problem and we have the following result on the asymptotic minimax risk.

Theorem 4. *An asymptotically optimal kernel function for the class $\mathcal{F}_L$ of bandlimited density functions of order $L>0$ is the Dirichlet kernel ψ_L^{D}, $L\in\mathbb{N}$,*

$$\psi_L^{\mathrm{D}}(x)=\sum_{\ell=0}^{L}(2\ell+1)\mathcal{U}_{2\ell}\left(\cos\frac{\omega(x)}{2}\right). \tag{9}$$

The corresponding asymptotic minimax risk is

$$\lim_{N\to\infty}\inf_{\psi\in L^2(\mathrm{SO}(3))}\sup_{f\in\mathcal{F}_L}\mathrm{MISE}(f_\psi^*)\cdot N=\|\psi_L^{\mathrm{D}}\|_2^2=\frac{(L+1)(2L+1)(2L+3)-3}{3}.$$

Proof. Let $f\in\mathcal{F}_L$. Plugging in the Dirichlet kernel ψ_L^{D} into (6) and making use of Parseval's identity (A.16) and the assumption $\hat{f}(0,0,0)=1$ we arrive at the estimate

$$\mathrm{MISE}(f_{\psi_L^{\mathrm{D}}}^*)=\frac{1}{N}(\|\psi_L^{\mathrm{D}}\|_2^2-\|f\|_2^2)\le\frac{1}{N}\left(\left(\sum_{\ell=0}^{L}(2\ell+1)^2\right)-1\right)=\frac{(L+1)(2L+1)(2L+3)-3}{3N}.$$

In order to show that this upper bound is asymptotically sharp we consider for a fixed sample size N a density function $f \in \mathcal{F}_L$ that maximizes the optimal MISE (8). Componentwise differentiation of (8) leads to

$$\hat{f}_\ell^2 = \frac{\sqrt{N}-1}{N-1}, \quad \ell = 1, \dots, L.$$

Since $\hat{f}_\ell^2$ decays to zero as N tends to infinity f can be guarantied to be nonnegative for N sufficiently large. By (8) we obtain

$$\mathrm{MISE}(f^*_{\psi_{f,N}}) = \sum_{\ell=1}^{L} (2\ell+1)^2 \frac{\frac{\sqrt{N}-1}{N-1}\left(1-\frac{\sqrt{N}-1}{N-1}\right)}{(N-1)\frac{\sqrt{N}-1}{N-1}+1} = \sum_{\ell=1}^{L} (2\ell+1)^2 \frac{\left(N-\sqrt{N}\right)^2}{N(N-1)^2},$$

and, hence,

$$\begin{aligned}
\lim_{N\to\infty} \inf_{\psi\in L^2(\mathrm{SO}(3))} \sup_{\tilde{f}\in\mathcal{F}_L} \mathrm{MISE}(\tilde{f}^*_\psi)\cdot N &\geq \lim_{N\to\infty} \inf_{\psi\in L^2(\mathrm{SO}(3))} \mathrm{MISE}(f^*_\psi)\cdot N \\
&= \lim_{N\to\infty} \mathrm{MISE}(f^*_{\psi_{f,N}})\cdot N \\
&= \lim_{N\to\infty} N\cdot \sum_{\ell=1}^{L} (2\ell+1)^2 \frac{\left(N-\sqrt{N}\right)^2}{N(N-1)^2} \\
&= \frac{(L+1)(2L+1)(2L+3)-3}{3}. \quad \square
\end{aligned}$$

Next we consider for $s, S > 0$ the smoothness classes $\mathcal{F}^2_{s,S}$ of density functions f with Sobolev norm $\|f\|_{2,s} < S$; cf. (A.18). Then because of Lemma 1, Theorem 18 and the assumption $\hat{\psi}(0) = 1$ the MISE can be bounded from above by

$$\mathrm{MISE}(f^*_\psi) \leq \sup_{\ell\in\mathbb{N}\setminus\{0\}} \frac{\left|1-\hat{\psi}(\ell)\right|^2}{\ell^s(\ell+1)^s} S^2 + N^{-1}\|\psi\|_2^2. \tag{10}$$

It turns out that by selecting a kernel function ψ that minimizes this upper bounds we obtain asymptotically optimal convergence rates.

Theorem 5. *Let $S > 0$ be sufficiently small. Then an asymptotically optimal kernel function with respect to the class $\mathcal{F}^2_{s,S}$ of functions with Sobolev norm of order $s > 0$ bounded by S is the Jackson type kernel $\psi^{\mathrm{J}}_{L,s}\colon \mathrm{SO}(3) \to \mathbb{R}$,*

$$\psi^{\mathrm{J}}_{L,s}(x) = 1 + \sum_{\ell=1}^{\lfloor L\rfloor} (2\ell+1)\left(1 - \frac{\ell^{s/2}(\ell+1)^{s/2}}{L^{s/2}(L+1)^{s/2}}\right) \mathcal{U}_{2\ell}\left(\cos\frac{\omega(x)}{2}\right), \tag{11}$$

with bandwidth

$$L = \left(\frac{2s}{3}\frac{NS^2}{C}\right)^{\frac{1}{2s+3}} \tag{12}$$

where $C = \left(\frac{4}{3} - \frac{8}{s+3} + \frac{4}{2s+3}\right)$. This kernel function asymptotically realizes the minimax risk

$$\lim_{N\to\infty} \inf_{\psi\in L^2(\mathrm{SO}(3))} \sup_{f\in\mathcal{F}^2_{s,S}} \mathrm{MISE}(f^*_\psi)\cdot N^{\frac{2s}{2s+3}} = \left(\left(\frac{2s}{3}\right)^{-\frac{2s}{2s+3}} + \left(\frac{2s}{3}\right)^{\frac{3}{2s+3}}\right) C^{\frac{2s}{2s+3}} S^{\frac{6}{2s+3}}.$$

Proof. See Appendix B. $\square$

For the specific case $s = 2$ the Jackson type kernel has the explicit representation

$$\psi^{\mathrm{J}}_{L,2}(\cos\omega) = \frac{(2L+3)\big((L-2)\cos\omega\sin 2L\omega - 3L\cos 2L\omega\sin\omega\big) + L(2L-1)\sin(2L+3)\omega}{8L(L+1)\sin^5\frac{\omega}{2}}.$$

The optimal rate of convergence, but with a worse constant, is also attained by the Dirichlet kernel $\psi^{\mathrm{D}}_{L_{\mathrm{opt}}}$, (9); see also [14]. For $s = 2$ the corresponding optimal parameter L_{opt} is given by

$$L^7_{\mathrm{opt}} \approx NS^2, \tag{13}$$

which leads to the asymptotic upper bound

$$\mathrm{MISE}(f^*_{\psi^{\mathrm{D}}_{L_{\mathrm{opt}}}}) \le L_{\mathrm{opt}}^{-4}S^2 + \frac{4}{3}(L_{\mathrm{opt}}^3 + \mathcal{O}(L_{\mathrm{opt}}^2))N^{-1} = \left(1+\frac{4}{3}\right)S^{6/7}N^{-4/7} + \mathcal{O}(N^{-5/7}) \approx \frac{7}{3}S^{6/7}N^{-4/7}.$$

Finally, we consider for $s > \frac{1}{2}$, $S > 0$ the smoothness class $\mathcal{F}^\infty_{s,S}$ of density functions f with polynomially decaying Fourier coefficients of order s; cf. (A.19). Then we have by Lemma 1, Theorem 18 and (A.16) the upper bound

$$\mathrm{MISE}(f^*_\psi) \le \sum_{\ell=1}^{\infty}\left(\frac{\left|1-\hat\psi(\ell)\right|^2}{\ell^s(\ell+1)^s}S^2 + N^{-1}(2\ell+1)^2\hat\psi(\ell)^2\right). \tag{14}$$

As in the just previous case of the Sobolev class we obtain asymptotically optimal convergence rates by selecting a kernel function ψ that minimizes this upper bound.

Theorem 6. *Let $S > 0$ be sufficiently small. Then an asymptotically optimal kernel function $\psi_{\mathcal{F}^\infty_{s,S},N}$ for the class $\mathcal{F}^\infty_{s,S}$ of density functions with polynomially decaying Fourier coefficients of order $s > \frac{1}{2}$ is the zonal function*

$$\psi_{\mathcal{F}^\infty_{s,S},N}(x) = \sum_{\ell=0}^{\infty}(2\ell+1)\frac{NS^2}{NS^2+(2\ell+1)^2\ell^s(\ell+1)^s}\mathcal{U}_{2\ell}\left(\cos\frac{\omega(x)}{2}\right). \tag{15}$$

This function asymptotically realizes the minimax risk

$$\lim_{N\to\infty}\inf_{\psi\in L^2(\mathrm{SO}(3))}\sup_{f\in\mathcal{F}^\infty_{s,S}}\mathrm{MISE}(f^*_\psi)\cdot N^{\frac{2s-1}{2s+2}} = \frac{2^{\frac{s-2}{s+1}}\pi}{(s+1)\sin\frac{3\pi}{2s+2}}S^{\frac{2}{2s+2}}.$$

Proof. See Appendix B. □

The results of Theorems 5 and 6 compare well with the inclusions $\mathcal{F}^\infty_{s+\varepsilon,S} \subset \mathcal{F}\ell^2_{s-\frac{1}{2},S'}$, with $\varepsilon, S, S' > 0$ as in (A.20), in the sense that for $\varepsilon \to 0$ we obtain for both classes the same optimal rates of convergence, namely, $N^{-\frac{2s-1}{2s+2}}$.

2.3. Nonnegative kernel functions

In many practical applications one is often interested in nonnegative kernel density estimators, i.e., in kernel density estimation with nonnegative kernel functions. In the Euclidean setting this restriction allows for a simple, asymptotically sharp, upper bound for the MISE the so called *asymptotic mean integrated squared error* (AMISE), cf. [40],

$$\mathrm{AMISE}(f^*_\psi) = \mu_2(\psi)^2\|\triangle f\|_2^2 + N^{-1}\|\psi\|_2^2 \tag{16}$$

where $f\colon \mathbb{R}^d \to \mathbb{R}$ is assumed to be a twice continuous differentiable function and

$$\mu_2(\psi) = \int_{\mathbb{R}^d}\|x\|_2^2\psi(x)\,\mathrm{d}x$$

denotes the second moment of the kernel function ψ. Using Theorem 18 we show an analogous upper bound for (10) in the setting of kernel density estimation on the rotation group with nonnegative zonal kernel functions. For formulating this result it is handy to make use of the notation $\psi(t) = \psi(x)$, $t \in [0,1]$, where $t = \cos\frac{\omega(x)}{2}$; cf. Appendix A.2.

Theorem 7. *Let $S > 0$, let $f \in \mathcal{F}^2_{2,S}(\mathrm{SO}(3))$ be a two times weakly differentiable density function and let $\psi \in L^2(\mathrm{SO}(3))$ be a nonnegative, zonal kernel function with $\hat\psi(0) = 1$. Then we have*

$$\mathrm{MISE}(f^*_\psi) \le \frac{1}{4}|1-\hat\psi(1)|^2\|f\|^2_{2,2} + N^{-1}\|\psi\|_2^2 = \mu_2(\psi)^2\|\tilde\triangle f\|_2^2 + N^{-1}\|\psi\|_2^2,$$

with

$$\mu_2(\psi) = \frac{4}{3\pi}\int_0^1(1-t^2)\psi(t)\sqrt{1-t^2}\,\mathrm{d}t$$

and with $\tilde\triangle$ denoting the Laplace–Beltrami operator on SO(3).

Proof. Using the decomposition of the MISE into a bias and a variance term, cf. Lemma 1, the left hand side inequality becomes a direct consequence of Theorem 18. For the right hand side equality we observe that for any nonnegative zonal function ψ with $\hat{\psi}(0) = 1$ we have by (A.14) and $\mathcal{U}_2(t) = 4t^2 - 1$,

$$\frac{1}{4}(1 - \hat{\psi}(1)) = \frac{1}{\pi}\int_0^1 \psi(t)\left(1 - \frac{1}{3}\mathcal{U}_2(t)\right)\sqrt{1-t^2}\,\mathrm{d}t = \frac{4}{3\pi}\int_0^1 \psi(t)(1-t^2)\sqrt{1-t^2}\,\mathrm{d}t$$

and, furthermore, that

$$\|f\|_{2,2}^2 = \|\tilde{\Delta} f\|_2^2. \quad \square$$

A popular nonnegative kernel function on SO(3) is the de la Vallée Poussin kernel $\psi_\kappa^{\mathrm{VP}}$, cf. [17], which is defined for any $\kappa \in \mathbb{N} \setminus \{0\}$ by its finite Chebyshev expansion

$$\psi_\kappa^{\mathrm{VP}}(x) = \frac{(\kappa+1)2^{2\kappa-1}}{\binom{2\kappa-1}{\kappa}}\left(\cos\frac{\omega(x)}{2}\right)^{2\kappa} = \binom{2\kappa+1}{\kappa}^{-1}\sum_{\ell=0}^{\kappa}(2\ell+1)\binom{2\kappa+1}{\kappa-\ell}\mathcal{U}_{2\ell}\left(\cos\frac{\omega(x)}{2}\right). \tag{17}$$

Choosing the parameter κ in an optimal way the de la Vallée Poussin kernel yields the optimal convergence rate $N^{-\frac{4}{7}}$ for the class $\mathcal{F}_{2,S}^2$.

Theorem 8. *Let $S > 0$ and $\mathcal{F}_{2,S}^2$ the class of density functions with bounded Sobolev norm of order 2. Then the optimal parameter κ_{opt} of the de la Vallée Poussin kernel $\psi_\kappa^{\mathrm{VP}}$ satisfies*

$$\kappa_{\mathrm{opt}}^7 \approx \frac{2^7}{9\pi}\|f\|_{2,2}^4 N^2 \tag{18}$$

which yields the MISE estimate

$$\sup_{f\in\mathcal{F}_{2,S}^2} \mathrm{MISE}(f^*_{\psi^{\mathrm{VP}}_{\kappa_{\mathrm{opt}}}}) \le 3.8 S^2 N^{-4/7}. \tag{19}$$

Proof. The L^2-norm of the de la Vallée Poussin kernel computes to

$$\begin{aligned}\|\psi_\kappa^{\mathrm{VP}}\|_2^2 &= \frac{4}{\pi}\int_0^1 \psi_\kappa^{\mathrm{VP}}(t)^2\sqrt{1-t^2}\,\mathrm{d}t = \frac{4}{\pi}\frac{(\kappa+1)2^{2\kappa-1}}{\binom{2\kappa-1}{\kappa}}\int_0^1 t^{4\kappa}\sqrt{1-t^2}\,\mathrm{d}t \\ &= \sqrt{\pi}\frac{\Gamma(\kappa+2)^2\Gamma\left(2\kappa+\frac{1}{2}\right)}{\Gamma\left(\kappa+\frac{1}{2}\right)^2\Gamma(2\kappa+2)}.\end{aligned}$$

Since we have asymptotically by the Stirling formula

$$\frac{\Gamma(\kappa+2)}{\Gamma\left(\kappa+\frac{1}{2}\right)} = \frac{\Gamma(\kappa+2)\Gamma(\kappa)2^{2\kappa-1}}{\Gamma(2\kappa)\sqrt{\pi}} = 2^{2\kappa}\frac{\kappa+1}{\sqrt{\pi}}\frac{\Gamma(\kappa+1)^2}{\Gamma(2\kappa+1)} \approx (\kappa+1)\sqrt{\kappa}$$

we obtain for the L^2-norm of the de la Vallée Poussin kernel the approximation

$$\|\psi_\kappa^{\mathrm{VP}}\|_2^2 = \sqrt{\pi}\frac{\Gamma(\kappa+2)^2\Gamma\left(2\kappa+\frac{1}{2}\right)}{\Gamma\left(\kappa+\frac{1}{2}\right)^2\Gamma(2\kappa+2)} \approx \sqrt{\frac{\pi}{2}}\frac{\sqrt{\kappa}(\kappa+1)^2}{2\kappa+1} \approx \sqrt{\frac{\pi}{8}}\kappa^{3/2}.$$

Since

$$1 - \hat{\psi}_\kappa^{\mathrm{VP}}(1) = 1 - \frac{\binom{2\kappa+1}{\kappa-1}}{\binom{2\kappa+1}{\kappa}} = 1 - \frac{\kappa!(\kappa+1)!}{(\kappa-1)!(\kappa+2)!} = 1 - \frac{\kappa}{\kappa+2} = \frac{2}{\kappa+2}$$

we obtain the upper bound

$$\mathrm{MISE}(f^*_{\psi_\kappa^{\mathrm{VP}}}) \le (\kappa+2)^{-2}\|f\|_{2,2}^2 + \sqrt{\frac{\pi}{8}}\kappa^{3/2}N^{-1}.$$

Minimization with respect to κ gives the stated results. $\square$

Similar to the Euclidean case kernel density estimation with nonnegative, zonal kernel functions ψ does not allow for better convergence rates than $N^{-\frac{4}{7}}$ independent of the smoothness of the function f. We have the following lower bound.

Theorem 9. *Let $\mathcal{F} \subset L^2(\mathrm{SO}(3))$ be any function class that contains the function $f(x) = 1 + D_1^{0,0}(x)$. Then we have the lower bound*

$$\inf_{\psi \geq 0} \sup_{f \in \mathcal{F}} \mathrm{MISE}(f_\psi^*) \geq N^{-4/7} - \frac{4}{3} N^{-1}.$$

Proof. From (A.5) we know that $f = 1 + D_1^{0,0}$ is indeed a density function. Let ψ be a nonnegative kernel function. Then we have by Lemma 19 a lower bound on $\|\psi\|_2$ and by Lemma 14 the upper bound $\hat{\psi}(1)^2 \leq 1$. Together with Lemma 1 and (A.17) this gives the following lower bound for the MISE of the kernel density estimator f_ψ^*

$$\begin{aligned} \mathrm{MISE}(f_\psi^*) &= \frac{(1-\hat{\psi}(1))^2}{3} + \frac{1}{N}\left(\|\psi\|_2^2 - 1 - \frac{\hat{\psi}(1)^2}{3}\right) \\ &\geq \frac{(1-\hat{\psi}(1))^2}{3} + \frac{64\sqrt{2}}{105N}(1-\hat{\psi}(1))^{-\frac{3}{2}} - \frac{4}{3}\frac{1}{N}. \end{aligned}$$

Minimizing the above expression with respect to $\hat{\psi}(1)$ results in

$$\mathrm{MISE}(f_\psi^*) \geq \frac{4}{3}\left(\frac{2}{5}\right)^{4/7}\left(\frac{7}{3}\right)^{3/7} N^{-4/7} - \frac{4}{3}N^{-1} > N^{-4/7} - \frac{4}{3}N^{-1}. \quad \square$$

For completeness we consider also another frequently applied nonnegative zonal kernel function on the rotation group, the Abel–Poisson kernel $\psi_\kappa^{\mathrm{AP}}$, cf. [17], which is defined for $\kappa \in (0, 1)$ by its Chebyshev series

$$\psi_\kappa^{\mathrm{AP}}(x) = \sum_{\ell=0}^{\infty} (2\ell+1)\kappa^{2\ell} \mathcal{U}_{2\ell}\left(\cos\frac{\omega(x)}{2}\right). \tag{20}$$

Computing its L^2-norm to

$$\|\psi_\kappa^{\mathrm{AP}}\|_2^2 = \sum_{\ell=0}^{\infty} (2\ell+1)^2\kappa^{4l} = \frac{1+6\kappa^4+\kappa^8}{(1-\kappa^4)^3},$$

Theorem 7 gives us for any $f \in \mathcal{F}_{2,S}^2$ the upper bound

$$\mathrm{MISE}(f_{\psi_\kappa^{\mathrm{AP}}}^*) \leq \frac{(1-\kappa^2)^2}{4}S^2 + \frac{1+6\kappa^4+\kappa^8}{(1-\kappa^4)^3}N^{-1} \approx (1-\kappa)^2 S^2 + \frac{1}{8}(1-\kappa)^{-3}N^{-1}.$$

Minimization with respect to κ yields

$$(1-\kappa_{\mathrm{opt}})^5 \approx \frac{3}{16}N^{-1}S^{-2} \tag{21}$$

and

$$\mathrm{MISE}(f_{\psi_{\kappa_{\mathrm{opt}}}^{\mathrm{AP}}}^*) \leq \left(\left(\frac{3}{16}\right)^{2/5} + \frac{1}{8}\left(\frac{3}{16}\right)^{-3/5}\right) S^{6/5}N^{-2/5} \approx 0.86 S^{6/5}N^{-2/5}.$$

Thus, the Abel–Poisson kernel seems not be well suited for kernel density estimation on the rotation group.

2.4. General lower bounds

In this section we consider the minmax risk (2) for the general setting, i.e., where the infimum is taken over all square integrable estimators. Since explicit constants are not known for this general case we just show that the rates of convergence are the same as for kernel density estimation. For the smoothness class $\mathcal{F}_{s,S}^2$ of density functions with bounded Sobolev norm this has been proven in [24]. We prove a similar result for the smoothness class $\mathcal{F}_{s,S}^\infty$ of density functions with polynomially decaying Fourier coefficients following the same lines. For completeness, we present here both proofs simultaneously. A general recipe for the derivation of lower bounds can be found in [35].

The idea is to restrict the density estimation problem to a finite subset $\mathcal{F}^* \subset \mathcal{F}^\infty_{s,S} \subset L^2(\mathrm{SO}(3))$, $s, S > 0$ by observing that for any sequence $\rho_N > 0$ and any constant $A > 0$

$$\rho_N^{-2} \inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}^\infty_{s,S}} \mathrm{MISE}(\mathcal{E}_N^f) \geq \rho_N^{-2} \inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}^*} \mathbb{E}\|f - \mathcal{E}_N^f\|_2^2 \geq A^2 \inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}^*} P_f(\|f - \mathcal{E}_N^f\|_2^2 \geq \rho_N^2 A^2). \tag{22}$$

In order to bound the right hand side from below the density functions in $\mathcal{F}^*$ have to be sufficiently separated in the L^2-norm and their Kullback divergences

$$K(f,g) = \int_{\mathrm{SO}(3)} f(x) \log \frac{f(x)}{g(x)}\, \mathrm{d}\lambda(x), \quad f, g \in \mathcal{F}^*$$

have to be sufficiently small. More precisely, we are going to apply Theorem 2.5 in [35].

Theorem 10. *Let $N > 0$ be the number of random samples and let $\mathcal{F}^*$ be a finite set of density functions on* SO(3) *such that there are constants A, ρ_N, $\alpha \in (0, \frac{1}{8})$ satisfying*

1. *the functions in $\mathcal{F}^*$ are $A\rho_N$ separated, i.e., for all $f, g \in \mathcal{F}^*$, $f \neq g$ we have*
$$\|f - g\|_2 \geq A\rho_N,$$
2. *the Kullback divergences between the functions in $\mathcal{F}^*$ are sufficiently small, i.e., there is a function $f_0 \in \mathcal{F}^*$, $f_0 > 0$ satisfying*
$$\frac{N}{|\mathcal{F}^*|} \sum_{f\in\mathcal{F}^*\setminus f_0} K(f, f_0) \leq \alpha \log \left|\mathcal{F}^*\right|.$$

Then

$$\inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}^*} P_f(\|f - \mathcal{E}_N^f\|_2 \geq \rho_N A) \geq \frac{\sqrt{|\mathcal{F}^*|}}{1 + \sqrt{|\mathcal{F}^*|}} \left(1 - 2\alpha - \sqrt{\frac{2\alpha}{\log|\mathcal{F}^*|}}\right) > 0.$$

As candidates for the finite set $\mathcal{F}^*$ we consider for $t \in \mathbb{R}$, $L \in \mathbb{N}$ and a constant $C > 0$ the sets

$$\mathcal{F}_{t,L} = \left\{ f(x) = 1 + CL^{-t-\frac{3}{2}} \sum_{\ell=L+1}^{2L} \sum_{k,k'=-\ell}^{\ell} \hat{f}(\ell,k,k')\sqrt{2\ell+1}\bar{D}_\ell^{k,k'} \mid \hat{f}(\ell,k,k') \in \{0,1\} \right\},$$

where the functions $\sqrt{2\ell+1}\bar{D}_l^{k,k'}$ are chosen as a real valued orthonormal basis in $\mathrm{Harm}_\ell(\mathrm{SO}(3))$. Since for every $L \in \mathbb{N}$, $s > 0$ and $f \in \mathcal{F}_{s,L}$,

$$\|f\|_{2,s}^2 = 1 + C^2 L^{-2s-3} \sum_{\ell=L+1}^{2L} \sum_{k,k'=-\ell}^{\ell} \ell^s(\ell+1)^s \left|\hat{f}(\ell,k,k')\right|^2 \leq \tilde{C}C^2$$

and

$$\|f\|_{\infty,s+\frac{1}{2}}^2 = \max\left\{1, C^2L^{-2s-3} \max_{\ell=L+1,\dots,2L} \sum_{k,k'=-\ell}^{\ell} \ell^{s+\frac{1}{2}}(\ell+1)^{s+\frac{1}{2}} \left|\hat{f}(\ell,k,k')\right|^2\right\} \leq \tilde{C}C^2$$

with a constant $\tilde{C}$ independent of L, the constant C can be chosen such that $\mathcal{F}_{s,L} \subset \mathcal{F}^2_{s,S}$ and $\mathcal{F}_{s,L} \subset \mathcal{F}^\infty_{s+\frac{1}{2},S}$. On the other hand we conclude from $\left|D_\ell^{k,k'}(x)\right| \leq 1$, cf. (A.5), that the constant C can be chosen such that all functions in $\mathcal{F}_{s,L}$ are nonnegative. It remains to compute the separation distance and the Kullback divergences of the functions in $\mathcal{F}_{t,L}$. For $f, g \in \mathcal{F}_{t,L}$ with $f \neq g$ we have

$$\|g - f\|_2 \geq CL^{-t-\frac{3}{2}}$$

and for $f_0 = 1 \in \mathcal{F}_{t,L}$ we have by Jensen's inequality

$$K(f, f_0) = \int_{\mathrm{SO}(3)} f(x) \log f(x)\, \mathrm{d}\lambda(x) \leq \int_{\mathrm{SO}(3)} (f(x) - 1)^2\, \mathrm{d}\lambda(x) \leq \tilde{C}L^{-2t} \tag{23}$$

with some constant $\tilde{C}$ independent of L. If we would apply Theorem 10 directly to the sets $\mathcal{F}_{t,L}$ the gained lower bound would be not optimal. However, by applying the lemma of Koo [23], we find a subset $\hat{\mathcal{F}}_{t,L}$ of $\mathcal{F}_{t,L}$ that has a larger separation distance, but almost the same cardinality.

Lemma 11. *Let $N \in \mathbb{N}, N > 8$. Then the largest subset $\Omega_N \subset \{0,1\}^N$ with separation distance $\frac{\sqrt{N}}{8}$, i.e., for all $x, y \in \Omega_N$ we have $\|x-y\|_2 \geq \frac{\sqrt{N}}{8}$, has the cardinality $\log|\Omega_N| - 1 \geq 0.27N$.*

Applying Koo's lemma to $\mathcal{F}_{t,L}$ we end up with a subset $\hat{\mathcal{F}}_{t,L} \subset \mathcal{F}_{t,L}$ with separation distance

$$\|g-f\|_2 \geq CL^{-t}, \quad f, g \in \hat{F}_{t,L},\ f \neq g,$$

but almost the same cardinality $\log\left|\hat{\mathcal{F}}_{t,L}\right| \geq 0.27L^3$. Now we are ready to prove the general lower bound.

Theorem 12. *For the smoothness class of density functions with Sobolev norm of order $s > 0$ bounded by $S > 0$ the minimax risk is bounded from below at rate $N^{-\frac{2s}{2s+3}}$, whereas for the smoothness class of density functions with polynomially decaying Fourier coefficients of order $s > \frac{1}{2}$ the minimax risk is bounded from below at rate $N^{-\frac{2s-1}{2s+2}}$, i.e., there are constants $C_1, C_2 > 0$ such that*

$$\lim_{N\to\infty} \inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}^2_{s,S}} \mathrm{MISE}(\mathcal{E}^f_N) \cdot N^{\frac{2s}{2s+3}} > C_1,$$

and

$$\lim_{N\to\infty} \inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}^\infty_{s,S}} \mathrm{MISE}(\mathcal{E}^f_N) \cdot N^{\frac{2s-1}{2s+2}} > C_2.$$

Proof. Let $N > 8, t > 0, C > 0$ and let $\hat{F}_{t,L}$ be the set defined just above the theorem with separation distance CL^{-t}. Setting $L = N^{\frac{1}{2t+3}}$ this separation distance becomes

$$\|f-g\|_2^2 \geq C^2L^{-2t} = C^2N^{-\frac{2t}{2t+3}}, \quad f, g \in \hat{F}_{t,L},\ f \neq g,$$

i.e. $\hat{F}_{t,L}$ satisfies the first condition of Theorem 10 with $\rho_N = N^{-\frac{2t}{2t+3}}$ and $A = C^2$. On the other hand we obtain from (23) for the Kullback divergence with $f_0 = 1$,

$$\frac{N}{\left|\hat{\mathcal{F}}_{t,L}\right|} \sum_{f\in\hat{\mathcal{F}}_{t,L}\setminus f_0} K(f, f_0) \leq N8C^2L^{-2s} = 8C^2L^3 \leq \frac{8C^2}{0.27} \log\left|\hat{\mathcal{F}}_{t,L}\right|.$$

By adjusting C such that $\frac{8C^2}{0.27} \leq \frac{1}{8}$ the second condition of Theorem 10 is satisfied. Consequently,

$$\inf_{\mathcal{E}_N} \sup_{f\in\mathcal{F}^*} P_f\left(\|f - \mathcal{E}^f_N\|_2 \geq N^{-\frac{2t}{2t+3}}C^2\right) \geq \tilde{C} > 0.$$

Thus, setting $t = s$ for the class $\mathcal{F}^2_{s,S}$ and $t = s - \frac{1}{2}$ for the class $\mathcal{F}^\infty_{s,S}$, the lower bounds for the minimax risk claimed in Theorem 10 follow from (22). □

3. Numerical experiments

In this section we are going to illustrate our theoretical findings by numerical experiments. The general concept is as follows:

1. Choose a test density function $f \in L^2(\mathrm{SO}(3))$.
2. Fix a kernel function $\psi \in L^2(\mathrm{SO}(3))$.
3. Draw a random sample from the distribution given by f of size $N \in \mathbb{N}$.
4. Compute the kernel density estimator f^*_ψ.
5. Compute the integrated squared error $\|f - f^*_\psi\|_2^2$.
6. Compute an estimate of the MISE by repeating M times the steps 3–5 and taking the mean value of the integrated squared errors.

As the test density function f we chose a linear combination of de la Vallée Poussin kernels $\psi^{\mathrm{VP}}_\kappa$, cf. (17), translated to two arbitrarily chosen locations of the rotation group,

$$f(x) = 0.2 + 0.7\psi^{\mathrm{VP}}_{90}(R_{e_2,\frac{\pi}{6}}x) + 0.1\psi^{\mathrm{VP}}_{350}(R_{e_1,\frac{4\pi}{9}}x), \tag{24}$$

where $R_{\eta,\omega}$ denotes the rotation about axis $\eta \in \mathbb{S}^2$ with angle ω. We have $f \in \mathcal{F}^2_{2,S}$ and $f \in \mathcal{F}^\infty_{2.5,S'}$ with $S = 4600$ and $S' = 5200$. Our numerical experiments showed that increasing the number or changing the type of the zonal functions $\psi^{\mathrm{VP}}_\kappa$ in (24) has only minor influence on our numerical results.

3.1. Drawing a random sample from a distribution on SO(3)

In this section we briefly describe how to draw a random sample of a radially symmetric density function on SO(3). How to draw random samples for other classes of distributions, e.g. the Bingham distribution, can be found in [5]. Let $f > 0$ be a strictly positive density function on $[0, 1]$. Then the corresponding cumulative distribution function

$$F(y) = \int_0^y f(x)\,\mathrm{d}x$$

defines a diffeomorphism

$$F\colon [0, 1] \to [0, 1]$$

and by the transformation rule we have for any integrable function $h\colon [0, 1] \to \mathbb{R}$,

$$\int_0^1 h(y)\,\mathrm{d}y = \int_0^1 h(F(x))f(x)\,\mathrm{d}x.$$

Hence, the distribution of f under F becomes the uniform distribution and we can draw a random sample from the distribution given by ψ by drawing a random sample of the uniform distribution on $[0, 1]$ and applying F^{-1} to it.

Let us now consider a zonal function ψ on the rotation group SO(3). Then ψ depends only on the rotational angle ω of a rotation $R_{\xi,\omega}$ about an arbitrary axis $\xi \in \mathbb{S}^2$ and we will write $\phi(\omega) = \psi(R_{\xi,\omega})$. Using the parametrization of the rotational group by axis and angle the integral of an integrable function $h\colon \mathrm{SO}(3) \to \mathbb{R}$ with respect to ψ may be decomposed as

$$\int_{\mathrm{SO}(3)} h(x)\psi(x)\,\mathrm{d}\lambda(x) = \int_0^\pi \int_{\mathbb{S}^2} h(R_{\xi,\omega})\,\mathrm{d}\sigma(\xi)\,\phi(\omega)\sin^2\frac{\omega}{2}\,\mathrm{d}\omega.$$

Hence, we can draw a random sample $X_n \in \mathrm{SO}(3)$, $n = 1, \dots, N$ of the distribution given by ψ by drawing a random sample $\xi_n \in \mathbb{S}^2$, $n = 1, \dots, N$ of the uniform distribution on the unit sphere and drawing a random sample $\omega_n \in [0, \pi]$, $n = 1, \dots, N$ of the distribution given by the density $f(\omega) = \phi(\omega)\sin^2\frac{\omega}{2}$ and setting $X_n = R_{\xi_n,\omega_n}$.

3.2. Numerical computation of the kernel density estimator

Since we want to check our results for large sample sizes, i.e., up to $N = 10^7$, we have to apply fast algorithms to compute the kernel density estimator. An algorithm that allows to evaluate the kernel density estimator (3) corresponding to N random samples at M arbitrarily chosen nodes with the numerical complexity $\mathcal{O}(N + M)$ is described in [17]. However, as we are only interested in the L^2-error $\|f_\psi^* - f\|_2$ we rest at computing the Fourier coefficients of f and f_ψ^* up to polynomial degree $L = 128$ and applying Parseval's identity.

Let $X_n \in \mathrm{SO}(3)$, $n = 1, \dots, N$ be a random sample and let $\psi \in L^2(\mathrm{SO}(3))$ be a zonal function with finite Chebyshev expansion (A.13). Then according to (A.4) and (A.12) the kernel density estimator f_ψ^* has the representation

$$\begin{aligned} f_\psi^*(x) &= \frac{1}{N}\sum_{n=1}^N \sum_{\ell=0}^L (2\ell+1)\hat{\psi}(\ell)\mathcal{U}_{2\ell}\left(\cos\frac{\omega(x^{-1}X_n)}{2}\right) \\ &= \frac{1}{N}\sum_{n=1}^N \sum_{\ell=0}^L (2\ell+1)\hat{\psi}(\ell)\sum_{k,k'=-\ell}^{\ell} \overline{D_\ell^{k,k'}(X_n)}D_\ell^{k,k'}(x). \end{aligned}$$

Hence, the Fourier coefficients of the kernel density estimator f_ψ^* are given by the sum

$$\hat{f}_\psi^*(l, k, k') = \frac{1}{N}\sum_{n=1}^N \hat{\psi}(\ell)\sqrt{2\ell+1}\overline{D_l^{k,k'}(X_n)}$$

which is essentially an adjoined Fourier transform on the rotation group SO(3). Algorithms for the fast Fourier transform on the rotation group as well as for its adjoined transform have been described in [25] for regular nodes and in [33] for arbitrary nodes. An implementation of the latter one is available as part of the NFFT library [20].

3.3. Numerical results

In our numerical experiments we estimated the MISE for sample sizes $N = 10^1$ up to $N = 10^7$. For a fixed sample size N we considered different kernel functions. On the one hand we applied the MISE optimal kernel function as defined in Theorem 3 and compared the numerical estimated MISE with the theoretical expression found in (8). On the other hand we used the formulas (12), (15), (13), (18), and (21) for the optimal parameters of the $\mathcal{F}_{2,S}^2$ optimal Jackson type kernel (11), the $\mathcal{F}_{2.5,S}^\infty$ optimal kernel (15), the Dirichlet kernel (9), the Abel–Poisson kernel (20), and the de la Valleé Poussin kernel (17) and computed the MISE for the kernel density estimator with respect to these kernel functions. Fig. 1 shows the Chebyshev

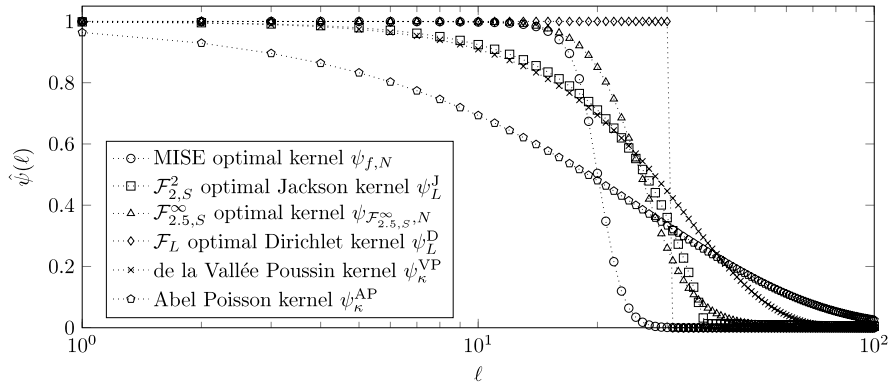

Fig. 1. This plot shows the Chebyshev coefficients of the kernel functions investigated throughout the numerical experiments. The kernel parameter has been chosen to be optimal with respect to the test function (24) and $N = 10^4$.

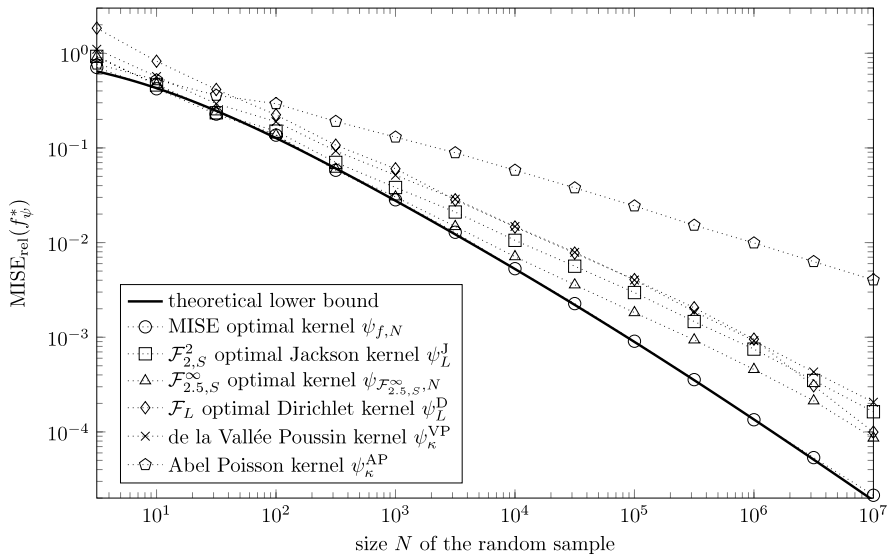

Fig. 2. This plot shows the MISE as a function of the number of random samples N and the kernel used for kernel density estimation. The theoretical bound as well as the MISE optimal kernel was computed according to Theorem 3. The parameters for the other kernel functions were chosen MISE optimal as specified in the formulas (12), (15), (13), (18), and (21).

coefficients of the kernel functions mentioned above with optimal kernel parameter for the specific choice of $N = 10^4$ random samples. In Fig. 2 the relative MISE

$$\mathrm{MISE}_{\mathrm{rel}}(f^*_\psi) = \frac{\mathrm{MISE}(f^*_\psi)}{\|f\|_2^2}$$

is plotted for the different kernel functions ψ.

Our numerical experiments show that the MISE for the optimal kernel almost perfectly fits our theoretical findings. This indicates that our approaches for generating the random sample and estimating the MISE work satisfactory. Furthermore, we observe for the $\mathcal{F}^2_{2,S}$ optimal, the $\mathcal{F}^\infty_{2.5,S}$ optimal, the Dirichlet kernel, and the de la Valleé Poussin kernel the predicted convergence rate $N^{-4/7}$ with a slightly better constant for the $\mathcal{F}^\infty_{2.5,S}$ optimal kernel function. As predicted we observe for the Abel–Poisson kernel the convergence rate $N^{-2/5}$. The more rapid convergence for the Dirichlet kernel starting with $N = 10^6$ is due to the fact that we worked with bandlimited functions.

3.4. Application to crystallographic texture analysis

The subject of crystallographic texture analysis is the microstructure of polycrystalline materials. A central question is the relative alignment of the crystals within the specimen. Thanks to the regular structure of the atomic lattice of crystals one can define for each crystal within the specimen a rotation $x \in \mathrm{SO}(3)$ that brings its atomic lattice in coincidence with a reference lattice fixed to the specimen. The alignment of a crystal within the specimen is discribed by a rotation $x \in \mathrm{SO}(3)$ that

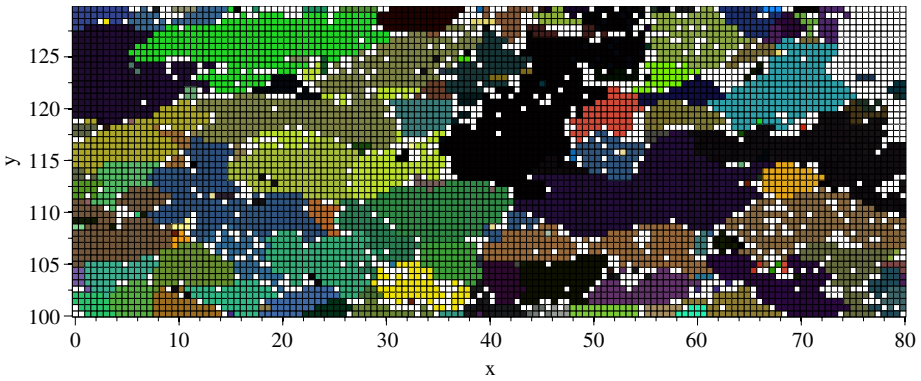

Fig. 3. The raw EBSD data. Each square corresponds to a single orientation measurement at the surface of the specimen. The color is computed by translating the rotation into Euler angles and assigning them to the RGB values. Empty squares are locations of measurement errors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

rotates some crystal fixed coordinate system into coincidence with some specimen fixed coordinate system. The rotational symmetries of the atomic lattice form a finite symmetry group $S \subset \mathrm{SO}(3)$ the so called *point group*. The *orientation* of a crystal is defined as the coset $xS \in \mathrm{SO}(3)/S$ of all symmetrically equivalent rotations.

Electron back scatter diffraction (EBSD) is a technique that uses an electron microscope to measure orientations $X_n S \in \mathrm{SO}(3)/S$, $n = 1, \dots, N$ at certain positions $p_x^n, p_y^n \in \mathbb{R}$ at the surface of a specimen [1,26]. In general, these positions are chosen to form a regular grid. In order to visualize such an orientation map each rotation is translated into Euler angles which in turn define the RGB values of a color. Fig. 3 shows a typical detail of such an orientation map. One easily identifies the crystals within the specimen which show up as connected regions of the same color.

Obviously, the orientation data are statistically dependent. This problem can be partly overcome by identifying the grain structure within the measurements, cf. [2], and selecting only one orientation per grain. Furthermore, the orientation data are usually affected by measurement errors. For simplicity, we assume here that all the orientation data X_n are exact and statistically independent, i.e., we completely disregard the spatial dependency. The impact of the absence of these simplification should be part of an upcoming paper.

An important characteristic of the crystallographic structure of a specimen is the so called *orientation density function* $f\colon \mathrm{SO}(3)/S \to \mathbb{R}$ which is defined as the relative frequency of orientations within the specimen by volume, i.e., for any measurable set $A \subset \mathrm{SO}(3)/S$ the integral $\int_A f(x)\,\mathrm{d}\lambda(x)$ is the volume portion of crystals within the specimen with an orientation in A. The orientation density function (ODF) of a specimen is the starting point for the calculation of several other texture characteristics and macroscopic properties of the specimen. Its determination from experimental measurements is a key problem in crystallographic texture analysis [7,29,16].

Estimating the ODF from individual crystal orientations has been discussed in many publications starting with [8]; see also [4,39] and the references therein. Assuming the orientation measurements to be a random sample of the unknown ODF f, a canonical estimator is the kernel density estimator [32,11]

$$f_\psi^*(xS) = \frac{1}{N}\sum_{n=1}^{N} \tilde{\psi}(X_n^{-1}x), \quad xS \in \mathrm{SO}(3)/S, \tag{25}$$

where $\psi \in L^2(\mathrm{SO}(3))$ is a zonal kernel function and

$$\tilde{\psi}(x) = \frac{1}{|S|}\sum_{s\in S} \psi(xs)$$

is its symmetrized version. Over the years there has been a lot of investigation in the texture community to find good kernel functions ψ and, most importantly, to determine the number of measurements which are necessary to estimate the true ODF f up to a given accuracy [3,6,39,27,10]. However, so far only results for specific model ODFs using numerical or experimental simulations are known.

In order to apply the results of our paper to crystallographic texture analysis we have to generalize them to the quotient $\mathrm{SO}(3)/S$, where $S \subset \mathrm{SO}(3)$ is a finite subgroup of cardinality $|S|$.

Theorem 13. *Let $f \in L^2(\mathrm{SO}(3)/S)$ be the true ODF, $X_1S, \dots, X_NS \in \mathrm{SO}(3)/S$ a corresponding random sample and let $\tilde{f} \in L^2(\mathrm{SO}(3))$ be defined by*

$$\tilde{f}(x) = f(xS), \quad x \in \mathrm{SO}(3).$$

Then we have for any zonal kernel function $\psi \in L^2(\mathrm{SO}(3))$ and its symmetrized version $\tilde{\psi} \in L^2(\mathrm{SO}(3))$,

$$\tilde{\psi}(x) = \frac{1}{|S|} \sum_{s \in S} \psi(xs),$$

the following representation of the MISE

$$\mathrm{MISE}(f_\psi^*) = \|\tilde{f} - \tilde{f} * \psi\|_2^2 + N^{-1}\Big(\|\tilde{\psi}\|_2^2 - \|\tilde{f} * \psi\|_2^2\Big).$$

Proof. First of all we observe that with $X_1, \dots, X_N \in \mathrm{SO}(3)$ being a random sample of $\tilde{f} \in L^2(\mathrm{SO}(3))$, the sequence $X_1 S, \dots, X_N S \in \mathrm{SO}(3)/S$ is a random sample of $f \in L^2(\mathrm{SO}(3)/S)$. Hence, we have for $x \in \mathrm{SO}(3)$,

$$\tilde{f}_{\tilde{\psi}}^*(x) = \frac{1}{N} \sum_{n=1}^{N} \tilde{\psi}(X_n^{-1} x) = f_\psi^*(xS)$$

and, consequently,

$$\begin{aligned}
\mathbb{E} f_\psi^*(xS) &= \mathbb{E} \tilde{f}_{\tilde{\psi}}^*(x) = \tilde{f} * \tilde{\psi}(x) \\
&= \frac{1}{|S|} \sum_{s \in S} \int_{\mathrm{SO}(3)} f(yS) \psi(y^{-1} xs) \, \mathrm{d}\lambda(y) \\
&= \frac{1}{|S|} \sum_{s \in S} \int_{\mathrm{SO}(3)} f(ysS) \psi(s^{-1} y^{-1} xs) \, \mathrm{d}\lambda(y) = \tilde{f} * \psi(x).
\end{aligned}$$

Together with Lemma 1 we obtain

$$\begin{aligned}
\mathrm{MISE}(f_\psi^*) &= \mathbb{E}\|f - f_\psi^*\|_2^2 = \mathbb{E}\|\tilde{f} - \tilde{f}_{\tilde{\psi}}^*\|_2^2 \\
&= \|\tilde{f} - \tilde{f} * \tilde{\psi}\|_2^2 + N^{-1}\Big(\|\tilde{\psi}\|_2^2 - \|\tilde{f} * \tilde{\psi}\|_2^2\Big) \\
&= \|\tilde{f} - \tilde{f} * \psi\|_2^2 + N^{-1}\Big(\|\tilde{\psi}\|_2^2 - \|\tilde{f} * \psi\|_2^2\Big). \quad \square
\end{aligned}$$

Observing further that for ψ sufficiently sharp, i.e.,

$$\int_{\mathrm{SO}(3)} \psi(sx)\psi(s'x) \, \mathrm{d}\lambda(x) \approx 0, \quad s, s' \in S, \ s \neq s',$$

we have

$$\|\tilde{\psi}\|_2^2 \approx |S|^{-1} \|\psi\|_2^2,$$

we conclude that all the results from Section 2 remain true for the quotient $\mathrm{SO}(3)/S$ if the number of random samples N is replaced by $\tilde{N} = |S| N$. As a consequence, the number of measurements N required for a given accuracy is the number of measurements $\tilde{N}$ required for the case without crystal symmetry divided by the number of symmetry elements $|S|$.

Since nonnegativity of the ODF is often required in texture analysis, we conclude from Section 2.3 that the de la Vallée Poussin kernel $\psi_\kappa^{\mathrm{VP}}$ is well suited for ODF estimation. Furthermore, if the smoothness of the ODF, i.e., $\|f\|_{2,2}$, is approximately known formula (18) gives the optimal kernel parameter and formula (19) can be used to determine the number of measurements necessary to achieve a given accuracy. If nothing is known about the smoothness of the function f adaptive estimation procedures have to be applied to determine the smoothness class simultaneously with the density estimate; cf. [35]. In this paper we mention only least squares cross validation [40], where the optimal kernel parameter κ_{opt} is derived from the data by

$$\kappa_{\mathrm{opt}} = \operatorname*{argmin}_\kappa \left(\|f_{\psi_\kappa^{\mathrm{VP}}}^*\|_2^2 + \frac{2}{(1 - N^{-1})} \left(\psi_\kappa^{\mathrm{VP}}(0) - \frac{1}{N} \sum_{m=1}^{N} f_{\psi_\kappa^{\mathrm{VP}}}^*(X_m) \right) \right). \tag{26}$$

A more throughout analysis of automatic bandwidth selection methods for kernel density estimation on the rotation group will be part of a forthcoming paper.

We complete our paper by considering a real world EBSD dataset of a ferrit specimen measured by I. Lischewski at the department of physical metallurgy and metal physics, Aachen, Germany. This dataset consists of 124.000 single orientations which subdivides into about 1500 crystals. A subset of these orientation data is plotted in Fig. 3. Using least squares cross validation (26) we estimate an optimal kernel parameter κ^* and apply the algorithm presented in Section 3.2 to efficiently evaluate the corresponding kernel density estimate $f_{\psi_{\kappa_{\mathrm{opt}}}^{\mathrm{VP}}}^*$. The common way to visualize the estimate $f_{\psi_{\kappa_{\mathrm{opt}}}^{\mathrm{VP}}}^*(x)$ of the ODF is

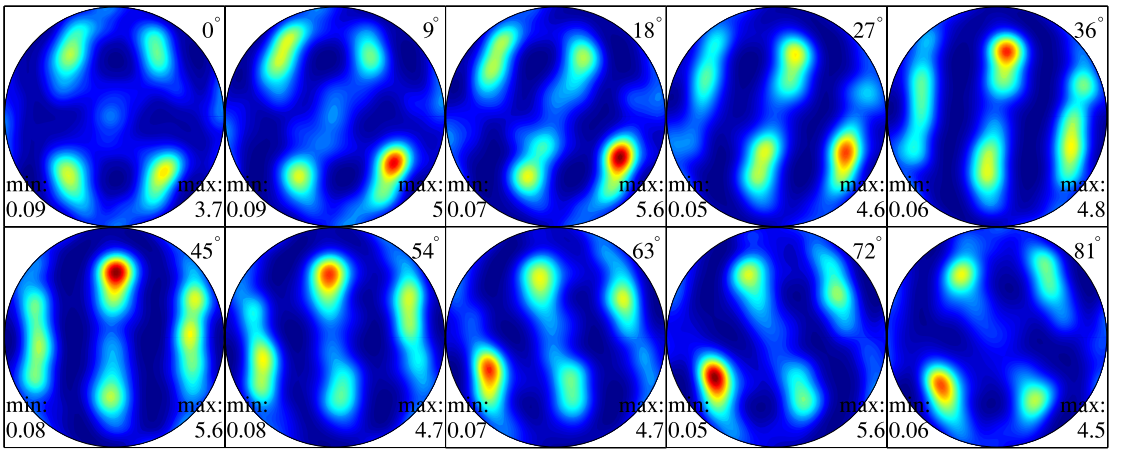

Fig. 4. Two-dimensional sections of the estimated ODF $f^*_{\psi_{\text{opt}}^{\text{VP}}}(x)$ with $\sigma_+ = \phi_1 + \phi_2$ fixed to the values $0°, 9°, \dots, 81°$ and the remaining parameters Φ and $\phi_1 - \phi_2$ describing in polar coordinates the location of a rotation $x \in \mathrm{SO}(3)$ within the plot, where ϕ_1, Φ, ϕ_2 are the Euler angles of $x \in \mathrm{SO}(3)$. Red colors indicate high values and blue colors low values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to plot it as a function of the Euler angles ϕ_1, Φ, ϕ_2. In order to sectioning this three dimensional domain we fix $\sigma_+ = \phi_1 + \phi_2$ and plot f with respect to the remaining parameters Φ and $\sigma_- = \phi_1 - \phi_2$. Such two dimensional sections are plotted in Fig. 4 where σ_+ is set to the values $0°, 9°, \dots, 81°$. More details about these, so called, sigma sections can be found in [28].

We conclude that the results presented in this paper allow to perform ODF estimation from EBSD data more efficiently. To make our results easily applicable, all algorithms described in this paper are freely available as part of the texture analysis toolbox MTEX [15,18].

Acknowledgments

The authors thank anonymous reviewers for their helpful comments which greatly improved the paper.

Appendix A. Harmonic analysis on the rotation group

We start by giving some basic notations and results on harmonic analysis on the rotation group SO(3). For a detailed introduction into harmonic analysis on SO(3), see [12,38,37]. By the rotation group SO(3) we denote the set of all orthogonal, three by three matrices with determinant one. Any such matrix $x \in \mathrm{SO}(3)$ can be interpreted as a rotation in the three dimensional Euclidean space about a certain axis of rotation $\xi \in \mathbb{S}^2$ and a certain rotational angle $\omega = \omega(x) = \arccos \frac{1}{2}(\mathrm{Tr}\, x - 1)$, where $\mathrm{Tr}\, x$ denotes the trace of x. Conversely, we denote for every unit vector $\xi \in \mathbb{S}^2$ and every angle $\omega \in [0, 2\pi]$ the matrix that acts as a rotation about ξ with angle ω by $R_{\xi,\omega} \in \mathrm{SO}(3)$.

Let $e_2 = (0, 1, 0)^t$, $e_3 = (0, 0, 1)^t$ and let $\alpha, \gamma \in [0, 2\pi)$, $\beta \in [0, \pi]$ be three angles. Then we define the Euler angle parametrization of the rotation group by the surjective mapping

$$(\alpha, \beta, \gamma) \mapsto x(\alpha, \beta, \gamma), \quad x(\alpha, \beta, \gamma) = R_{e_3,\alpha} R_{e_2,\beta} R_{e_3,\gamma}.$$

Note that the Euler angle parametrization is not unique in the identity, i.e., for all $\alpha \in [0, 2\pi]$, $R_{e_2,0} = R_{e_3,\alpha} R_{e_2,0} R_{e_3,-\alpha}$. Since SO(3) is a compact topological group it possesses a unique Haar measure λ such that $\lambda(\mathrm{SO}(3)) = 1$. In terms of Euler angles the Haar measure has the representation

$$\lambda(A) = \int_{\mathrm{SO}(3)} 1_A(x)\, \mathrm{d}\lambda(x) = \frac{1}{8\pi^2} \int_0^{2\pi} \int_0^{\pi} \int_o^{2\pi} 1_A(x(\alpha, \beta, \gamma))\, \mathrm{d}\alpha \sin\beta\, \mathrm{d}\beta\, \mathrm{d}\gamma$$

where A is an open subset of SO(3) and 1_A denotes the corresponding indicator function.

A.1. Harmonic functions

Our major function space will be the space of square integrable functions $L^2(\mathrm{SO}(3))$ on the rotation group endowed with the inner product

$$\langle f_1, f_2 \rangle = \int_{\mathrm{SO}(3)} f_1(x) \overline{f_2(x)}\, \mathrm{d}\lambda(x)$$

and the corresponding norm $\|f\|_2 = \sqrt{\langle f, f\rangle}$. An important function system on the rotation group is formed by the so called *Wigner-D functions*, cf. [37],

$$D_\ell^{k,k'}(x(\alpha,\beta,\gamma)) = \mathrm{e}^{-\mathrm{i}k\alpha}\mathrm{e}^{-\mathrm{i}k'\gamma} d_\ell^{k,k'}(\cos\beta), \quad \ell \in \mathbb{N},\ k,k' = -\ell,\dots,\ell, \tag{A.1}$$

with *Wigner-d functions* $d_\ell^{k,k'}: [-1,1] \to \mathbb{R}$, $\ell \in \mathbb{N}$, $k,k' = -\ell,\dots,\ell$,

$$d_\ell^{k,k'}(t) = \frac{(-1)^{\ell-k}}{2^\ell}\sqrt{\frac{(\ell+k)!}{(\ell-k')!(\ell+k')!(\ell-k)!}}\sqrt{\frac{(1-t)^{k'-k}}{(1+t)^{k+k'}}}\frac{\partial^{\ell-k}}{\partial t^{\ell-k}}\frac{(1+t)^{k'+\ell}}{(1-t)^{k'-\ell}}. \tag{A.2}$$

The Wigner-*D* functions can be characterized as the matrix elements of the left regular representation of the group SO(3) in $L^2(\mathbb{S}^2)$, i.e., for a certain orthonormal basis of spherical harmonics $\mathcal{Y}_\ell^k \in L^2(\mathbb{S}^2)$, $\ell = 0,\dots,\infty$, $k = -\ell,\dots,\ell$, they satisfy the representation properties

$$D_\ell^{k,k'}(x) = \left\langle \mathcal{Y}_\ell^{k'}(x^{-1}\cdot), \mathcal{Y}_\ell^k \right\rangle = \frac{1}{4\pi}\int_{\mathbb{S}^2} \mathcal{Y}_\ell^{k'}(x^{-1}\cdot\eta)\overline{\mathcal{Y}_\ell^k(\eta)}\,\mathrm{d}\sigma(\eta), \quad x \in \mathrm{SO}(3), \tag{A.3}$$

and

$$D_\ell^{k,k'}(xy) = \sum_{j=-\ell}^{\ell} D_\ell^{k,j}(x) D_\ell^{j,k'}(y), \quad x,y \in \mathrm{SO}(3). \tag{A.4}$$

In particular, (A.3) implies for all $k,k' = -\ell,\dots,\ell$, $x \in \mathrm{SO}(3)$,

$$\left|D_\ell^{k,k'}(x)\right| \le 1. \tag{A.5}$$

As a consequence of the Peter–Weyl Theorem [38, Section 3.3] the Wigner-*D* functions are orthogonal, i.e.,

$$\left\langle D_{\ell_1}^{k_1,k_1'}, D_{\ell_2}^{k_2,k_2'}\right\rangle = \frac{1}{8\pi^2}\int_{\mathrm{SO}(3)} D_{\ell_1}^{k_1,k_1'}(x)\overline{D_{\ell_2}^{k_2,k_2'}(x)}\ \mathrm{d}\lambda(x) = \frac{1}{2\ell_1+1}\delta_{k_1k_2}\delta_{k_1'k_2'}\delta_{\ell_1\ell_2}, \tag{A.6}$$

$\ell_1, \ell_2 \in \mathbb{N}$, $k,k' = -\ell,\dots,\ell$, and form a basis of $L^2(\mathrm{SO}(3))$. In particular, any function $f \in L^2(\mathrm{SO}(3))$ has a unique series expansion in terms of Wigner-*D* functions

$$f = \sum_{\ell=0}^{\infty}\sum_{k=-\ell}^{\ell}\sum_{k'=-\ell}^{\ell} \hat{f}(\ell,k,k')\sqrt{2\ell+1}D_\ell^{k,k'} \tag{A.7}$$

with Fourier coefficients $\hat{f}(\ell,k,k')$ given by the integral

$$\hat{f}(\ell,k,k') = \left\langle f, \sqrt{2\ell+1}D_\ell^{k,k'}\right\rangle. \tag{A.8}$$

Parseval's identity yields

$$\|f\|_2^2 = \sum_{\ell=0}^{\infty}\sum_{k,k'=-\ell}^{\ell}\left|\hat{f}(\ell,k,k')\right|^2. \tag{A.9}$$

Additionally, a complete system of rotational invariant and irreducible subspaces is given by

$$\mathrm{Harm}_\ell(\mathrm{SO}(3)) = \mathrm{span}\left\{D_\ell^{k,k'} \mid k,k' = -\ell,\dots,\ell\right\}$$

which satisfy

$$L^2(\mathrm{SO}(3)) = \mathrm{clos}\bigoplus_{\ell=0}^{\infty}\mathrm{Harm}_\ell(\mathrm{SO}(3)),$$

where clos denotes the closure in $L^2(\mathrm{SO}(3))$. Let $f, h \in L^2(\mathrm{SO}(3))$ be two square integrable functions on SO(3). Then their convolution

$$f * h(x) = \int_{\mathrm{SO}(3)} f(y)h(y^{-1}x)\,\mathrm{d}\lambda(y)$$

defines a function in $L^2(\mathrm{SO}(3))$ and we have the well known identity of its Fourier coefficients [25]

$$\widehat{f*h}(\ell,k,k') = \frac{1}{\sqrt{2\ell+1}}\sum_{j=-\ell}^{l}\hat{f}(\ell,k,j)\hat{h}(\ell,j,k'), \quad \ell \in \mathbb{N},\ k,k' = -\ell,\dots,\ell. \tag{A.10}$$

A.2. Zonal functions

A function $\psi : \mathrm{SO}(3) \to \mathbb{C}$ is called *zonal* if and only if it satisfies for all $x, y \in \mathrm{SO}(3)$

$$\psi(x) = \psi(yxy^{-1}).$$

Since for any $x \in \mathrm{SO}(3)$ the set of rotations $\{ yxy^{-1} \mid y \in \mathrm{SO}(3) \}$ can be identified with the set of all rotations $y \in \mathrm{SO}(3)$ having rotation angle $\omega(y) = \omega(x)$, a zonal function ψ can be written as a function of $t = \cos \frac{\omega(x)}{2}$. As long as it does not cause any confusion we write for the latter function

$$\psi(t) = \psi(x),$$

where x is an arbitrary rotation with $\cos \frac{\omega(x)}{2} = t$. Moreover, we have for $\psi \in L^2(\mathrm{SO}(3))$

$$\|\psi\|_2^2 = \frac{1}{8\pi^2} \int_{\mathrm{SO}(3)} |\psi(x)|^2 \, \mathrm{d}\lambda(x) = \frac{4}{\pi} \int_0^1 |\psi(t)|^2 \sqrt{1-t^2} \, \mathrm{d}t, \tag{A.11}$$

i.e., $t \mapsto \psi(t)$ is a function in $L^2([0,1], \sqrt{1-t^2}\,\mathrm{d}t)$.

By the Peter–Weyl Theorem the subspace of zonal functions in $L^2(\mathrm{SO}(3))$ is spanned by the characters $\chi_\ell,\ \ell \in \mathbb{N}$,

$$\chi_\ell(x) = \sum_{k=-\ell}^{\ell} D_\ell^{k,k}(x) = \mathcal{U}_{2\ell}\left(\cos \frac{\omega(x)}{2}\right) = \frac{\sin \frac{2\ell+1}{2}\omega(x)}{\sin \frac{\omega(x)}{2}}, \tag{A.12}$$

where $\mathcal{U}_\ell$ denotes the Chebyshev polynomials of second kind and degree $\ell \in \mathbb{N}$. In particular, the subspace of zonal functions in $\mathrm{Harm}_\ell(\mathrm{SO}(3))$ is one dimensional and any zonal function $\psi \in L^2(\mathrm{SO}(3))$ has a Chebyshev expansion of the form

$$\psi = \sum_{\ell=0}^{\infty} \hat{\psi}(\ell)(2\ell+1)\chi_\ell(x). \tag{A.13}$$

The Chebyshev coefficients $\hat{\psi}(\ell)$ are given by

$$\hat{\psi}(\ell) = \frac{4}{\pi} \frac{1}{2\ell+1} \int_0^1 \psi(t)\mathcal{U}_{2\ell}(t)\sqrt{1-t^2}\,\mathrm{d}t \tag{A.14}$$

and are related to the Fourier coefficients $\hat{\psi}(\ell, k, k'),\ \ell \in \mathbb{N}, k, k' = -\ell, \ldots, \ell$ of ψ by

$$\hat{\psi}(\ell, k, k') = \begin{cases} \sqrt{2\ell+1}\hat{\psi}(\ell), & k = k', \\ 0, & k \neq k'. \end{cases} \tag{A.15}$$

In particular, the Chebyshev coefficients satisfy Parseval's identity

$$\|\psi\|_2^2 = \sum_{\ell=0}^{\infty} (2\ell+1)^2 \, |\hat{\psi}(\ell)|^2. \tag{A.16}$$

As a special case of the convolution formula (A.10) the convolution of a function $f \in L^2(\mathrm{SO}(3))$ with a zonal function $\psi \in L^2(\mathrm{SO}(3))$ has the Fourier coefficients

$$\widehat{f * \psi}(\ell, k, k') = \hat{f}(\ell, k, k')\hat{\psi}(\ell), \quad \ell \in \mathbb{N},\ k, k' = -\ell, \ldots, \ell. \tag{A.17}$$

We will need also the following estimate on the Fourier coefficients of nonnegative functions on the rotation group.

Lemma 14. *Let $f \in L^2(\mathrm{SO}(3))$, $f \neq 0$, be an almost everywhere nonnegative function. Then we have for all $\ell \in \mathbb{N} \setminus \{0\}$,*

$$\frac{1}{(2\ell+1)^2} \sum_{k,k'=-\ell}^{\ell} \left| \hat{f}(\ell, k, k') \right|^2 < \hat{f}(0,0,0)^2.$$

For zonal functions $\psi \in L^2(\mathrm{SO}(3))$ with $\psi \geq 0$, almost everywhere, and $\hat{\psi}(0) > 0$ the above inequality simplifies to

$$\hat{\psi}(\ell)^2 < \hat{\psi}(0)^2, \quad \ell \in \mathbb{N} \setminus \{0\}.$$

Proof. Without loss of generality we may assume $\hat{f}(0,0,0) = \int_{\mathrm{SO}(3)} f(x)\,\mathrm{d}\lambda(x) = 1$. Together with the assumptions $f \geq 0$ almost everywhere this implies that there are an open set $A \subset \mathrm{SO}(3)$ and an $\varepsilon > 0$ such that $f(x) \geq \varepsilon$ for all $x \in A$. On

the other hand the characters $\chi_\ell(x) = \mathcal{U}_{2\ell}(\cos\frac{\omega(x)}{2})$, $\ell = 1, \dots, \infty$ are polynomials in $\cos\frac{\omega(x)}{2}$ and, hence, for every fixed $x \in \mathrm{SO}(3)$ we have $f * \chi_\ell(x) \neq \chi_\ell(xy^{-1})$ for almost all $y \in A$. From this we conclude

$$\begin{aligned} 0 &< \int_{\mathrm{SO}(3)} \big(f * \chi_\ell(x) - \chi_\ell(xy^{-1})\big)^2 f(y)\,\mathrm{d}\lambda(y) \\ &= (f * \chi_\ell)^2(x) - 2(f * \chi_\ell)^2(x) + f * \chi_\ell^2(x) = (f * \chi_\ell^2)(x) - (f * \chi_\ell)^2(x). \end{aligned}$$

Integration over SO(3) and making use of Fubini's Theorem, (A.13) and (A.17) result in

$$\begin{aligned} 0 &< \int_{\mathrm{SO}(3)} (f * \chi_\ell^2)(x) - (f * \chi_\ell)^2(x)\,\mathrm{d}\lambda(x) \\ &= \int_{\mathrm{SO}(3)} \int_{\mathrm{SO}(3)} \chi_\ell(xy^{-1})^2 f(y)\,\mathrm{d}\lambda(y)\,\mathrm{d}\lambda(x) - \|f * \chi_\ell\|_2^2 \\ &= \|\chi_\ell\|_2^2 - \|f * \chi_\ell\|_2^2 \\ &= 1 - \frac{1}{(2\ell+1)^2} \sum_{k,k'=-\ell}^{\ell} \left|\hat{f}(\ell, k, k')\right|^2. \end{aligned}$$

The second assertion is a direct consequence of (A.15). □

A.3. Sobolev spaces and integral means

In order to quantify the smoothness of functions on the rotation group we define weighted Sobolev spaces. Let $s \geq 0$ and let $f \in L^2(\mathrm{SO}(3))$ with Fourier coefficients $\hat{f}(\ell, k, k')$, $\ell \in \mathbb{N}$, $k, k' = -\ell, \dots, \ell$. Then we define Sobolev semi-norms of f by

$$\|f\|_{2,s}^2 = \sum_{\ell=1}^{\infty} \sum_{k,k'=-\ell}^{\ell} \ell^s (\ell+1)^s \left|\hat{f}(\ell, k, k')\right|^2,$$

$$\|f\|_{\infty,s}^2 = \sup_{\ell \in \mathbb{N}\setminus\{0\}} \sum_{k,k'=-\ell}^{\ell} \ell^s (\ell+1)^s \left|\hat{f}(\ell, k, k')\right|^2,$$

and consider for some $S > 0$ the smoothness class of probability densities

$$\mathcal{F}_{s,S}^2 = \left\{ f \in L^2(\mathrm{SO}(3)) \mid f \geq 0,\ \int f(x)\,\mathrm{d}\lambda(x) = 1,\ \text{and } \|f\|_{2,s} < S \right\}, \tag{A.18}$$

with finite Sobolev norm and the smoothness class of probability densities

$$\mathcal{F}_{s,S}^\infty = \left\{ f \in L^2(\mathrm{SO}(3)) \mid f \geq 0,\ \int f(x)\,\mathrm{d}\lambda(x) = 1,\ \text{and } \|f\|_{\infty,s} < S \right\} \tag{A.19}$$

with polynomially decaying Fourier coefficients. Since for any $s \geq \frac{1}{2}$ and $\varepsilon > 0$,

$$\begin{aligned} \sup_{\ell \in \mathbb{N}\setminus\{0\}} \sum_{k,k'=-\ell}^{\ell} \ell^{s-\frac{1}{2}} (\ell+1)^{s-\frac{1}{2}} \left|\hat{f}(\ell, k, k')\right|^2 &\leq \sum_{\ell=1}^{\infty} \sum_{k,k'=-\ell}^{\ell} \ell^{s-\frac{1}{2}} (\ell+1)^{s-\frac{1}{2}} \left|\hat{f}(\ell, k, k')\right|^2 \\ &\leq \left(\sum_{\ell=1}^{\infty} \ell^{-\frac{1}{2}-\varepsilon} (\ell+1)^{-\frac{1}{2}-\varepsilon} \right) \sup_{\tilde{\ell} \in \mathbb{N}\setminus\{0\}} \sum_{k,k'=-\tilde{\ell}}^{\tilde{\ell}} \tilde{\ell}^{s+\varepsilon} (\tilde{\ell}+1)^{s+\varepsilon} \left|\hat{f}(\tilde{\ell}, k, k')\right|^2 \end{aligned}$$

we have for $S > 0$ and $S' = S \sum_{\ell=1}^{\infty} \ell^{-\frac{1}{2}-\varepsilon} (\ell+1)^{-\frac{1}{2}-\varepsilon}$,

$$\mathcal{F}_{s+\varepsilon,S}^\infty \subset \mathcal{F}_{s-\frac{1}{2},S'}^2 \subset \mathcal{F}_{s-\frac{1}{2},S'}^\infty. \tag{A.20}$$

The final goal of this section is to derive an estimate for the approximation error $\|f - f * \psi\|_2^2$ for functions $f \in \mathcal{F}_{s,S}^2$ and $f \in \mathcal{F}_{s,S}^\infty$ and a zonal function $\psi \in L^2(\mathrm{SO}(3))$ with $\hat{\psi}(0) = 1$. To this end we consider the integral means

$$\tau_t f(x) = \frac{1}{4\pi} \int_{\mathbb{S}^2} f(x R_{\xi, 2\arccos t})\,\mathrm{d}\sigma(\xi), \quad t \in [0, 1],\ x \in \mathrm{SO}(3),$$

where $R_{\xi, 2\arccos t} \in \mathrm{SO}(3)$ is the rotation about $\xi \in \mathbb{S}^2$ with angle $2\arccos t \in [0, \pi]$ and σ is the spherical surface measure. The value $\tau_t f(x)$ represents the mean of the function f along all rotations y that are at distance $2\arccos t$ from x. In analogy to the spherical Funk–Hecke formula, cf. [30, Theorem 6], we have the following result on integral means of the Wigner-D functions.

Lemma 15. *Let $t \in [0,1]$, $\ell \in \mathbb{N}$, and $k, k' = -\ell, \dots, \ell$. Then we have*

$$\tau_t D_\ell^{k,k'} = \frac{1}{2\ell+1} \mathcal{U}_{2\ell}(t) D_\ell^{k,k'}.$$

Proof. First of all we recognize that $\tau_t D_\ell^{k,k'}$ may be rewritten by using (A.4) as

$$\begin{aligned}
\tau_t D_\ell^{k,k'}(x) &= \frac{1}{4\pi} \int_{\mathbb{S}^2} D_\ell^{k,k'}(xR_{\xi,2\arccos t}) \,\mathrm{d}\sigma(\xi) \\
&= \sum_{j=-\ell}^{\ell} D_\ell^{k,j}(x) \frac{1}{4\pi} \int_{\mathbb{S}^2} D_\ell^{j,k'}(R_{\xi,2\arccos t}) \,\mathrm{d}\sigma(\xi) \\
&= \sum_{j=-\ell}^{\ell} D_\ell^{k,j}(x) \frac{1}{8\pi^2} \int_{\mathrm{SO}(3)} D_\ell^{j,k'}(R_{y\xi,2\arccos t}) \,\mathrm{d}\lambda(y) \\
&= \sum_{j=-\ell}^{\ell} D_\ell^{k,j}(x) \frac{1}{8\pi^2} \int_{\mathrm{SO}(3)} D_\ell^{j,k'}(yR_{\xi,2\arccos t}y^{-1}) \,\mathrm{d}\lambda(y)
\end{aligned}$$

where the last two terms are independent from the specific choice of $\xi \in \mathbb{S}^2$. Since the last integral defines a zonal function with respect to $R_{\xi,2\arccos t}$ that is contained in $\mathrm{Harm}_\ell(\mathrm{SO}(3))$ we obtain

$$\frac{1}{8\pi^2} \int_{\mathrm{SO}(3)} D_\ell^{j,k'}(yR_{\xi,2\arccos t}y^{-1}) \,\mathrm{d}\lambda(y) = \begin{cases} \dfrac{1}{2\ell+1}\mathcal{U}_{2\ell}(t), & \text{if } j = k', \\ 0, & \text{if } j \neq k' \end{cases}$$

and consequently

$$\tau_t D_\ell^{k,k'}(x) = \frac{1}{2\ell+1} \mathcal{U}_{2\ell}(t) D_\ell^{k,k'}(x). \quad \square$$

Next we proceed as in [41] and show that the family of integral means τ_t, $t \in [0,1]$ defines an approximation process as $t \to 1$. To this end we need the following estimate.

Lemma 16. *Let $\omega \in [0,\pi]$ and $\ell \in \mathbb{N}$, $\ell > 0$. Then*

$$\frac{(2\ell+1) - \frac{\sin(2\ell+1)\omega}{\sin\omega}}{(2\ell+1)\ell(\ell+1)} \leq \frac{3 - \frac{\sin 3\omega}{\sin\omega}}{6}.$$

Proof. Let $\omega \in [0,\pi]$ and $\ell \in \mathbb{N}$, $\ell > 0$. Using the cosine expansion of the Dirichlet kernel and the Fejér kernel, cf. [9],

$$\frac{\sin(2\ell+1)\omega}{\sin\omega} = 1 + 2\sum_{k=1}^{\ell} \cos 2k\omega, \qquad \frac{\sin^2 \ell\omega}{\sin^2\omega} = \ell + 2\sum_{k=1}^{\ell-1} (\ell-k)\cos 2k\omega,$$

we obtain

$$\begin{aligned}
(2\ell+1) - \frac{\sin(2\ell+1)\omega}{\sin\omega} &= 2\sum_{k=1}^{\ell}(1-\cos 2k\omega) = 4(\sin^2\omega)\sum_{k=1}^{\ell}\frac{\sin^2 k\omega}{\sin^2\omega} \\
&= 4(\sin^2\omega)\left(\frac{\ell(\ell+1)}{2} + 2\sum_{k=1}^{\ell}\sum_{m=1}^{k-1}(k-m)\cos 2m\omega\right) \\
&= 2(\sin^2\omega)\left(\ell(\ell+1) + 2\sum_{m=1}^{\ell-1}(\ell-m)(\ell+1-m)\cos 2m\omega\right).
\end{aligned}$$

In particular, we have

$$\begin{aligned}
&(2\ell+1)\ell(\ell+1)\left(3 - \frac{\sin 3\omega}{\sin\omega}\right) - 6\left((2\ell+1) - \frac{\sin(2\ell+1)\omega}{\sin\omega}\right) \\
&= 4(\sin^2\omega)\left((2\ell+1)\ell(\ell+1) - 3\ell(\ell+1) - 6\sum_{m=1}^{\ell-1}(\ell-m)(\ell+1-m)\cos 2m\omega\right) \\
&\geq 4(\sin^2\omega)\left(2(\ell-1)\ell(\ell+1) - 6\sum_{m=1}^{\ell-1}(\ell-m)(\ell+1-m)\right) = 0. \quad \square
\end{aligned}$$

Lemma 17. *Let $S > 0, f \in \mathcal{F}^2_{2,S}$ and $t \in [0, 1]$. Then*

$$\|f - \tau_t f\|_2 \le \frac{2}{3}(1 - t^2)S.$$

Proof. For $f \in \mathcal{F}^2_{2,S}$ and $t \in [0, 1]$ we have by Lemma 15 and the normalization $\mathcal{U}_0(t) = 1$,

$$\begin{aligned}\|f - \tau_t f\|_2^2 &= \sum_{\ell=0}^{\infty} \sum_{k,k'=-\ell}^{\ell} \left(1 - \frac{\mathcal{U}_{2\ell}(t)}{2\ell+1}\right)^2 \left|\hat{f}(\ell, k, k')\right|^2 \\ &\le \left(\sup_{\ell\in\mathbb{N}\setminus\{0\}} \frac{(2\ell+1) - \mathcal{U}_{2\ell}(t)}{\ell(\ell+1)(2\ell+1)}\right)^2 \sum_{\ell=1}^{\infty} \sum_{k,k'=-\ell}^{\ell} \ell^2(\ell+1)^2 \left|\hat{f}(\ell, k, k')\right|^2.\end{aligned}$$

From Lemma 16 we know that for all $\omega \in [0, \pi]$ and $\ell \in \mathbb{N}$, $\ell > 0$,

$$\frac{(2\ell+1) - \mathcal{U}_{2\ell}(\cos\omega)}{\ell(\ell+1)(2\ell+1)} = \frac{(2\ell+1) - \frac{\sin(2\ell+1)\omega}{\sin\omega}}{(2\ell+1)\ell(\ell+1)} \le \frac{3 - \frac{\sin 3\omega}{\sin\omega}}{6} = \frac{3 - \mathcal{U}_2(\cos\omega)}{6},$$

and, hence,

$$\|f - \tau_t f\|_2^2 \le \left(\frac{3 - \mathcal{U}_2(t)}{6}\right)^2 S^2 = \left(\frac{2 - 2t^2}{3}\right)^2 S^2. \quad \square$$

We conclude our remarks on harmonic analysis on the rotation group by giving the promised approximation result on $\|f - f * \psi\|_2^2$.

Theorem 18. *Let $\psi \in L^2(\mathrm{SO}(3))$ be a zonal function with $\hat{\psi}(0) = 1$. Then we have for any $s, S > 0, f \in \mathcal{F}^2_{s,S}$ the inequality*

$$\|f - f * \psi\|_2^2 \le \sup_{\ell\in\mathbb{N}\setminus\{0\}} \frac{\left|1 - \hat{\psi}(\ell)\right|^2}{\ell^s(\ell+1)^s} S^2,$$

and for any $s > \frac{1}{2}, f \in \mathcal{F}^{\infty}_{s,S}$ the inequality

$$\|f - f * \psi\|_2^2 \le \sum_{\ell=1}^{\infty} \frac{\left|1 - \hat{\psi}(\ell)\right|^2}{\ell^s(\ell+1)^s} S^2.$$

If $f \in \mathcal{F}^2_{2,S}$ and $\psi \ge 0$, then the above estimate simplifies to

$$\|f - f * \psi\|_2 \le \frac{1}{2}\big(1 - \hat{\psi}(1)\big)S. \tag{A.21}$$

Proof. By (A.7) and (A.17) the convolution $f * \psi$ has the Fourier expansion

$$f * \psi = \sum_{\ell=0}^{\infty} \sum_{k,k'=-\ell}^{\ell} \hat{\psi}(\ell)\hat{f}(\ell, k, k')\sqrt{2\ell+1} D_\ell^{k,k'}.$$

Using Parseval's identity (A.9) and the assumption $\hat{\psi}(0) = 1$ we obtain the approximation error

$$\begin{aligned}\|f - f * \psi\|_2^2 &= \sum_{\ell=0}^{\infty} \sum_{k,k'=-\ell}^{\ell} \left|1 - \hat{\psi}(\ell)\right|^2 \left|\hat{f}(\ell, k, k')\right|^2 \\ &= \sum_{\ell=1}^{\infty} \sum_{k,k'=-\ell}^{\ell} \frac{\left|1 - \hat{\psi}(\ell)\right|^2}{\ell^s(\ell+1)^s} \ell^s(\ell+1)^s \left|\hat{f}(\ell, k, k')\right|^2 \\ &\le \sup_{\ell\in\mathbb{N}\setminus\{0\}} \frac{\left|1 - \hat{\psi}(\ell)\right|^2}{\ell^s(\ell+1)^s} \|f\|_{2,s}^2\end{aligned}$$

and, analogously,

$$\|f - f * \psi\|_2^2 \le \sum_{\ell \in \mathbb{N}\setminus\{0\}} \frac{\left|1 - \hat{\psi}(\ell)\right|^2}{\ell^s(\ell+1)^s} \|f\|_{\infty,s}^2.$$

The last sum is finite since we assumed $s > \frac{1}{2}$.

The proof for the case that ψ is nonnegative was adapted from [41]. First, we notice that for zonal functions ψ the convolution $f * \psi$ can be written in terms of the integral means τ_t,

$$\begin{aligned} f * \psi(x) &= \int_{\mathrm{SO}(3)} f(y)\psi(y^{-1}x)\,\mathrm{d}\lambda(y) = \int_{\mathrm{SO}(3)} f(xy)\psi(y)\,\mathrm{d}\lambda(y) \\ &= \frac{1}{\pi^2}\int_0^1 \int_{\mathbb{S}^2} f(xR_{\xi,2\arccos t})\psi(R_{\xi,2\arccos t})\,\mathrm{d}\sigma(\xi)\sqrt{1-t^2}\,\mathrm{d}t \\ &= \frac{4}{\pi}\int_0^1 \tau_t f(x)\psi(t)\sqrt{1-t^2}\,\mathrm{d}t. \end{aligned}$$

Together with $\hat{\psi}(0) = 1$ this allows us to write

$$\|f - f * \psi\|_2 = \left\| \frac{4}{\pi}\int_0^1 (f - \tau_t f)\psi(t)\sqrt{1-t^2}\,\mathrm{d}t \right\|_2.$$

Since $\psi \ge 0$ we may apply Jensen's inequality to interchange norm and integral. Using the estimate of the norm $\|f - \tau_t f\|_2$ found in Lemma 17 and formula (A.14) for the Chebyshev coefficients of ψ we obtain

$$\begin{aligned} \|f - f * \psi\|_2 &\le \frac{4}{\pi}\int_0^1 \|f - \tau_t f\|_2\, \psi(t)\sqrt{1-t^2}\,\mathrm{d}t \\ &\le \frac{4}{\pi}\int_0^1 \|f\|_{2,2}\left(\frac{1}{2} - \frac{1}{6}\mathcal{U}_2(t)\right)\psi(t)\sqrt{1-t^2}\,\mathrm{d}t \\ &= \frac{1}{2}\big(1 - \hat{\psi}(1)\big)\,\|f\|_{2,2}\,. \end{aligned}$$

It should be noted that because of $\psi \ge 0$ and Lemma 14 we have $1 - \hat{\psi}(1) > 0$. □

Eq. (A.21) of Theorem 18 allows us to find a lower bound of the L^2-norm of a nonnegative zonal function in dependency of its first Chebyshev coefficient.

Lemma 19. *Let $\psi \in L^2(\mathrm{SO}(3))$ be a nonnegative zonal function with $\hat{\psi}(0) = 1$. Then we have*

$$\|\psi\|_2^2 \ge \frac{64\sqrt{2}}{105}(1 - \hat{\psi}(1))^{-\frac{3}{2}}.$$

Proof. By Theorem 18 we have for any function $f \in L^2(\mathrm{SO}(3))$ with $\|f\|_{2,s} < \infty$,

$$\|f - f * \psi\|_2^2 \le \frac{(1 - \hat{\psi}(1))^2}{4} \sum_{\ell=0}^{\infty} \sum_{k,k'=-\ell}^{\ell} \ell^2(\ell+1)^2 \left|\hat{f}(\ell,k,k')\right|^2.$$

Setting for some $\ell \in \mathbb{N}$, $f = \sqrt{2\ell+1}D_\ell^{k,k'}$ we obtain

$$(1 - \hat{\psi}(\ell))^2 \le \frac{(1 - \hat{\psi}(1))^2}{4}\ell^2(\ell+1)^2$$

and, hence,

$$\hat{\psi}(\ell) \ge 1 - \frac{(1 - \hat{\psi}(1))}{2}\ell(\ell+1).$$

The right hand side is nonnegative whenever $L := \sqrt{\frac{2}{1-\hat{\psi}(1)}} - 1 \ge \ell$. Applying Parseval's identity (A.16) we obtain

$$\|\psi\|_2^2 \ge \sum_{\ell=0}^{L}(2\ell+1)^2\hat{\psi}(\ell)^2 \ge \sum_{\ell=0}^{\sqrt{\frac{2}{1-\hat{\psi}(1)}}-1}(2\ell+1)^2\left(1 - \frac{(1-\hat{\psi}(1))}{2}\ell(\ell+1)\right)^2 \ge \frac{64\sqrt{2}}{105}(1-\hat{\psi}(1))^{-\frac{3}{2}}. \quad \square$$

Appendix B. Proofs of Section 2

Proof of Lemma 1. First of all, we note that the mean of the kernel density estimator f_ψ^* may be written as

$$\mathbb{E}f_\psi^*(x) = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\psi(X_n^{-1}x) = \int_{\mathrm{SO}(3)} f(y)\psi(y^{-1}x)\,\mathrm{d}\lambda(y) = f*\psi(x).$$

Inserting the mean of the kernel density estimator f_ψ^* into the definition of the MISE and applying Fubini's Theorem we obtain

$$\begin{aligned}\mathbb{E}\|f-f_\psi^*\|_2^2 &= \mathbb{E}\|(f-\mathbb{E}f_\psi^*)-(f_\psi^*-\mathbb{E}f_\psi^*)\|_2^2\\ &= \mathbb{E}\|f-f*\psi\|^2 + \mathbb{E}\|f_\psi^*-\mathbb{E}f_\psi^*\|_2^2 - 2\mathbb{E}\left\langle f-\mathbb{E}f_\psi^*, f_\psi^*-\mathbb{E}f_\psi^*\right\rangle\\ &= \|f-f*\psi\|^2 + \mathbb{E}\|f_\psi^*-\mathbb{E}f_\psi^*\|_2^2.\end{aligned}$$

Using the independence of the random sample X_n, we have for any $x\in\mathrm{SO}(3)$

$$\begin{aligned}\mathbb{E}(\mathbb{E}f_\psi^*(x)-f_\psi^*(x))^2 &= \frac{1}{N^2}\sum_{n=1}^{N}\mathbb{E}(\mathbb{E}\psi(X_n^{-1}x)-\psi(X_n^{-1}x))^2\\ &= \frac{1}{N^2}\sum_{n=1}^{N}\mathbb{E}(\psi(X_n^{-1}x)^2-(\mathbb{E}\psi(X_n^{-1}x))^2)\\ &= \frac{1}{N}\left((f*\psi^2)(x)-(f*\psi)^2(x)\right),\end{aligned}$$

and, hence, the variation term on the right hand side of the previous sum yields

$$\begin{aligned}\mathbb{E}\|f_\psi^*-\mathbb{E}f_\psi^*\|_2^2 &= \int_{\mathrm{SO}(3)}\mathbb{E}(\mathbb{E}f_\psi^*(x)-f_\psi^*(x))^2\,\mathrm{d}\lambda(x)\\ &= \frac{1}{N}\int_{\mathrm{SO}(3)}(f*\psi^2)(x)-(f*\psi)^2(x)\,\mathrm{d}\lambda(x)\\ &= \frac{1}{N}\int_{\mathrm{SO}(3)}\int_{\mathrm{SO}(3)}\psi^2(y^{-1}x)f(y)\,\mathrm{d}\lambda(y)\,\mathrm{d}\lambda(x) - \frac{1}{N}\|f*\psi\|_2^2\\ &= \frac{1}{N}\|\psi\|_2^2 - \frac{1}{N}\|f*\psi\|_2^2. \quad \square\end{aligned}$$

Proof of Theorem 3. Since $f\geq 0$, we have by Lemma 14 for all $\ell\in\mathbb{N}$ the inequality $0\leq\hat{f}_\ell^2\leq 1$. Hence, each summand in (6) is a quadratic polynomial with respect to $\hat{\psi}(\ell)$ with minimum at

$$\hat{\psi}_{f,N}(\ell) = \frac{N\hat{f}_\ell^2}{(N-1)\hat{f}_\ell^2+1}.$$

Since, $f\in L^2(\mathrm{SO}(3))$ we have by Parseval's identity (A.9) that

$$\sum_{\ell=0}^{\infty}(2\ell+1)^2\hat{f}_\ell^2 = \sum_{\ell=0}^{\infty}\sum_{k,k'=-\ell}^{\ell}\left|\hat{f}(\ell,k,k')\right|^2 = \|f\|_2^2<\infty,$$

which shows that the Chebyshev coefficients $\hat{\psi}_{f,N}(\ell)\leq N\hat{f}_\ell^2$ are absolutely summable and, hence, square summable with respect to the weights $(2\ell+1)^2$. In particular, using Parseval's identity for zonal functions (A.16) we conclude that the Chebyshev coefficients $\hat{\psi}_{f,N}(\ell)$, $\ell\in\mathbb{N}$ define a zonal function $\psi_{f,N}\in L^2(\mathrm{SO}(3))$ such that the MISE of the corresponding kernel density estimator is optimal. A direct calculation of $\mathrm{MISE}(f_{\psi_{f,N}}^*)$ shows (8).

Since, we assumed $f\in L^2(\mathrm{SO}(3))$ to be not constant and nonnegative almost everywhere we conclude from Lemma 14 that there is a polynomial degree $\ell_0\in\mathbb{N}\setminus\{0\}$ such that $0<\hat{f}_{\ell_0}^2<1$. Hence, there is a constant $C>0$ such that

$$\mathrm{MISE}(f_{\psi_{f,N}}^*)\geq(2\ell_0+1)^2\frac{\hat{f}_{\ell_0}^2(1-\hat{f}_{\ell_0}^2)}{(N-1)\hat{f}_{\ell_0}^2+1} = (2\ell_0+1)^2\frac{\hat{f}_{\ell_0}^2(1-\hat{f}_{\ell_0}^2)}{\frac{N-1}{N}\hat{f}_{\ell_0}^2+N^{-1}}N^{-1}\geq CN^{-1}. \quad \square$$

Proof of Theorem 5. We start by deriving an asymptotic expression for the L^2-norm of the Jackson type kernel $\psi^{\mathrm{J}}_{L,s}$,

$$\begin{aligned}\|\psi^{\mathrm{J}}_{L,s}\|_2^2 &= \sum_{\ell=0}^{\lfloor L\rfloor}(2\ell+1)^2\left(1-\frac{\ell^{s/2}(\ell+1)^{s/2}}{L^{s/2}(L+1)^{s/2}}\right)^2\\ &\le \mathcal{O}(L^2)+4\sum_{\ell=0}^{\lfloor L\rfloor}\ell^2\left(1-\frac{\ell^s}{L^s}\right)^2\\ &= \mathcal{O}(L^2)+4\int_0^L \ell^2-2\frac{\ell^{2+s}}{L^s}+\frac{\ell^{2+2s}}{L^{2s}}\,\mathrm{d}\ell = \mathcal{O}(L^2)+CL^3,\end{aligned}$$

and, analogously,

$$\|\psi^{\mathrm{J}}_{L,s}\|_2^2 \ge \mathcal{O}(L^2)+4\sum_{\ell=0}^{\lfloor L\rfloor}(\ell+1)^2\left(1-\frac{(\ell+1)^s}{(L+1)^s}\right)^2 = \mathcal{O}(L^2)+CL^3, \tag{B.1}$$

where $C=4\left(\frac{1}{3}-\frac{2}{s+3}+\frac{1}{2s+3}\right)=\frac{8s^2}{3(s+3)(2s+3)}$. Inserting this asymptotic expression into (10) we can bound the MISE for any function $f\in F^2_{s,S}$ by

$$\mathrm{MISE}(f^*_{\psi^{\mathrm{J}}_{L,s}})\le L^{-2s}S^2+(CL^3+\mathcal{O}(L^2))N^{-1},$$

which is asymptotically minimized by

$$L^{2s+3}_{\mathrm{opt}}=\frac{2s}{3}C^{-1}S^2N.$$

Hence, we arrive at the upper bound

$$\lim_{N\to\infty}\inf_{\psi\in L^2(\mathrm{SO}(3))}\sup_{f\in\mathcal{F}^2_{s,S}}\mathrm{MISE}(f_\psi)\cdot N^{\frac{2s}{2s+3}}\le\left(\left(\frac{2s}{3}\right)^{-\frac{2s}{2s+3}}+\left(\frac{2s}{3}\right)^{\frac{3}{2s+3}}\right)C^{\frac{2s}{2s+3}}S^{\frac{6}{2s+3}}.$$

Next we want to show that this upper bound is strict. Therefore we consider an arbitrary kernel function ψ with Chebyshev coefficients satisfying $0\le\hat\psi(\ell)\le 1$ for all $\ell=1,\dots,\infty$. Let $\ell_0\in\mathbb{N}$ and $L\in\mathbb{R}$, $L>0$ such that

$$\ell_0=\operatorname*{argmax}_{\ell\in\mathbb{N}\setminus\{0\}}\frac{1-\hat\psi(\ell)}{\ell^{s/2}(\ell+1)^{s/2}}$$

and

$$L^{-s/2}(L+1)^{-s/2}=\frac{1-\hat\psi(\ell_0)}{\ell_0^{s/2}(\ell_0+1)^{s/2}}. \tag{B.2}$$

Then the Chebyshev coefficients of the Jackson type kernel $\psi^{\mathrm{J}}_{L,s}$ satisfy for all $\ell=1,\dots,L$,

$$\hat\psi^{\mathrm{J}}_{L,s}(\ell)=1-\ell^{s/2}(\ell+1)^{s/2}\frac{1-\hat\psi(\ell_0)}{\ell_0^{s/2}(\ell_0+1)^{s/2}}\le 1-\ell^{s/2}(\ell+1)^{s/2}\frac{1-\hat\psi(\ell)}{\ell^{s/2}(\ell+1)^{s/2}}=\hat\psi(\ell),$$

where equality is attained for $\ell=\ell_0$. In particular, we have by Parseval's equality (A.16) $\|\psi^{\mathrm{J}}_{L,s}\|_2\le\|\psi\|_2$.

For $S>0$ sufficiently small the function $f=1+S\ell_0^{-s/2}(\ell_0+1)^{-s/2}\sqrt{2\ell+1}D^{0,0}_{\ell_0}\in\mathcal{F}^2_{s,S}$ is nonnegative and we have by Parseval's equality (A.9), (A.17) and (B.2)

$$\|f\|_2^2=1+S^2\ell_0^{-s}(\ell_0+1)^{-s}\le 1+S^2$$

and

$$\|f-f*\psi\|_2^2=\hat f(\ell_0,0,0)^2(1-\hat\psi(\ell_0))^2=S^2\ell_0^{-s}(\ell_0+1)^{-s}\frac{\ell_0^s(\ell_0+1)^s}{L^s(L+1)^s}=S^2L^{-s}(L+1)^{-s}.$$

Hence, we obtain by Lemma 1, the fact $\|f*\psi\|_2\le\|f\|_2\le 1+S^2$, and (B.1) the following lower bound for the MISE of the function f,

$$\begin{aligned}\mathrm{MISE}(f^*_\psi) &\ge \|f-f*\psi\|_2^2+N^{-1}\left(\|\psi\|_2^2-\|f\|_2^2\right)\\ &\ge S^2L^{-s}(L+1)^{-s}+N^{-1}\left(\|\psi^{\mathrm{J}}_{L,s}\|_2^2-1-S^2\right)\\ &\ge S^2(L+1)^{-2s}+N^{-1}(CL^3+\mathcal{O}(L^2)).\end{aligned}$$

Minimizing the last term with respect to L we see that the lower bound coincides asymptotically with the upper bound. □

Proof of Theorem 6. Let $f \in \mathcal{F}_{s,S}^{\infty}$ and $\psi \in L^2(\mathrm{SO}(3))$ be a zonal kernel function. Setting the partial derivative of the upper bound (14) for each Chebyshev coefficient $\hat{\psi}(\ell)$, $\ell \in \mathbb{N} \setminus \{0\}$ to zero we obtain that the optimal kernel function $\psi_{\mathcal{F}_{s,S}^{\infty},N}$ is defined by the Chebyshev coefficients

$$\hat{\psi}_{\mathcal{F}_{s,S}^{\infty},N}(\ell) = \frac{NS^2}{NS^2 + (2\ell+1)^2\ell^s(\ell+1)^s}, \quad \ell \in \mathbb{N}. \tag{B.3}$$

Given $s > \frac{1}{2}$ the corresponding Chebyshev series converges in $L^2(\mathrm{SO}(3))$. Plugging in the kernel function $\psi_{\mathcal{F}_{s,S}^{\infty},N}$ into the upper bound we obtain

$$\begin{aligned}\mathrm{MISE}(f^*_{\psi_{\mathcal{F}_{s,S}^{\infty},N}}) &\le \sum_{\ell=1}^{\infty}\left(\frac{(1-\hat{\psi}_{\mathcal{F}_{s,S}^{\infty},N}(\ell))^2}{\ell^s(\ell+1)^s}S^2 + (2\ell+1)^2\hat{\psi}_{\mathcal{F}_{s,S}^{\infty},N}(\ell)^2N^{-1}\right)\\ &= \sum_{\ell=1}^{\infty}\frac{(2\ell+1)^4\ell^s(\ell+1)^sS^2 + (2\ell+1)^2NS^4}{\left(NS^2+(2\ell+1)^2\ell^s(\ell+1)^s\right)^2}.\end{aligned}$$

Next we make use of the fact

$$\left|(\ell+1)^s - \ell^s\right| \le \mathcal{O}(\ell^{s-1})$$

and of the integral (cf. (3.241.5) in [19])

$$\int_0^{\infty}\frac{x^{a-1}}{\left(A+x^b\right)^2}\,\mathrm{d}x = A^{\frac{a}{b}-2}\frac{\pi(b-a)}{b^2\sin\frac{a\pi}{b}}, \quad a \le 2b,\ A > 0$$

to conclude

$$\begin{aligned}\mathrm{MISE}(f^*_{\psi_{\mathcal{F}_{s,S}^{\infty},N}}) &\le \int_0^{\infty}\frac{(16\ell^{2s+4}+\mathcal{O}(\ell^{2s+3}))S^2 + (4\ell^2+\mathcal{O}(\ell))NS^4}{\left(NS^2+4\ell^{2s+2}\right)^2}\,\mathrm{d}\ell\\ &= \frac{2^{\frac{s-2}{s+1}}\pi}{(s+1)\sin\frac{3\pi}{2s+2}}S^{\frac{6}{2s+2}}N^{-\frac{2s-1}{2s+2}} + \mathcal{O}\left(N^{-\frac{2s}{2s+2}}\right).\end{aligned}$$

In order to prove that the upper bound is asymptotically sharp we consider the function $f \in \mathcal{F}_{s,S}^{\infty}$,

$$f = 1 + S\sum_{\ell=1}^{\infty}\ell^{-s/2}(\ell+1)^{-s/2}\sqrt{2\ell+1}D_{\ell}^{0,0},$$

which is a density function for S sufficiently small. Then for any zonal kernel function $\psi \in L^2(\mathrm{SO}(3))$ the bias term becomes

$$\|f - f*\psi\|_2^2 = \sum_{\ell=1}^{\infty}(1-\hat{\psi}(\ell))^2S^2\ell^{-s}(\ell+1)^{-s} = S^2\sum_{\ell=1}^{\infty}\frac{(1-\hat{\psi}(\ell))^2}{\ell^s(\ell+1)^s},$$

and, hence, the MISE evaluates to

$$\mathrm{MISE}(f^*_{\psi}) = \sum_{\ell=1}^{\infty}\left(\frac{(1-\hat{\psi}(\ell))^2}{\ell^s(\ell+1)^s}S^2 + \frac{\hat{\psi}(\ell)^2}{N}\left((2\ell+1)^2 - \frac{S^2}{\ell^s(\ell+1)^s}\right)\right).$$

Minimizing the MISE with respect to $\hat{\psi}(\ell)$ we obtain a similar result as in (B.3), i.e., with $(2\ell+1)^2$ replaced by $(2\ell+1)^2 - \frac{S^2}{\ell^s(\ell+1)^s}$,

$$\hat{\psi}(\ell) = \frac{NS^2}{NS^2 + \left((2\ell+1)^2 - \frac{S^2}{\ell^s(\ell+1)^s}\right)\ell^s(\ell+1)^s} = \frac{NS^2}{(N-1)S^2 + (2\ell+1)^2\ell^s(\ell+1)^s} \quad \ell \in \mathbb{N}.$$

Performing a similar calculation of the MISE as above shows that the upper bound is asymptotically sharp. □

References

[1] B.L. Adams, S.I. Wright, K. Kunze, Orientation imaging: the emergence of a new microscopy, J. Metall. Mater. Trans. A 24 (1993) 819–831.
[2] F. Bachmann, R. Hielscher, H. Schaeben, Grain detection from 2D and 3D EBSD data—specification of the MTEX algorithm, Ultramicroscopy 111 (2011) 1720–1733.
[3] T. Baudin, J. Jura, R. Penelle, J. Pospiech, Estimation of the minimum grain number for the orientation distribution function calculated from individual orientation measurements on Fe–3Si and Ti–4Al–6V alloys, J. Appl. Crystallogr. 28 (1995) 582–589.

[4] T. Baudin, R. Penelle, Determination of the total texture function from individual orientation measurements by electron backscattering pattern, Metall. Trans. A 24 (1993) 2299–2311.
[5] M.A. Bingham, S.B. Vardeman, D.J. Nordman, Bayes one-sample and one-way random effects analyses for 3-D orientations with application to materials science, Bayesian Anal. 4 (2009) 607–630.
[6] N. Bozzolo, F. Gerspach, G. Sawina, F. Wagner, Accuracy of orientation distribution function determination based on EBSD data—a case study of a recrystallized low alloyed Zr sheet, J. Microsc. 227 (2007) 275–283.
[7] H.J. Bunge, Texture Analysis in Material Science, Butterworths, 1982.
[8] H.-J. Bunge, F. Haessner, Three-dimensional orientation distribution function of crystals in cold-rolled copper, J. Appl. Phys. 39 (1968) 5503–5514.
[9] P.L. Butzer, R.J. Nessel, Fourier Analysis and Approximation, Volume 1, Birkhäuser, 1971.
[10] P. Eisenlohr, F. Roters, Selecting sets of discrete orientations for accurate texture reconstruction, Comput. Mater. Sci. 42 (2008) 670–678.
[11] O. Engler, G. Gottstein, J. Pospiech, J. Jura, Statistics, evaluation and representation of single grain orientation measurements, Mater. Sci. Forum 157–162 (1994) 259–274.
[12] I.M. Gelfand, R.A. Minlos, Z.Y. Shapiro, Representations of the Rotation and Lorentz Groups and their Applications, Pergamon Press, Oxford, 1963.
[13] D.M. Healy Jr., H. Hendriks, P.T. Kim, Spherical deconvolution, J. Multivariate Anal. 67 (1998) 1–22.
[14] H. Hendriks, Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansion, Ann. Statist. 18 (1990) 832–849.
[15] R. Hielscher, MTEX 3.0—a texture calculation toolbox. http://mtex.googlecode.com.
[16] R. Hielscher, D. Potts, J. Prestin, H. Schaeben, M. Schmalz, The Radon transform on SO(3): a Fourier slice theorem and numerical inversion, Inverse Problems 24 (2008) 025011.
[17] R. Hielscher, J. Prestin, A. Vollrath, Fast summation of functions on SO(3), Math. Geosci. 42 (2010) 773–794.
[18] R. Hielscher, H. Schaeben, A novel pole figure inversion method: specification of the *MTEX* algorithm, J. Appl. Crystallogr. 41 (6) (2008) 1024–1037.
[19] A. Jeffrey, D. Zwillinger (Eds.), 3–4 Definite integrals of elementary functions, in: Table of Integrals, Series, and Products (Sixth Edition), sixth ed., Academic Press, San Diego, 2000, pp. 243–614.
[20] J. Keiner, S. Kunis, D. Potts, Using NFFT3—a software library for various nonequispaced fast Fourier transforms, ACM Trans. Math. Software 36 (2009) 1–30. Article 19.
[21] P.T. Kim, Deconvolution density estimation on SO(N), Ann. Statist. 26 (1998) 1083–1102.
[22] P.T. Kim, J.Y. Koo, Z.M. Luo, Weyl eigenvalue asymptotics and sharp adaptation on vector bundles, J. Multivariate Anal. 100 (9) (2009) 1962–1978.
[23] J.Y. Koo, Optimal rates of convergence for nonparametric statistical inverse problems, Ann. Statist. 21 (1993) 590–599.
[24] J.Y. Koo, P.T. Kim, Asymptotic miinmax bounds for stochastic deconvolution over groups, IEEE Trans. Inform. Theory 54 (2008) 289–298.
[25] P.J. Kostelec, D.N. Rockmore, FFTs on the rotation group, J. Fourier Anal. Appl. 14 (2008) 145–179.
[26] K. Kunze, S.I. Wright, B.L. Adams, D.J. Dingley, Advances in automatic EBSP single orientation measurements, Textures Microstruct. 20 (1993) 41–54.
[27] V. Luzin, Optimization of texture measurements. III. Statistical relevance of ODF represented by individual orientations, Mater. Sci. Forum 273–275 (1998) 107–112.
[28] S. Matthies, K. Helming, K. Kunze, On the representation of orientation distributions in texture analysis by sigma-sections. II. Consideration of crystal and sample symmetry, examples, Phys. Status Solidi (B) 157 (1990) 489–507.
[29] S. Matthies, G. Vinel, K. Helmig, Standard Distributions in Texture Analysis, Volume 1, Akademie-Verlag, Berlin, 1987.
[30] C. Müller, Spherical Harmonics, Springer, Aachen, 1966.
[31] B. Pelletier, Kernel density estimation on Riemannian manifolds, Statist. Probab. Lett. 73 (2005) 297–304.
[32] J. Pospiech, K. Lücke, The rolling textures of copper and [alpha]-brasses discussed in terms of the orientation distribution function, Acta Metall. 23 (1975) 997–1007.
[33] D. Potts, J. Prestin, A. Vollrath, A fast algorithm for nonequispaced Fourier transforms on the rotation group, Numer. Algorithms 52 (2009) 355–384.
[34] M. Reiß, Asymptotic equivalence for nonparametric regression with multivariate and random design, Ann. Statist. 36 (2008) 1957–1982.
[35] A. Tsybakov, Introduction to Nonparametric Estimation, Springer Verlag, Berlin, 2009.
[36] A.W. van der Vaart, Asymptotic Statistics, Cambridge University Press, Cambridge, 1998.
[37] D. Varshalovich, A. Moskalev, V. Khersonskii, Quantum Theory of Angular Momentum, World Scientific Publishing, Singapore, 1988.
[38] N. Vilenkin, Special Functions and the Theory of Group Representations, Amer. Math. Soc., Providence, RI, USA, 1968.
[39] F. Wagner, S. Matthies, O. Van Landuyt, Processing individual orientation data to calculate ODFs, Mater. Sci. Forum 273–275 (1998) 89–98.
[40] M.P. Wand, M.C. Jones, Kernel Smoothing, Chapman & Hall/CRC, 1995.
[41] M. Wehrens, Legendre–Transformationsmethoden und Approximation von Funktionen auf der Einheitskugel im R^3, Dissertation, Institut für Mathematik, Technische Hochschule Aachen, 1980.