TECHNISCHE UNIVERSITÄT
CHEMNITZ

Kernel Density Estimation on the Rotation Group

R. Hielscher

Preprint 2010-7

Fakultät für Mathematik

Impressum:
Herausgeber:
Der Dekan der
Fakultät für Mathematik
an der Technischen Universität Chemnitz
Sitz:
Reichenhainer Straße 39
09126 Chemnitz
Postanschrift:
09107 Chemnitz
Telefon: (0371) 531-22000
Telefax: (0371) 531-22009
E-Mail: dekanat@mathematik.tu-chemnitz.de
Internet:
http://www.tu-chemnitz.de/mathematik/
ISSN 1614-8835 (Print)

Kernel Density Estimation on the Rotation Group

Ralf Hielscher

June 11, 2010

Abstract

We are concerned with kernel density estimation on the rotation group $\mathrm{SO}(3)$ and the corresponding mean integrated squared error. We give lower and upper bounds for different function classes and derive optimal kernel functions. Furthermore, we consider approximations to the mean integrated squared error that depend on certain Sobolev norms of the density function and analyze them with respect to asymptotic behavior and optimal kernel functions. We compare our optimal kernels functions to families of kernel functions commonly used for kernel density estimation on the rotation group. Finally, we give a fast algorithm for the computation of the kernel density estimator for large sampling sets and verify our theoretical findings by numerical experiments.

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. While for Euclidean domains and for directional and spherical data there exists a rich literature (see [9, 18, 17] and the references therein), the statistical properties of the kernel density estimator for more specific domains where investigated in less detail, e.g. in [2, 11, 1].

In this paper we are concerned with kernel density estimation on the rotation group $\mathrm{SO}(3)$. 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 of a specimen from electron back scattering diffraction (EBSD) data [8, 13]. Furthermore, recently algorithms have been developed that allow for the fast evaluation of the kernel density estimator on the rotation group for very large sampling sets [4].

Let λ be the Haar measure on $\mathrm{SO}(3)$ and let $X_1, \ldots, X_N \in \mathrm{SO}(3)$ be a random sample corresponding to a square integrable, probability density function $f \in L^2(\mathrm{SO}(3))$ with respect to λ. Then for any square integrable, zonal function $\psi \in L^2(\mathrm{SO}(3))$ the *kernel density estimator* $f^*_\psi \colon \mathrm{SO}(3) \to \mathbb{R}$ of f given the random sample $X_1, \ldots, X_N \in \mathrm{SO}(3)$ is defined as

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

In our paper we are interested in the *mean integrated squared error* (MISE),

$$\mathrm{MISE}(f_\psi^*) = \mathbb{E}\,\|f - f_\psi^*\|^2 = \mathbb{E}\int_{\mathrm{SO}(3)} |f(x) - f^*(x)|^2 \,\mathrm{d}\lambda(x)$$

as a measure for the mean discrepancy between the probability density function f and the kernel density estimator f_ψ^*.

In the groundbreaking papers of Hendriks [2] and Pelletier [11] the authors gave asymptotic upper bounds for the MISE for specific kernel functions ψ in a much more general setting of d–dimensional Riemannian manifolds. More specifically, it was shown by Hendriks that the Dirichlet kernel $\psi_L(x) = \sum_{\ell=0}^{L} D_\ell(x)$, where D_ℓ denote an eigenbasis of the Laplace–Beltrami operator on the manifold, yields the estimate

$$\inf_{L\in\mathbb{N}} \mathrm{MISE}(f_{\psi_L}^*) \le CN^{-\frac{2s}{2s+d}},$$

where the density function f is assumed to be $s > d/2$ times differentiable with square integrable derivatives and the constant C is independent of N. This result was extended by Pelletier to nonnegative kernel functions $\psi_\alpha(x, X_n) = \psi(\frac{d(x,X_n)}{\alpha})$ that are derived from a nonnegative kernel functions $\psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ with vanishing first moment, finite second moment and supremum at 0 by dilating it according to the Rimannian distance. Assuming, furthermore, that f is two times differentiable with square integrable derivatives Pelletier proved the estimate

$$\inf_{\alpha\in\mathbb{R}} \mathrm{MISE}(f_{\psi_\alpha}^*) \le CN^{-\frac{4}{4+d}}.$$

The purpose of this paper is specialize and extend these results to the specific domain of the rotation group $\mathrm{SO}(3)$, i.e., we want to provide lower and upper bounds with explicite constants for a broader class of kernel functions. The findings of our paper are twofold. As a first result we show in Theorem 6 that under the very mild condition $f \in L^2(\mathrm{SO}(3))$ it exists a well defined optimal zonal kernel function ψ_{opt} such that the corresponding MISE is minimal and satisfies

$$\mathrm{MISE}(f_{\psi_{\mathrm{opt}}}^*) \to 0$$

as the number of random samples N tends to infinity. For the case of bandlimited density functions f we prove in Theorem 7 the existence of a constant $C > 0$ such that

$$CN^{-1} < \mathrm{MISE}(f_{\psi_{\mathrm{opt}}}^*) \le \frac{(L+1)(1+2L)(3+2L)}{3(1+\sqrt{N})^2}.$$

In the case that the Fourier coefficients of f decay polynomial with order $s > \frac{1}{2}$ we prove in Theorem 8 for the optimal MISE the upper bound

$$\mathrm{MISE}(f_{\psi_{\mathrm{opt}}}^*) \le \left(\frac{4^4}{3}\,\|f\|_{\infty,s}^{\frac{6}{2s+2}} + \frac{\pi}{2+2s}\right) N^{-\frac{2s-1}{2s+2}}.$$

Unfortunately, the optimal kernel function ψ_{opt} is given in terms the Fourier coefficients of the unknown density function f. Consequently, it is of limited interest for practical applications.

Therefore, one considers asymptotically strict upper bounds of the MISE that depend on some Sobolev norm of the density function f and looks for optimal kernel functions with respect to these upper bounds. In our paper we generalize in Theorem 9 the notion of the *asymptotic mean integrated squared error* (AMISE) known from the Euclidean setting and extend it in Theorem 10 to general Sobolev spaces. More precisely, we prove for $s > 0$ times weakly differentiable functions f an upper bound $\mathrm{AMISE}_{2,s}(f^*_{\psi_{\mathrm{opt}}})$ of the MISE and show in Theorem 11 that it possesses an optimal kernel functions ψ_{opt} which satisfies

$$\mathrm{MISE}(f^*_{\psi_{\mathrm{opt}}}) \leq \mathrm{AMISE}_{2,s}(f^*_{\psi_{\mathrm{opt}}}) = C\, \|f\|_s^{\frac{6}{2s+3}}\, N^{-\frac{2s}{2s+3}} + \mathcal{O}(N^{-\frac{2s+1}{2s+3}})$$

with constant $C = \left(\left(\frac{2s}{3}\right)^{-\frac{2s}{2s+3}} + \left(\frac{2s}{3}\right)^{\frac{3}{2s+3}} \right) \left(\frac{4}{3} - \frac{8}{s+3} + \frac{4}{2s+3} \right)^{\frac{2s}{2s+3}}$. For the case that the Fourier coefficients of f decay polynomially of order $s > \frac{1}{2}$ we find an upper bound $\mathrm{AMISE}_{\infty,s}(f^*_\psi)$ of the MISE and derive in Theorem 12 an optimal kernel function ψ_{opt} which satisfies

$$\mathrm{MISE}(f^*_{\psi_{\mathrm{opt}}}) \leq \mathrm{AMISE}_{\infty,s}(f^*_{\psi_{\mathrm{opt}}}) = \frac{2^{\frac{2-s}{s+1}}\pi}{(s+1)\sin\frac{3\pi}{2s+2}}\, \|f\|_{\infty,s}^{\frac{2}{2s+2}}\, N^{-\frac{2s-1}{2s+2}} + \mathcal{O}(N^{-\frac{2s}{2s+2}}).$$

We consider also some important families of zonal functions on the rotation group: the Dirichlet kernel, the Jackson kernel, the Abel Poisson kernel, the de la Valle'e Poussin kernel, which are commonly used in practical applications. For these families we compute the AMISE - optimal parameters and compare the corresponding asymptotic behavior of the AMISE with our optimal kernel functions.

We complete our paper with some numerical experiments to verify our theoretical findings. In order to evaluate the kernel sum (1) for large sampling sets, i.e., for $N \approx 10^7$, and to compute its Fourier coefficients we apply the nonequispaced fast Fourier transform on $\mathrm{SO}(3)$, [12] and follow the ideas in [4].

Most of our results we derived by extending the approaches in [2] and [6] and making extensively use of harmonic analysis on the rotation group. In particular, we proved in Lemma 5 an inequality for the Fourier coefficients of nonnegative functions on $\mathrm{SO}(3)$ and generalized in Theorem 4 an approximation result for the convolution with nonnegative functions from [19] for the rotation group. The authors are optimistic that the applied techniques might be useful for other domains then the rotation group $\mathrm{SO}(3)$ as well.

2 Harmonic Analysis on the Rotation Group

We start by giving some basic notations and results on harmonic analysis on the rotation group $\mathrm{SO}(3)$. By the rotation group $\mathrm{SO}(3)$ 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

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

Since, $\mathrm{SO}(3)$ 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 $\mathrm{SO}(3)$ and 1_A denotes the corresponding indicator function.

2.1 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 $\|f\| = \sqrt{\langle f, f \rangle}$. An important function system on the rotation group is formed by the so called *Wigner-D functions* (cf. [15])

$$D_\ell^{k,k'}(x(\alpha, \beta, \gamma)) = \mathrm{e}^{-\mathrm{i}k\alpha}\mathrm{e}^{-\mathrm{i}k'\gamma} d_\ell^{k,k'}(\cos\beta), \quad \ell \in \mathbb{N},\ k, k' = -\ell, \ldots, \ell, \tag{2}$$

with *Wigner-d functions*

$$d_\ell^{k,k'}(x) = \frac{(-1)^{\ell-k}}{2^\ell} \sqrt{\frac{(\ell+k)!}{(\ell-k')!(\ell+k')!(\ell-k)!}} \sqrt{\frac{(1-x)^{k'-k}}{(1+x)^{k+k'}}} \frac{\mathrm{d}^{\ell-k}}{\mathrm{d}x^{\ell-k}} \frac{(1+x)^{k'+\ell}}{(1-x)^{k'-\ell}}. \tag{3}$$

The Wigner-D functions can be characterized as the matrix elements of the left regular representation of the group $\mathrm{SO}(3)$ in $L^2(\mathbb{S}^2)$, i.e., they satisfy the representation property

$$D_\ell^{k,k'}(xy) = \sum_{j=-\ell}^{\ell} D_\ell^{k,j}(x) D_\ell^{j,k'}(y). \tag{4}$$

As a consequence of the Peter - Weyl Theorem [16, Sect. 3.3] the Wigner-D functions are orthogonal, i.e.,

$$\left\langle D_{\ell_1}^{k_1,k_1'}, D_{\ell_2}^{k_2,k_2'} \right\rangle = \frac{1}{8\pi^2} \int_{\mathrm{SO}(3)} D_{\ell_1}^{k_1,k_1'}(x)\overline{D_{\ell_2}^{k_2,k_2'}(x)}\ \mathrm{d}\lambda(x) = \frac{1}{2\ell+1} \delta_{k_1 k_2} \delta_{k_1' k_2'} \delta_{\ell_1 \ell_2}, \tag{5}$$

$\ell \in \mathbb{N}$, $k, k' = -\ell, \ldots, \ell$, 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

$$f = \sum_{\ell=0}^{\infty} \sum_{k=-\ell}^{\ell} \sum_{k'=-\ell}^{\ell} \hat{f}(\ell, k, k') \sqrt{2\ell+1} D_\ell^{k,k'} \tag{6}$$

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

$$\hat{f}(\ell,k,k') = \left\langle f, \sqrt{2\ell+1} D_\ell^{k,k'} \right\rangle. \tag{7}$$

The Parsevall identity yields

$$\|f\|^2 = \sum_{\ell=0}^{\infty} \sum_{k,k'=-\ell}^{\ell} \left| \hat{f}(l,k,k') \right|^2. \tag{8}$$

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

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

which satisfy

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

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

$$f * h(x) = \frac{1}{8\pi^2} \int_{\mathrm{SO}(3)} f(y) h(y^{-1}x) \,\mathrm{d}y$$

defines a function in $L^2(\mathrm{SO}(3))$ and we have the well known identity of its Fourier coefficients [7]

$$\widehat{f*h}(\ell,k,k') = \frac{1}{\sqrt{2\ell+1}} \sum_{j=-\ell}^{l} \hat{f}(\ell,k,j)\hat{h}(\ell,j,k'), \quad \ell \in \mathbb{N},\ k,k' = -\ell,\ldots,\ell. \tag{9}$$

2.2 Zonal Functions

A function $\psi\colon \mathrm{SO}(3) \to \mathbb{C}$ 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 $\psi \in L^2(\mathrm{SO}(3))$

$$\|\psi\|^2 = \frac{1}{8\pi^2} \int_{\mathrm{SO}(3)} |\psi(x)|^2 \,\mathrm{d}\lambda(x) = \frac{2}{\pi} \int_{-1}^{1} |\psi(t)|^2 \sqrt{1-t^2} \,\mathrm{d}t, \tag{10}$$

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

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

$$\chi_\ell(x) = \sum_{k=-\ell}^{\ell} D_\ell^{k,k}(x) = \mathcal{U}_{2\ell}(\cos\tfrac{\omega(x)}{2}) = \frac{\sin\frac{2\ell+1}{2}\omega(x)}{\sin\frac{\omega(x)}{2}},$$

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

$$\psi(x) \sim \sum_{\ell=0}^{\infty} \hat{\psi}(\ell)(2\ell+1)U_{2\ell}\Big(\cos\tfrac{\omega(x)}{2}\Big),$$

where $\sim$ indicates that the convergence of the series is meant in the L^2–norm. The Chebyshev coefficients $\hat{\psi}(\ell)$ are given by

$$\hat{\psi}(\ell) = \frac{2}{\pi}\frac{1}{2\ell+1}\int_{-1}^{1}\psi(t)\mathcal{U}_{2\ell}(t)\sqrt{1-t^2}\,\mathrm{d}t \tag{11}$$

and satisfy the Parsevall identity

$$\|\psi\|^2 = \sum_{\ell=0}^{\infty}(2\ell+1)^2\,|\hat{\psi}(\ell)|^2. \tag{12}$$

As a special case of the convolution formulae (9) 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

$$\widehat{f * \psi}(\ell,k,k') = \hat{f}(\ell,k,k')\hat{\psi}(\ell), \quad \ell \in \mathbb{N},\ k,k' = -\ell,\dots,\ell. \tag{13}$$

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

Lemma 1. *Let $f \in L^2(\mathrm{SO}(3))$ be an almost everywhere nonnegative function with $\hat{f}(0) > 0$. Then we have for all $\ell \in \mathbb{N}\backslash\{0\}$,*

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

Proof. Since the characters $\chi_\ell(x) = \mathcal{U}_{2\ell}(\cos\frac{\omega(x)}{2})$, $\ell = 1,\dots,\infty$ are not constant at any open subset of $\mathrm{SO}(3)$ we have for all $f \geq 0$, $\hat{f}(0) > 0$ and all $x \in \mathrm{SO}(3)$,

$$0 < \int_{\mathrm{SO}(3)}\left(f * \chi_\ell(x) - \chi_\ell(xy^{-1})\right)^2 f(y)\,\mathrm{d}\lambda(y) = (\chi_\ell^2 * f)(x) - (\chi_\ell * f)^2(x).$$

Integration over $\mathrm{SO}(3)$ results in

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

□

2.3 Sobolev Spaces and Integral Means

In order to quantify the smoothness of functions on the rotation group we define weighted Sobolev spaces. Let $s \in \mathbb{R}$, $s \geq 0$, $L \in \mathbb{N}$ and let

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

be a band-limited function with Fourier coefficients $\hat{f}(l,k,k') \in \mathbb{C}$. Then we define weighted Sobolev semi–norms of f by

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

and denote by $F\ell_{2,s}(\mathrm{SO}(3))$ and $F\ell_{\infty,s}(\mathrm{SO}(3))$ the closure of the space of band-limited functions in $L^2(\mathrm{SO}(3))$ with respect to the corresponding Sobolev semi–norm. Since for any $s \geq \frac{1}{2}$ and $\varepsilon > 0$,

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

we have

$$F\ell_{\infty,s+\varepsilon}(\mathrm{SO}(3)) \subset F\ell_{2,s-\frac{1}{2}}(\mathrm{SO}(3)) \subset F\ell_{\infty,s-\frac{1}{2}}(\mathrm{SO}(3)). \tag{14}$$

The final goal of this section is to derive an estimate for the approximation error $\|f - f * \psi\|^2$ for functions $f \in F\ell_{2,s}(\mathrm{SO}(3))$ or $f \in F\ell_{\infty,s}(\mathrm{SO}(3))$ and a radially symmetric kernel function $\psi \in L^2(\mathrm{SO}(3))$. To this end we consider the integral means

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

where $R_{\xi,2\arccos t} \in \mathrm{SO}(3)$ is the rotation about $\xi \in \mathbb{S}^2$ with angle $\arccos t \in [0, 2\pi)$ 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 t from x. In analogy to the spherical Funck Hecke formula (cf. [10, Theorem 6]) we have the following result on integral means of the Wigner-D functions.

Lemma 2. *Let $t \in [0,1]$, $\ell \in \mathbb{N}$, and $k, k' = -\ell, \ldots, \ell$. 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 (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

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

and consequently

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

□

Next we proceed as in [19] and show that the family of integral means τ_t, $t \in (-1,1)$ defines an approximation process as $t \to 1$.

Lemma 3. *Let $f \in \mathcal{F}\ell_{2,2}(\mathrm{SO}(3))$ and $t \in [-1,1]$. Then*

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

Proof. For $f \in \mathcal{F}\ell_{2,2}(\mathrm{SO}(3))$ and $t \in [-1,1]$ we have

$$\begin{aligned}\|f - \tau_t f\|^2 &= \sum_{\ell=0}^{\infty} \sum_{k,k'=-\ell}^{\ell} \Big(1 - \frac{\mathcal{U}_{2\ell}(t)}{2\ell+1}\Big)^2 \left|\hat{f}(\ell,k,k')\right|^2 \\ &\le \Big(\sup_{l \in \mathbb{N}\setminus\{0\}} \frac{(2\ell+1)-\mathcal{U}_{2\ell}(t)}{\ell(\ell+1)(2\ell+1)}\Big)^2 \sum_{\ell=1}^{\infty} \sum_{k,k'=-\ell}^{\ell} \ell^2(\ell+1)^2 \left|\hat{f}(\ell,k,k')\right|^2 \\ &= \left(\frac{3-\mathcal{U}_2(t)}{6}\right)^2 \sum_{\ell=1}^{\infty} \sum_{k,k'=-\ell}^{\ell} \ell^2(\ell+1)^2 \left|\hat{f}(\ell,k,k')\right|^2 .\end{aligned}$$

□

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

Theorem 4. *Let $\psi \in L^2(\mathrm{SO}(3))$ be a zonal function with the Chebyshev expansion*

$$\psi(x) = 1 + \sum_{\ell=1}^{\infty} (2\ell+1)\hat{\psi}(\ell)\mathcal{U}_{2\ell}\Big(\cos\frac{\omega(x)}{2}\Big).$$

Then we have for any $s > 0$, $f \in F\ell_{2,s}(\mathrm{SO}(3))$ the inequality

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

and for any $s > \frac{1}{2}$, $f \in F\ell_{\infty,s}(\mathrm{SO}(3))$ the inequality

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

If $f \in F\ell_{2,2}(\mathrm{SO}(3))$ and $\psi \ge 0$. Then the above estimate simplifies to

$$\|f - f * \psi\|^2 \le \frac{1}{4}\left|1 - \hat{\psi}(1)\right|^2 \|f\|_{2,2}^2\,. \tag{15}$$

Proof. By (13) the convolution $f * \psi$ has the Fourier expansion

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

Hence, we obtain for the approximation error

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

and analogously

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

The proof for the more specific case that ψ is nonnegative we adapted from [19]. By (11) and Lemma 3 we have

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

□

3 Direct MISE Estimates

Let $f \in L^2(\mathrm{SO}(3))$ be a probability density function and let Ω be an arbitrary probability space. We consider a list of $N \in \mathbb{N}$, independently, identically distributed random variables $X_1, \ldots, X_N \colon \Omega \to \mathrm{SO}(3)$ which share the same distribution as the probability density function f, i.e., for any fixed $\omega \in \Omega$ the list $X_n(\omega) \in \mathrm{SO}(3)$, $n = 1, \ldots, N$ is a random sample of the probability distribution corresponding to f. Our objective is to estimate f from the random sample $X_n(\omega)$, $n = 1, \ldots, N$.

Let $\psi \colon \mathrm{SO}(3) \to \mathbb{R}$ be a zonal function with $\hat{\psi}(0) = 1$. Then the *kernel density estimator* of the probability density function f given the family of random variables X_n, $n = 1, \ldots, N$ is defined as

$$f_\psi^*(x) = N^{-1} \sum_{n=1}^{N} \psi(X_n^{-1} x).$$

Since the kernel density estimator f_ψ^* is a function of the random sample X_n, $n = 1, \ldots, N$ it is a random variable on Ω and one can ask for the mean integrated squared error

$$\mathrm{MISE} = \mathbb{E}\,\|f - f_\psi^*\|^2 = \mathbb{E}\int_{\mathrm{SO}(3)} |f(x) - f^*(x)|^2 \,\mathrm{d}\lambda(x)$$

between the probability density function f and the kernel density estimator f_ψ^*. The following well known decomposition result of the MISE into a bias term and a variance term (see e.g. [14]) actually holds true for the much more general setting of locally compact groups.

Lemma 5. *Let $\psi \in L^2(\mathrm{SO}(3))$ be a zonal function, $f \in L^2(\mathrm{SO}(3))$ a density function on $\mathrm{SO}(3)$, and $f_\psi^* \in L^2(\mathrm{SO}(3))$ its kernels density estimator. Then the mean squared integrated error allows for the decomposition*

$$\mathbb{E}\,\|f - f_\psi^*\|^2 = \mathbb{E}\,\|f - \mathbb{E}f_\psi^*\|^2 + \mathbb{E}\,\|f_\psi^* - \mathbb{E}f_\psi^*\|^2 \tag{16}$$

into a bias term

$$\mathbb{E}\,\|f - \mathbb{E}f_\psi^*\|^2 = \|f - f * \psi\|^2 \tag{17}$$

and a variance term

$$\mathbb{E}\,\|f_\psi^* - \mathbb{E}f_\psi^*\|^2 = N^{-1}\Big(\|\psi\|^2 - \|f * \psi\|^2\Big). \tag{18}$$

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

$$\lim_{N\to\infty} f_\psi^* = f * \psi. \tag{19}$$

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

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

Inserting the mean of the kernel density estimator f_ψ^* into the definition of the MISE we obtain

$$\begin{aligned}
\mathbb{E}\,\|f - f^*\|^2 &= \int_{\mathrm{SO}(3)} \mathbb{E}(f(x) - f^*(x))^2 \,\mathrm{d}\lambda(x) \\
&= \int_{\mathrm{SO}(3)} \mathbb{E}(f(x) - \mathbb{E}f^*(x))^2 + \mathbb{E}(\mathbb{E}f^*(x) - f^*(x))^2 \,\mathrm{d}\lambda(x) \\
&= \int_{\mathrm{SO}(3)} (f(x) - \mathbb{E}f^*(x))^2 + \mathbb{E}(\mathbb{E}f^*(x) - f^*(x))^2 \,\mathrm{d}\lambda(x).
\end{aligned}$$

The bias term on the left hand side of the sum elaborates to

$$\int_{\mathrm{SO}(3)} (f(x) - \mathbb{E}f^*(x))^2 \,\mathrm{d}\lambda(x) = \int_{\mathrm{SO}(3)} (f - f * \psi(x))^2 \,\mathrm{d}\lambda(x) = \|f - f * \psi\|^2 .$$

Using the independence of the random sample X_n, we calculate for the right hand side

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

Hence, the variation term on the right hand side of the sum yields

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

□

Next we follow the idea of [2] and consider the MISE in terms of the Fourier coefficients of the density function f and the Chebyshev coefficients of the kernel function ψ. Therefore, we abbreviate the Fourier coefficients of the function $f \in L^2(\mathrm{SO}(3))$,

$$f = 1 + \sum_{\ell=1}^{\infty}\sum_{k,k'=-\ell}^{\ell} \hat{f}(\ell,k,k')D_\ell^{k,k'},$$

by

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

and write the kernel function $\psi \in L^2(\mathrm{SO}(3))$ as its Chebyshev series

$$\psi(x) \sim 1 + \sum_{\ell=1}^{\infty}(2\ell+1)\hat{\psi}(\ell)\mathcal{U}_{2\ell}(\tfrac{\omega(x)}{2}).$$

Applying the Parseval identities (8), (12) and the convolution formula (13) we obtain the following representation of the mean integrated squared error

$$\mathrm{MISE} = \sum_{\ell=1}^{\infty}(2\ell+1)^2\hat{f}_\ell^2(1-\hat{\psi}(\ell))^2 + \frac{(2\ell+1)^2}{N}\hat{\psi}(\ell)^2\left(1-\hat{f}_\ell^2\right). \tag{20}$$

Since the MISE completely decomposes with respect to the rotational invariant subspaces in $L^2(\mathrm{SO}(3))$ we can find an optimal kernel by minimizing each summand with respect to $\hat{\psi}(\ell)$, separately.

Theorem 6. *Let $f \in L^2(\mathrm{SO}(3))$ be a probability density function and $N \in \mathbb{N}$ the size of a corresponding random sample. Then the MISE of the kernel density estimator $f^*_{\psi_{opt}}$ with respect to the zonal function*

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

is optimal compare to the MISE of the kernel density estimator f^*_ψ with respect to any other zonal function $\psi \in L^2(\mathrm{SO}(3))$, i.e.,

$$\mathrm{MISE}(f^*_{\psi_{opt}}) \leq \mathrm{MISE}(f^*_\psi),$$

and we have

$$\mathrm{MISE}(f^*_{\psi_{opt}}) = \sum_{\ell=0}^{\infty} (2\ell+1)^2 \frac{\hat{f}_\ell^2 (1 - \hat{f}_\ell^2)}{(N-1)\hat{f}_\ell^2 + 1}. \tag{21}$$

In particular, we have for $N \to \infty$

$$\lim_{N \to \infty} \mathrm{MISE}(f^*_{\psi(opt)}) = 0.$$

Proof. Since $f \geq 0$, we have by Lemma 1 for all $\ell \in \mathbb{N}$ the inequality $0 \leq \hat{f}_\ell^2 \leq 1$. Hence, each summand in (20) is a quadratic polynomial with respect to $\hat{\psi}(\ell)$ with minimum at

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

From $f \in L^2(\mathrm{SO}(3))$ we know by the Parseval identity (8)

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

and, hence,

$$\sum_{\ell=0}^{\infty} (2\ell+1) \frac{N \hat{f}_\ell^2}{(N-1)\hat{f}_\ell^2 + 1} \leq \infty.$$

Using the Parseval identity for zonal functions (12) we conclude that the Chebyshev coefficients $\hat{\psi}_{\mathrm{opt}}(\ell)$, $\ell = 0, 1, \ldots$ define a zonal function $\psi_{\mathrm{opt}} \in L^2(\mathrm{SO}(3))$ such that the MISE of the corresponding kernel density estimator is optimal. Direct calculation of $\mathrm{MISE}(f^*_{\psi_{\mathrm{opt}}})$ shows (21). $\square$

It should be noted that the optimal kernel is in general not applicable in real world applications since it requires the Fourier coefficients of the unknown density function f. However, the optimal kernel provides a useful tool to analyze best possible convergence rates of any kernel density estimator. The next Theorem shows, that the best possible convergence rate of the MISE for density functions $f \neq 1$ is $\mathcal{O}(N^{-1})$. Moreover, this convergence rate is attained if f is a bandlimited function.

Theorem 7. *Let $f \in L^2(\mathrm{SO}(3))$ be a not constant, probability density function. Then there is a constant $C > 0$ such that*

$$CN^{-1} \leq \mathrm{MISE}(f^*_{\psi_{opt}}).$$

Let f be furthermore bandlimited, i.e., there is a $L \in \mathbb{N}$ such that $\hat{f}_\ell = 0$ for all $\ell > L$. Then

$$\mathrm{MISE}(f^*_{\psi_{opt}}) \leq \frac{(L+1)(1+2L)(3+2L)}{3(1+\sqrt{N})^2}.$$

Proof. Again we apply Lemma 1 and observe that because of $f \in L^2(\mathrm{SO}(3))$ is nonnegative almost everywhere and not constant, there is a polynomial degree $\ell_0 \in \mathbb{N}\backslash\{0\}$ such that $0 < \hat{f}^2_{\ell_0} < 1$. Hence, there is a $C > 0$ such that

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

In order to prove the upper bound we consider the summands in (21) and look for their maximizer $\hat{f}_{\ell,\max}$. Direct computation shows

$$\hat{f}_{\ell,\max} = \frac{1}{\sqrt{1+\sqrt{N}}}.$$

Inserting this into (21) we arrive at

$$\mathrm{MISE}(f^*_{\psi_{\mathrm{opt}}}) = \sum_{\ell=0}^{L}(2\ell+1)^2 \frac{\hat{f}^2_{\ell,\max}(1-\hat{f}^2_{\ell,\max})}{(N-1)\hat{f}^2_{\ell,\max}+1} = \sum_{\ell=0}^{L} \frac{(2\ell+1)^2}{(1+\sqrt{N})^2} = \frac{(L+1)(1+2L)(3+2L)}{3(1+\sqrt{N})^2}.$$

□

In the case of not bandlimited functions f the decay rate of the MISE depends on the decay rate of the Fourier coefficients of f.

Theorem 8. *Let $s > \frac{1}{2}$ and $f \in \mathcal{F}\ell_{\infty,s}$. Then*

$$\mathrm{MISE}(f^*_{\psi_{opt}}) \leq \left(\frac{4^4}{3}\|f\|_{\infty,s}^{\frac{6}{2s+2}} + \frac{\pi}{2+2s}\right) N^{-\frac{2s-1}{2s+2}}.$$

Proof. First of all we note that because of $f \in \mathcal{F}\ell_{\infty,s}$ we have

$$\hat{f}^2_\ell \leq \ell^{-s}(\ell+1)^{-s}(2\ell+1)^{-2}\|f\|^2_{\infty,s} \leq \frac{1}{4}(\ell+1)^{-(2s+2)}\|f\|^2_{\infty,s}.$$

Let $L_0 = 4\left(\|f\|^2_{\infty,s}\sqrt{1+\sqrt{N}}\right)^{1/(2s+2)} - 1$ be the smallest polynomial degree ℓ such that

$$\frac{1}{4}(\ell+1)^{-(2s+2)}\|f\|^2_{\infty,s} \leq \frac{1}{\sqrt{1+\sqrt{N}}}$$

for all $\ell > L_0$. Then the MISE with respect to the optimal kernel function ψ_{opt} is

$$\begin{aligned}
\operatorname{MISE}(f^*_{\psi_{\text{opt}}}) &= \sum_{\ell=0}^{L_0}(2\ell+1)^2\frac{\hat{f}_\ell^2(1-\hat{f}_\ell^2)}{(N-1)\hat{f}_\ell^2+1} + \sum_{\ell=L_0+1}^{\infty}(2\ell+1)^2\frac{\hat{f}_\ell^2(1-\hat{f}_\ell^2)}{(N-1)\hat{f}_\ell^2+1}\\
&\leq \sum_{\ell=0}^{L_0}\frac{(2\ell+1)^2}{(1+\sqrt{N})^2} + \sum_{\ell=L_0+1}^{\infty}(2\ell+1)^2\frac{\frac{1}{4}(\ell+1)^{-(2s+2)}}{\frac{1}{4}(N-1)(\ell+1)^{-(2s+2)}+1}\\
&\leq \sum_{\ell=0}^{L_0}\frac{(2\ell+1)^2}{(1+\sqrt{N})^2} + \sum_{\ell=L_0}^{\infty}\frac{(\ell+1)^2}{\frac{1}{4}(N-1)+(\ell+1)^{2s+2}}\\
&\leq \frac{4}{3}\frac{(L_0+1)^3}{(1+\sqrt{N})^2} + \int_0^\infty \frac{\ell^2}{\frac{1}{4}(N-1)+\ell^{2s+2}}\\
&\leq \frac{4^4}{3}\,\|f\|_{\infty,s}^{\frac{6}{2s+2}}\,\frac{\sqrt{1+\sqrt{N}}^{3/(2s+2)}}{(1+\sqrt{N})^2} + \frac{\pi}{\sin(\frac{3\pi}{2+2s})(2+2s)}N^{-\frac{2s-1}{2s+2}}\\
&\leq \left(\frac{4^4}{3}\,\|f\|_{\infty,s}^{\frac{6}{2s+2}} + \frac{\pi}{2+2s}\right)N^{-\frac{2s-1}{2s+2}}.
\end{aligned}$$

□

4 AMISE Estimates

As we have already mentioned in the previous section the MISE, and in particular, the MISE optimal kernel function are not feasible for practical applications of kernel density estimation since they require to know the Fourier coefficients of the unknown density function f. Hence, one is interested in asymptotically strict upper bounds of the MISE that depend only on some Sobolev norm of the unknown density function f that can be estimated from the data. Those upper bounds are commonly called *asymptotic mean integrated squared error* (AMISE) (cf. [17]). In the Euclidean setting, i.e., $f\colon \mathbb{R}^d \to \mathbb{R}$ they are of the form

$$\operatorname{AMISE}(f^*_\psi) = \mu_2(\psi)^2\,\|\triangle f\|^2 + N^{-1}\,\|\psi\|^2 \tag{22}$$

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

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

denotes the second moment of the kernel function ψ.

As a generalization of the AMISE for the Euclidean setting we define the *asymptotic mean integrated squared error* (AMISE) on the rotation group for a density function $f \in F\ell_{2,2}(\mathrm{SO}(3))$ and a nonnegative zonal function $\psi \in L^2(\mathrm{SO}(3))$ as

$$\operatorname{AMISE}(f^*_\psi) = \mu_2(\psi)^2\,\|\tilde{\triangle} f\|^2 + N^{-1}\,\|\psi\|^2, \tag{23}$$

with

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

and with $\tilde{\triangle}$ denoting the Laplace–Beltrami operator on $\mathrm{SO}(3)$. According to (10) the weight $\sqrt{1-t^2}$ is the canonical weight when working with zonal functions on the rotation group $\mathrm{SO}(3)$. Next we show that the AMISE is indeed an upper bound for the MISE.

Theorem 9. *Let $f \in F\ell_{2,2}(\mathrm{SO}(3))$ and let $\psi \in L^2(\mathrm{SO}(3))$ be a nonnegative, zonal kernel function with Chebyshev coefficients $\hat{\psi}(\ell)$, $\ell \in \mathbb{N}$. Then we have*

$$\mathrm{MISE}(f_\psi^*) \leq \frac{1}{4}\left|1-\hat{\psi}(1)\right|^2 \|f\|_{2,2}^2 + N^{-1}\|\psi\|^2 = \mathrm{AMISE}(f_\psi^*).$$

Proof. By Lemma 5 we have

$$\mathrm{MISE}(f_\psi^*) = \mathbb{E}\,\|f-f_\psi^*\|^2 = \|f-f*\psi\|^2 + N^{-1}\Big(\|\psi\|^2 - \|f*\psi\|^2\Big) \tag{24}$$

and the left hand side inequality becomes a direct consequence of Theorem 4.

For the right hand side equality we observe that for any nonnegative kernel function ψ with $\hat{\psi}(0) = 1$ we have

$$1-\hat{\psi}(1) = \frac{2}{\pi}\int_{-1}^{1}\psi(t)(1-\frac{1}{3}\mathcal{U}_{2\ell}(t))\sqrt{1-t^2}\,\mathrm{d}t = \frac{8}{3\pi}\int_{-1}^{1}\psi(t)(1-t^2)\sqrt{1-t^2}\,\mathrm{d}t$$

and, furthermore, that

$$\|f\|_{2,2}^2 = \|\tilde{\triangle} f\|^2.$$

□

Next we generalize the AMISE to nonnegative kernel functions and other smoothness classes of f. Applying again Theorem 4 to (24) we end up with the following estimates.

Theorem 10. *Let $s > 0$, $f \in F\ell_{2,s}(\mathrm{SO}(3))$ and let $\psi \in L^2(\mathrm{SO}(3))$ be a zonal kernel function with Chebyshev coefficients $\hat{\psi}(\ell)$, $\ell \in \mathbb{N}$. Then we have*

$$\mathrm{MISE}(f_\psi^*) \leq \sup_{\ell\in\mathbb{N}\setminus\{0\}} \frac{|1-\hat{\psi}(\ell)|^2}{\ell^s(\ell+1)^s}\|f\|_{2,s}^2 + N^{-1}\|\psi\|^2 =: \mathrm{AMISE}_{2,s}(f_\psi^*).$$

For $s > \frac{1}{2}$ and $f \in F\ell_{\infty,s}(\mathrm{SO}(3))$ we have

$$\mathrm{MISE}(f_\psi^*) \leq \sum_{\ell\in\mathbb{N}\setminus\{0\}} \frac{|1-\hat{\psi}(\ell)|^2}{\ell^s(\ell+1)^s}\|f\|_{\infty,s}^2 + N^{-1}\|\psi\|^2 =: \mathrm{AMISE}_{\infty,s}(f_\psi^*).$$

For these generalized AMISE we now prove the existence of optimal kernel functions and give asymptotic lower and upper bounds. Let us start with the $\mathrm{AMISE}_{2,s}$.

Theorem 11. *Let $s > 0$, $f \in \mathcal{F}\ell_{2,s}(\mathrm{SO}(3))$ a probability density function, and $N \in \mathbb{N}$ the size of a corresponding random sample. Then the $\mathrm{AMISE}_{2,s}$ of the kernel density estimator $f^*_{\psi_L}$ with respect to the Jackson type kernel $\phi_L \colon \mathrm{SO}(3) \to \mathbb{R}$, $L \in \mathbb{R}$, $L > 0$,*

$$\phi_L = 1 + \sum_{\ell=1}^{\lfloor L \rfloor} (2\ell+1) \left(1 - \frac{\ell^{s/2}(\ell+1)^{s/2}}{L^{s/2}(L+1)^{s/2}}\right) \mathcal{U}_{2\ell} \tag{25}$$

is optimal compared to the $\mathrm{AMISE}_{2,s}$ of the kernel density estimator f^*_ψ with respect to any other zonal function ψ, i.e.,

$$\mathrm{AMISE}_{2,s}(f^*_\psi) \geq \min_{L \in \mathbb{R}_+} \mathrm{AMISE}_{2,s}(f^*_{\psi_L}).$$

The optimal bandwidth L_* of the Jackson type kernel is approximately

$$L_*^{2s+3} = \frac{2s}{3} C^{-1} N \left\| f \right\|_{2,s}^2, \tag{26}$$

which satisfies

$$\begin{aligned} A \left\| f \right\|_s^{\frac{6}{2s+3}} N^{-\frac{2s}{2s+3}} + \mathcal{O}(N^{-\frac{2s+1}{2s+3}}) &= \min_{L \in \mathbb{R}_+} \mathrm{AMISE}_{2,s}(f^*_{\phi_L}) \\ &\leq \mathrm{AMISE}_{2,s}(f^*_{\phi_{L_*}}) = A \left\| f \right\|_s^{\frac{6}{2s+3}} N^{-\frac{2s}{2s+3}} + \mathcal{O}(N^{-\frac{2s+1}{2s+3}}) \end{aligned}$$

with constants $C = \left(\frac{4}{3} - \frac{8}{s+3} + \frac{4}{2s+3}\right)$ and $A = \left((\frac{2s}{3})^{-\frac{2s}{2s+3}} + (\frac{2s}{3})^{\frac{3}{2s+3}}\right) C^{\frac{2s}{2s+3}}$.

Proof. First of all we may assume that for an optimal kernel function ψ the Chebyshev coefficients satisfy $0 \leq \hat{\psi}(\ell) \leq 1$ for all $l \in \mathbb{N}$. Let $L \in \mathbb{R}$, $L > 0$ such that

$$L^{-s/2}(L+1)^{-s/2} = \sup_{\ell \in \mathbb{N} \setminus \{0\}} \frac{|1 - \hat{\psi}(\ell)|}{\ell^{s/2}(\ell+1)^{s/2}}.$$

Then, the zonal function $\phi_L \in L^2(\mathrm{SO}(3))$ given by

$$\hat{\phi}_L(\ell) = \max\left\{0, 1 - \frac{\ell^{s/2}(\ell+1)^{s/2}}{L^{s/2}(L+1)^{s/2}}\right\}$$

satisfies $\left\|\psi\right\|_2 \geq \left\|\phi_L\right\|_2$ and

$$\sup_{\ell \in \mathbb{N} \setminus \{0\}} \frac{|1 - \hat{\psi}(\ell)|}{\ell^{s/2}(\ell+1)^{s/2}} = L^{-s/2}(L+1)^{-s/2} = \sup_{\ell \in \mathbb{N} \setminus \{0\}} \frac{|1 - \hat{\phi}_L(\ell)|}{\ell^{s/2}(\ell+1)^{s/2}}.$$

Consequently,

$$\mathrm{AMISE}_{2,s}(f^*_\psi) \geq \mathrm{AMISE}_{2,s}(f^*_{\phi_L}).$$

Next we are going to find estimates for the AMISE for the Jackson kernel ϕ_L. In order to compute the L^2 - norm of the Jackson kernel we need that there is a nonnegative number $\alpha \in \mathbb{R}$ such that

$$\sum_{\ell=0}^{\lfloor L\rfloor} \frac{\ell^{s+1}(\ell+1)^{s+1}}{L^s(L+1)^s} \le \sum_{\ell=0}^{\lfloor L\rfloor} \frac{(\ell+1)^{2s+2}}{(L+1)^{2s}} \le \int_0^{\lfloor L\rfloor+1} \frac{(\ell+1)^{2s+2}}{(L+1)^{2s}}\,\mathrm{d}\ell \le \frac{1}{2s+3}(L+\alpha)^3$$

and

$$\sum_{\ell=0}^{\lfloor L\rfloor} \frac{\ell^{s+1}(\ell+1)^{s+1}}{L^s(L+1)^s} \ge \sum_{\ell=0}^{\lfloor L\rfloor} \frac{\ell^{2s+2}}{L^{2s}} \ge \int_0^{\lfloor L\rfloor} \frac{\ell^{2s+2}}{L^{2s}}\,\mathrm{d}\ell \ge \frac{1}{2s+3}(L-\alpha)^3.$$

As a result we obtain that there is a constant $\beta > 0$ such that for all integer $L \in \mathbb{N}$,

$$\begin{aligned}
\|\phi_L\|^2 &= \sum_{\ell=0}^{\lfloor L\rfloor}(2\ell+1)^2\left(1-\frac{\ell^{s/2}(\ell+1)^{s/2}}{L^{s/2}(L+1)^{s/2}}\right)^2 \\
&\ge 4\sum_{\ell=0}^{\lfloor L\rfloor}\left(\ell(\ell+1) - 2\frac{\ell^{s/2+1}(\ell+1)^{s/2+1}}{L^{s/2}(L+1)^{s/2}} + \frac{\ell^{s+1}(\ell+1)^{s+1}}{L^s(L+1)^s}\right) \\
&\ge 4\left(\frac{1}{3}(L-\alpha)^3 - \frac{2}{s+3}(L+\alpha)^3 + \frac{1}{2s+1}(L-\alpha)^3\right) \\
&\ge 4\left(\frac{1}{3} - \frac{2}{s+3} + \frac{1}{2s+3}\right)(L-\beta)^3
\end{aligned}$$

and, analogously,

$$\|\phi_L\|^2 \le 4\left(\frac{1}{3} - \frac{2}{s+3} + \frac{1}{2s+3}\right)(L+\beta)^3.$$

Hence, we have for the AMISE

$$(L+1)^{-2s}\|f\|_s^2 + C(L-\beta)^3N^{-1} \le \mathrm{AMISE}_{2,s}(f^*_{\phi_L}) \le L^{-2s}\|f\|_s^2 + C(L+\beta)^3N^{-1},$$

with $C = 4\left(\frac{1}{3} - \frac{2}{s+3} + \frac{1}{2s+3}\right)$. An approximate minimizer of the left and the right hand side is

$$L_*^{2s+3} = \frac{2s}{3}C^{-1}N\|f\|^2.$$

Moreover, the true minimizer of the left and ride hand sides are contained in the interval $[L_* - \beta, L_* + \beta]$. For the corresponding AMISE we obtain

$$\begin{aligned}
A\|f\|_s^{\frac{6}{2s+3}} N^{-\frac{2s}{2s+3}} + \mathcal{O}(N^{-\frac{2s+1}{2s+3}}) &\le \min_{L\in\mathbb{R}_+} \mathrm{AMISE}_{2,s}(f^*_{\phi_L}) \\
&\le \mathrm{AMISE}_{2,s}(f^*_{\phi_{L_*}}) = A\|f\|_s^{\frac{6}{2s+3}} N^{-\frac{2s}{2s+3}} + \mathcal{O}(N^{-\frac{2s+1}{2s+3}})
\end{aligned}$$

with constant $A = \left((\frac{2s}{3})^{-\frac{2s}{2s+3}} + (\frac{2s}{3})^{\frac{3}{2s+3}}\right) C^{\frac{2s}{2s+3}}$. $\square$

Next we consider the AMISE$_{\infty,s}$.

Theorem 12. *Let $s > \frac{1}{2}$, $f \in \mathcal{F}\ell_{\infty,s}$ and N the size of the random sample. Then the* AMISE$_{\infty,s}$ *optimal kernel function is given by*

$$\psi_{opt}(x) = \sum_{\ell=0}^{\infty}(2\ell+1)\frac{N\|f\|_{\infty,s}^2}{N\|f\|_{\infty,s}^2 + (2\ell+1)^2\ell^s(\ell+1)^s}\mathcal{U}_{2\ell}\left(\cos\frac{\omega(x)}{2}\right), \tag{27}$$

i.e., for any kernel function ϕ we have

$$\mathrm{AMISE}_{\infty,s}(f_\phi^*) \geq \mathrm{AMISE}_{\infty,s}(f_{\psi_{opt}}^*).$$

The corresponding AMISE is asymptotically

$$\mathrm{AMISE}_{\infty,s}(f_{\psi_{opt}}^*) = \frac{2^{\frac{2-s}{s+1}}\pi}{(s+1)\sin\frac{3\pi}{2s+2}}\|f\|_{\infty,s}^{\frac{2}{2s+2}} N^{-\frac{2s-1}{2s+2}} + \mathcal{O}(N^{-\frac{2s}{2s+2}}). \tag{28}$$

Proof. Using once again the Parsevals identities (8) and (12) we obtain for the AMISE$_{\infty,s}$ the Fourier space representation

$$\mathrm{AMISE}_{\infty,s}(f_\psi^*) = \sum_{\ell=1}^{\infty}\frac{(1-\hat\psi(\ell))^2}{\ell^s(\ell+1)^s}\|f\|_{\infty,s}^2 + (2\ell+1)^2\hat\psi(\ell)^2N^{-1}.$$

Setting the partial derivative for each Chebyshev coefficient $\hat\psi(\ell)$, $\ell \in \mathbb{N}$ to zero we obtain, that the optimal kernel function ψ_{opt} is defined by the Chebyshev coefficients

$$\hat\psi_{\mathrm{opt}}(\ell) = \frac{N\|f\|_{\infty,s}^2}{N\|f\|_{\infty,s}^2 + (2\ell+1)^2\ell^s(\ell+1)^s},\ \ell \in \mathbb{N}.$$

Given $s > \frac{1}{2}$ the corresponding Chebyshev series converges in $L^2(\mathrm{SO}(3))$.

For the optimal AMISE we obtain the upper bound

$$\begin{aligned}
\mathrm{AMISE}_{\infty,s}(f_{\psi_{\mathrm{opt}}}^*) &= \sum_{\ell=0}^{\infty}\frac{(1-\hat\psi_{\mathrm{opt}}(\ell))^2}{\ell^s(\ell+1)^s}\|f\|_{\infty,s}^2 + (2\ell+1)^2\hat\psi_{\mathrm{opt}}(\ell)^2N^{-1}\\
&= \sum_{\ell=0}^{\infty}\frac{(2\ell+1)^4\ell^s(\ell+1)^s\|f\|_{\infty,s}^2 + (2\ell+1)^2N\|f\|_{\infty,s}^4}{\left(N\|f\|_{\infty,s}^2 + (2\ell+1)^2\ell^s(\ell+1)^s\right)^2}\\
&\leq \int_0^\infty \frac{16(\ell+1)^{2s+4}\|f\|_{\infty,s}^2 + 4(\ell+1)^2N\|f\|_{\infty,s}^4}{\left(N\|f\|_{\infty,s}^2 + 4\ell^{2s+2}\right)^2}\,\mathrm{d}l + \mathcal{O}(N^{-1})\\
&= \frac{2^{\frac{2-s}{s+1}}\pi}{(s+1)\sin\frac{3\pi}{2s+2}}\|f\|_{\infty,s}^{\frac{6}{2s+2}} N^{-\frac{2s-1}{2s+2}} + \mathcal{O}(N^{-\frac{2s}{2s+2}})
\end{aligned}$$

and analogously the lower bound. □

The results of Theorem 11 and Theorem 12 compare well with the inclusions (14) in the sense that for $f \in \mathcal{F}\ell_{\infty,s+\varepsilon}(\mathrm{SO}(3)) \subset \mathcal{F}\ell_{2,s-\frac{1}{2}}(\mathrm{SO}(3))$ we have for both generalized asymptotic mean squared errors $\mathrm{AMISE}_{\infty,s}(f^*_{\psi_{\mathrm{opt}}}) = \mathcal{O}(N^{-\frac{2s-1}{2s+2}})$ and $\tilde{\mathrm{AMISE}}_{2,s-\frac{1}{2}}(f^*_{\psi_{\mathrm{opt}}}) = \mathcal{O}(N^{-\frac{2s-1}{2s+2}})$, each with respect to the corresponding optimal kernel function.

Next we review some important zonal functions on the rotation group according to their feasibility for kernel density estimation. In particular, we derive formulas for the optimal kernel parameter depending on the smoothness of the function f and the number of random samples and investigate the behavior of the asymptotic mean integrated error as the number of random samples N tends to infinity. Some more zonal functions which might be useful for kernel density estimation on the rotation group can be found in [3]. Because of its practical relevance we restrict ourselves to two times differentiable functions, i.e., to the case $s = 2$.

4.1 The Dirichlet Kernel

Let us start with the Dirichlet kernel which was already studied in [2] for the more general setting of Riemannian manifolds. On the rotation group $\mathrm{SO}(3)$ the Dirichlet kernel $\psi_L \in L^2(\mathrm{SO}(3))$ is defined by its Chebyshev series

$$\psi_L(t) = \sum_{\ell=0}^{L}(2\ell+1)\mathcal{U}_{2\ell}(t).$$

The $\mathrm{AMISE}_{2,2}$ of the kernel density estimator $f^*_{\psi_L}$ corresponding to the Dirichlet kernel computes to

$$\begin{aligned}\mathrm{AMISE}_{2,2}(f^*_{\psi_L}) &= (L+1)^{-2}(L+2)^{-2}\left\|f\right\|_{2,2}^2 + \sum_{\ell=0}^{L}(2\ell+1)^2 N^{-1}\\ &= (L+1)^{-2}(L+2)^{-2}\left\|f\right\|_{2,2}^2 + \frac{1}{3}(L+1)(2L+1)(2L+3)N^{-1}\\ &\approx L^{-4}\left\|f\right\|_{2,2}^2 + \frac{4}{3}L^3 N^{-1}.\end{aligned}$$

From this we obtain an approximately AMISE optimal parameter L_*

$$L_*^7 \approx N\left\|f\right\|_{2,2}^2 \tag{29}$$

which leads to the AMISE

$$\mathrm{AMISE}_{2,2}(f^*_{\psi_{L_*}}) \approx \frac{7}{3}\left\|f\right\|_{2,2}^{6/7} N^{-4/7}.$$

Comparing the convergence rate of the kernel density estimator using the Dirichlet kernel with the convergence rates derived in Theorem 11 we conclude that the Dirichlet kernel has an asymptotically optimal convergence rate. However, the constant involved is approximately two times larger compared to the constant for the optimal Jackson type kernel.

4.2 The de la Vallee Poussin Kernel

The de la Vallee Poussin kernel ψ_κ is a nonnegative zonal function depending on a parameter $\kappa \in \mathbb{N}\backslash\{0\}$ that has the finite Chebyshev expansion

$$\psi_\kappa(t) = \frac{(\kappa+1)2^{2\kappa-1}}{\binom{2\kappa-1}{\kappa}} t^{2\kappa} = \binom{2\kappa+1}{\kappa}^{-1} \sum_{\ell=0}^{\kappa} (2\ell+1)\binom{2\kappa+1}{\kappa-\ell} \mathcal{U}_{2\ell}(t).$$

The L^2–norm of the de la Vallee Poussin kernel computes to

$$\|\psi_\kappa\|^2 = \frac{2}{\pi}\int_{-1}^{1} \psi_\kappa(t)^2 \sqrt{1-t^2}\,\mathrm{d}t = \frac{2}{\pi}\frac{(\kappa+1)2^{2\kappa-1}}{\binom{2\kappa-1}{\kappa}} \int_{-1}^{1} t^{4\kappa}\sqrt{1-t^2}\,\mathrm{d}t$$
$$= \sqrt{\pi}\frac{\Gamma(\kappa+2)^2\Gamma(2\kappa+\frac{1}{2})}{\Gamma(\kappa+\frac{1}{2})^2\Gamma(2\kappa+2)}.$$

Since we have asymptotically by the Stirling formula

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

we obtain for the L^2–norm of the de la Vallee Poussin kernel the approximation

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

Since

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

the AMISE$(f^*(\psi_\kappa))$ is approximately

$$\text{AMISE}(f^*_{\psi_\kappa}) \approx (\kappa+2)^{-2}\|f\|_{2,2}^2 + \frac{\pi}{8}\kappa^{3/2}N^{-1}.$$

An approximation of the optimal parameter is

$$\kappa_*^7 \approx \frac{2^7}{9\pi}\|f\|_{2,2}^4 N^2. \tag{30}$$

The corresponding optimal AMISE is

$$\text{AMISE}(f^*_{\psi_{\kappa_*}}) \approx 3.8\,\|f\|_{2,2}^{6/7} N^{-4/7}.$$

As we have seen in Theorem 11 this is the optimal decay rate of the AMISE error we can expect for $s = 2$.

4.3 The Abel Poisson Kernel

The Abel Poisson kernel $\psi_\kappa \in L^2(\mathrm{SO}(3))$ is a nonnegative zonal function depending on a parameter $\kappa \in (0,1)$ which is defined by its Chebyshev series

$$\psi_\kappa(t) = \sum_{\ell=0}^{\infty}(2\ell+1)\kappa^{2\ell}\mathcal{U}_{2\ell}(t),$$

and has L^2-norm

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

Since the Abel Poisson kernel is nonnegative we compute for $f \in \mathcal{F}\ell_{2,2}(\mathrm{SO}(3))$ the $\mathrm{AMISE}(f^*(\psi_\kappa))$ approximately for $\kappa \to 1$,

$$\begin{aligned}\mathrm{AMISE}(f^*_{\psi_\kappa}) &= \frac{(1-\kappa^2)^2}{4}\|f\|^2_{2,2} + \frac{1+6\kappa^4+\kappa^8}{(1-\kappa^4)^3}N^{-1} \\ &\approx (1-\kappa)^2\|f\|^2_{2,2} + \frac{1}{8}(1-\kappa)^{-3}N^{-1}.\end{aligned}$$

The corresponding optimal parameter κ_* is approximately

$$(1-\kappa_*)^5 \approx 48N^{-1}\|f\|^{-2}_{2,2} \tag{31}$$

which yields the AMISE

$$\mathrm{AMISE}(f^*_{\psi_{\kappa_*}}) \approx 0.37\,\|f\|^{6/5}_{2,2}\,N^{-2/5}.$$

Comparing the convergence rate of the kernel density estimator using the Abel Poisson kernel with the convergence rates derived in Theorem 11 we conclude that the Abel Poisson kernel is not well suited for kernel density estimation, except for small values of N.

5 Numerical Experiments

In this section we are going to verify our theoretical findings by numerical experiments. The general concept of our numerical experiments 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 $\|f - f^*_\psi\|^2$.

6. Compute an estimate of the MISE by repeating M times the steps 2 to 4 and taking the mean value of the integrated squared errors.

As the test density function f we chose a linear combination of de la Vallee Poussin kernels ψ_L translated to two arbitrarily chosen locations of the rotation group,

$$f(x) = 0.2 + 0.7\psi_{90}(R_{e_2,30^\circ}x) + 0.1\psi_{350}(R_{e_1,80^\circ}x). \tag{32}$$

The advantage of choosing this specific test function is that drawing a random sample and computing the kernel density estimator becomes numerical feasible as we explain in the subsequent sections. Our Numerical experiments showed, that increasing the number or changing type of the kernel functions ψ_L has only minor influence to our numerical results.

5.1 Drawing a random sample from a distribution on SO(3)

Let $\psi > 0$ be a strictly positive density function on $[0,1]$. Then the corresponding cumulative distribution function

$$\Psi(y) = \int_0^y \psi(x)\,\mathrm{d}x$$

defines a diffeomorphism

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

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

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

Hence, the distribution of ψ under Ψ 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 Ψ^{-1} to it.

Let us now consider a zonal function ψ on the rotation group $\mathrm{SO}(3)$. Then ψ depends only on the rotational angle ω of a rotation $R_{\xi.\omega}$ about an arbitrary axis ξ and we will write $\phi(\omega) = \psi(R_{\xi,\omega})$. Using the parameterization 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

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

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

5.2 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 (1) corresponding to N random samples at M arbitrarily chosen nodes with the numerical complexity $\mathcal{O}(N + M)$ is described in [4]. However, as we are only interested in the L^2–error $\|f_\psi^* - f\|$ we rest at computing the Fourier coefficients of f and f_ψ^* up to a sufficient large order and applying Parsevals identity.

Let $\psi \in L^2(\mathrm{SO}(3))$ be a zonal function with finite Chebyshev expansion

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

and let $X_n \in \mathrm{SO}(3)$, $n = 1, \ldots, N$ be a random sample. Then the kernel density estimator has the representation

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

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

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

which is essentially an adjoined Fourier transform on the rotation group $\mathrm{SO}(3)$. Algorithms for the fast Fourier transform on the rotation group at arbitrary nodes as well as for its adjoined transform has been described in [12] and are available as part of the NFFT library [5].

5.3 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 we considered different kernel functions. On the one hand we applied the MISE optimal kernel function as defined in Theorem 6 and compared the numerical estimated MISE with the theoretical expression found in (21). On the other hand we used the formulae (26), (27), (29), (30), and (31) for the optimal parameters of the $\mathrm{AMISE}_{2,2}$ optimal kernel, the $\mathrm{AMISE}_{\infty,2.5}$ optimal kernel, the Dirichlet kernel, the Abel Poisson kernel, and the de la Valleé Poussin kernel and computed the MISE for the kernel density estimator with respect to these kernel functions. Figure 1 shows the Chebyshev coefficients of the kernel functions mentioned above with optimal kernel parameter for the specific choice of $N = 10^4$ random samples. In Figure 2 the relative MISE

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

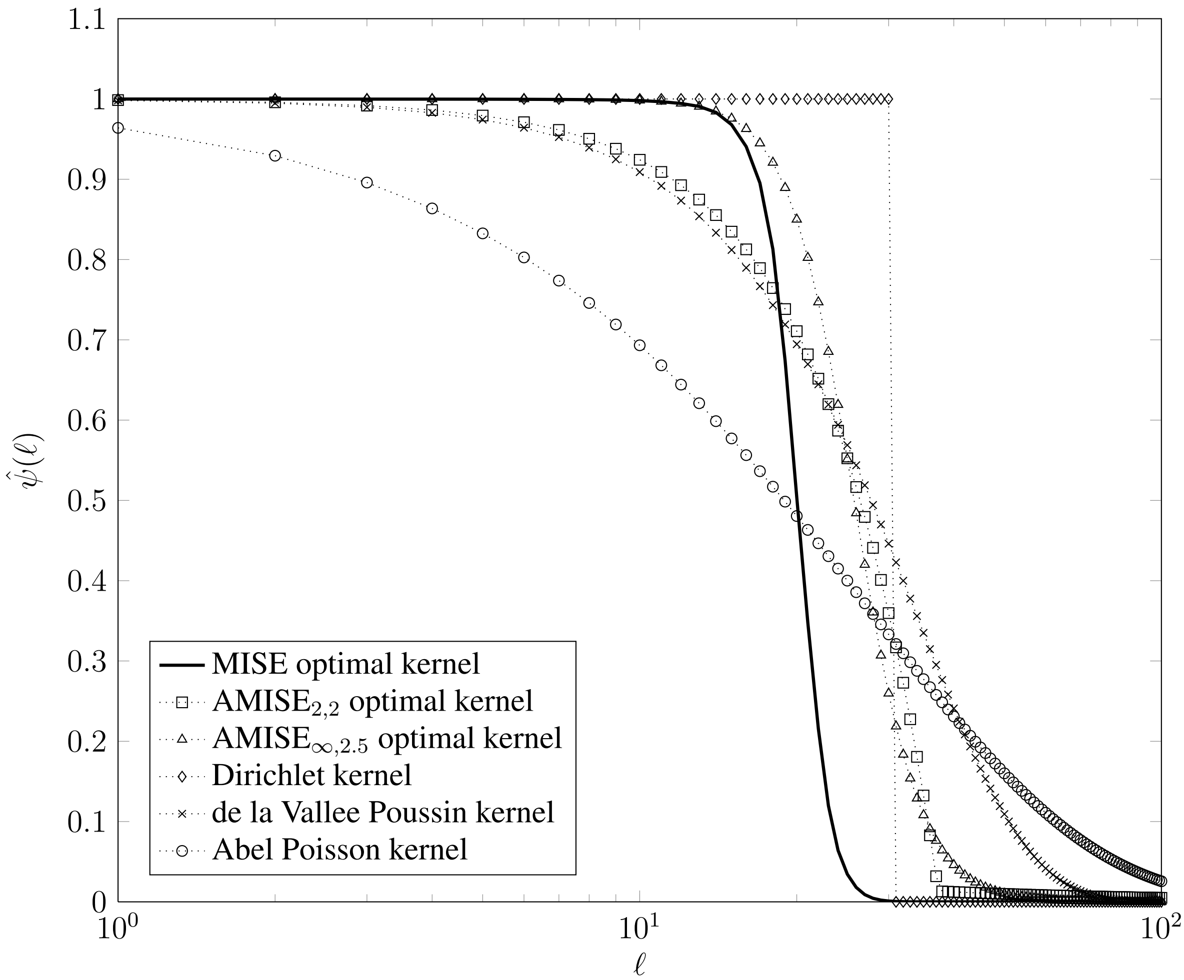

Figure 1: This plot shows the Chebyshev coefficients of the kernel functions investigated throughout the numerical experiments. The kernel parameter has been chosen to be optimal with respect to the test function (32) and $N = 10^4$.

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 AMISE$_{2,2}$ optimal, the AMISE$_{\infty,2.5}$ 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 AMISE$_{\infty,2.5}$ 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.

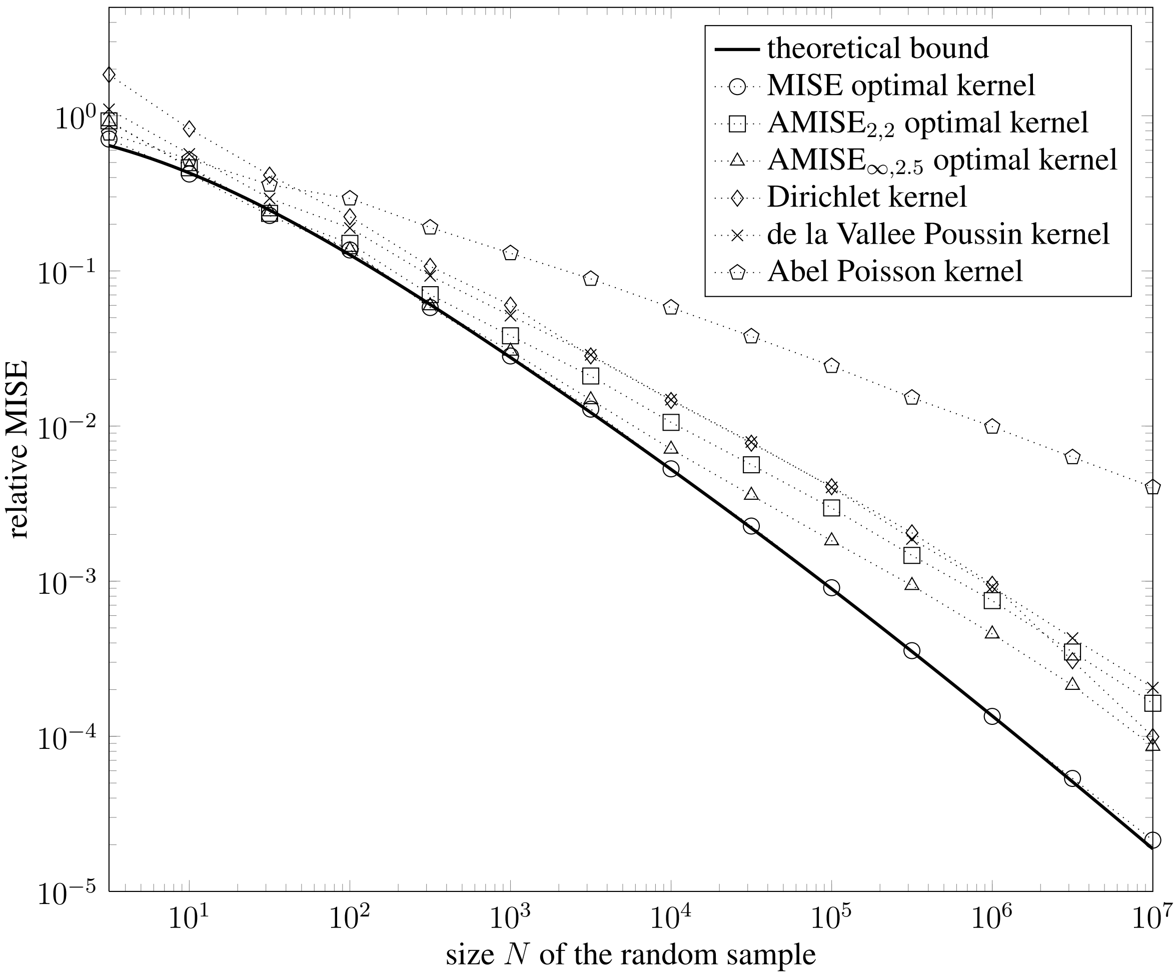

Figure 2: This plot shows the MISE as a function of the number of random samples N and the kernel used for kernel density estimation. The theoretical bound as well as the MISE optimal kernel where computed according to Theorem 6. The parameters for the other kernel functions where chosen AMISE optimal as specified in the formulae (26), (27), (29), (30), and (31).

References

[1] Y. Chikuse. *Statistics on Special Manifolds*. Lecture Notes in Statistics. Springer, New York, 2002.

[2] H. Hendriks. Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansion. *Ann. Statist.*, 18:832–849, 1990.

[3] R. Hielscher. The Radon Transform on the Rotation Group – Inversion and Application to Texture Analysis. Dissertation, Department of Mathematics, Technical University Bergakademie Freiberg, 2007. URL `https://fridolin.tu-freiberg.de/archiv/pdf/MathematikHielscherRalf361401.pdf`.

[4] R. Hielscher, J. Prestin, and A. Vollrath. Fast summation of functions on SO(3). *Math. Geosci.*, accepted, 2010.

[5] J. Keiner, S. Kunis, and D. Potts. Using NFFT3 - a software library for various nonequispaced fast Fourier transforms. *ACM Trans. Math. Software*, 36:Article 19, 1 – 30, 2009.

[6] P. T. Kim. Deconvolution density estimation on SO(N). *Ann. Statist.*, 26:1083–1102, 1998.

[7] P. J. Kostelec and D. N. Rockmore. FFTs on the rotation group. *J. Fourier Anal. Appl.*, 14: 145 – 179, 2008.

[8] K. Kunze, S. I. Wright, B. L. Adams, and D. J. Dingley. Advances in automatic EBSP single orientation measurements. *Textures and Microstructures*, 20:41–54, 1993.

[9] K. Mardia. *Statistics of Directional Data*. Academic Press, London, 1972.

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

[11] B. Pelletier. Kernel density estimation on Riemannian manifolds. *Statist. Probab. Lett.*, 73: 297–304, 2005.

[12] D. Potts, J. Prestin, and A. Vollrath. A fast algorithm for nonequispaced Fourier transforms on the rotation group. *Numer. Algorithms*, 52:355 – 384, 2009.

[13] K. G. van den Boogaart. *Statistics for individual crystallographic orientation measurements*. PhD thesis, TU Freiberg, 2001.

[14] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, Cambridge, 1998.

[15] D. Varshalovich, A. Moskalev, and V. Khersonski. *Quantum Theory of Angular Momentum*. World Scientific Publishing, Singapore, 1988.

[16] N. Vilenkin. *Special Functions and the Theory of Group Representations*. Amer. Math. Soc., Providence, RI, USA, 1968.

[17] M. P. Wand and M. C. Jones. *Kernel Smoothing*. Chapman & Hall/CRC, 1995.

[18] G. Watson. *Statistics on Spheres*. Wiley, New York, 1983.

[19] M. Wehrens. Legendre - transformationsmethoden und approximation von funktionen auf der einheitskugel im R^3. Dissertation, Institut für Mathematik, Technische Hochschule Aachen, 1980.