SFB 823

Optimal designs for regression with spherical data

Holger Dette, Maria Konstantinou, Kirsten Schorning, Josua Gösmann

Discussion Paper

Nr. 20/2017

SFB 823

Optimal designs for regression with spherical data

Holger Dette, Maria Konstantinou, Kirsten Schorning, Josua Gösmann

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum, Germany

Abstract

In this paper optimal designs for regression problems with spherical predictors of arbitrary dimension are considered. Our work is motivated by applications in material sciences, where crystallographic textures such as the missorientation distribution or the grain boundary distribution (depending on a four dimensional spherical predictor) are represented by series of hyperspherical harmonics, which are estimated from experimental or simulated data.

For this type of estimation problems we explicitly determine optimal designs with respect to Kiefers Φ_p-criteria and a class of orthogonally invariant information criteria recently introduced in the literature. In particular, we show that the uniform distribution on the m-dimensional sphere is optimal and construct discrete and implementable designs with the same information matrices as the continuous optimal designs. Finally, we illustrate the advantages of the new designs for series estimation by hyperspherical harmonics, which are symmetric with respect to the first and second crystallographic point group.

1 Introduction

Regression problems with a predictor of spherical nature arise in various fields such as geology, crystallography, astronomy (cosmic microwave background radiation data), the calibration of electromagnetic motion-racking systems or the representation of spherical viruses [see Chapman et al. (1995), Zheng et al. (1995), Chang et al. (2000), Schaeben and van den Boogaart (2003), Genovese et al. (2004), Shin et al. (2007) among many others] and their parametric and nonparametric estimation has found considerable attention in the literature.

Several methods for estimating a spherical regression function nonparameterically have been proposed in the literature. Di Marzio et al. (2009, 2014) investigate kernel type methods, while spherical splines have been considered by Wahba (1981) and Alfed et al. (1996). A frequently used technique is that of series estimators based on spherical harmonics [see Abrial et al. (2008) for example], which - roughly speaking - generalise estimators of a regression function on the line based on Fourier series to data on the sphere. Alternative series estimators have been proposed by Narcowich et al. (2006), Baldi et al. (2009) and Monnier (2011) who suggest to use spherical wavelets (needlets) in situations where better localisation properties are required. Most authors consider the 2-dimensional sphere $\mathbb{S}_3$ in $\mathbb{R}^3$ as they are interested in the development of statistical methodology for concrete applications such as earth and planetary sciences.

On the other hand, regression models with spherical predictors with a dimension larger than three have also found considerable attention in the literature, mainly in physics, chemistry and material sciences. Here predictors on the unit sphere

$$\mathbb{S}_m = \{x \in \mathbb{R}^m : ||x||_2 = 1\},$$

with $m > 3$ and series expansions in terms of the so called *hyperspherical harmonics* are considered. These functions form an orthonormal system with respect to the uniform distribution on the sphere $\mathbb{S}_m$ and have been, for example, widely used to solve the Schroedinger equation by reducing the problem to a system of coupled ordinary differential equations in a single variable [see for example Avery and Wen (1982) or Krivec (1998) among many others]. Further applications in this field can be found in Meremianin (2009), who proposed the use of hyperspherical harmonics for the representation of the wave function of the hydrogen atom in the momentum space. Similarly, Lombardi et al. (2016) suggested to represent the potential energy surfaces (PES) of atom-molecule or molecular dimers interactions in terms of a series of four-dimensional hyperspherical harmonics. Their method consists in fitting a certain number of points of the PES, previously determined, selected on the basis of geometrical and physical characteristics of the system. The resulting potential energy function is suitable to serve as a PES for molecular dynamics simulations. Hosseinbor et al. (2013) applied four-dimensional hyperspherical harmonics in medical imaging and estimated the coefficients in the corresponding series expansion via least squares methods to analyse brain subcortical structures. A further important application of series expansions appears in material sciences, where crystallographic textures as quaternion distributions are represented by means of series expansions based on (symmetrized) hyperspherical harmonics [see Bunge (1993), Zheng et al. (1995), Mason and Schuh (2008) and Mason (2009) among many others].

It is well known that a carefully designed experiment can improve the statistical inference in regression analysis substantially, and numerous authors have considered the problem of constructing optimal designs for various regression models [see, for example, the monographs of Fedorov (1972), Silvey (1980) and Pukelsheim (2006)]. On the other hand, despite of its importance, the problem of constructing optimal or efficient designs for least squares (or alternative) estimation of the coefficients in series expansions based on hyperspherical harmonics has not found much interest in the statistical literature, in particular if the dimension m is large. The case $m = 2$ corresponding to Fourier regression models has been discussed intensively [see Karlin and Studden (1966), page 347, Lau and Studden (1985), Kitsos et al. (1988) and Dette and Melas (2003) among many others]. Furthermore, optimal designs for series estimators in terms of spherical harmonics (that is, for $m = 3$) have been determined by Dette et al. (2005) and Dette and Wiens (2009), however, to the best of our knowledge no results are available for hyperspherical harmonics if the dimension of the predictor is larger than 3.

In the present paper we consider optimal design problems for regression models with a spherical predictor of dimension $m > 3$ and explicitly determine optimal designs for series estimators in hyperspherical harmonic expansions. In Section 2 we introduce some basic facts about optimal design theory and hyperspherical harmonics, which will be required for the results presented in this paper. Analytic solutions of the optimal design problem are given in Section 3.1, where we determine optimal designs with respect to all Kiefer's Φ_p-criteria [see Kiefer (1974)] as well as with respect to a class of optimality criteria recently introduced by Harman (2004). As it turns out the approximate optimal designs are absolute continuous distributions on the sphere and thus cannot be directly implemented in practice. Therefore, in Section 3.2 we provide discrete designs with the same information matrices as the continuous optimal designs. To achieve this we construct new Gaussian quadrature formulas for integration on the sphere, which are of own interest. In Section 4 we investigate the performance of the optimal designs determined in Section 3.2 when they are used in typical applications in material sciences. Here energy functions are represented in terms of series of *symmetrized hyperspherical harmonics* which are obtained as well as defined as linear combinations of the hyperspherical harmonics such that the symmetry of a crystallographic point group is reflected in the energy function. It is demonstrated that the derived designs have very good efficiencies (for the first crystallographic point group the design is in fact D-optimal). Finally, a proof of a technical result can be found in Appendix A.

The results obtained in this paper provide a first step towards the solution of optimal design problems for regression models with spherical predictors if the dimension is $m > 3$ and offer a deeper understanding of the general mathematical structure of hyperspherical harmonics, which so far were only considered in the cases $m = 2$ and $m = 3$.

2 Optimal designs and hyperspherical harmonics

2.1 Optimal design theory

We consider the linear regression model

$$E[Y|x] = f^T(x)\boldsymbol{c}\ ;\ x \in \mathcal{X}, \tag{2.1}$$

where $f^T(x) = (f_1(x), \ldots, f_D(x))$ is a vector of linearly independent regression functions, $\boldsymbol{c} \in \mathbb{R}^D$ is the vector of unknown parameters, x denotes a real-valued covariate which varies in a compact design space, say $\mathcal{X}$ (which will be $\mathbb{S}_m$ in later sections), and different observations are assumed to be independent with the same variance, say $\sigma^2 > 0$. Following Kiefer (1974) we define an approximate design as a probability measure ξ on the set $\mathcal{X}$ (more precisely on its Borel field). If the design ξ has finite support with masses w_i at the points x_i $(i = 1, \ldots, k)$ and n observations can be made by the experimenter, this means that the quantities $w_i n$ are rounded to integers, say n_i, satisfying $\sum_{i=1}^k n_i = n$, and the experimenter takes n_i observations at each location x_i $(i = 1, \ldots, k)$. The information matrix of the least squares estimator is defined by

$$M(\xi) = \int_{\mathcal{X}} f(x) f^T(x) d\xi(x), \tag{2.2}$$

[see Pukelsheim (2006)] and measures the quality of the design ξ as the matrix $\frac{\sigma^2}{n} M^{-1}(\xi)$ can be considered as an approximation of the covariance matrix $\sigma^2(X^TX)^{-1}$ of the least squares estimator in the corresponding linear model $\boldsymbol{Y} = X\boldsymbol{c} + \varepsilon$. Similarly, if the main interest is the estimation of s linear combinations $K^T\boldsymbol{c}$, where $K \in \mathbb{R}^{D\times s}$ is a given matrix of rank $s \le D$, the covariance matrix of the least squares estimator for these linear combinations is given by $\frac{\sigma^2}{n}(K^T(X^TX)^-K)$, where $(X^TX)^-$ denotes the generalized inverse of the matrix X^TX and it is assumed that range$(K) \subset$ range(X^TX). The corresponding analogue of its inverse for an approximate design ξ satisfying the range inclusion range$(K) \subset$ range$(M(\xi))$ is given by (up to the constant $\frac{\sigma^2}{n}$)

$$C_K(\xi) = (K^T M^-(\xi) K)^{-1}\ . \tag{2.3}$$

It follows from Pukelsheim (2006), Section 8.3, that for each design ξ there always exists a design $\bar{\xi}$ with at most $s(s+1)/2$ support points such that $C_K(\xi) = C_K(\bar{\xi})$. An optimal design maximises an appropriate functional of the matrix $M(\xi)$ and numerous criteria have been proposed in the literature to discriminate between competing designs [see Pukelsheim (2006)]. Throughout this paper we consider Kiefer's Φ_p-criteria, which are defined for $-\infty \le p < 1$ as

$$\Phi_p(\xi) = \left(\text{tr}\left\{\left(C_K(\xi)\right)^p\right\}\right)^{1/p} = \left(\text{tr}\left\{\left(K^T M^-(\xi) K\right)^{-p}\right\}\right)^{1/p}\ . \tag{2.4}$$

Following Kiefer (1974), a design ξ^* is called Φ_p-optimal for estimating the linear combinations $K^T\boldsymbol{c}$ if ξ^* maximises the expression $\Phi_p(\xi)$ among all approximate designs ξ for which $K^T\boldsymbol{c}$ is estimable, that is, range$(K) \subset$ range$(M(\xi))$. This family of optimality criteria includes the well-known criteria of D-, E- and A-optimality corresponding to the cases $p = 0$, $p = -\infty$ and $p = -1$, respectively.
Moreover, we consider a generalised version of the criterion of E-optimality introduced by Harman (2004) [see also Filov et al. (2011)]. For the information matrix $M(\xi)$ let $\lambda(M(\xi)) = (\lambda_1(M(\xi)), \ldots, \lambda_D(M(\xi)))^T$ be the vector of the eigenvalues of $M(\xi)$ in non-decreasing order. Then, for $s \in 1, \ldots, D$, we define $\Phi_{E_s}(\xi)$ by the sum of the s-th smallest eigenvalues of $M(\xi)$, that is,

$$\Phi_{E_s}(\xi) = \sum_{i=1}^{s} \lambda_i(M(\xi)). \tag{2.5}$$

For a fixed $s \in \{1, \ldots, D\}$ we call a design ξ^* Φ_{E_s}-optimal if it maximises the term $\Phi_{E_s}(\xi)$ among all approximate designs ξ.

In general, the determination of Φ_p-optimal designs and of Φ_{E_s}-optimal designs in an explicit form is a very difficult task and the corresponding optimal design problems have only been solved in rare circumstances [see for example Cheng (1987), Dette and Studden (1993), Pukelsheim (2006), p.241, and Harman (2004)]. In the following discussion we will explicitly determine Φ_p-optimal designs for regression models which arise from a series expansion of a function on the m-dimensional sphere $\mathbb{S}_m$ in terms of hyperspherical harmonics. It turns out that the Φ_p-optimal designs are also Φ_{E_s}-optimal for an appropriate choice of s.
We introduce the hyperspherical harmonics next.

2.2 Hyperspherical harmonics

Assume that the design space is given by the m-dimensional sphere $\mathbb{S}_m = \{x \in \mathbb{R}^m : ||x||_2 = 1\}$. The hyperspherical harmonics are functions of $m - 1$ dimensionless variables, namely the hyperangles, which describe the points $x = (x_1, \ldots, x_m)^T \in \mathbb{S}_m$ on the hypersphere by the equations

$$\begin{aligned} x_1 &= \cos\theta_1, \qquad (2.6)\\ x_2 &= \sin\theta_1 \cos\theta_2,\\ x_3 &= \sin\theta_1 \sin\theta_2 \cos\theta_3,\\ &\vdots \qquad\qquad \vdots\\ x_{m-1} &= \sin\theta_1 \ldots \sin\theta_{m-2} \cos\phi,\\ x_m &= \sin\theta_1 \ldots \sin\theta_{m-2} \sin\phi, \end{aligned}$$

where $\theta_i \in [0, \pi]$ for all $i = 1, \dots, m-2$, $\phi \in [-\pi, \pi]$ [see, for example, Andrews et al. (1991) or Meremianin (2009)]. As noted by Dokmanić and Petrinović (2010), this choice of coordinates is not unique but rather a matter of convenience since it is a natural generalisation of the spherical polar coordinates in $\mathbb{R}^3$.
In the literature, hyperspherical harmonics are given explicitly in a complex form (see, for example, Vilenkin (1968) and Avery and Wen (1982)). Following the notation in Avery and Wen (1982), they are defined as

$$\tilde{Y}_{\lambda,\boldsymbol{\mu_{m-3}},\pm\mu_{m-2}}(\boldsymbol{\theta_{m-2}}, \phi) = \tilde{A}_{\lambda,\boldsymbol{\mu_{m-2}}} \prod_{i=1}^{m-2} \Big[C_{\mu_{i-1}-\mu_i}^{\mu_i + \frac{m-i-1}{2}}(\cos\theta_i)(\sin\theta_i)^{\mu_i} \Big] e^{\pm i \mu_{m-2}\phi},$$

where $\boldsymbol{\theta_{m-2}} = (\theta_1, \dots, \theta_{m-2})$, $\boldsymbol{\mu_k} = (\mu_1, \dots, \mu_k)$ for $k = m-2, m-3$, and

$$\tilde{A}_{\lambda,\boldsymbol{\mu_{m-2}}} = \frac{1}{\sqrt{2\pi}} \prod_{i=1}^{m-2} \Big[\frac{2^{2\mu_i+m-i-3}(\mu_{i-1}-\mu_i)!(2\mu_{i-1}+m-i-1)\Gamma^2(\mu_i + \frac{m-i-1}{2})}{\pi(\mu_{i-1}+\mu_i+m-i-2)!} \Big]^{1/2},$$

is a normalising constant, $\lambda := \mu_0 \geq \mu_1 \geq \mu_2 \geq \ldots \geq \mu_{m-2} \geq 0$ are a set of integers and the functions

$$C_{\mu_{i-1}-\mu_i}^{\mu_i + \frac{m-i-1}{2}}(x),$$

are the Gegenbauer polynomials (of degree $\mu_{i-1} - \mu_i \in \mathbb{N}_0$ with parameter $\mu_i + \frac{m-i-1}{2}$), which are orthogonal with respect to the measure

$$(1-x^2)^{\mu_i + (m-i-1)/2 - 1/2} I_{[-1,1]}(x) dx,$$

(here $I_A(x)$ denotes the indicator function of the set A). The complex hyperspherical functions are orthogonal to their corresponding complex conjugate and form an orthonormal basis of the space of square integrable functions with respect to the uniform distribution on the sphere

$$L^2(\mathbb{S}_m) = \{ f : \mathbb{S}_m \to \mathbb{C} \mid \int_{\mathbb{S}_m} |f(x)|^2 dx < \infty \}.$$

In fact the constants $\tilde{A}_{\lambda,\boldsymbol{\mu_{m-2}}}$ are chosen based on this property [see, for example, Avery and Wen (1982) for more details].

However, as mentioned in Mason and Schuh (2008), expansions of real-valued functions on the sphere are easier to handle in terms of real hyperspherical harmonics which are obtained from the complex hyperspherical harmonics via the linear transformations

$$
\begin{aligned}
Y_{\lambda,\boldsymbol{\mu_{m-3}},\mu_{m-2}}(\boldsymbol{\theta_{m-2}},\phi) &= \frac{(-1)^{\mu_{m-2}}[\tilde{Y}_{\lambda,\boldsymbol{\mu_{m-3}},\mu_{m-2}} + \tilde{Y}_{\lambda,\boldsymbol{\mu_{m-3}},-\mu_{m-2}}]}{\sqrt{2}} \\
&= A_{\lambda,\boldsymbol{\mu_{m-3}}} \prod_{i=1}^{m-3} \Big[C_{\mu_{i-1}-\mu_i}^{\mu_i+\frac{m-i-1}{2}}(\cos\theta_i)(\sin\theta_i)^{\mu_i} \Big] \\
&\quad \times B_{\mu_{m-3},\mu_{m-2}} P_{\mu_{m-3}}^{\mu_{m-2}}(\cos\theta_{m-2})\cos(\mu_{m-2}\phi), \\
Y_{\lambda,\boldsymbol{\mu_{m-3}},-\mu_{m-2}}(\boldsymbol{\theta_{m-2}},\phi) &= \frac{(-i)(-1)^{\mu_{m-2}}[\tilde{Y}_{\lambda,\boldsymbol{\mu_{m-3}},\mu_{m-2}} - \tilde{Y}_{\lambda,\boldsymbol{\mu_{m-3}},-\mu_{m-2}}]}{\sqrt{2}} \\
&= A_{\lambda,\boldsymbol{\mu_{m-3}}} \prod_{i=1}^{m-3} \Big[C_{\mu_{i-1}-\mu_i}^{\mu_i+\frac{m-i-1}{2}}(\cos\theta_i)(\sin\theta_i)^{\mu_i} \Big] \\
&\quad \times B_{\mu_{m-3},\mu_{m-2}} P_{\mu_{m-3}}^{\mu_{m-2}}(\cos\theta_{m-2})\sin(\mu_{m-2}\phi), \\
Y_{\lambda,\boldsymbol{\mu_{m-3}},0}(\boldsymbol{\theta_{m-2}},\phi) &= (-1)^{\mu_{m-2}}\tilde{Y}_{\lambda,\boldsymbol{\mu_{m-3}},0} \\
&= A_{\lambda,\boldsymbol{\mu_{m-3}}} \prod_{i=1}^{m-3} \Big[C_{\mu_{i-1}-\mu_i}^{\mu_i+\frac{m-i-1}{2}}(\cos\theta_i)(\sin\theta_i)^{\mu_i} \Big] \\
&\quad \times \frac{B_{\mu_{m-3},0}}{\sqrt{2}} P_{\mu_{m-3}}^{0}(\cos\theta_{m-2}),
\end{aligned} \tag{2.7}
$$

where

$$
A_{\lambda,\boldsymbol{\mu_{m-3}}} = \prod_{i=1}^{m-3} \Big[\frac{2^{2\mu_i+m-i-3}(\mu_{i-1}-\mu_i)!(2\mu_{i-1}+m-i-1)\Gamma^2(\mu_i+\frac{m-i-1}{2})}{\pi(\mu_{i-1}+\mu_i+m-i-2)!} \Big]^{1/2}, \tag{2.8}
$$

$$
B_{\mu_{m-3},\mu_{m-2}} = \Big[\frac{2(2\mu_{m-3}+1)(\mu_{m-3}-\mu_{m-2})!}{4\pi(\mu_{m-3}+\mu_{m-2})!} \Big]^{1/2}, \tag{2.9}
$$

and $P_{\mu_{m-3}}^{\mu_{m-2}}(\cos\theta_{m-2})$ is the associated Legendre polynomial which can be expressed in terms of a Gegenbauer polynomial via

$$
(-1)^{\mu_{m-2}} \frac{(2\mu_{m-2})!}{2^{\mu_{m-2}}(\mu_{m-2})!} (\sin\theta_{m-2})^{\mu_{m-2}} C_{\mu_{m-3}-\mu_{m-2}}^{\mu_{m-2}+\frac{1}{2}}(\cos\theta_{m-2}) = P_{\mu_{m-3}}^{\mu_{m-2}}(\cos\theta_{m-2}).
$$

It is easy to check that in the case of $\mathbb{R}^3$, the expressions in (2.7), (2.8) and (2.9) give the well known spherical harmonics involving only the associated Legendre polynomial [see Chapter 9 in Andrews et al. (1991) for more details].

The real hyperspherical harmonics defined in (2.7), (2.8) and (2.9) preserve the orthogonality properties of complex hyperspherical harmonics proven in Avery and Wen (1982). In other words, the real hyperspherical harmonics form an orthonormal basis of the Hilbert space

$$
L^2(\mathbb{S}_m, d\Omega_m) = \Big\{ g : \mathbb{S}_m \to \mathbb{R} \mid \int |g(\boldsymbol{\theta_{m-2}},\phi)|^2 d\Omega_m < \infty \Big\},
$$

that is,

$$\int Y_{\lambda,\boldsymbol{\mu_{m-3}},\mu_{m-2}}(\boldsymbol{\theta_{m-2}},\phi) Y_{\lambda',\boldsymbol{\mu_{m-3}'},\mu'_{m-2}}(\boldsymbol{\theta_{m-2}},\phi)\, d\Omega_m = \delta_{\lambda\lambda'} \prod_{i=1}^{m-2} \delta_{\mu_i\mu'_i}, \tag{2.10}$$

where

$$d\Omega_m = (\sin\theta_1)^{m-2} d\theta_1 (\sin\theta_2)^{m-3} d\theta_2 \dots (sin\theta_{m-2}) d\theta_{m-2} d\phi,$$

is the element of solid angle.

We now consider the linear regression model (2.1), where the vector of regression functions is obtained by a truncated expansion of a function $g \in L^2(\mathbb{S}_m, d\Omega_m)$ of order, say, d in terms of hyperspherical harmonics, that is,

$$\sum_{\lambda=0}^{d} \sum_{\mu_1=0}^{\lambda} \dots \sum_{\mu_{m-2}=-\mu_{m-3}}^{\mu_{m-3}} c_{\lambda,\boldsymbol{\mu_{m-3}},\mu_{m-2}} Y_{\lambda,\boldsymbol{\mu_{m-3}},\mu_{m-2}}(\boldsymbol{\theta_{m-2}},\phi).$$

Consequently, we obtain form (2.1) (using the coordinates $\boldsymbol{\theta_{m-2}} = (\theta_1,\dots,\theta_{m-2})$, ϕ)

$$E[Y|\boldsymbol{\theta_{m-2}},\phi] = f_d^T(\boldsymbol{\theta_{m-2}},\phi)\boldsymbol{c}, \tag{2.11}$$

where

$$\begin{aligned} f_d(\boldsymbol{\theta_{m-2}},\phi) \;\; &= \;\; \big(Y_{0,0,\dots,0}(\boldsymbol{\theta_{m-2}},\phi), Y_{1,0,\dots,0}(\boldsymbol{\theta_{m-2}},\phi), Y_{1,1,0,\dots,0}(\boldsymbol{\theta_{m-2}},\phi), \\ &\qquad \dots, Y_{1,1,\dots,1,-1}(\boldsymbol{\theta_{m-2}},\phi), \dots, Y_{d,d,\dots,d}(\boldsymbol{\theta_{m-2}},\phi)\big)^T, \end{aligned}$$

is the vector of hyperspherical harmonics of order d and the vector of parameters is given by

$$\boldsymbol{c} = (c_{0,0,\dots,0}, c_{1,0,\dots,0}, c_{1,1,0,\dots,0}, \dots, c_{1,1,\dots,1,-1}, \dots, c_{d,d,\dots,d})^T .$$

Note that the dimension of the vectors f_d and c is

$$D := \sum_{\lambda=0}^{d} \sum_{\mu_1=0}^{\lambda} \dots \sum_{\mu_{m-3}=0}^{\mu_{m-4}} \sum_{\mu_{m-2}=-\mu_{m-3}}^{\mu_{m-3}} 1 \;=\; \sum_{\lambda=0}^{d} \frac{(m+2\lambda-2)(\lambda+m-3)!}{\lambda!(m-2)!}, \tag{2.12}$$

where the expression for the sums over the μ_i's $(i = 1,\dots,m-2)$ is obtained from Avery and Wen (1982).

3 Φ_p- and Φ_{E_s}- optimal designs for hyperspherical harmonics

3.1 Optimal designs with a Lebesgue density

In this section we determine Φ_p-optimal designs for estimating the parameters in a series expansion of a function defined on the unit sphere $\mathbb{S}_m$. The corresponding regression model is defined by (2.11) and as mentioned in Section 2.1 a Φ_p-optimal (approximate) design maximises the criterion (2.4) in the class of all probability measures ξ on the set $[0,\pi]^{m-2} \times [-\pi,\pi]$ satisfying the range inclusion $\mathrm{range}(K) \subset \mathrm{range}(M(\xi))$, where the information matrix $M(\xi)$ is given by

$$M(\xi) = \int_{-\pi}^{\pi} \int_{0}^{\pi} \dots \int_{0}^{\pi} f_d(\boldsymbol{\theta_{m-2}}, \phi) f_d^T(\boldsymbol{\theta_{m-2}}, \phi)\, d\xi(\boldsymbol{\theta_{m-2}}, \phi).$$

We are interested in finding a design that is efficient for the estimation of the Fourier coefficients corresponding to the $s(k)$ hyperspherical harmonics

$$Y_{k,0,\dots,0}, Y_{k,1,0,\dots,0}, \dots, Y_{k,\dots,k,-k}, \dots, Y_{k,\dots,k,k},$$

where

$$s(k) = \frac{(m+2k-2)(k+m-3)!}{k!(m-2)!}, \tag{3.1}$$

and $k \in \{0, \dots, d\}$ denotes a given level of resolution. To relate this to the definition of the Φ_p-optimality criteria, let $q \in \mathbb{N}_0$, $0 \le k_0 < k_1 < \dots < k_q \le d$ and $0_{k,l}$ be the $s(k) \times s(l)$ matrix with all entries equal to 0. Define the matrix

$$K^T = (K_{j,l})_{j=0,\dots,q}^{l=0,\dots,d}, \tag{3.2}$$

where

$$K_{j,l} = \begin{cases} 0_{k_j,l} & l \ne k_j \\ I_{s(k_j)} & l = k_j \end{cases}, \tag{3.3}$$

I_a denotes the $a \times a$ identity and $0_{a,b}$ is an $a \times b$ matrix with all entries equal to 0. Note that $K \in \mathbb{R}^{D \times s}$ where $D = \sum_{\lambda=0}^{d} s(\lambda)$ is defined in (2.12), and that $K^T \boldsymbol{c} \in \mathbb{R}^s$ defines a vector with

$$s = \sum_{j=0}^{q} s(k_j) = \sum_{j=0}^{q} \frac{(m+2k_j-2)(k_j+m-3)!}{k_j!(m-2)!}, \tag{3.4}$$

components, that is

$$\left\{ c_{k_j,\mu_1,\dots,\mu_{m-2}} | k_j \ge \mu_1 \ge \dots \ge \mu_{m-3}; -\mu_{m-3} \le \mu_{m-2} \le \mu_{m-3}; j = 0, \dots, q \right\}, \tag{3.5}$$

$(s \le D)$. The following theorem shows that the uniform distribution on the hypersphere is Φ_{E_s}- and Φ_p-optimal for estimating the parameters $K^T\boldsymbol{c}$ (for any $-\infty \le p < 1$) .

Theorem 3.1. *Let $p \in [-\infty, 1)$, $0 \le k_0 < k_1 < \dots < k_q \le d$ be given indices and denote by $K \in \mathbb{R}^{D \times s}$ the matrix defined by (3.2) and (3.3). Consider the design given by the uniform distribution on the hypersphere, that is,*

$$\begin{aligned} \xi^* &= \xi^*(d\theta_1, \dots, d\theta_{m-2}, d\phi) = \frac{d\Omega_m}{\tilde{\Omega}} \\ &= \frac{1}{\tilde{\Omega}} (\sin\theta_1)^{m-2} d\theta_1 (\sin\theta_2)^{m-3} d\theta_2 \dots (\sin\theta_{m-2}) d\theta_{m-2} d\phi, \end{aligned} \tag{3.6}$$

where $\tilde{\Omega}$ is a normalising constant given by

$$\tilde{\Omega} = \int_{-\pi}^{\pi}\int_{0}^{\pi}\dots\int_{0}^{\pi}(\sin\theta_1)^{m-2}d\theta_1(\sin\theta_2)^{m-3}d\theta_2\dots(sin\theta_{m-2})d\theta_{m-2}d\phi = \frac{N_m}{(m-2)!!}, \tag{3.7}$$

$N_m = 2(m-2)!!\pi^{m/2}/\Gamma(m/2)$ *and the double factorial* $n!!$ *for* $n \in \mathbb{N}$ *is defined by*

$$n!! = \begin{cases} \prod_{k=1}^{\frac{n}{2}}(2k) & n \text{ is even} \\ \prod_{k=1}^{\frac{n+1}{2}}(2k-1) & n \text{ is odd} \end{cases}.$$

(i) The information matrix of ξ^* *is given by* $M(\xi^*) = \frac{1}{\tilde{\Omega}}I_D$*, where* D *is defined in* (2.12).

(ii) The design ξ^* *is* Φ_p*-optimal for estimating the linear combination* $K^T\boldsymbol{c}$ *in the regression model* (2.11).

(iii) Let $s = \sum_{i=0}^{q} s(k_i)$ *be the number of considered hyperspherical harmonics where* $s(k_i)$ *is defined by* (3.1). *Then the design* ξ^* *defined by* (3.6) *is also* Φ_{E_s}*-optimal.*

Proof. We note that the explicit expression for the normalising constant $\tilde{\Omega}$ in (3.7) is given in equation (30) in Wen and Avery (1985). Let ξ^* denote the design corresponding to the density defined by (3.6) and (3.7). Then due to the orthonormality property of the real hyperspherical harmonics, given in equation (2.10), it follows that

$$M(\xi^*) = \frac{1}{\tilde{\Omega}}I_D, \tag{3.8}$$

where $\tilde{\Omega}$ is defined in equation (3.7). This proves part (i) of the Theorem.
For a proof of (ii) let $p > -\infty$. According to the general equivalence theorem in Pukelsheim (2006), Section 7.20, the measure ξ^* is Φ_p-optimal if and only if the inequality

$$f_d^T(\boldsymbol{\theta_{m-2}},\phi)\tilde{\Omega}K\Big(K^T\tilde{\Omega}K\Big)^{-p-1}K^T\tilde{\Omega}f_d(\boldsymbol{\theta_{m-2}},\phi) \le \operatorname{tr}\Big\{\Big(K^T\tilde{\Omega}K\Big)^{-p}\Big\}, \tag{3.9}$$

holds for all $\boldsymbol{\theta_{m-2}} \in [0,\pi]^{m-2}$ and $\phi \in [-\pi,\pi]$.

From the definition of the matrix K given in equations (3.2) and (3.3) we have that $K^TK = I_s$ where $s = \sum_{j=0}^{q} s(k_j)$ and $s(k_j)$ is given in (3.1). Therefore, condition (3.9) reduces to

$$s \ge \sum_{j=0}^{q}\sum_{\mu_1=0}^{k_j}\dots\sum_{\mu_{m-3}=0}^{\mu_{m-4}}\sum_{\mu_{m-2}=-\mu_{m-3}}^{\mu_{m-3}}\tilde{\Omega}\big(Y_{k_j,\boldsymbol{\mu_{m-3}},\mu_{m-2}}(\boldsymbol{\theta_{m-2}},\phi)\big)^2. \tag{3.10}$$

Now the right-hand side can be simplified observing the sum rule for real hyperspherical harmonics, that is

$$\sum_{\mu_1=0}^{\lambda}\cdots\sum_{\mu_{m-2}=-\mu_{m-3}}^{\mu_{m-3}}\Big(Y_{\lambda,\boldsymbol{\mu_{m-3}},\mu_{m-2}}(\boldsymbol{\theta_{m-2}},\phi)\Big)^2=\frac{(m+2\lambda-2)(m-4)!!(\lambda+m-3)!}{N_m\lambda!(m-3)!}, \tag{3.11}$$

where the constant N_m is given by

$$N_m=\frac{2(m-2)!!\pi^{m/2}}{\Gamma\left(\frac{m}{2}\right)}=\begin{cases}(2\pi)^{m/2} & m \text{ is even}\\ 2(2\pi)^{(m-2)/2} & m \text{ is odd}\end{cases}, \tag{3.12}$$

(see Avery and Wen (1982)). Therefore, the right-hand side of (3.10) becomes

$$\sum_{j=0}^{q}\frac{N_m}{(m-2)!!}\frac{(m+2k_j-2)(m-4)!!(k_j+m-3)!}{N_mk_j!(m-3)!}=\sum_{j=0}^{q}\frac{(m+2k_j-2)(k_j+m-3)!}{k_j!(m-2)!}=s,$$

where the last equality follows from the definition of s in (3.4). Consequently, the right-hand side and left-hand side of (3.10) coincide, which proves that the design ξ^* corresponding to the density defined by (3.6) and (3.7) is Φ_p-optimal for any $p\in(-\infty,1)$ and any matrix K of the form (3.2) and (3.3). The remaining case $p=-\infty$ follows from Lemma 8.15 in Pukelsheim (2006), which completes the proof of part (ii).
For a proof of part (iii) let $\mathrm{diag}(\gamma_1,\dots,\gamma_D)$ denote a diagonal matrix with entries $\gamma_1,\dots,\gamma_D$ and let

$$\partial\Phi_{E_s}(\xi^*)=\Big\{\mathrm{diag}(\gamma_1,\dots,\gamma_D)\in\mathbb{R}^{D\times D}\Big|\ \gamma_1,\dots,\gamma_D\in[0,1],\sum_{k=1}^{D}\gamma_k=s\Big\}, \tag{3.13}$$

denote the subgradient of Φ_{E_s}. Then it follows from Theorem 4 of Harman (2004), that the design ξ^* is Φ_{E_s}-optimal if and only if there exists a matrix $\Gamma\in\partial\Phi_{E_s}(\xi^*)$ such that the inequality

$$f_d^T(\boldsymbol{\theta_{m-2}},\phi)\Gamma f_d(\boldsymbol{\theta_{m-2}},\phi)\leq\sum_{k=1}^{s}\lambda_k(M(\xi^*)), \tag{3.14}$$

holds for all $\boldsymbol{\theta_{m-2}}\in[0,\pi]^{m-2}$ and $\phi\in[-\pi,\pi]$.
We now set $\Gamma=KK^T$ where K is defined by the equations (3.2) and (3.3). Therefore Γ is a diagonal matrix with entries 0 or 1, and

$$\mathrm{tr}(\Gamma)=\mathrm{tr}(KK^T)=\mathrm{tr}(K^TK)=\mathrm{tr}(I_s)=s,$$

that is, the matrix Γ is contained in the subgradient $\partial\Phi_{E_s}(\xi^*)$. Using this matrix in (3.14) the left-hand side of the inequality reduces to

$$f_d^T(\boldsymbol{\theta_{m-2}},\phi)\Gamma f_d(\boldsymbol{\theta_{m-2}},\phi)=\sum_{j=0}^{q}\sum_{\mu_1=0}^{k_j}\cdots\sum_{\mu_{m-3}=0}^{\mu_{m-4}}\sum_{\mu_{m-2}=-\mu_{m-3}}^{\mu_{m-3}}\big(Y_{k_j,\boldsymbol{\mu_{m-3}},\mu_{m-2}}(\boldsymbol{\theta_{m-2}},\phi)\big)^2,$$

and part (i) yields for the right hand side of the inequality

$$\sum_{k=1}^{s} \lambda_k(M(\xi^*)) = \tfrac{s}{\tilde{\Omega}},$$

where $\tilde{\Omega}$ is defined by (3.7). Consequently, the inequality (3.14) is equivalent to (3.10), which has been proved in the proof of part (ii). This completes the proof of Theorem 3.1. □

3.2 Discrete Φ_p- and Φ_{E_s}- optimal designs

While the result of the previous section provides a very elegant solution to the Φ_p-optimal design problem from a mathematical point of view, the derived designs ξ^* cannot be directly implemented as the optimal probability measure is absolute continuous. In practice, if $n \in \mathbb{N}$ observations are available to estimate the parameters in the linear regression model (2.11), one has to specify a number, say k, of different points $(\boldsymbol{\theta^1_{m-2}}, \phi^1), \ldots, (\boldsymbol{\theta^k_{m-2}}, \phi^k) \in [0,\pi]^{m-2} \times [-\pi,\pi]$ defining by (2.6) the locations on the sphere where observations should be taken, and relative frequencies n_j/n defining the proportion of observations taken at each point ($\sum_{j=1}^k n_j = n$). The maximisation of the function (2.4) in the class of all measures of this type yields a non-linear and non-convex discrete optimisation problem, which is usually intractable.

Therefore, for the construction of optimal or (at least) efficient designs we proceed as follows. Due to Caratheodory's theorem [see, for example, Silvey (1980)] there always exists a probability measure ξ on the set $[0,\pi]^{m-2} \times [-\pi,\pi]$ with at most $D(D+1)/2$ support points such that the information matrices of ξ and ξ^* coincide, that is,

$$M(\xi) = M(\xi^*) = \frac{1}{\tilde{\Omega}} I_D. \tag{3.15}$$

We now identify such a design ξ assigning at the points $\{(\boldsymbol{\theta^j_{m-2}}, \phi^j)\}_{j=1}^k$ the weights $\{\boldsymbol{\omega^j}\}_{j=1}^k = \{(\omega_1^j, \omega_2^j, \ldots, \omega_{m-2}^j, \omega_\phi^j)\}_{j=1}^k$ such that the identity (3.15) is satisfied, where we simultaneously try to keep the number k of support points "small". The numbers n_j specifying the numbers of repetitions at the different experimental conditions in the concrete experiment are finally obtained by rounding the numbers $n\boldsymbol{\omega^j}$ to integers [see, for example, Pukelsheim and Rieder (1992)]. We begin with an auxiliary result about Gauss quadrature which is of independent interest and is proven in the appendix.

Lemma 3.1. *Let a be a positive and integrable weight function on the interval $[-1,1]$ with $\tilde{a} = \int_{-1}^1 a(x)\,dx$, and let $-1 \le x_1 < x_2 < \ldots < x_r \le 1$ denote $r \in \mathbb{N}$ points with corresponding positive weights $\omega_1, \ldots, \omega_r$ ($\sum_{j=1}^r \omega_j = 1$). Then the points x_i and weights*

ω_i generate a quadrature formula of degree $z \geq r$, that is

$$\int_{-1}^{1} a(x)x^\ell \, dx = \tilde{a} \sum_{j=1}^{r} \omega_j x_j^\ell, \qquad \ell = 0, \ldots, z, \tag{3.16}$$

if and only if the following two conditions are satisfied:

(A) The polynomial $V_r(x) = \prod_{j=1}^{r}(x - x_j)$ is orthogonal with respect to the weight function $a(x)$ to all polynomials of degree $z - r$, that is,

$$\int_{-1}^{1} V_r(x)a(x)x^\ell \, dx = 0, \qquad \ell = 0, \ldots, z - r. \tag{3.17}$$

(B) The weights ω_j are given by

$$\omega_j = \frac{1}{\tilde{a}} \int_{-1}^{1} a(x)\ell_j(x) \, dx \quad j = 1, \ldots, r, \tag{3.18}$$

where $\ell_j(x) = \prod_{k=1, k \neq j}^{r} \frac{x - x_k}{x_k - x_j}$ denotes the jth Lagrange interpolation polynomial with nodes $x_1, \ldots, x_r$.

In the following, we use Lemma 3.1 for $z = 2d$ and the weight function

$$a(x) = (1 - x^2)^{(m-i-2)/2}.$$

Note that the Gegenbauer polynomials $C_r^{(m-i-1)/2}(x)$ are orthogonal with respect to the weight function $a(x) = (1 - x^2)^{(m-i-2)/2}$ on the interval $[-1, 1]$ [see Andrews et al. (1991), p. 302]. Hence the r roots of $C_r^{(m-i-1)/2}(x)$ have multiplicity 1, are real and located in the interval $(-1, 1)$. As condition (3.17) is satisfied for $a(x) = (1 - x^2)^{(m-i-2)/2}$, they define together with the corresponding (positive) weights in (3.18) a Gaussian quadrature formula. Therefore, it follows that for any $r \in \{d + 1, \ldots, 2d\}$ there exists at least one quadrature formula $\{x_i^j, \omega_i^j\}_{j=1}^r$ for every $i = 1, \ldots, m - 2$, such that (3.16) holds with $a(x) = (1 - x^2)^{(m-i-2)/2}$. We consider quadrature formulas of this type and define the designs

$$\zeta_i = \begin{pmatrix} \theta_1^i & \ldots & \theta_r^i \\ \omega_1^i & \ldots & \omega_r^i \end{pmatrix}, \tag{3.19}$$

on $[0, \pi]$, where

$$\theta_j^i = \arccos x_j^i \qquad i = 1, \ldots, m - 2; \quad j = 1, \ldots, r. \tag{3.20}$$

Similarly we define for any $t \in \mathbb{N}$ and any $\beta \in (-\frac{t+1}{t}\pi, -\pi]$ a design $\nu = \nu(\beta, t)$ on the interval $[-\pi, \pi]$ by

$$\nu = \nu(\beta, t) = \begin{pmatrix} \phi^1 & \ldots & \phi^t \\ \frac{1}{t} & \ldots & \frac{1}{t} \end{pmatrix}, \tag{3.21}$$

where the points ϕ^j are given by

$$\phi^j = \beta + \frac{2\pi j}{t}, \qquad j = 1, \dots, t. \tag{3.22}$$

The following theorem shows that designs of the form

$$\zeta_1 \otimes \dots \otimes \zeta_{m-2} \otimes \nu, \tag{3.23}$$

are Φ_p- as well as Φ_{E_s}-optimal designs.

Theorem 3.2. *Let $p \in [-\infty, 1)$, $0 \le k_0 < k_1 < \dots < k_q \le d$ and K be a matrix defined by (3.2) and (3.3). For any $t \ge 2d+1$ and any $r \in \{d+1, \dots, 2d\}$, the design $\zeta_1 \otimes \dots \otimes \zeta_{m-2} \otimes \nu$ defined in (3.23) is Φ_p-optimal for estimating the coefficients $K^T \boldsymbol{c}$. Moreover, if $s = \sum_{i=0}^{q} s(k_j)$ is the number of considered hyperspherical harmonics defined in (3.4), then for any $t \ge 2d+1$ and any $r \in \{d+1, \dots, 2d\}$, the design $\zeta_1 \otimes \dots \otimes \zeta_{m-2} \otimes \nu$ defined in (3.23) is Φ_{E_s}-optimal.*

Proof. The assertion can be established by showing the identity

$$M(\zeta_1 \otimes \dots \otimes \zeta_{m-2} \otimes \nu) = \frac{1}{\tilde{\Omega}} I_D, \tag{3.24}$$

where the dimension D is defined in (2.12). Let

$$\begin{aligned} \psi(\phi) &= (\psi_{-d}(\phi), \psi_{-d+1}(\phi), \dots, \psi_d(\phi))^T \\ &= (\sqrt{2}\sin(d\phi), \dots, \sqrt{2}\sin(\phi), 1, \sqrt{2}\cos(\phi), \dots, \sqrt{2}\cos(d\phi))^T. \end{aligned}$$

Then the real hyperspherical harmonics defined in (2.7), (2.8) and (2.9) can be rewritten as

$$\begin{aligned} Y_{\lambda, \boldsymbol{\mu_{m-3}}, \mu_{m-2}}(\boldsymbol{\theta_{m-2}}, \phi) &= \prod_{i=1}^{m-3} \tilde{\gamma}_{\mu_{i-1}, \mu_i} \prod_{i=1}^{m-3} \Big[C_{\mu_{i-1}-\mu_i}^{\mu_i + \frac{m-i-1}{2}} (\cos\theta_i)(\sin\theta_i)^{\mu_i} \Big] \\ &\quad \times \gamma_{\mu_{m-3}, \mu_{m-2}} P_{\mu_{m-3}}^{|\mu_{m-2}|}(\cos\theta_{m-2}) \psi_{\mu_{m-2}}(\phi), \end{aligned}$$

where the constants $\tilde{\gamma}_{\mu_{i-1}, \mu_i}$ and $\gamma_{\mu_{m-3}, \mu_{m-2}}$ are defined by

$$\tilde{\gamma}_{\mu_{i-1}, \mu_i} = \Big[\frac{2^{2\mu_i + m - i - 3} (\mu_{i-1} - \mu_i)! (2\mu_{i-1} + m - i - 1) \Gamma^2(\mu_i + \frac{m-i-1}{2})}{\pi (\mu_{i-1} + \mu_i + m - i - 2)!} \Big]^{1/2},$$

and

$$\gamma_{\mu_{m-3}, \mu_{m-2}} = \Big[\frac{(2\mu_{m-3} + 1)(\mu_{m-3} - |\mu_{m-2}|)!}{4\pi (\mu_{m-3} + |\mu_{m-2}|)!} \Big]^{1/2}.$$

Therefore, the identity (3.24) is equivalent to the system of equations

$$\int Y_{\lambda,\boldsymbol{\mu_{m-3}},\mu_{m-2}}(\boldsymbol{\theta_{m-2}},\phi)Y_{\lambda',\boldsymbol{\mu_{m-3}}',\mu'_{m-2}}(\boldsymbol{\theta_{m-2}},\phi)\,d(\zeta_1\otimes\ldots\otimes\zeta_{m-2})(\boldsymbol{\theta_{m-2}})\,d\nu(\phi)$$
$$=\prod_{i=1}^{m-3}\tilde{\gamma}_{\mu_{i-1},\mu_i}\prod_{i=1}^{m-3}\tilde{\gamma}_{\mu'_{i-1},\mu'_i}\gamma_{\mu_{m-3},\mu_{m-2}}\gamma_{\mu'_{m-3},\mu'_{m-2}}$$
$$\times\int_{-\pi}^{\pi}\int_0^{\pi}\ldots\int_0^{\pi}\prod_{i=1}^{m-3}\left[C_{\mu_{i-1}-\mu_i}^{\mu_i+\frac{m-i-1}{2}}(\cos\theta_i)(\sin\theta_i)^{\mu_i}\right]P_{\mu_{m-3}}^{|\mu_{m-2}|}(\cos\theta_{m-2})\psi_{\mu_{m-2}}(\phi)$$
$$\prod_{i=1}^{m-3}\left[C_{\mu'_{i-1}-\mu'_i}^{\mu'_i+\frac{m-i-1}{2}}(\cos\theta_i)(\sin\theta_i)^{\mu'_i}\right]P_{\mu'_{m-3}}^{|\mu'_{m-2}|}(\cos\theta_{m-2})\psi_{\mu'_{m-2}}(\phi)\,d\zeta_1(\theta_1)\ldots\,d\zeta_{m-2}(\theta_{m-2})\,d\nu(\phi)$$
$$=\frac{1}{\tilde{\Omega}}\delta_{\lambda\lambda'}\delta_{\mu_1\mu'_1}\ldots\delta_{\mu_{m-2}\mu'_{m-2}},$$

where

$$\lambda,\lambda'=0,\ldots,d;\ \mu_1=0,\ldots,\lambda;\ \ldots\ ;\ \mu_{m-3}=0,\ldots,\mu_{m-4};\ \mu_{m-2}=-\mu_{m-3},\ldots,\mu_{m-3};$$
$$\mu'_1=0,\ldots,\lambda'\ ;\ldots\ ;\ \mu'_{m-3}=0,\ldots,\mu'_{m-4};\ \mu'_{m-2}=-\mu'_{m-3},\ldots,\mu'_{m-3}.$$

Note that

$$\tilde{\Omega}=2\pi\prod_{i=1}^{m-2}\int_0^{\pi}(\sin\theta_i)^{m-i-1}\,d\theta_i=2\pi\prod_{i=1}^{m-2}\int_{-1}^{1}(1-x^2)^{\frac{m-i-2}{2}}\,dx,$$

and that $a(x)=(1-x^2)^{\frac{m-i-2}{2}}$ is the weight function defining each of the quadrature formulas for $i=1,\ldots,m-2$.

Consequently, by Fubini's theorem the system above is satisfied if the following equations hold

$$\int_{-\pi}^{\pi}\psi_{\mu_{m-2}}(\phi)\psi_{\mu'_{m-2}}(\phi)\,d\nu(\phi)=\delta_{\mu_{m-2}\mu'_{m-2}},\tag{3.25}$$

$(\mu_{m-2},\mu'_{m-2}=-d,\ldots,d)$

$$\gamma_{\mu_{m-3},\mu_{m-2}}\gamma_{\mu'_{m-3},\mu'_{m-2}}\int_0^{\pi}P_{\mu_{m-3}}^{|\mu_{m-2}|}(\cos\theta_{m-2})P_{\mu'_{m-3}}^{|\mu'_{m-2}|}(\cos\theta_{m-2})\,d\zeta_{m-2}(\theta_{m-2})$$
$$=\frac{1}{2\pi\int_{-1}^{1}1\,dx}\delta_{\mu_{m-3}\mu'_{m-3}}\delta_{\mu_{m-2}\mu'_{m-2}},\tag{3.26}$$

$(\mu_{m-3},\mu'_{m-3}=0,\ldots,d;\mu_{m-2}=0,\ldots,\mu_{m-3};\mu'_{m-2}=0,\ldots,\mu'_{m-3})$ and for each $i=1,\ldots,m-3$

$$\tilde{\gamma}_{\mu_{i-1},\mu_i}\tilde{\gamma}_{\mu'_{i-1},\mu'_i}\int_0^{\pi}C_{\mu_{i-1}-\mu_i}^{\mu_i+\frac{m-i-1}{2}}(\cos\theta_i)(\sin\theta_i)^{\mu_i}C_{\mu'_{i-1}-\mu'_i}^{\mu'_i+\frac{m-i-1}{2}}(\cos\theta_i)(\sin\theta_i)^{\mu'_i}\,d\zeta_i(\theta_i)$$
$$=\frac{1}{\int_{-1}^{1}(1-x^2)^{\frac{m-i-2}{2}}\,dx}\delta_{\mu_{i-1}\mu'_{i-1}}\delta_{\mu_i\mu'_i},\tag{3.27}$$

$(\mu_{i-1}, \mu'_{i-1} = 0, \dots, d; \mu_i = 0, \dots, \mu_{i-1}; \mu'_i = 0, \dots, \mu'_{i-1})$.

It is well known [see Pukelsheim (2006)] that equation (3.25) is satisfied for measures of the form (3.21). Hence in what follows we can restrict ourselves to the case $\mu_{m-2} = \mu'_{m-2}$.

Now the integrand in equation (3.26) is a polynomial of degree $\mu_{m-3} + \mu'_{m-3} \leq 2d$. Furthermore, since ζ_{m-2} corresponds to a quadrature formula for $a(x) = 1$ that integrates polynomials of degree $2d$ exactly, we have from Lemma 3.1 for $z = 2d$ and $a(x) = 1$ that

$$\int_0^\pi P_{\mu_{m-3}}^{|\mu_{m-2}|}(\cos\theta_{m-2}) P_{\mu'_{m-3}}^{|\mu'_{m-2}|}(\cos\theta_{m-2})\, d\zeta_{m-2}(\theta_{m-2})$$

$$= \sum_{j=1}^r \omega_{m-2}^j P_{\mu_{m-3}}^{|\mu_{m-2}|}(x_{m-2}^j) P_{\mu'_{m-3}}^{|\mu_{m-2}|}(x_{m-2}^j) = \frac{1}{2}\int_{-1}^1 P_{\mu_{m-3}}^{|\mu_{m-2}|}(x) P_{\mu'_{m-3}}^{|\mu_{m-2}|}(x)\, dx.$$

From Andrews et al. (1991) p.457 we have that

$$\int_{-1}^1 \left[P_{\mu_{m-3}}^{|\mu_{m-2}|}(x)\right]^2 dx = \frac{2(\mu_{m-3} + |\mu_{m-2}|)!}{(2\mu_{m-3}+1)(\mu_{m-3} - |\mu_{m-2}|)!} = \frac{2}{4\pi\gamma_{\mu_{m-3},\mu_{m-2}}\gamma_{\mu_{m-3},\mu_{m-2}}}.$$

Therefore,

$$\int_0^\pi P_{\mu_{m-3}}^{|\mu_{m-2}|}(\cos\theta_{m-2}) P_{\mu'_{m-3}}^{|\mu'_{m-2}|}(\cos\theta_{m-2})\, d\zeta_{m-2}(\theta_{m-2}) = \frac{1}{4\pi}\frac{\delta_{\mu_{m-3}\mu'_{m-3}}}{\gamma_{\mu_{m-3},\mu_{m-2}}\gamma_{\mu'_{m-3},\mu_{m-2}}},$$

since associated Legendre polynomials are orthogonal on $[-1,1]$. This implies equation (3.26) and in what follows we can restrict ourselves to the case $\mu_{m-3} = \mu'_{m-3}$.

For establishing the system of equations (3.27), we begin with establishing the equation for $i = m-3$, that is,

$$\tilde{\gamma}_{\mu_{m-4},\mu_{m-3}}\tilde{\gamma}_{\mu'_{m-4},\mu_{m-3}} \int_0^\pi C_{\mu_{m-4}-\mu_{m-3}}^{\mu_{m-3}+1}(\cos\theta_{m-3})(\sin\theta_{m-3})^{\mu_{m-3}} C_{\mu'_{m-4}-\mu_{m-3}}^{\mu_{m-3}+1}(\cos\theta_{m-3})$$

$$(\sin\theta_{m-3})^{\mu_{m-3}}\, d\zeta_{m-3}(\theta_{m-3}) = \frac{1}{\int_{-1}^1 (1-x^2)^{\frac{1}{2}}\, dx}\delta_{\mu_{m-4}\mu'_{m-4}}. \tag{3.28}$$

The integrand is a polynomial of degree $2\mu_{m-3} + \mu_{m-4} - \mu_{m-3} + \mu'_{m-4} - \mu_{m-3} = \mu_{m-4} + \mu'_{m-4} \leq 2d$. Also since ζ_{m-3} corresponds to a quadrature formula for $a(x) = \sqrt{1-x^2}$ that integrates polynomials of degree $2d$ exactly, it follows from Lemma 3.1 for $z = 2d$ and

$a(x) = \sqrt{1-x^2}$ that

$$\begin{aligned}
&\int_0^\pi (\sin\theta_{m-3})^{2\mu_{m-3}} C^{\mu_{m-3}+1}_{\mu_{m-4}-\mu_{m-3}}(\cos\theta_{m-3}) C^{\mu_{m-3}+1}_{\mu'_{m-4}-\mu_{m-3}}(\cos\theta_{m-3})\, d\zeta_{m-3}(\theta_{m-3}) \\
&= \sum_{j=1}^{r} \omega^j_{m-3}(1-(x^j_{m-3})^2)^{\mu_{m-3}} C^{\mu_{m-3}+1}_{\mu_{m-4}-\mu_{m-3}}(x^j_{m-3}) C^{\mu_{m-3}+1}_{\mu'_{m-4}-\mu_{m-3}}(x^j_{m-3}) \\
&= \frac{1}{\int_{-1}^1 \sqrt{1-x^2}\,dx} \int_{-1}^1 \sqrt{1-x^2}(1-x^2)^{\mu_{m-3}} C^{\mu_{m-3}+1}_{\mu_{m-4}-\mu_{m-3}}(x) C^{\mu_{m-3}+1}_{\mu'_{m-4}-\mu_{m-3}}(x)\, dx \\
&= \frac{1}{\pi/2} \int_{-1}^1 (1-x^2)^{\mu_{m-3}+\frac{1}{2}} C^{\mu_{m-3}+1}_{\mu_{m-4}-\mu_{m-3}}(x) C^{\mu_{m-3}+1}_{\mu'_{m-4}-\mu_{m-3}}(x)\, dx.
\end{aligned}$$

From Andrews et al. (1991), Corollary 6.8.4, we have that

$$\begin{aligned}
&\int_{-1}^1 (1-x^2)^{\mu_{m-3}+\frac{1}{2}} \left[C^{\mu_{m-3}+1}_{\mu_{m-4}-\mu_{m-3}}(x)\right]^2 dx \\
&= \frac{\pi}{2} \frac{(\mu_{m-4}+\mu_{m-3}+1)!}{2^{2\mu_{m-3}}(\mu_{m-3}!)^2(\mu_{m-4}+1)(\mu_{m-4}-\mu_{m-3})!} = \frac{1}{\tilde{\gamma}_{\mu_{m-4},\mu_{m-3}}\tilde{\gamma}_{\mu_{m-4},\mu_{m-3}}}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\int_0^\pi (\sin\theta_{m-3})^{2\mu_{m-3}} C^{\mu_{m-3}+1}_{\mu_{m-4}-\mu_{m-3}}(\cos\theta_{m-3}) C^{\mu_{m-3}+1}_{\mu'_{m-4}-\mu_{m-3}}(\cos\theta_{m-3})\, d\zeta_{m-3}(\theta_{m-3}) \\
&= \frac{2}{\pi} \frac{\delta_{\mu_{m-4}\mu'_{m-4}}}{\tilde{\gamma}_{\mu_{m-4},\mu_{m-3}}\tilde{\gamma}_{\mu'_{m-4},\mu_{m-3}}}
\end{aligned}$$

since Gegenbauer polynomials $C^{\mu_{m-3}+1}_{\mu_{m-4}-\mu_{m-3}}(x)$ are orthogonal with respect to $(1-x^2)^{\mu_{m-3}+1/2}$ on the interval $[-1,1]$. This implies (3.28) and in what follows we can restrict ourselves to the case $\mu_{m-4} = \mu'_{m-4}$.

It remains to show that if (3.27) holds for $i = k+1$, that is, if

$$\begin{aligned}
&\tilde{\gamma}_{\mu_k,\mu_{k+1}}\tilde{\gamma}_{\mu'_k,\mu'_{k+1}} \int_0^\pi C^{\mu_{k+1}+\frac{m-k-2}{2}}_{\mu_k-\mu_{k+1}}(\cos\theta_{k+1})(\sin\theta_{k+1})^{\mu_{k+1}} C^{\mu'_{k+1}+\frac{m-k-2}{2}}_{\mu'_k-\mu'_{k+1}}(\cos\theta_{k+1}) \\
&(\sin\theta_{k+1})^{\mu'_{k+1}}\, d\zeta_{k+1}(\theta_{k+1}) = \frac{1}{\int_{-1}^1 (1-x^2)^{\frac{m-k-3}{2}}\,dx} \delta_{\mu_k\mu'_k}\delta_{\mu_{k+1}\mu'_{k+1}}, \qquad (3.29)
\end{aligned}$$

then (3.27) holds for $i = k$, that is,

$$\begin{aligned}
&\tilde{\gamma}_{\mu_{k-1},\mu_k}\tilde{\gamma}_{\mu'_{k-1},\mu'_k} \int_0^\pi C^{\mu_k+\frac{m-k-1}{2}}_{\mu_{k-1}-\mu_k}(\cos\theta_k)(\sin\theta_k)^{\mu_k} C^{\mu'_k+\frac{m-k-1}{2}}_{\mu'_{k-1}-\mu'_k}(\cos\theta_k) \\
&(\sin\theta_k)^{\mu'_k}\, d\zeta_k(\theta_k) = \frac{1}{\int_{-1}^1 (1-x^2)^{\frac{m-k-2}{2}}\,dx} \delta_{\mu_{k-1}\mu'_{k-1}}\delta_{\mu_k\mu'_k}. \qquad (3.30)
\end{aligned}$$

Note that we use somewhat a "backward induction step" since $\mu_k \geq \mu_{k+1}$.

Now since (3.29) holds, for proving (3.30) we can restrict ourselves to the case $\mu_k = \mu_k'$. The integrand in (3.30) is a polynomial of degree $2\mu_k + \mu_{k-1} - \mu_k + \mu_{k-1}' - \mu_k = \mu_{k-1} + \mu_{k-1}' \leq 2d$. Furthermore, since ζ_k corresponds to a quadrature formula for $a(x) = (1-x^2)^{\frac{m-k-2}{2}}$ that integrates polynomials of degree $2d$ exactly, we have from Lemma 3.1 for $z = 2d$ and $a(x) = (1-x^2)^{\frac{m-k-2}{2}}$ that

$$\begin{aligned}
&\int_0^\pi (\sin\theta_k)^{2\mu_k} C_{\mu_{k-1}-\mu_k}^{\mu_k+\frac{m-k-1}{2}}(\cos\theta_k) C_{\mu_{k-1}'-\mu_k}^{\mu_k+\frac{m-k-1}{2}}(\cos\theta_k)\, d\zeta_k(\theta_k) \\
&= \sum_{j=1}^r \omega_j^k (1-(x_k^k)^2)^{\mu_k} C_{\mu_{k-1}-\mu_k}^{\mu_k+\frac{m-k-1}{2}}(x_j^k) C_{\mu_{k-1}'-\mu_k}^{\mu_k+\frac{m-k-1}{2}}(x_j^k) \\
&= \frac{1}{\int_{-1}^1 (1-x^2)^{\frac{m-k-2}{2}}\, dx} \int_{-1}^1 (1-x^2)^{\mu_k+\frac{m-k-2}{2}} C_{\mu_{k-1}-\mu_k}^{\mu_k+\frac{m-k-1}{2}}(x) C_{\mu_{k-1}'-\mu_k}^{\mu_k+\frac{m-k-1}{2}}(x)\, dx \\
&= \frac{1}{\int_{-1}^1 (1-x^2)^{\frac{m-k-2}{2}}\, dx} \frac{\delta_{\mu_{k-1}\mu_{k-1}'}}{\tilde{\gamma}_{\mu_{k-1},\mu_k} \tilde{\gamma}_{\mu_{k-1}',\mu_k}},
\end{aligned}$$

since Gegenbauer polynomials $C_{\mu_{k-1}-\mu_k}^{\mu_k+\frac{m-k-1}{2}}(x)$ are orthogonal with respect to $(1-x^2)^{\mu_k+\frac{m-k-2}{2}}$ on interval $[-1,1]$ and

$$\begin{aligned}
&\int_{-1}^1 (1-x^2)^{\mu_k+\frac{m-k-2}{2}} \left[C_{\mu_{k-1}-\mu_k}^{\mu_k+\frac{m-k-1}{2}}(x) \right]^2 dx \\
&= \frac{\pi(\mu_{k-1}+\mu_k+m-k-2)!}{2^{2\mu_k+m-k-3}(\mu_{k-1}-\mu_k)!(2\mu_{k-1}+m-k-1)\Gamma^2\left(\mu_k+\frac{m-k-1}{2}\right)} = \frac{1}{\tilde{\gamma}_{\mu_{k-1},\mu_k}\tilde{\gamma}_{\mu_{k-1},\mu_k}}.
\end{aligned}$$

[see again Andrews et al. (1991) Corollary 6.8.4]. This implies (3.30) and by induction the system of equations (3.27) is established which completes the proof of the theorem. □

Example 3.1. To illustrate our approach we consider the dimension $m = 4$ and a series expansion of order $d = 4$. By Theorem 3.2 with $r = d + 1 = 5$ we have to consider the weight functions

$$a_1(x) = (1-x^2)^{1/2}, \quad a_2(x) = 1.$$

The corresponding Gegenbauer polynomials are given by

$$C_5^1(x) = 32x^5 - 32x^3 + 6x, \quad C_5^{1/2}(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x),$$

and we obtain the following discrete optimal design $\zeta_1^* \otimes \zeta_2^* \otimes \nu^*$ given by

$$\zeta_1^* = \begin{pmatrix} \frac{\pi}{6} & \frac{\pi}{3} & \frac{\pi}{2} & \frac{2\pi}{3} & \frac{5\pi}{6} \\ \frac{1}{12} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & \frac{1}{12} \end{pmatrix},$$

$$\zeta_2^* = \begin{pmatrix} \arccos(x_2) & \arccos(x_1) & \frac{\pi}{2} & \arccos(-x_1) & \arccos(-x_2) \\ \frac{322-13\sqrt{70}}{1800} & \frac{322+13\sqrt{70}}{1800} & \frac{64}{225} & \frac{322+13\sqrt{70}}{1800} & \frac{322-13\sqrt{70}}{1800} \end{pmatrix}, \quad (3.31)$$

$$\nu^* = \begin{pmatrix} -\frac{7}{9}\pi & -\frac{5}{9}\pi & -\frac{3}{9}\pi & -\frac{1}{9}\pi & \frac{1}{9}\pi & \frac{3}{9}\pi & \frac{5}{9}\pi & \frac{7}{9}\pi & \pi \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{pmatrix},$$

where $x_1 = \frac{1}{3}\sqrt{\frac{1}{7}(35 - 2\sqrt{70})}$ and $x_2 = \frac{1}{3}\sqrt{\frac{1}{7}(35 + 2\sqrt{70})}$. By Theorem 3.2 this design is Φ_p- and Φ_{E_s}-optimal.

We now compare the optimal design $\zeta_1^* \otimes \zeta_2^* \otimes \nu^*$ with two uniform designs $\hat{\zeta}_1 \otimes \hat{\zeta}_2 \otimes \hat{\nu}$ and $\tilde{\zeta}_1 \otimes \tilde{\zeta}_2 \otimes \tilde{\nu}$, where the marginal distributions of these designs are given by

$$\hat{\zeta}_1 = \begin{pmatrix} 0 & \frac{\pi}{4} & \frac{\pi}{2} & \frac{3\pi}{4} & \pi \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}, \hat{\zeta}_2 = \begin{pmatrix} 0 & \frac{\pi}{4} & \frac{\pi}{2} & \frac{3\pi}{4} & \pi \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}, \hat{\nu} = \nu^*, \quad (3.32)$$

and

$$\tilde{\zeta}_1 = \begin{pmatrix} \frac{\pi}{6} & \frac{\pi}{3} & \frac{\pi}{2} & \frac{2\pi}{3} & \frac{5\pi}{6} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix},$$

$$\tilde{\zeta}_2 = \begin{pmatrix} \arccos(x_2) & \arccos(x_1) & \frac{\pi}{2} & \arccos(-x_1) & \arccos(-x_2) \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}, \tilde{\nu} = \nu^*, \quad (3.33)$$

respectively. Note that the design $\hat{\zeta}_1 \otimes \hat{\zeta}_2 \otimes \hat{\nu}$ defined by (3.32) corresponds to a uniform distribution on a grid in $[0, \pi] \times [0, \pi] \times [-\pi, \pi]$, while the design $\tilde{\zeta}_1 \otimes \tilde{\zeta}_2 \otimes \tilde{\nu}$ in (3.33) is an equidistant version of the optimal design $\zeta_1^* \otimes \zeta_2^* \otimes \nu^*$. In particular, it uses the same support points as the optimal design.

To compare the uniform designs with the optimal design $\zeta_1^* \otimes \zeta_2^* \otimes \nu^*$ obtained by Theorem 3.2 we consider the efficiency

$$\mathrm{eff}(\zeta_1 \otimes \zeta_2 \otimes \nu) = \frac{\Phi(\zeta_1 \otimes \zeta_2 \otimes \nu)}{\Phi(\zeta_1^* \otimes \zeta_2^* \otimes \nu^*)},$$

where Φ is either the D-, E- or Φ_{E_s}-optimality criterion.

We focus on the estimation of $K^T \boldsymbol{c}$ where we fix $q = 1$, $k_0 = 0$ and $k_1 = 4$ and K is a block matrix of the form (3.2) with appropriate blocks given by (3.3). For the case of Φ_{E_s}-optimality we set $s = s(k_0) + s(k_1) = 26$. The D-, E- and Φ_{E_s}-efficiencies of the designs $\hat{\zeta}_1 \otimes \hat{\zeta}_2 \otimes \hat{\nu}$ and $\tilde{\zeta}_1 \otimes \tilde{\zeta}_2 \otimes \tilde{\nu}$ are presented in Table 1.

For the modified optimal design $\tilde{\zeta}_1 \otimes \tilde{\zeta}_2 \otimes \tilde{\nu}$ (with the same support points as the optimal design) we observe a good D-efficiency, however the Φ_{E_s}- and the E-efficiencies are substantially smaller (54.58% and 49.56%, repectively). The uniform design $\hat{\zeta}_1 \otimes \hat{\zeta}_2 \otimes \hat{\nu}$ performs worse with respect to the all considered criteria which shows that this uniform design is inefficient in applications.

	D-efficiency	E-efficiency	Φ_{E_s}-efficiency
$\hat{\zeta}_1 \otimes \hat{\zeta}_2 \otimes \hat{\nu}$	40.64	3.18	11.27
$\tilde{\zeta}_1 \otimes \tilde{\zeta}_2 \otimes \tilde{\nu}$	91.05	49.56	54.58

Table 1: *Efficiencies (in %) for the the uniform designs* $\hat{\zeta}_1 \otimes \hat{\zeta}_2 \otimes \hat{\nu}$ *and* $\tilde{\zeta}_1 \otimes \tilde{\zeta}_2 \otimes \tilde{\nu}$ *defined in* (3.32) *and* (3.33).

4 Symmetrized hyperspherical harmonics

In the previous example we have already shown that the use of the optimal designs yields a substantially more accurate statistical inference in series estimation with hyperspherical harmonics. In this section we consider a typical application of these functions (more precisely of linear combinations of hyperspherical harmonics) in material sciences and demonstrate some advantages of the new designs in this context. Due to space limitations we are not able to provide the complete background on the representations of crystallographic texture however, we explain the main ideas and refer to Bunge (1993), Mason and Schuh (2008) and Patala et al. (2012) for further explanation. Some helpful background with more details can also be found in the monograph of Marinucci and Peccati (2011).

Example 4.1. We begin with a brief discussion of the case $m = 2$ which - although not relevant for applications in material sciences - is very helpful for understanding the main idea behind the construction of symmetrized hyperspherical harmonics. In this case the Fourier basis

$$\Big\{\frac{1}{\sqrt{2}}, \cos(x), \sin(x), \cos(2x), \sin(2x), \ \ldots\Big\},$$

is a complete orthonormal system in the Hilbert space $L^2([0, 2\pi))$ with the common inner product $\langle f, g\rangle = \frac{1}{\pi}\int_0^{2\pi} f(x)g(x)dx$. The aim is now to construct an orthonormal basis for the subspace of functions in $L^2([0, 2\pi))$, which are invariant with respect to the rotation group $\{R_0, R_{\pi/2}, R_\pi, R_{3\pi/2}\}$ defined by

$$R_a : \begin{cases} [0, 2\pi) \to [0, 2\pi) \\ x \mapsto x + a \mod (2\pi) \end{cases} \quad (a \in \{0, \pi/2, \pi, 3/2\pi\}), \tag{4.1}$$

that is $f(\cdot) = f(R_a^{-1}(\cdot))$ (or equivalently $f(\cdot) = f(R_a(\cdot))$) for all $a \in \{0, \pi/2, \pi, 3/2\pi\}$. For this purpose consider the trigonometric polynomial

$$f(x) = \frac{a_0}{\sqrt{2}} + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \sin(kx) = \sum_{k=0}^{\infty} c_k^T \cdot Y_k(x) \ , \tag{4.2}$$

where the vectors c_k and $Y_k(x)$ are defined by

$$c_k = (a_k, b_k)^T \quad \text{and} \quad Y_k(x) = \begin{cases} (1/\sqrt{2}, 0)^T & k = 0, \\ (\cos(kx), \sin(kx))^T & \text{otherwise} \end{cases},$$

respectively, and assume that the function f is invariant with respect to the rotation group $\{R_0, R_{\pi/2}, R_\pi, R_{3\pi/2}\}$, that is,

$$\sum_{k=0}^{\infty} c_k^T \cdot Y_k(x) = f(x) = f(R_a(x)) = \sum_{k=0}^{\infty} c_k^T \cdot Y_k(R_a(x)) = \sum_{k=0}^{\infty} c_k^T \cdot D_k(a) \cdot Y_k(x),$$

where the matrices D_k are defined by

$$D_k(a) = \begin{pmatrix} \cos(ka) & -\sin(ka) \\ \sin(ka) & \cos(ka) \end{pmatrix},$$

where we have used the addition formulas for the trigonometric functions. This means that f is invariant under R_a if and only if c_k is an eigenvector for the eigenvalue 1 of $D_k(a)^T$. Because $\{R_0, R_{\pi/2}, R_\pi, R_{3\pi/2}\}$ is generated by $R_{\pi/2}$, it suffices to consider the case $a = \pi/2$. It is now easy to see that only the matrices $D_{4\ell}(\pi/2)^T = I_2$ have the eigenvalue 1, which is of multiplicity 2 with corresponding eigenvectors $c_{4\ell} = (\beta, 0)^T$, $\tilde{c}_{4\ell} = (0, \gamma)^T$ $(\beta, \gamma \in \mathbb{R} \setminus \{0\})$. Consequently, a complete orthonormal basis of the subset of all functions in $L^2([0, 2\pi))$, which are invariant with respect to the rotation group $\{R_0, R_{\pi/2}, R_\pi, R_{3\pi/2}\}$, is obtained by choosing $\beta = \gamma = 1$, which yields the linear combinations $\{c_{4\ell}^T Y_{4\ell}(x), \tilde{c}_{4\ell}^T Y_{4\ell}(x)\}_{\ell=0,1,\dots}$ given by

$$\Big\{\frac{1}{\sqrt{2}},\ \cos(4x),\ \sin(4x),\ \cos(8x),\ \sin(8x), \dots\Big\}.$$

In applications in material sciences the dimension is $m = 4$ and the groups under consideration are much more complicated and induce crystal symmetries. For example, Mason and Schuh (2008) define representations of crystallographic textures as quaternion distributions (this corresponds to the case $m = 4$ in our notation) by series expansions in terms of hyperspherical harmonics to reflect sample and crystal symmetries such that the resulting expansions are more efficient. For this purpose they define the symmetrized hyperspherical harmonics as specific linear combinations of real hyperspherical harmonics which remain invariant under rotations corresponding to the simultaneous application of a crystal symmetry and sample symmetry operation. The exact definition of the symmetrized hyperspherical harmonics is complicated and requires sophisticated arguments from representation theory [see Sections 2 - 4 in Mason and Schuh (2008)], but - in principle - it follows essentially the same arguments as described in Example 4.1.

More precisely, the groups induced by the crystal symmetry, sample symmetry operation and the level of resolution λ, define $N(\lambda)$ *symmetrized hyperspherical harmonics* of the form

$$\dot{\dot{Z}}_{\lambda}^{\eta}(\theta_1,\theta_2,\phi) = \sum_{\mu_1=0}^{\lambda} \sum_{\mu_2=-\mu_1}^{\mu_1} \alpha_{\lambda,\mu_1,\mu_2}^{\eta} Y_{\lambda,\mu_1,\mu_2}(\theta_1,\theta_2,\phi), \quad \eta = 1,\ldots,N(\lambda) \ , \tag{4.3}$$

where the coefficients $\alpha_{\lambda,\mu_1,\mu_2}^{\eta}$ are well defined and can be determined form the symmetry properties. A list of the first at least 30 symmetrized hyperspherical harmonics polynomials for the different 11 point groups can be found in the online supplement of Mason and Schuh (2008). If the coefficients are standardized appropriately, the symmetrized hyperspherical harmonics also form a complete orthonormal system, that is,

$$\int_{-\pi}^{\pi}\int_{0}^{\pi}\int_{0}^{\pi} \dot{\dot{Z}}_{\lambda}^{\eta}(\theta_1,\theta_2,\phi)\dot{\dot{Z}}_{\lambda'}^{\eta'}(\theta_1,\theta_2,\phi)\sin^2\theta_1\sin\theta_1 d\theta_1 d\theta_2 d\phi = \delta_{\lambda\lambda'}\delta_{\eta\eta'} \ .$$

Moreover, any square integrable function g that satisfies the same requirement of crystal and sample symmetry can be uniquely represented as a linear combination of these symmetrized hyperspherical harmonics in the form

$$g(\theta_1,\theta_2,\phi) = \sum_{\lambda=0,2,\ldots,\infty} \sum_{\eta=1}^{N(\lambda)} c_{\lambda,\eta} \dot{\dot{Z}}_{\lambda}^{\eta}(\theta_1,\theta_2,\phi). \tag{4.4}$$

Patala et al. (2012) obtained estimates of the missorientation distribution function by fitting experimentally measured missorientation data to a linear combination of symmetrized hyperspherical harmonics, while Patala and Schuh (2013) used truncated series to obtain estimates of the grain boundary distribution from simulated data As these experiments are very expensive and the simulations are very time consuming it is of particular importance to obtain good designs for the estimation by series of hypershperical harmonics. Therefore, we now consider the linear regression model (2.1) where the vector of regression functions is obtained by the truncated expansion of the function g of order d, that is,

$$\sum_{\lambda=0,2,\ldots,d} \sum_{\eta=1}^{N(\lambda)} c_{\lambda,\eta} \dot{\dot{Z}}_{\lambda}^{\eta}(\theta_1,\theta_2,\phi), \tag{4.5}$$

and investigate the performance of the designs determined in Section 3 in models of the form (4.5). Due to space restrictions we concentrate on the case $d = 4$ and on the symmetrized hyperspherical harmonics for samples with orthorhombic symmetry and crystal symmetry corresponding to the crystallographic point groups 1 and 2. Similar results for expansions of higher order and different crystallographic point groups can be obtained following along the same lines.

For the crystallographic point group 1 there are 11 symmetrized hyperspherical harmonics up to order $d = 4$ which can be obtained from the online supplement of Mason and Schuh (2008) and are given by

$$\begin{aligned}
\dot{\ddot{Z}}_0^1 &= Y_{0,0,0} \\
\dot{\ddot{Z}}_4^1 &= \sqrt{\tfrac{2}{5}}\quad Y_{4,0,0} + \sqrt{\tfrac{7}{20}}\quad Y_{4,4,0} + \sqrt{\tfrac{1}{4}}\quad Y_{4,4,4} \\
\dot{\ddot{Z}}_4^2 &= \sqrt{\tfrac{2}{5}}\quad Y_{4,1,0} - \sqrt{\tfrac{1}{10}}\quad Y_{4,3,0} - \sqrt{\tfrac{1}{2}}\quad Y_{4,4,-4} \\
\dot{\ddot{Z}}_4^3 &= \sqrt{\tfrac{2}{5}}\quad Y_{4,1,1} + \sqrt{\tfrac{3}{80}}\quad Y_{4,3,1} - \sqrt{\tfrac{1}{16}}\quad Y_{4,3,3} + \sqrt{\tfrac{7}{16}}\quad Y_{4,4,-1} + \sqrt{\tfrac{1}{16}}\quad Y_{4,4,-3} \\
\dot{\ddot{Z}}_4^4 &= \sqrt{\tfrac{2}{5}}\quad Y_{4,1,-1} + \sqrt{\tfrac{3}{80}}\quad Y_{4,3,-1} + \sqrt{\tfrac{1}{16}}\quad Y_{4,3,-3} - \sqrt{\tfrac{7}{16}}\quad Y_{4,4,1} + \sqrt{\tfrac{1}{16}}\quad Y_{4,4,3} \\
\dot{\ddot{Z}}_4^5 &= \sqrt{\tfrac{4}{7}}\quad Y_{4,2,0} + \sqrt{\tfrac{5}{28}}\quad Y_{4,4,0} - \sqrt{\tfrac{1}{4}}\quad Y_{4,4,4} \qquad (4.6) \\
\dot{\ddot{Z}}_4^6 &= \sqrt{\tfrac{2}{7}}\quad Y_{4,2,1} - \sqrt{\tfrac{5}{16}}\quad Y_{4,3,-1} + \sqrt{\tfrac{3}{16}}\quad Y_{4,3,-3} + \sqrt{\tfrac{3}{112}}\quad Y_{4,4,1} + \sqrt{\tfrac{3}{16}} Y_{4,4,3} \\
\dot{\ddot{Z}}_4^7 &= \sqrt{\tfrac{2}{7}}\quad Y_{4,2,-1} + \sqrt{\tfrac{5}{16}}\quad Y_{4,3,1} + \sqrt{\tfrac{3}{16}}\quad Y_{4,3,3} + \sqrt{\tfrac{3}{112}}\quad Y_{4,4,-1} - \sqrt{\tfrac{3}{16}} Y_{4,4,-3} \\
\dot{\ddot{Z}}_4^8 &= \sqrt{\tfrac{4}{7}}\quad Y_{4,2,2} - \sqrt{\tfrac{3}{7}}\quad Y_{4,4,2} \\
\dot{\ddot{Z}}_4^9 &= \sqrt{\tfrac{2}{7}}\quad Y_{4,2,-2} - \sqrt{\tfrac{1}{2}}\quad Y_{4,3,2} - \sqrt{\tfrac{3}{14}}\quad Y_{4,4,-2} \\
\dot{\ddot{Z}}_4^{10} &= Y_{4,3,-2}.
\end{aligned}$$

Note that the 26 functions $(Y_{0,0,0}, Y_{4,0,0}, \ldots, Y_{4,3,3}, Y_{4,4,-4}, \ldots, Y_{4,4,4})^T$ define 11 symmetrized hyperspherical harmonics. Consequently, considering the symmetries of the crystallographic group 1 the vector of regression functions in model (2.1) is of the form

$$f_1^T = (\dot{\ddot{Z}}_0^1, \dot{\ddot{Z}}_4^1, \ldots, \dot{\ddot{Z}}_4^{10})^T. \qquad (4.7)$$

To illustrate the symmetries induced by the crystallographic group in the symmetrized hyperspherical harmonics we use a visualization described by Mason and Schuh (2008). For a fixed hyperangle $(\theta_1, \theta_2, \phi)$, the functional value of $\dot{\ddot{Z}}_4^j(\theta_1, \theta_2, \phi)$ is presented by using a projection of the hyperangle to an appropriate two-dimensional disk. More precisely, we project the hyperangle $(\theta_1, \theta_2, \phi)$ onto a two-dimensional disk by

$$P(\theta_1, \theta_2, \phi) = \begin{pmatrix} x_1(\theta_1, \theta_2, \phi) \\ x_2(\theta_1, \theta_2, \phi) \end{pmatrix} = \begin{pmatrix} R(\theta_1, \theta_2)\cos(\phi) \\ R(\theta_1, \theta_2)\sin(\phi) \end{pmatrix}, \qquad (4.8)$$

where the function $R(\theta_1, \theta_2)$ is given by

$$R(\theta_1, \theta_2) = (3/2)^{1/3}(\theta_1 - \sin(\theta_1)\cos(\theta_1))^{1/3}\sqrt{2(1 - |\cos(\theta_2)|)}.$$

For instance the angle $(\theta_1, \theta_2, \phi) = (\frac{11}{48}\pi, \frac{\pi}{4}, \pi)$ is projected onto the point $(x_1, x_2) = (-0.5323, 0)$. In Figure 4 we display the value (represented by an appropriate color)

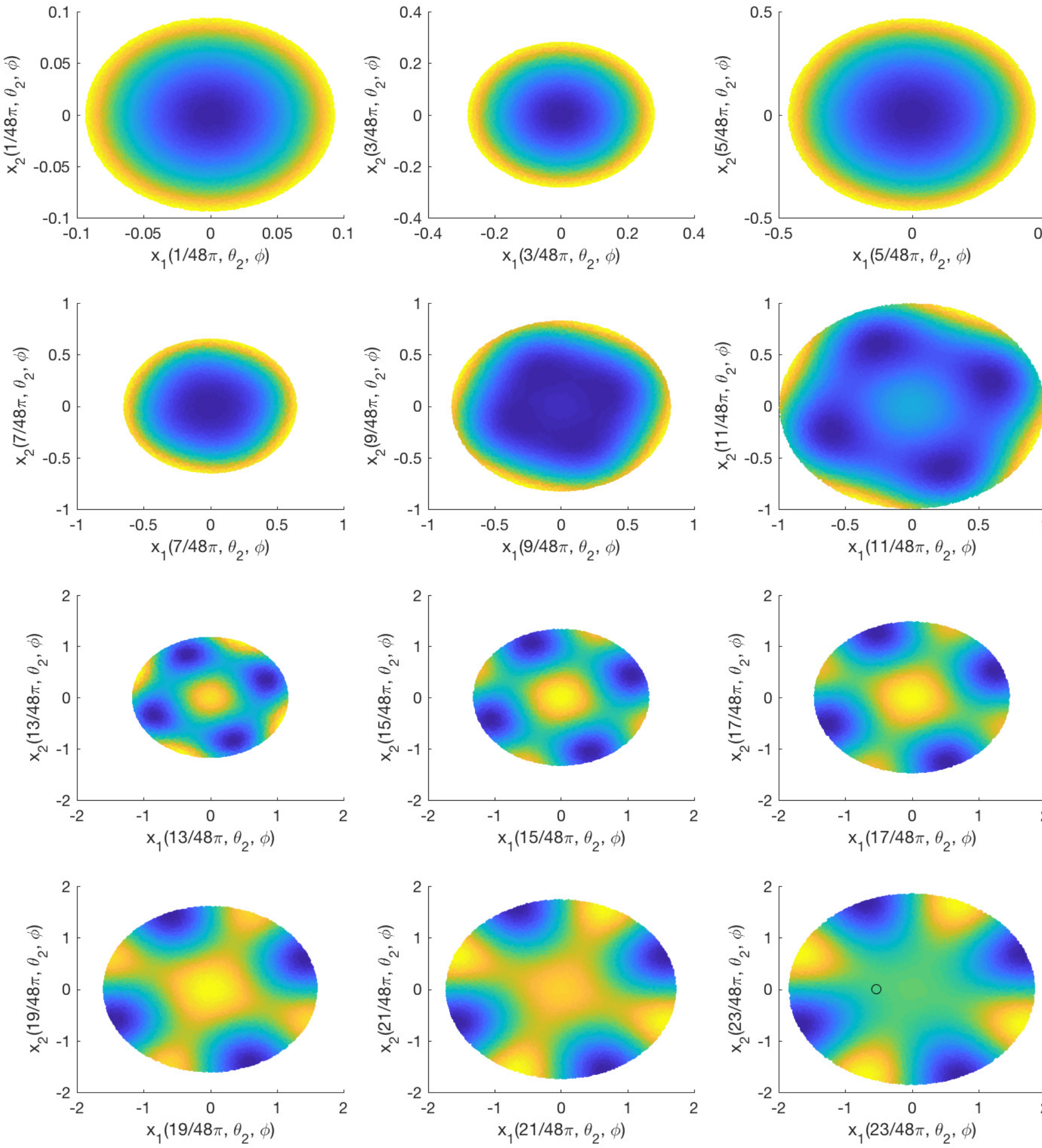

Figure 1: *Visualization of the symmetrized hyperspherical harmonic* $\dot{Z}_4^2$ *of the crystallo-grahic point group* 1 *using the projection of the hyperangles onto the two-dimensional disk given by* (4.8)*. For each panel,* θ_1 *is fixed to a value in* $\{\frac{1}{48}\pi, \frac{3}{48}\pi, \ldots, \frac{11}{48}\pi\}$*, while* (θ_2, ϕ) *vary in* $[0, \frac{\pi}{2}] \times [-\pi, \pi]$.

of the symmetrized harmonic $\dot{Z}_4^2$ of the crystallographic group 1 as a function of the coordinates $(x_1(\theta_1, \theta_2, \phi), x_2(\theta_1, \theta_2, \phi))$. In each of the twelve panels of Figure 4, θ_1 is fixed to one of the values $\frac{1}{48}\pi, \frac{3}{48}\pi, \ldots, \frac{11}{48}\pi$, while the angles θ_2 and ϕ vary between $[0, \pi/2]$ and $[-\pi, \pi]$, respectively. For instance, the value of $\dot{Z}_4^j(\theta_1, \theta_2, \phi)$ at the hyperangle $(\theta_1, \theta_2, \phi) = (\frac{11}{48}\pi, \frac{\pi}{4}, \pi)$ is presented in the bottom right panel in a light green color (see the black circle in the bottom right panel of Figure 4).

We now investigate the efficiency of the optimal design for the estimation of the coefficients in the regression model (2.1) with the hyperspherical harmonics up to order $d = 4$, that is, the vector of regression functions is given by

$$f^T = (Y_{0,0,0}, Y_{1,0,0}, \ldots, Y_{1,1,1}, \ldots, Y_{4,4,-4}, \ldots, Y_{4,4,4})^T.$$

The optimal design for this model has been determined in Example 3.1 and a tedious calculation shows that the design $\zeta_1^* \otimes \zeta_2^* \otimes \nu^*$ defined in (3.31) satisfies the general equivalence theorem in Section 7.20 of Pukelsheim (2006). Consequently, this design is also Φ_p-optimal in the regression model (2.1), where the vector of regression functions is given by the symmetrized hyperspherical harmonics defined in (4.7), which correspond to the crystallographic point group 1.
For the crystallographic point group 2 there are 7 symmetrized hyperspherical harmonics up to order $d = 4$ consisting of a subset of the functions given in (4.6). These symmetrized hyperspherical harmonics define a linear regression model of the form (2.1), where the vector of regression functions f_2 is given by

$$f_2^T = (\dot{Z}_0^1, \dot{Z}_4^1, \dot{Z}_4^2, \dot{Z}_4^5, \dot{Z}_4^8, \dot{Z}_4^9, \dot{Z}_4^{10})^T. \tag{4.9}$$

In the case of the crystallographic point group 2 the design $\zeta_1^* \otimes \zeta_2^* \otimes \nu^*$ defined by (3.31) is not Φ_p-optimal. However, using particle swarm optimization [see Clerc (2006) for details] we determined the D-efficiency of the design $\zeta_1^* \otimes \zeta_2^* \otimes \nu^*$ numerically which is given by 81%. We also investigate the performance of the designs $\hat{\zeta}_1 \otimes \hat{\zeta}_2 \otimes \hat{\nu}$ and $\tilde{\zeta}_1 \otimes \tilde{\zeta}_2 \otimes \tilde{\nu}$ defined in equation (3.32) and (3.33) of Example 3.1. The D-efficiencies of these two designs are given by 59.38% and 74.59%, respectively. Recall that the latter design uses the same support points as the optimal design. Our calculations show that the design $\zeta_1^* \otimes \zeta_2^* \otimes \nu^*$ in (3.31) provides reasonable efficiencies for estimating the coefficients in the regression model (2.1) with symmetrized hyperspherical harmonics with respect to the crystallographic point group 2, whereas the uniform design $\hat{\zeta}_1 \otimes \hat{\zeta}_2 \otimes \hat{\nu}$ should not be used in this case.

A A Technical Result

A.1 Proof of Lemma 3.1

Assume that conditions (A) and (B) are satisfied and let $Q(x)$ be an arbitrary polynomial of degree z. The polynomial Q can be represented in the form

$$Q(x) = P(x)V_r(x) + R(x),$$

where $V_r(x) = \prod_{j=1}^r (x - x_j)$ is of degree r, the polynomial $P(x)$ is of degree $z - r$ and the polynomial $R(x)$ is of degree less than r. Since $x_1, \dots, x_r$ are the zeros of $V_r(x)$ we have that $Q(x_j) = R(x_j)$ for all $j = 1, \dots, r$ and furthermore, because the degree of $R(x)$ is at most $r - 1$, it can be represented as

$$R(x) = \sum_{j=1}^r \ell_j(x) R(x_j).$$

Then from conditions (A) and (B) we obtain that

$$\begin{aligned} \int_{-1}^1 a(x)Q(x)\,dx &= \int_{-1}^1 a(x)R(x)\,dx = \sum_{j=1}^r R(x_j) \int_{-1}^1 a(x)\ell_j(x)\,dx \\ &= \sum_{j=1}^r R(x_j)\tilde{a}\omega_j = \tilde{a} \sum_{j=1}^r \omega_j Q(x_j). \end{aligned}$$

Using $Q(x) = x^\ell$, $\ell = 0, \dots, z$ yields the identities in (3.16) and for $\ell = 0$ we obtain the expression $\tilde{a} = \int_{-1}^1 a(x)\,dx$.

Now assume that (3.16) is valid. For $\ell = 0, \dots, z - r$ we have that

$$\int_{-1}^1 V_r(x)a(x)x^\ell\,dx = \tilde{a} \sum_{j=1}^r \omega_j V_r(x_j) x_j^\ell = 0,$$

which gives condition (A). By noting that $\ell_j(x_k) = \delta_{jk}$ we get that

$$\frac{1}{\tilde{a}} \int_{-1}^1 a(x)\ell_j(x)\,dx = \sum_{k=1}^r \omega_k \ell_j(x_k) = \sum_{k=1}^r \omega_k \delta_{jk} = \omega_j,$$

which gives condition (B).

Acknowledgements

This work has been supported in part by the Collaborative Research Center "Statistical modelling of nonlinear dynamic processes" (SFB 823, Teilprojekt A1, C2) of the German

Research Foundation (DFG). H. Dette was partially supported by a grant from the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM107639. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors would like to thank Rebecca Janisch for providing some background information about crystallographic textures.

References

Abrial, P., Moudden, Y., Starck, J.-L., Fadili, J., Delabrouille, J., and Nguyen, M. K. (2008). CMB data analysis and sparsity. *Statistical Methodology*, 5(4):289–298.

Alfed, P., Neamtu, M., and Schumaker, L. L. (1996). Fitting scattered data on sphere-like surfaces using spherical splines. *Journal of Computational and Applied Mathematics*, 73(1-2):5–43.

Andrews, G., Askey, R., and Roy, R. (1991). *Special Functions.* Cambrdge Univ. Press.

Avery, J. and Wen, Z.-Y. (1982). Angular integrations in m-dimensional spaces and hyperspherical harmonics. *International Journal of Quantum Chemistry*, 22:717–738.

Baldi, P., Kerkyacharian, G., Marinucci, D., and Picard, D. (2009). Asymptotics for spherical needlets. *Annals of Statistics*, 37(3):1150–1171.

Bunge, H. J. (1993). *Texture Analysis in Material Science: Mathematical Methods, 1st. ed.* Cuvillier Verlag, Göttingen.

Chang, T., Ko, D., Royer, J.-Y., and Lu, J. (2000). Regression techniques in plate tectonics. *Statistical Science*, 15(4):342–356.

Chapman, G. R., Chen, C., and Kim, P. T. (1995). Assessing geometric integrity through spherical regression techniques. *Statistica Sinica*, 5:173–220.

Cheng, C. S. (1987). An application of the Kiefer-Wolfowitz equivalence theorem to a problem in Hadamard transform optics. *Annals of Statistics*, 15(4):1593–1603.

Clerc, M. (2006). *Particle Swarm Optimization.* Iste Publishing Company, London.

Dette, H. and Melas, V. B. (2003). Optimal designs for estimating individual coefficients in Fourier regression models. *Annals of Statistics*, 31(5):1669–1692.

Dette, H., Melas, V. B., and Pepelyshev, A. (2005). Optimal designs for 3D shape analysis with spherical harmonic descriptors. *Annals of Statistics*, 33:2758–2788.

Dette, H. and Studden, W. J. (1993). Geometry of e-optimality. *Annals of Statistics*, 21(1):416–433.

Dette, H. and Wiens, D. P. (2009). Robust designs for 3D shape analysis with spherical harmonic descriptors. *Statistica Sinica*, 19:83–102.

Di Marzio, M., Panzera, A., and Taylor, C. C. (2009). Local polynomial regression for circular predictors. *Statistics & Probability Letters*, 79(1):2066–2075.

Di Marzio, M., Panzera, A., and Taylor, C. C. (2014). Nonparametric regression for spherical data. *Journal of the American Statistical Association*, 109(506):748–763.

Dokmanić, I. and Petrinović, D. (2010). Convolution on the n-sphere with application to PDF modeling. *IEEE Transactions on Signal Processing*, 58(3):1157–1170.

Fedorov, V. V. (1972). *Theory of Optimal Experiments.* Academic Press, New York.

Filov, L., Harman, R., and Klein, T. (2011). Approximate e-optimal designs for the model of spring balance weighing with a constant bias. *Journal of Statistical Planning and Inference*, 141(7):2480 – 2488.

Genovese, C. R., Miller, C. J., Nichol, R. C., Arjunwadkar, M., and Wasserman, L. (2004). Nonparametric inference for the cosmic microwave background. *Statistical Science*, 19(2):308–321.

Harman, R. (2004). Minimal efficiency of designs under the class of orthogonally invariant information criteria. *Metrika*, 60(2):137–153.

Hosseinbor, A. P., Chung, M. K., Schaefer, S. M., van Reekum, C. M., Peschke-Schmitz, L., Sutterer, M., Alexander, A. L., and Davidson, R. J. (2013). 4D hyperspherical harmonic (HyperSPHARM) representation of multiple disconnected brain subcortical structures. *Medical Image Computing and Computer-Assisted Intervention*, 16:598–605.

Karlin, S. and Studden, W. J. (1966). *Tchebysheff Systems: With Application in Analysis and Statistics*. Wiley, New York.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, 2(5):849–879.

Kitsos, C. P., Titterington, D. M., and Torsney, B. (1988). An optimal design problem in rhythmometry. *Biometrics*, 44:657–671.

Krivec, R. (1998). Hyperspherical-harmonics methods for few-body problems. *Few-Body Systems*, 25:199–238.

Lau, T.-S. and Studden, W. J. (1985). Optimal designs for trigonometric and polynomial regression using canonical moments. *Annals of Statistics*, 13(1):383–394.

Lombardi, A., Palazzetti, F., Aquilanti, V., Grossi, G., Albernaz, A. F., Barreto, P. R. P., and Cruz, A. C. P. S. (2016). Spherical and hyperspherical harmonic representation of van der Waals aggregates. *International Conference of Computational Methods in Sciences and Engineering*, 1790(1):020005.

Marinucci, D. and Peccati, G. (2011). *Random Fields on the Sphere: Representation, Limit Theorems and Cosmological Applications*. London Mathematical Society Lecture Note Series. Cambridge University Press.

Mason, J. K. (2009). The relationship of the hyperspherical harmonics to $SO(3)$, $(SO4)$ and orientation distribution functions. *Acta Crystallographica Sec. A.*, 65:259–266.

Mason, J. K. and Schuh, C. A. (2008). Hyperspherical harmonics for the representation of crystallographic texture. *Acta Materialia*, 56(20):6141–6155.

Meremianin, A. V. (2009). Hyperspherical harmonics with arbitrary arguments. *Journal of Mathematical Physics*, 50(1):013526.

Monnier, J.-B. (2011). Nonparametric regression on the hyper-sphere with uniform design. *TEST*, 20(2):412–446.

Narcowich, F. J., Petrushev, P., and Ward, J. D. (2006). Localized tight frames on spheres. *SIAM Journal on Mathematical Analysis*, 38:574–594.

Patala, S., Mason, J. K., and Schuh, C. A. (2012). Improved representations of misorientation information for grain boundary science and engineering. *Progress in Materials Science*, 57(8):1383 – 1425.

Patala, S. and Schuh, C. A. (2013). Representation of single-axis grain boundary functions. *Acta Materialia*, 61(8):3068 – 3081.

Pukelsheim, F. (2006). *Optimal Design of Experiments*. SIAM, Philadelphia.

Pukelsheim, F. and Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika*, 79:763–770.

Schaeben, H. and van den Boogaart, K. G. (2003). Spherical harmonics in texture analysis. *Tectonophysics*, 370:253–268.

Shin, H. H., Takahara, G. K., and Murdoch, D. J. (2007). Optimal designs for calibration of orientations. *Canadian Journal of Statistics*, 35(3):365–380.

Silvey, S. D. (1980). *Optimal design*. Chapman and Hall, London, London-New York.

Vilenkin, N. J. (1968). *Special Functions and the Theory of Group Representations*, volume 22. Translations of Mathematical Monographs, American Mathematical Society, Providence, RI.

Wahba, G. (1981). Spline interpolation and smoothing on the sphere. *SIAM Journal on Scientific and Statistical Computing*, 2(1):5–16.

Wen, Z.-Y. and Avery, J. (1985). Some properties of hyperspherical harmonics. *Math. Phys.*, 29:396–403.

Zheng, A., Doerschuk, P. C., and Johnson, J. E. (1995). Determination of three-dimensional low-resolution viral structure from solution x-ray scattering data. *Biophysical Journal*, 69(2):619–639.