

Sensitivity Analysis on the Sphere and a Spherical ANOVA Decomposition

✉ Laura Weidensager
Faculty of Mathematics
Chemnitz University of Technology
D-09107 Chemnitz, Germany
`laura.weidensager@math.tu-chemnitz.de`

December 29, 2025

Abstract

We establish sensitivity analysis on the sphere. We present formulas that allow us to decompose a function $f: \mathbb{S}^d \rightarrow \mathbb{R}$ into a sum of terms $f_{\mathbf{u}, \boldsymbol{\xi}}$. The index $\mathbf{u}$ is a subset of $\{1, 2, \dots, d+1\}$, where each term $f_{\mathbf{u}, \boldsymbol{\xi}}$ depends only on the variables with indices in $\mathbf{u}$. In contrast to the classical analysis of variance (ANOVA) decomposition, we additionally use the decomposition of a function into functions with different parity, which adds the additional parameter $\boldsymbol{\xi}$. The natural geometry on the sphere naturally leads to the dependencies between the input variables. Using certain orthogonal basis functions for the function approximation, we are able to model high-dimensional functions with low-dimensional variable interactions.

Keywords – ANOVA decomposition, high-dimensional sphere, orthogonal polynomials

Mathematics Subject Classification (2020) – 41A10, 41A63, 65D15, 65D40

1 Introduction

Functions on the sphere arise in diverse applications, ranging from geophysics to quantum mechanics. To efficiently analyze and approximate such functions, a deep understanding of their internal structure is essential. The ANOVA (analysis of variance) decomposition provides a remarkably elegant framework for understanding the structure of multivariate functions. In the classical setting, where the domain is a tensor product space, the decomposition exhibits a particularly transparent form: component functions corresponding to different variable subsets are orthogonal, and the hierarchical interaction structure can be characterized cleanly. However, when the domain is the high-dimensional sphere, this clarity gets lost. The lack of a natural product structure complicates both the definition and computation of ANOVA components, making the extension of classical results to spherical domains highly nontrivial.

In this paper, we develop a spherical ANOVA decomposition that allows us to split the total variance of the function into contributions of specific groups of variables. This approach not only yields valuable insights into variable importance but also forms the theoretical foundation for adaptive approximation schemes.

The key difference to the classical decomposition is that we do not have a tensor product structure of the domain here, which makes modeling more difficult and forces the introduction of additional parameters and structures. We found that the decomposition of the function into terms of different parity is a key ingredient for the spherical ANOVA decomposition.

Our main contributions

- We propose a spherical ANOVA decomposition for the uniform distribution on the sphere $\mathbb{S}^d$. Due to the structure of the sphere, we have to additionally distinguish between even and odd terms and therefore introduce a parity vector.

- We introduce reduced orthogonal bases to describe the lower dimensional terms, which we use to implement high-dimensional function approximation on the sphere. Additionally, we define Sobol indices on the sphere in order to apply sensitivity analysis for the spherical ANOVA decomposition.
- We show in numerical examples that the spherical ANOVA decomposition captures the expected non-zero spherical ANOVA terms, similar to how one would expect the terms from the classical ANOVA decomposition.

However, extending the ANOVA decomposition from a tensor product domain to the sphere leads to theoretical challenges arising from the geometry of the domain. In Section 2.4, we recall this classical ANOVA decomposition. The intrinsic curvature of the sphere automatically induces variable interactions. Even if the function has a simple structure in cartesian coordinates, the manifold’s constraints enforce coupling between variables, effectively creating artificial interactions. Second, the natural algebraic dependence of the variables (i.e., $\|\mathbf{x}\| = 1$) implies that the inputs cannot vary independently. Consequently, it is generally impossible to construct a decomposition consisting of strictly orthogonal terms in the classical sense.

The analysis of functions on the sphere, the ball and the simplex are related. Here, we focus on the analysis of the sphere, but the key idea can be transferred to the unit ball and the simplex. Due to the different geometry on the unit ball and the simplex, the dependencies between the ANOVA terms must be considered carefully. So, we leave the considerations of our approach on the ball and the simplex open for future work.

Related work

Several alternative approaches for dimension reduction and approximation on the sphere exist. Geometric approaches such as those described in [4] attempt to approximate the data using optimal (rotated) submanifolds. Global polynomial approximation methods for scattered data typically require a number of samples exponential in d , as described in [9].

Employing the Dirac measure for high-dimensional approximation leads to the anchored ANOVA decomposition, see for example [18]. Due to the geometric properties on the sphere, it is not possible to generalize this decomposition to the sphere.

In the context of sensitivity analysis, quantifying the influence of groups of variables remains a challenging task when inputs are correlated. As recently discussed by [19], defining the explicit contribution of subsets of variables requires careful consideration in the presence of dependence. Recent advances in global sensitivity analysis for dependent variables also include kernel-embedding techniques and Shapley effects, [5] as well as random Fourier features, [28].

To date, no dedicated ANOVA decomposition for the sphere has been established in the literature. The approach proposed in this work fills this gap and aims to identify the importance of the original physical variables and their specific interactions.

Structure of this work

The next section is a preliminary in which we recall a collection of known concepts, definitions and variable substitutions that we will use in this paper. We generalize the projection operator from the classical ANOVA decomposition in Section 3.1. With this, we introduce an ANOVA operator in Section 3.2. We show properties of this operator, related to the parity, in Section 3.3.

In Section 3.4, we finally introduce the spherical ANOVA decomposition, for which we show orthogonality and redundancy relations in Section 3.5. In Section 3.6, we highlight the problems that naturally arise with even functions and show how we can fix them. In Section 4, we introduce the orthogonal basis on the lower dimensional balls, which we use to describe the spherical ANOVA terms. How we proceed numerically is described in more detail in Section 5 and finally in Section 6 we apply our algorithms to some test functions to calculate the spherical ANOVA decomposition of these functions.

Notation

We introduce some notation, which we will use in this paper. We use bold face for vectors and matrices.

The parameter d is always the ambient dimension. We denote the sphere and ball of radius r by

$$\begin{aligned}\mathbb{S}_r^d &= \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\|_2 = r\}, \\ \mathbb{B}_r^{d+1} &= \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\|_2 \leq r\}.\end{aligned}$$

As usual, we omit the parameter r for the unit sphere $\mathbb{S}^d$ and unit ball $\mathbb{B}^{d+1}$. Let the surface measure on $\mathbb{S}^d$ be denoted by μ_d and we assume it is normalized such that

$$\int_{\mathbb{S}^d} d\mu_d(\mathbf{x}) = \frac{2\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})} =: \omega_d.$$

Furthermore, $\mathbb{S}^0 := \{-1, 1\}$. Let $[d+1] = \{1, 2, \dots, d+1\}$. For every index $\mathbf{u} \subseteq [d+1]$ we denote the corresponding partial vector of the vector $\mathbf{x}$ by $\mathbf{x}_{\mathbf{u}} = (x_i)_{i \in \mathbf{u}}$. For every subset $\mathbf{u} \subset [d+1]$ we denote the complement by $\mathbf{u}^c := [d+1] \setminus \mathbf{u}$. We compose a vector from 2 (or more) parts by writing $\mathbf{c} = (\mathbf{a}_{\mathbf{u}}, \mathbf{b}_{\mathbf{u}^c})$, which means for the vector $\mathbf{c}$ that $\mathbf{c}_{\mathbf{u}} = \mathbf{a}_{\mathbf{u}}$ and $\mathbf{c}_{\mathbf{u}^c} = \mathbf{b}_{\mathbf{u}^c}$. We use the usual notation for L_2 -spaces and the corresponding scalar products on the domain $\mathbb{D} \in \{\mathbb{B}^{d+1}, \mathbb{S}^d\}$,

$$\begin{aligned}\|f\|_{L_2(\mathbb{D})} &= \left(\int_{\mathbb{D}} |f(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2}, \quad L_2(\mathbb{D}) = \{f: \mathbb{D} \rightarrow \mathbb{R} \mid \|f\|_{L_2(\mathbb{D})} < \infty\}, \\ \langle f, g \rangle_{\mathbb{D}} &= \int_{\mathbb{D}} f(\mathbf{x})g(\mathbf{x})d\mathbf{x}.\end{aligned}$$

2 Preliminaries

In this section we collect basic concepts and results which we will use later to construct a spherical ANOVA decomposition.

2.1 Variable substitutions on the sphere $\mathbb{S}^d$ and the ball $\mathbb{B}^d$

We collect some substitutions which we will use later here. Splitting the integral over the ball into two integrals over the sphere and the radius by using polar coordinates leads to the following basic lemma.

Lemma 2.1. *For a function $g: \mathbb{B}^d \rightarrow \mathbb{R}$ the integral on the d -dimensional ball can be decomposed into*

$$\int_{\mathbb{B}^d} g(\mathbf{x}) d\mathbf{x} = \int_0^1 \int_{\mathbb{S}^{d-1}} g(r\mathbf{z}) r^{d-1} d\mu_{d-1}(\mathbf{z}) dr.$$

2.1.1 Variable substitutions on the sphere

Let us decompose the surface measure μ_d by the following. For every $\mathbf{x} \in \mathbb{S}^d$ and $\emptyset \neq \mathbf{u} \subset [d+1]$ we can write as $1 = \|\mathbf{x}_{\mathbf{u}}\|^2 + \|\mathbf{x}_{\mathbf{u}^c}\|^2$. From [6, Lemma A.5.4.] and Lemma 2.1 follows for $|\mathbf{u}| \geq 2$ and $\mathbf{x}_{\mathbf{u}} \in \mathbb{S}^{|\mathbf{u}|-1}$,

$$\begin{aligned}\int_{\mathbb{S}^{|\mathbf{v}|+|\mathbf{u}|-1}} g(\mathbf{x}) d\mu_{|\mathbf{v}|+|\mathbf{u}|-1}(\mathbf{x}) &= \int_{\mathbb{B}^{|\mathbf{v}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{v}}\|^2}^{|\mathbf{u}|-2} \int_{\mathbb{S}^{|\mathbf{u}|-1}} g\left(\mathbf{x}_{\mathbf{v}}, \sqrt{1 - \|\mathbf{x}_{\mathbf{v}}\|^2} \mathbf{x}_{\mathbf{u}}\right) d\mu_{|\mathbf{u}|-1}(\mathbf{x}_{\mathbf{u}}) d\mathbf{x}_{\mathbf{v}} \\ &= \int_0^1 \int_{\mathbb{S}^{|\mathbf{v}|-1}} \sqrt{1 - r^2}^{|\mathbf{u}|-2} r^{|\mathbf{v}|-1} \int_{\mathbb{S}^{|\mathbf{u}|-1}} g(r\mathbf{z}_{\mathbf{v}}, \sqrt{1 - r^2} \mathbf{x}_{\mathbf{u}}) d\mu_{|\mathbf{u}|-1}(\mathbf{x}_{\mathbf{u}}) d\mu_{|\mathbf{v}|-1}(\mathbf{z}_{\mathbf{v}}) dr.\end{aligned}\tag{2.1}$$

For the case where $\mathbf{u} = i$ we write for $\mathbf{x}_{\mathbf{v} \cup i} \in \mathbb{S}^{|\mathbf{v}|}$, $\mathbf{x}_{\mathbf{v} \cup i} = (\sqrt{1 - r^2} \mathbf{y}_{\mathbf{v}}, r)$, where $\mathbf{x}_{\mathbf{v}} \in \mathbb{S}^{|\mathbf{v}|-1}$ and $r \in [-1, 1]$. It follows that

$$d\mu_{|\mathbf{v}|}(\mathbf{x}_{\mathbf{v} \cup i}) = \sqrt{1 - r^2}^{|\mathbf{v}|-2} dr d\mu_{|\mathbf{v}|-1}(\mathbf{x}_{\mathbf{v}}).$$

Changing the variables $\mathbf{x}_{\mathbf{v} \cup i} \rightarrow (\sqrt{1 - r^2} \mathbf{y}_{\mathbf{v}}, r)$ gives

$$\int_{\mathbb{S}^{|\mathbf{v}|}} g(\mathbf{x}) d\mu_{|\mathbf{v}|}(\mathbf{x}) = \int_{-1}^1 \int_{\mathbb{S}^{|\mathbf{v}|-1}} g\left(\sqrt{1 - r^2} \mathbf{y}_{\mathbf{v}}, r\right) \sqrt{1 - r^2}^{|\mathbf{v}|-2} d\mu_{|\mathbf{v}|-1}(\mathbf{y}_{\mathbf{v}}) dr.\tag{2.2}$$

We use (2.1) and obtain in the special case $\mathbf{x} = (\mathbf{x}_{\mathbf{u}^c}, \mathbf{x}_{\mathbf{u}}) = (r\mathbf{z}_{\mathbf{u}^c}, \mathbf{x}_{\mathbf{u}}) \in \mathbb{S}^d$ with $d = |\mathbf{u}| + |\mathbf{u}^c| - 1$, $\mathbf{z}_{\mathbf{u}^c} \in \mathbb{S}^{|\mathbf{u}^c|-1}$

$$\begin{aligned} \int_{\mathbb{S}^d} g(\mathbf{x}) d\mu_d(\mathbf{x}) &= \int_{\mathbb{B}^{|\mathbf{u}^c|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u}^c}\|^2}^{|\mathbf{u}|-2} \int_{\mathbb{S}^{|\mathbf{u}|-1}} g\left(\mathbf{x}_{\mathbf{u}^c}, \sqrt{1 - \|\mathbf{x}_{\mathbf{u}^c}\|^2} \mathbf{x}_{\mathbf{u}}\right) d\mu_{|\mathbf{u}|-1}(\mathbf{x}_{\mathbf{u}}) d\mathbf{x}_{\mathbf{u}^c} \quad (2.3) \\ &= \int_0^1 \int_{\mathbb{S}^{|\mathbf{u}^c|-1}} \sqrt{1 - r^2}^{|\mathbf{u}|-2} r^{|\mathbf{u}^c|-1} \int_{\mathbb{S}^{|\mathbf{u}|-1}} g(r\mathbf{z}_{\mathbf{u}^c}, \sqrt{1 - r^2} \mathbf{x}_{\mathbf{u}}) d\mu_{|\mathbf{u}|-1}(\mathbf{x}_{\mathbf{u}}) d\mu_{|\mathbf{u}^c|-1}(\mathbf{z}_{\mathbf{u}^c}) dr. \end{aligned}$$

2.1.2 Integrals on spheres with radius r

For the surface integral for a sphere with radius r , we have for $\mathbf{y} = r\mathbf{x}$, $\mathbf{x} \in \mathbb{S}^d$, $\mathbf{y} \in \mathbb{S}_r^d$,

$$\int_{\mathbb{S}_r^d} h(\mathbf{y}) d\mu_d(\mathbf{y}) = r^d \int_{\mathbb{S}^d} h(r\mathbf{x}) d\mu_d(\mathbf{x}). \quad (2.4)$$

We use for $h: \mathbb{S}^d \rightarrow \mathbb{R}$ the following decomposition of the integral on the sphere,

$$\begin{aligned} \int_{\mathbb{S}^d} h(\mathbf{x}) d\mu_d(\mathbf{x}) &\stackrel{(2.3)}{=} \int_{\mathbb{B}^{|\mathbf{u}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u}}\|^2}^{|\mathbf{u}^c|-2} \int_{\mathbb{S}^{|\mathbf{u}^c|-1}} h\left(\mathbf{x}_{\mathbf{u}}, \sqrt{1 - \|\mathbf{x}_{\mathbf{u}}\|^2} \mathbf{x}_{\mathbf{u}^c}\right) d\mu_{|\mathbf{u}^c|-1}(\mathbf{x}_{\mathbf{u}^c}) d\mathbf{x}_{\mathbf{u}} \\ &\stackrel{(2.4)}{=} \int_{\mathbb{B}^{|\mathbf{u}|}} \frac{1}{\sqrt{1 - \|\mathbf{x}_{\mathbf{u}}\|^2}^2} \int_{\mathbb{S}^{|\mathbf{u}^c|-1}} h(\mathbf{x}_{\mathbf{u}}, \mathbf{x}_{\mathbf{u}^c}) d\mu_{|\mathbf{u}^c|-1}(\mathbf{x}_{\mathbf{u}^c}) d\mathbf{x}_{\mathbf{u}}. \quad (2.5) \end{aligned}$$

Similarly, the integral on a sphere with different radius,

$$\begin{aligned} \int_{\mathbb{S}_r^d} h(r\mathbf{x}) d\mu_d(r\mathbf{x}) &\stackrel{(2.4)}{=} r^d \int_{\mathbb{S}^d} h(r\mathbf{x}) d\mu_d(\mathbf{x}) \\ &\stackrel{(2.5)}{=} r^d \int_{\mathbb{B}^{|\mathbf{u}|}} \frac{1}{\sqrt{1 - \|\mathbf{x}_{\mathbf{u}}\|^2}^2} \int_{\mathbb{S}^{|\mathbf{u}^c|-1}} h(r\mathbf{x}) d\mu_{|\mathbf{u}^c|-1}(\mathbf{x}_{\mathbf{u}^c}) d\mathbf{x}_{\mathbf{u}}. \quad (2.6) \end{aligned}$$

For balls with different radius, we have

$$\int_{\mathbb{B}_r^d} h(\mathbf{y}) d\mathbf{y} = r^d \int_{\mathbb{B}^d} h(r\mathbf{x}) d\mathbf{x}. \quad (2.7)$$

2.2 Spherical harmonics

A polynomial $p: \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ is called **homogeneous of degree ℓ** if

$$p(r\mathbf{x}) = r^\ell p(\mathbf{x}) \quad \text{for all } r > 0, \mathbf{x} \in \mathbb{R}^{d+1}.$$

Let $\mathcal{P}_\ell^d$ denote the space of real homogeneous polynomials of degree ℓ . A polynomial is called **harmonic** if it satisfies the Laplace equation

$$\Delta p = 0, \quad \Delta = \sum_{i=1}^{d+1} \frac{\partial^2}{\partial x_i^2}.$$

For $\ell = 0, 1, \dots$, we denote the linear spaces of real harmonic polynomials of degree ℓ by

$$\mathcal{H}_\ell^d := \{p \in \mathcal{P}_\ell^d \mid \Delta p = 0\}.$$

The restriction of a homogeneous harmonic polynomial p of degree ℓ to the sphere $\mathbb{S}^d$ is called a **spherical harmonic of degree ℓ** . We denote the corresponding space by

$$\mathcal{H}_\ell(\mathbb{S}^d) := \{p|_{\mathbb{S}^d} : p \text{ homogeneous, harmonic, and of degree } \ell\},$$

see for example [6, 7] for more details about these spaces. The spaces $\mathcal{H}_\ell(\mathbb{S}^d)$ form an orthogonal decomposition

$$L_2(\mathbb{S}^d) = \bigoplus_{\ell=0}^{\infty} \mathcal{H}_\ell(\mathbb{S}^d),$$

and any $f \in L_2(\mathbb{S}^d)$ admits the spherical harmonic expansion

$$f(\mathbf{x}) = \sum_{\ell=0}^{\infty} \sum_{m=1}^{\dim(\mathcal{H}_\ell(\mathbb{S}^d))} \hat{f}_{\ell,m} Y_{\ell,m}(\mathbf{x}),$$

where $\{Y_{\ell,m}\}_{m=1}^{\dim(\mathcal{H}_\ell(\mathbb{S}^d))}$ is an orthonormal basis of $\mathcal{H}_\ell(\mathbb{S}^d)$ and $\dim(\mathcal{H}_\ell(\mathbb{S}^d))$ denotes its dimension.

Remark 2.2 (Polynomial representatives and coordinate dependence). Each spherical harmonic $Y_{\ell,m} \in \mathcal{H}_\ell(\mathbb{S}^d)$ is the restriction of a homogeneous harmonic polynomial $P_{\ell,m} : \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ of degree ℓ , i.e.

$$Y_{\ell,m}(\mathbf{x}) = P_{\ell,m}(\mathbf{x})|_{\mathbb{S}^d}.$$

The polynomial $P_{\ell,m}$ can be expressed as a linear combination of monomials

$$x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_{d+1}^{\alpha_{d+1}}, \quad |\boldsymbol{\alpha}| := \alpha_1 + \cdots + \alpha_{d+1} = \ell,$$

whose coefficients are determined by the harmonicity condition $\Delta P_{\ell,m} = 0$. Thus, every spherical harmonic of degree ℓ involves monomials of total degree ℓ in the Cartesian coordinates. In particular, the index ℓ corresponds to the total polynomial degree, while the secondary index m enumerates the independent harmonic combinations of these monomials, forming a basis of the space

$$\mathcal{H}_\ell(\mathbb{S}^d) = \text{span}\{P_{\ell,m}(\mathbf{x})|_{\mathbb{S}^d} : 1 \leq m \leq \dim(\mathcal{H}_\ell(\mathbb{S}^d))\}.$$

Consequently, each $Y_{\ell,m}$ generally depends on all coordinates $x_1, \dots, x_{d+1}$, though with specific parity and symmetry relations determined by the monomial exponents $\boldsymbol{\alpha}$.

For $d = 2$, $\dim(\mathcal{H}_\ell(\mathbb{S}^d)) = 2$. In polar coordinates $(x_1, x_2) = (r \cos \theta, r \sin \theta)$ of $\mathbb{R}^2$, this basis is given by

$$Y_\ell^{(1)} = r^\ell \cos(\ell\theta), \quad Y_\ell^{(2)} = r^\ell \sin(\ell\theta). \quad (2.8)$$

2.3 Decomposition into odd and even parts

Let the domain be $\mathbb{D} \in \{\mathbb{S}^d, \mathbb{B}^{d+1}\}$. We analyze the parity properties of the underlying functions. In the one-dimensional case, a function f is defined as **odd** if $f(-x) = -f(x)$ and as **even** if $f(-x) = f(x)$ almost everywhere. Then, an arbitrary function f is decomposed into its even and odd parts by

$$f(x) = f_e(x) + f_o(x), \quad f_e(x) = \frac{1}{2}(f(x) + f(-x)), \quad f_o(x) = \frac{1}{2}(f(x) - f(-x)).$$

Similarly, we do this for a function on $\mathbb{S}^d$ or $\mathbb{B}^{d+1}$ by using the parity vector $\boldsymbol{\xi} = (\xi_i)_{i=1}^{d+1} \in \{0, 1\}^{d+1}$ and defining

$$f(\mathbf{x}) = \sum_{\boldsymbol{\xi} \in \{0,1\}^{d+1}} [\Xi_{\boldsymbol{\xi}} f](\mathbf{x}), \quad (2.9)$$

$$\Xi_{\boldsymbol{\xi}} : L_2(\mathbb{D}) \rightarrow L_2(\mathbb{D}) \quad \text{with} \quad [\Xi_{\boldsymbol{\xi}} f](\mathbf{x}) := \frac{1}{2^{d+1}} \sum_{\mathbf{k} \in \{-1,1\}^{d+1}} \left(\underbrace{\prod_{i=1}^{d+1} k_i^{\xi_i}}_{\in \{-1,1\}} \right) f(\mathbf{k} \odot \mathbf{x})$$

where $\xi_i = 1$ belongs to an odd function in direction i and $\xi_i = 0$ belongs to an even function in direction i . In (2.9) the construction of the terms is exactly such that one factor (-1) is added for every $i \in [d+1]$, for which $k_i = -1$ and $\xi_i = 1$.

Definition 2.3. We denote the vector $\boldsymbol{\xi}$ as the **parity** of the function g , if for all $i \in [d+1]$, for almost all $\mathbf{x} \in \mathbb{D}$,

$$\begin{aligned} \xi_i = 0 &\Leftrightarrow g(\mathbf{x}) = g(\mathbf{x}_{[d+1] \setminus i}, -x_i), \\ \xi_i = 1 &\Leftrightarrow g(\mathbf{x}) = -g(\mathbf{x}_{[d+1] \setminus i}, -x_i). \end{aligned}$$

For $f \in L_2(\mathbb{D})$ we define the corresponding function spaces

$$L_2(\boldsymbol{\xi}) = \{f \in L_2(\mathbb{D}) \mid f \text{ has parity } \boldsymbol{\xi}\}.$$

The spaces $L_2(\boldsymbol{\xi})$ are orthogonal, i.e.

$$L_2(\mathbb{D}) = \bigoplus_{\boldsymbol{\xi} \in \{0,1\}^{d+1}} L_2(\boldsymbol{\xi}).$$

2.4 The classical ANOVA decomposition

The classical ANOVA (analysis of variance) decomposition was introduced by Hoeffding [13] and has been studied in many contexts, see for example [8, 16, 34, 24, 18] and references therein. Often this model is also called high-dimensional model representation, see also [29] for an overview. The ANOVA decomposition decomposes a d -variate function into 2^d ANOVA terms where each term belongs to a subset $\mathbf{u} \subseteq [d]$ of coordinate indices. Each single term only depends on the variables in the corresponding subset and the number of these variables is the order of the ANOVA term. The ANOVA decomposition is a valuable tool for understanding the structure and behavior of functions and enables to focus on the most significant factors when constructing approximation models. It has extensively been used for the analysis of quasi-Monte Carlo methods for integration, see e.g. [3, 20, 23, 33, 11] and references therein. Applications to sparse grids can be found in [2, 10, 12, 14].

Let $\mathbb{D}^d \in \{\mathbb{T}^d, \mathbb{R}^d, [0, 1]^d\}$ be a tensor product domain, for a function $f: \mathbb{D}^d \rightarrow \mathbb{R}$ the **ANOVA decomposition** (Analysis of variance) is given by

$$f(\mathbf{x}) = f_\emptyset + \sum_{i=1}^d f_{\{i\}}(x_i) + \sum_{i \neq j=1}^d f_{\{i,j\}}(x_i, x_j) + \cdots + f_{[d]}(\mathbf{x}) = \sum_{\mathbf{u} \subseteq [d]} f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}). \quad (2.10)$$

It is common to refer to the terms $f_{\mathbf{u}}$ where $|\mathbf{u}| = q$ collectively as the **order- q terms**. In general, there are many possibilities for a decomposition (2.10). Indeed, an arbitrary choice of $f_{\mathbf{u}}$ for all $|\mathbf{u}| < d$ can be accommodated by taking $f_{[d]}$ to be f minus all the other terms. But with some additional conditions the decomposition is unique. Typically, one demands orthogonality of the involved terms with respect to some density μ . Examples for the application of the classical ANOVA decomposition can be found in:

- For periodic functions on the torus $\mathbb{T}^d$, trigonometrical polynomials are used as basis in [26] and wavelet function in [22].
- On the cube $[0, 1]$ the half-period cosine basis was studied in [27].
- For the use of random Fourier features for the generalized ANOVA decomposition on $\mathbb{R}^d$ see [28].

The classical ANOVA decomposition has the following properties

- The density $\mu(\mathbf{x}) = \prod_{i=1}^d \mu_i(x_i)$ is a tensor product density.
- For the definition of the terms a projection operator is used, which integrates the function over all variables not in $\mathbf{u}$ by

$$P_{\mathbf{u}}f = \left[\prod_{i \in \mathbf{u}} P_i \right] f = \int_{\mathbb{D}^{d \setminus |\mathbf{u}|}} f(\mathbf{x}) \, d \left(\prod_{i \in [d] \setminus \mathbf{u}} \mu_i(x_i) \right). \quad (2.11)$$

- The ANOVA terms are defined by

$$f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) = \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{v}|} P_{\mathbf{v}}f. \quad (2.12)$$

- All terms are orthogonal: $0 = \langle f_{\mathbf{u}}, f_{\mathbf{v}} \rangle_{\mu}$ for all $\mathbf{u} \neq \mathbf{v}$. This is a necessary condition for sensitivity analysis, where the total variance of the function is the sum of the variances of the ANOVA terms.

For a generalization of the classical decomposition to some non-tensor product densities see for example [15, 21, 30, 28]. The needed assumption there is that the domain is grid-closed to maintain at least some tensor product structures. But this assumption is not fulfilled on the sphere. For this reason, the aim of this paper is to define a spherical ANOVA decomposition for functions $f: \mathbb{S}^d \rightarrow \mathbb{R}$ on the sphere, whereby we want to retain as many properties of the classical decomposition as possible. For the following reasons we can not apply the classical or generalized ANOVA decomposition directly to the sphere:

- Using the classical decomposition on $\mathbb{R}^{d+1}$ or $[0, 1]^{d+1}$ is not even with the generalized ANOVA decomposition possible, since $\mathbb{S}^d$ is not a grid-closed domain in $\mathbb{R}^{d+1}$. Problems arise at the intersection points of the axes with the sphere.
- Using spherical coordinates would lead to the loss of interpretability of the variables, which is a key advantage of the ANOVA decomposition.
- For tensor product domains, the domain of the lower dimensional functions $f_{\mathbf{u}}$ are the lower-dimensional tensor product domains. For functions on the sphere this is not possible. If $f_{\mathbf{u}}: \mathbb{S}^{|\mathbf{u}|-1} \rightarrow \mathbb{R}$, these functions have to be extended to higher dimensional spheres. We will later define the spherical terms $f_{\mathbf{u}}: \mathbb{B}^{|\mathbf{u}|-1} \rightarrow \mathbb{R}$, defined on lower dimensional balls, such that they are automatically expanded to the sphere $\mathbb{S}^d$.
- For the classical ANOVA decomposition, one uses an orthonormal basis in L_2 and then finds a connection between the basis functions and the ANOVA terms. This works well because of the tensor product structure of the basis in the d -dimensional function space L_2 . In contrast, on the sphere $\mathbb{S}^d$, there is no basis of tensor product form, which can be used for this purpose.
- Defining the projection operator by $P_j(f) = f(x_1, \dots, x_{j-1}, c_j, x_{j+1}, \dots, x_d)$ for some anchor point $\mathbf{c} = (c_j)_{j=1}^d$ leads to the anchored ANOVA decomposition, see for example [18]. Due to the geometric properties on the sphere, it is not possible to generalize this anchored decomposition to the sphere. On the unit ball $\mathbb{B}^{d+1}$, the anchored ANOVA decomposition with anchor point $\mathbf{0}$ can be calculated equally to the tensor product domain.

While proposing a spherical ANOVA decomposition in this paper, we will also highlight the differences and challenges in comparison to the classical ANOVA decomposition.

3 A spherical ANOVA decomposition and its properties

Analyzing high-dimensional functions on the sphere poses a significant challenge, as separating local dependencies from global structures is often non-trivial. The ANOVA decomposition provides a systematic framework for this task by hierarchically decomposing the function into main effects and higher-order interactions. This enables us not only to quantify the influence of individual variables but also to understand the underlying coupling structure of the function.

Denote $\mathcal{D} := \{\mathbf{u} \subseteq [d+1] \mid |\mathbf{u}| \neq d\}$. For a function $f: \mathbb{S}^d \rightarrow \mathbb{R}$ we will study a decomposition of the form

$$f(\mathbf{x}) = f_{\emptyset} + \sum_{q=1}^{d-1} \sum_{|\mathbf{u}|=q} f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) + f_{d+1}(\mathbf{x}) = \sum_{\mathbf{u} \in \mathcal{D}} f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}),$$

where the component functions $f_{\mathbf{u}}: \mathbb{B}^{|\mathbf{u}|} \rightarrow \mathbb{R}$ are functions on lower-dimensional balls and $f_{d+1}: \mathbb{S}^d \rightarrow \mathbb{R}$. The domain of the functions $f_{\mathbf{u}}$ is the lower-dimensional ball $\mathbb{B}^{|\mathbf{u}|}$, but naturally with this decomposition, the terms $f_{\mathbf{u}}$ are also defined on $\mathbb{S}^d$. We do not need terms for the indices $|\mathbf{u}| = d$, since due to the normalization $\|\mathbf{x}\| = 1$, the last variable would be specified, except for the sign, from fixing all other variables. We treat all these parts of the function f in the very last term $f_{[d+1]}$.

We propose to combine the spherical ANOVA decomposition (3) with the decomposition (2.9) into even and odd parts. We will always use the index $\mathbf{u}$ for the spherical ANOVA terms and the index ξ for the parity of the term. Remember that $\xi_i = 1$ belongs to an odd function in direction x_i and $\xi_i = 0$ to an even function in direction x_i . Since the function $f_{\mathbf{u}}$ does not depend on the variables x_i for $i \in \mathbf{u}^c$, in the term $f_{\mathbf{u}}$ there can be only terms with parity vector ξ with $\xi_{\mathbf{u}^c} = \mathbf{0}$. For that reason we introduce the index-set

$$\Upsilon_{\mathbf{u}} := \{\xi \in \{0, 1\}^{d+1} \mid \xi_{\mathbf{u}^c} = \mathbf{0}\},$$

where $|\Upsilon_{\mathbf{u}}| = 2^{|\mathbf{u}|}$. We study in this paper the decomposition of the form

$$f(\mathbf{x}) = \sum_{\mathbf{u} \in \mathcal{D}} \sum_{\xi \in \Upsilon_{\mathbf{u}}} f_{\mathbf{u}, \xi}(\mathbf{x}_{\mathbf{u}}). \quad (3.1)$$

In this decomposition we have in total

$$\sum_{q=0}^{d-1} \binom{d+1}{q} 2^q + 2^{d+1} = \sum_{q=0}^{d+1} \binom{d+1}{q} 2^q - 2^d (d+1) = 3^{d+1} - 2^d (d+1) \quad (3.2)$$

spherical ANOVA terms.

Example 3.1. For the case $d = 2$, we have to include the $15 = 3^3 - 2^2 \cdot 3$ terms

$f_{\emptyset, \mathbf{0}},$	order 0
$f_{\{1\}, (1,0,0)^\top}, f_{\{1\}, (0,0,0)^\top}, f_{\{2\}, (0,1,0)^\top}, f_{\{2\}, (0,0,0)^\top}, f_{\{3\}, (0,0,1)^\top}, f_{\{3\}, (0,0,0)^\top},$	order 1
$f_{\{1,2,3\}, (0,0,0)^\top}, f_{\{1,2,3\}, (0,0,1)^\top}, f_{\{1,2,3\}, (0,1,0)^\top}, f_{\{1,2,3\}, (0,1,1)^\top},$	order 3
$f_{\{1,2,3\}, (1,0,0)^\top}, f_{\{1,2,3\}, (1,0,1)^\top}, f_{\{1,2,3\}, (1,1,0)^\top}, f_{\{1,2,3\}, (1,1,1)^\top}$	order 3

Note that we do not need the two-dimensional terms $f_{\mathbf{u}}$ for $|\mathbf{u}| = 2$.

In the following subsection we introduce an operator $P_{\mathbf{u}}$, which we will use in Section 3.4 to define the spherical ANOVA terms $f_{\mathbf{u}, \xi}$.

3.1 The projection operator

For an index $\mathbf{u} \in \mathcal{D} \setminus [d+1]$ and $\mathbf{y} \in \mathbb{B}^{|\mathbf{u}|}$ denote the *fiber*

$$\{\mathbf{x} \in \mathbb{S}^d \subset \mathbb{R}^{d+1} : \mathbf{x}_{\mathbf{u}} = \mathbf{y}\} \quad (3.3)$$

and the corresponding integrals $m_{\mathbf{u}} : \mathbb{B}^{|\mathbf{u}|} \rightarrow \mathbb{R}$

$$m_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}}) := \int_{\mathbb{S}^{\lfloor \mathbf{u}^c \rfloor - 1}} \frac{1}{\sqrt{1 - \|\mathbf{y}_{\mathbf{u}}\|^2}} \mathrm{d}\mu_{d-|\mathbf{u}|}(\mathbf{x}_{\mathbf{u}^c}),$$

where $\mu_{d-|\mathbf{u}|}(\mathbf{x}_{\mathbf{u}^c})$ is the surface measure of the $d - |\mathbf{u}|$ -dimensional sphere of radius $\sqrt{1 - \|\mathbf{y}_{\mathbf{u}}\|^2} = \|\mathbf{x}_{\mathbf{u}^c}\|$.

We study the uniform distribution μ_d on the sphere. Due to (2.4), the corresponding conditional distribution on the lower dimensional balls $\mathbb{B}^{|\mathbf{u}|}$, is the distribution

$$\begin{aligned} m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}}) &= \begin{cases} \sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}^{d-|\mathbf{v}|} \int_{\mathbb{S}^{d-|\mathbf{v}|}} \mathrm{d}\mu_{d-|\mathbf{v}|}(\mathbf{x}) & \text{if } |\mathbf{v}| \leq d-1 \\ 0 & \text{if } |\mathbf{v}| = d, \\ 1 & \text{if } \mathbf{v} = [d+1], \end{cases} \\ &= \begin{cases} \omega_{d-|\mathbf{v}|} \sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}^{d-|\mathbf{v}|} & \text{if } |\mathbf{v}| \leq d-1, \\ 0 & \text{if } |\mathbf{v}| = d, \\ 1 & \text{if } \mathbf{v} = [d+1], \\ \omega_d & \text{if } \mathbf{v} = \emptyset. \end{cases} \end{aligned}$$

Definition 3.2 (Projection operator). We define for $\mathbf{u} \in \mathcal{D} \setminus [d+1]$ the projection operator $P_{\mathbf{u}} : L_2(\mathbb{S}^d) \rightarrow L_2(\mathbb{B}^{|\mathbf{u}|})$ by

$$\begin{aligned} [P_{\mathbf{u}} f](\mathbf{y}_{\mathbf{u}}) &:= \frac{1}{m_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}})} \int_{\mathbb{S}^{\lfloor \mathbf{u}^c \rfloor - 1}} \frac{f(\mathbf{x})}{\sqrt{1 - \|\mathbf{y}_{\mathbf{u}}\|^2}} \mathrm{d}\mu_{|\mathbf{u}^c|-1}(\mathbf{x}_{\mathbf{u}^c}) \\ &= \frac{1}{\omega_{d-|\mathbf{u}|}} \int_{\mathbb{S}^{\lfloor \mathbf{u}^c \rfloor - 1}} f\left(\mathbf{y}_{\mathbf{u}}, \sqrt{1 - \|\mathbf{y}_{\mathbf{u}}\|^2} \mathbf{x}_{\mathbf{u}^c}\right) \mathrm{d}\mu_{|\mathbf{u}^c|-1}(\mathbf{x}_{\mathbf{u}^c}). \end{aligned} \quad (3.4)$$

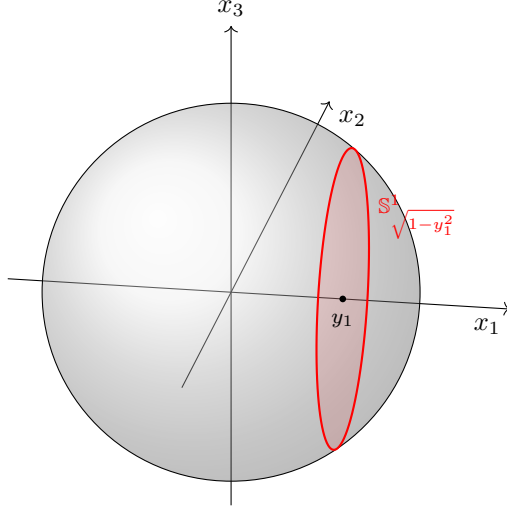


Figure 1: Visualization of $P_{\{1\}}f$ for $d = 2$: Integration over the red circle with fixed $x_1 = y_1$.

The operators $P_{\mathbf{u}}$ calculate the mean of the function f on the fiber (3.3). Note that for $|\mathbf{u}| = d$ this operator is not defined and for $\mathbf{u} = [d+1]$ the operator $P_{[d+1]}$ is the identity operator. For one-dimensional $\mathbf{u} = \{i\}$, the operator $P_{\mathbf{u}}$ is

$$[P_{\{i\}}f](y_i) = \frac{1}{\omega_{d-1}\sqrt{1-y_i^2}} \int_{\frac{\mathbb{S}^{d-1}}{\sqrt{1-y_i^2}}} f(\mathbf{x}) d\mu_{d-1}(\mathbf{x}_{[d+1]\setminus i}) = \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} f\left(y_i, \sqrt{1-y_i^2} \mathbf{x}_{\mathbf{u}^c}\right) d\mu_{d-1}(\mathbf{x}_{\mathbf{u}^c}).$$

For $d = 2$, the operator $P_{\{1\}}$ is illustrated in Figure 1.

Here we see the difference to the classical ANOVA decomposition: In the tensor product setting the projection operator has also tensor product structure, i.e. $P_{\mathbf{u}}$ is a product of one-dimensional P_i , see (2.11). For the operator (3.4) on the sphere, we do not have such a structure, but nevertheless the operator $P_{\mathbf{u}}$ has some properties, which we will use to define a spherical ANOVA decomposition. First, the sequential application of the projection operator is the same the operator applied with the smaller index set, as the following lemma shows.

Lemma 3.3. *The operators $P_{\mathbf{u}}$ fulfill for $\mathbf{v} \subset \mathbf{u}$,*

$$P_{\mathbf{v}}P_{\mathbf{u}}f = P_{\mathbf{v}}f.$$

Proof. Due to the definition,

$$\begin{aligned} [P_{\mathbf{v}}P_{\mathbf{u}}f](\mathbf{y}_{\mathbf{u}}) &= \frac{1}{m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}})} \int_{\frac{\mathbb{S}^{|\mathbf{v}^c|-1}}{\sqrt{1-\|\mathbf{y}_{\mathbf{v}}\|^2}}} \frac{1}{m_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}})} \int_{\frac{\mathbb{S}^{|\mathbf{u}^c|-1}}{\sqrt{1-\|\mathbf{y}_{\mathbf{u}}\|^2}}} f(\mathbf{x}) d\mu_{d-|\mathbf{u}|}(\mathbf{x}_{\mathbf{u}^c}) d\mu_{d-|\mathbf{v}|}(\mathbf{x}_{\mathbf{v}^c}) \\ &= \frac{1}{m_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}})} \int_{\frac{\mathbb{S}^{|\mathbf{u}^c|-1}}{\sqrt{1-\|\mathbf{y}_{\mathbf{u}}\|^2}}} \frac{1}{m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}})} \int_{\frac{\mathbb{S}^{|\mathbf{v}^c|-1}}{\sqrt{1-\|\mathbf{y}_{\mathbf{v}}\|^2}}} f(\mathbf{x}) d\mu_{d-|\mathbf{v}|}(\mathbf{x}_{\mathbf{v}^c}) d\mu_{d-|\mathbf{u}|}(\mathbf{x}_{\mathbf{u}^c}) \\ &= \frac{1}{m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}})} \int_{\frac{\mathbb{S}^{|\mathbf{v}^c|-1}}{\sqrt{1-\|\mathbf{y}_{\mathbf{v}}\|^2}}} f(\mathbf{x}) d\mu_{d-|\mathbf{v}|}(\mathbf{x}_{\mathbf{v}^c}) = [P_{\mathbf{v}}f](\mathbf{y}_{\mathbf{v}}). \end{aligned}$$

This finishes the proof. $\square$

Applying the operator $P_{\mathbf{a}}$ to a function, which depends only on the variables in $\mathbf{u}$ leads to the following lemma. This will be helpful in Section 3.2 to define integral conditions for the ANOVA operator.

Lemma 3.4. *If $\mathbf{u} \in \mathcal{D} \setminus \emptyset$ and for $\mathbf{a} \in \mathcal{D}$ with $\mathbf{a} \subset \mathbf{u}$, the operator $P_{\mathbf{a}}$ applied to a function $g_{\mathbf{u}}: \mathbb{B}^{|\mathbf{u}|} \rightarrow \mathbb{R}$, which is naturally also a function $g_{\mathbf{u}}: \mathbb{S}^d \rightarrow \mathbb{R}$, can also be written as*

$$[P_{\mathbf{a}}g_{\mathbf{u}}](\mathbf{y}_{\mathbf{a}}) = \frac{\omega_{d-|\mathbf{u}|}}{\omega_{d-|\mathbf{a}|}} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2}^{d-|\mathbf{u}|-1} g_{\mathbf{u}}\left(\mathbf{y}_{\mathbf{a}}, \sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2} \mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\right) d\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}.$$

For arbitrary $\mathbf{v} \in \mathcal{D}$, which does not necessarily has to be a subset of $\mathbf{u}$, denote $\mathbf{a} = \mathbf{u} \cap \mathbf{v}$. Then

$$[P_{\mathbf{v}}g_{\mathbf{u}}](\mathbf{y}_{\mathbf{v}}) = \frac{\omega_{|\mathbf{u} \cup \mathbf{v}|^c|-1}}{\omega_{d-|\mathbf{v}|}} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2}^{|\mathbf{u} \cup \mathbf{v}|^c|-2} g_{\mathbf{u}} \left(\mathbf{y}_{\mathbf{a}}, \sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2} \mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \right) d\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}.$$

Proof. We start with the first case where $\mathbf{a} \subset \mathbf{u}$, and use the definition of the operator $P_{\mathbf{a}}$ applied to the function $g_{\mathbf{u}}$,

$$\begin{aligned} [P_{\mathbf{a}}g_{\mathbf{u}}](\mathbf{y}_{\mathbf{a}}) &= \frac{1}{m_{\mathbf{a}}(\mathbf{y}_{\mathbf{a}})} \int_{\frac{\mathbb{S}^{d-|\mathbf{a}|}}{\sqrt{1-\|\mathbf{y}_{\mathbf{a}}\|^2}}} g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) d\mu_{d-|\mathbf{a}|}(\mathbf{x}_{\mathbf{a}^c}) \quad \text{set } r = \sqrt{1 - \frac{\|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2}{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}} = \sqrt{\frac{1 - \|\mathbf{x}_{\mathbf{u}}\|^2}{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}}, \\ &\stackrel{(2.6)}{=} \frac{1}{m_{\mathbf{a}}(\mathbf{y}_{\mathbf{a}})} \sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}^{d-|\mathbf{a}|} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \frac{1}{r} g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) \int_{\mathbb{S}_r^{|\mathbf{u}^c|-1}} 1 d\mu_{d-|\mathbf{a}|} \left(\frac{\mathbf{x}_{\mathbf{u}^c}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}} \right) d \frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}} \\ &\stackrel{(2.7)}{=} \frac{1}{\omega_{d-|\mathbf{a}|}} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \frac{1}{r} g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) \omega_{d-|\mathbf{u}|} r^{d-|\mathbf{u}|} d \frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}} \\ &= \frac{\omega_{d-|\mathbf{u}|}}{\omega_{d-|\mathbf{a}|}} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} r^{d-|\mathbf{u}|-1} g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) d \frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}} \\ &= \frac{\omega_{d-|\mathbf{u}|}}{\omega_{d-|\mathbf{a}|}} \sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}^{-d+|\mathbf{u}|+1} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u}}\|^2}^{d-|\mathbf{u}|-1} g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) d \frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}} \\ &= \frac{\omega_{d-|\mathbf{u}|}}{\omega_{d-|\mathbf{a}|}} \sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2}^{-d+|\mathbf{u}|+1} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \sqrt{(1 - \|\mathbf{y}_{\mathbf{a}}\|^2)(1 - \|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2)}^{d-|\mathbf{u}|-1} g_{\mathbf{u}} \left(\mathbf{y}_{\mathbf{a}}, \sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2} \mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \right) d\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \\ &= \frac{\omega_{d-|\mathbf{u}|}}{\omega_{d-|\mathbf{a}|}} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2}^{d-|\mathbf{u}|-1} g_{\mathbf{u}} \left(\mathbf{y}_{\mathbf{a}}, \sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2} \mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \right) d\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \\ &= \frac{\omega_{d-|\mathbf{u}|}}{\omega_{d-|\mathbf{a}|}} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2}^{d-|\mathbf{u}|-1} g_{\mathbf{u}} \left(\mathbf{y}_{\mathbf{a}}, \sqrt{1 - \|\mathbf{y}_{\mathbf{a}}\|^2} \mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \right) d\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}. \end{aligned}$$

For arbitrary $\mathbf{v} \in \mathcal{D}$, which does not necessarily has to be a subset of $\mathbf{u}$, Let $\mathbf{a} = \mathbf{u} \cap \mathbf{v}$,

$$\begin{aligned} [P_{\mathbf{v}}g_{\mathbf{u}}](\mathbf{y}_{\mathbf{v}}) &= \frac{1}{m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}})} \int_{\frac{\mathbb{S}^{|\mathbf{v}^c|-1}}{\sqrt{1-\|\mathbf{y}_{\mathbf{v}}\|^2}}} g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) d\mu_{|\mathbf{v}^c|-1}(\mathbf{x}_{\mathbf{v}^c}), \quad \text{set } r := \sqrt{1 - \frac{\|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2}{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}}, \\ &\stackrel{(2.6)}{=} \frac{\sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}^{|\mathbf{v}^c|-1}}{m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}})} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \frac{1}{r} \int_{\mathbb{S}_r^{|\mathbf{v} \cup \mathbf{u}|^c|-1}} g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) d\mu_{|\mathbf{u} \cup \mathbf{v}|^c|-1} \left(\frac{\mathbf{x}_{(\mathbf{u} \cup \mathbf{v})^c}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}} \right) d \left(\frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}} \right) \\ &= \frac{\sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}^{|\mathbf{v}^c|-1}}{m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}})} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \frac{g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}})}{r} \int_{\mathbb{S}_r^{|\mathbf{v} \cup \mathbf{u}|^c|-1}} d\mu_{|\mathbf{u} \cup \mathbf{v}|^c|-1} \left(\frac{\mathbf{x}_{(\mathbf{u} \cup \mathbf{v})^c}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}} \right) d \left(\frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}} \right) \\ &= \frac{\omega_{|\mathbf{u} \cup \mathbf{v}|^c|-1}}{m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}})} \sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}^{|\mathbf{v}^c|-1} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) r^{|\mathbf{u} \cup \mathbf{v}|^c-2} d \left(\frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}}{\sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}} \right) \\ &= \frac{\omega_{|\mathbf{u} \cup \mathbf{v}|^c|-1}}{m_{\mathbf{v}}(\mathbf{y}_{\mathbf{v}})} \sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2}^{|\mathbf{v}^c|-1} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2}^{|\mathbf{u} \cup \mathbf{v}|^c-2} g_{\mathbf{u}} \left(\mathbf{y}_{\mathbf{a}}, \sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2} \mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \right) d\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \\ &= \frac{\omega_{|\mathbf{u} \cup \mathbf{v}|^c|-1}}{\omega_{d-|\mathbf{v}|}} \sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{a}|}} \sqrt{1 - \|\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}\|^2}^{|\mathbf{u} \cup \mathbf{v}|^c-2} g_{\mathbf{u}} \left(\mathbf{y}_{\mathbf{a}}, \sqrt{1 - \|\mathbf{y}_{\mathbf{v}}\|^2} \mathbf{x}_{\mathbf{u} \setminus \mathbf{a}} \right) d\mathbf{x}_{\mathbf{u} \setminus \mathbf{a}}. \end{aligned}$$

This finishes the proof. $\square$

3.2 The ANOVA operator and the integral conditions

Every function $f \in L_2(\mathbb{S}^d)$ can be decomposed arbitrarily in the form (3), if we choose arbitrary terms for $f_{\mathbf{u}}$ for $|\mathbf{u}| \leq d-1$ and simply choose the remainder that is missing from f as the term $f_{[d+1]}(\mathbf{x})$. For classical ANOVA decomposition, orthogonality is required to get a unique decomposition. This goes hand in hand with the condition that certain mean values of the ANOVA terms are zero. In the following we propose similar integral conditions also for the spherical ANOVA decomposition.

Definition 3.5. For a function $f \in L_2(\mathbb{S}^d)$, we introduce ANOVA operators $\mathcal{A}_{\mathbf{u}}$, which calculates for every $\mathbf{u} \in \mathcal{D} \setminus [d+1]$ the parts

$$\begin{aligned}\mathcal{A}_{\emptyset} f &= P_{\emptyset} f = \frac{1}{\omega_d} \int_{\mathbb{S}^d} f(\mathbf{x}) \, d\mu_d(\mathbf{x}), \\ [\mathcal{A}_{\mathbf{u}} f](\mathbf{x}_{\mathbf{u}}) &= [P_{\mathbf{u}} f](\mathbf{x}_{\mathbf{u}}) - \sum_{\mathbf{v} \subset \mathbf{u}} [\mathcal{A}_{\mathbf{v}} f](\mathbf{x}_{\mathbf{v}}) \quad \text{if } 1 \leq |\mathbf{u}| \leq d-1, \\ [\mathcal{A}_{[d+1]} f](\mathbf{x}) &= f(\mathbf{x}) - \sum_{\mathbf{v} \in \mathcal{D} \setminus [d+1]} [\mathcal{A}_{\mathbf{v}} f](\mathbf{x}_{\mathbf{v}}).\end{aligned}$$

In this definition the terms $\mathcal{A}_{\mathbf{u}} f: \mathbb{B}^{|\mathbf{u}|} \rightarrow \mathbb{R}$ depend only on the variables $\mathbf{x}_{\mathbf{u}}$.

Another formula for the terms, which follows by induction directly from the iterative definition of the operator $\mathcal{A}_{\mathbf{u}}$ is

$$\mathcal{A}_{\mathbf{u}} f = \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{v}|} P_{\mathbf{v}} f. \quad (3.5)$$

This is similar to the definition of the classical ANOVA decomposition, see Section 2.4. Additionally, all terms sum up to the original function, and we obtain the decomposition on the sphere,

$$f(\mathbf{x}) = \sum_{\mathbf{u} \in \mathcal{D}} [\mathcal{A}_{\mathbf{u}} f](\mathbf{x}_{\mathbf{u}}). \quad (3.6)$$

A closer look at Definition 3.5 and equation (3.6) reveals that, for the definition to be well-defined, for any function f consisting of a single term, $f = \mathcal{A}_{\mathbf{u}} f$, it must hold that

$$\mathcal{A}_{\mathbf{a}} f = \begin{cases} f & \text{if } \mathbf{u} = \mathbf{a}, \\ 0 & \text{otherwise.} \end{cases}$$

This automatically means by (3.5) that for $\mathbf{a} \neq \emptyset$ and a function f with $f = \mathcal{A}_{\mathbf{u}} f$,

$$[P_{\mathbf{a}} \mathcal{A}_{\mathbf{u}} f](\mathbf{y}_{\mathbf{a}}) = \begin{cases} [\mathcal{A}_{\mathbf{u}} f](\mathbf{y}_{\mathbf{u}}) & \text{if } \mathbf{a} \supseteq \mathbf{u} \\ 0 & \text{in all other cases.} \end{cases} \quad (3.7)$$

The terms defined in Definition 3.5 are orthogonal with respect to the uniform measure on the sphere $\mathbb{S}^d$, which the following theorem shows.

Theorem 3.6. For $\mathbf{u}, \mathbf{v} \in \mathcal{D}$ and $\mathbf{v} \neq \mathbf{u}$ the terms $\mathcal{A}_{\mathbf{u}} f$ and $\mathcal{A}_{\mathbf{v}} f$, defined in Definition 3.5, are orthogonal with respect to the surface measure on $\mathbb{S}^d$,

$$\langle \mathcal{A}_{\mathbf{u}} f, \mathcal{A}_{\mathbf{v}} f \rangle_{\mathbb{S}^d} = 0.$$

Proof. Set $\mathbf{w} = \mathbf{u} \cap \mathbf{v}$. If $\mathbf{w} = \emptyset$, we skip the following first step and proceed with the inner integral. Otherwise let without loss of generality $\mathbf{u} \setminus \mathbf{w} \neq \emptyset$, then

$$\begin{aligned}\langle \mathcal{A}_{\mathbf{u}} f, \mathcal{A}_{\mathbf{v}} f \rangle_{\mathbb{S}^d} &= \int_{\mathbb{S}^d} [\mathcal{A}_{\mathbf{u}} f](\mathbf{x}_{\mathbf{u}}) [\mathcal{A}_{\mathbf{v}} f](\mathbf{x}_{\mathbf{v}}) \, d\mu_d(\mathbf{x}) \\ &\stackrel{(2.5)}{=} \int_{\mathbb{B}^{|\mathbf{w}|}} \frac{1}{\sqrt{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}} \int_{\mathbb{S}^{|\mathbf{w}^c|-1}} [\mathcal{A}_{\mathbf{u}} f](\mathbf{x}_{\mathbf{u}}) [\mathcal{A}_{\mathbf{v}} f](\mathbf{x}_{\mathbf{v}}) \, d\mu_{d-|\mathbf{w}|}(\mathbf{x}_{\mathbf{w}^c}) \, d\mathbf{x}_{\mathbf{w}}.\end{aligned}$$

For the inner integral we have

$$\begin{aligned}&\int_{\mathbb{S}^{|\mathbf{w}^c|-1}} \frac{[\mathcal{A}_{\mathbf{u}} f](\mathbf{x}_{\mathbf{u}}) [\mathcal{A}_{\mathbf{v}} f](\mathbf{x}_{\mathbf{v}})}{\sqrt{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}} \, d\mu_{d-|\mathbf{w}|}(\mathbf{x}_{\mathbf{w}^c}), \quad b := \sqrt{1 - \frac{\|\mathbf{x}_{\mathbf{u} \setminus \mathbf{w}}\|^2}{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}} = \sqrt{\frac{1 - \|\mathbf{x}_{\mathbf{u}}\|^2}{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}}, \\ &\stackrel{(2.6)}{=} \sqrt{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}^{|\mathbf{w}^c|-1} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{w}|}} \frac{1}{b} \int_{\mathbb{S}_b^{|\mathbf{u}^c|-1}} [\mathcal{A}_{\mathbf{u}} f](\mathbf{x}_{\mathbf{u}}) [\mathcal{A}_{\mathbf{v}} f](\mathbf{x}_{\mathbf{v}}) \, d\mu_{|\mathbf{u}^c|-1} \left(\frac{\mathbf{x}_{\mathbf{u}^c}}{\sqrt{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}} \right) d \frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{w}}}{\sqrt{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}} \\ &= \sqrt{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}^{|\mathbf{w}^c|-1} \int_{\mathbb{B}^{|\mathbf{u} \setminus \mathbf{w}|}} \frac{1}{b} [\mathcal{A}_{\mathbf{u}} f](\mathbf{x}_{\mathbf{u}}) \int_{\mathbb{S}_b^{|\mathbf{u}^c|-1}} [\mathcal{A}_{\mathbf{v}} f](\mathbf{x}_{\mathbf{v}}) \, d\mu_{|\mathbf{u}^c|-1} \left(\frac{\mathbf{x}_{\mathbf{u}^c}}{\sqrt{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}} \right) d \frac{\mathbf{x}_{\mathbf{u} \setminus \mathbf{w}}}{\sqrt{1 - \|\mathbf{x}_{\mathbf{w}}\|^2}}.\end{aligned}$$

Again, we proceed with the inner integral by

$$\begin{aligned} \int_{\mathbb{S}_b^{|u^c|-1}} [\mathcal{A}_v f](x_v) d\mu_{|u^c|-1} \left(\frac{x_{u^c}}{\sqrt{1 - \|x_w\|^2}} \right) &= \frac{1}{\sqrt{1 - \|x_w\|^2}^{|u^c|-1}} \int_{\mathbb{S}_{\sqrt{1 - \|x_w\|^2}}^{|u^c|-1}} [\mathcal{A}_v f](x_v) d\mu_{|u^c|-1}(x_{u^c}) \\ &= \frac{1}{\sqrt{1 - \|x_w\|^2}^{|u^c|-1} m_u(y_u)} [P_u[\mathcal{A}_v f]](x_w) = 0. \end{aligned}$$

These terms are zero due to (3.7). $\square$

For the classical ANOVA decomposition with tensor product structure, it is enough to require that $P_i \mathcal{A}_u f = 0$ for all $i \in u$ due to the tensor product structure (2.11) of the operator P_u . On the sphere, these conditions are slightly more complicated since the projection operator P_u is not the product of one-dimensional projections.

Lemma 3.7. *The operators $\mathcal{A}_u$ in Definition 3.5 fulfill condition (3.7) if the integral conditions*

$$0 = \int_{\mathbb{B}^{|u \setminus a|}} \sqrt{1 - \|x_{u \setminus a}\|^2}^{d-|u|-1} [\mathcal{A}_u f] \left(y_a, \sqrt{1 - \|y_a\|^2} x_{u \setminus a} \right) dx_{u \setminus a} \quad (3.8)$$

for all $a \subset u, y_a \in \mathbb{B}^{|a|}$

$$0 = \sqrt{1 - \|y_a\|^2 - r^2} \int_{\mathbb{B}^{|u \setminus a|}} \sqrt{1 - \|x_{u \setminus a}\|^2}^m [\mathcal{A}_u f] \left(y_a, \sqrt{1 - \|y_a\|^2 - r^2} x_{u \setminus a} \right) dx_{u \setminus a}, \quad (3.9)$$

for all $a \subset u, y_a \in \mathbb{B}^{|a|}, 0 \leq r \leq \sqrt{1 - \|y_a\|^2}, m = -1, 0, 1, \dots, d - |u| - 2$

are fulfilled.

Proof. In Lemma 3.4 we applied the projection operator P_a to a function, which depends only on the variables x_u . The function $\mathcal{A}_u f$ is such a function. Hence, this lemma follows directly from Lemma 3.4. $\square$

These conditions are similar to the integral conditions for the classical ANOVA decomposition, where we fix one y_a in $f_u(x_u)$ and integrate over all other variables. These integrals then vanish for all y_a . Shortly, the ANOVA terms have to fulfill (3.7). For the classical ANOVA decomposition due to the tensor product structure (2.11), it is enough to demand $P_{\{i\}} f_u$ for all $i \in u$. In the spherical case, this is much more complicated and we have to demand the integral conditions in Lemma 3.7. As we will see in the next subsection, only odd functions $\mathcal{A}_u f$ fulfill these conditions.

3.3 The parity of the terms $\mathcal{A}_u f$

In Definition 3.5 we defined operators $\mathcal{A}_u$, which split a spherical function into lower dimensional terms. In Lemma 3.7 we showed integral conditions that the terms $\mathcal{A}_u f$ have to fulfill. In fact, only odd functions fulfill these conditions.

Theorem 3.8. *Let a function $g \in L_2(\mathbb{B}^n)$, where $n = |u| \leq d - 1$. That the function g fulfills the integral conditions (3.8) and (3.9) is equivalent to the function g being odd in every direction.*

Proof. If the function g is odd, it easily follows that (3.8) and (3.9) hold true, so one direction follows directly.

For the other direction, let $g \in L_2(\mathbb{B}^n)$ be arbitrary and fulfill the integral conditions (3.8) and (3.9). Then, decompose g into its even and odd parts g_ξ like in (2.9).

Let us fix $i \in \{1, \dots, n\}$. Then, (3.9) for $y_{[n] \setminus i}$ and $m = 0$ states that

$$0 = \sqrt{1 - \|y_{[n] \setminus i}\|^2 - r^2} \int_{-1}^1 g \left(y_{[n] \setminus i}, \sqrt{1 - \|y_{[n] \setminus i}\|^2 - r^2} x_i \right) dx_i$$

for all $0 \leq r \leq \sqrt{1 - \|\mathbf{y}_{[n]\setminus i}\|^2}$. Using the decomposition (2.9) of g into odd and even parts, in this integral of the terms $g_{\boldsymbol{\xi}}$ with $\xi_i = 1$ are zero, such that we have

$$\begin{aligned} 0 &= \sqrt{1 - \|\mathbf{y}_{[n]\setminus i}\|^2 - r^2} \int_{-1}^1 \sum_{\substack{\boldsymbol{\xi} \in \{0,1\}^n \\ \xi_i = 0}} g_{\boldsymbol{\xi}} \left(\mathbf{y}_{[n]\setminus i}, \sqrt{1 - \|\mathbf{y}_{[n]\setminus i}\|^2 - r^2} x_i \right) dx_i \\ &= \int_{-\sqrt{1 - \|\mathbf{y}_{[n]\setminus i}\|^2 - r^2}}^{\sqrt{1 - \|\mathbf{y}_{[n]\setminus i}\|^2 - r^2}} \sum_{\substack{\boldsymbol{\xi} \in \{0,1\}^n \\ \xi_i = 0}} g_{\boldsymbol{\xi}} (\mathbf{y}_{[n]\setminus i}, x_i) dx_i. \end{aligned}$$

That means, that the function $G := \sum_{\substack{\boldsymbol{\xi} \in \{0,1\}^n \\ \xi_i = 0}} g_{\boldsymbol{\xi}}$ is orthogonal to all step functions of the form $\mathbf{1}_{[-r,r]}(x_i)$. The symmetric indicator functions of arbitrary step functions are a linear combination of the indicator functions of the form $\mathbf{1}_{[-r,r]}(t)$. Furthermore, it follows that G is also orthogonal to all step functions on $\mathbb{B}^n$, which are supported on a tensor-product domain and are even in x_i direction. Since step functions are dense in $L_2(\mathbb{B}^n)$, it follows that

$$\langle G, h \rangle_{L_2(\mathbb{B}^n)} = 0 \quad \text{for all even } h \in L_2(\mathbb{B}^n) \text{ with } h(\mathbf{x}) = h(\mathbf{x}_{[n]\setminus i}, x_i) \text{ a.e.}$$

Hence G must vanish almost everywhere. Inserting the definitions of the functions $g_{\boldsymbol{\xi}}$ from (2.9) into G gives then

$$\begin{aligned} 0 &= \sum_{\substack{\boldsymbol{\xi} \in \{0,1\}^n \\ \xi_i = 0}} g_{\boldsymbol{\xi}}(\mathbf{x}) = \sum_{\substack{\boldsymbol{\xi} \in \{0,1\}^n \\ \xi_i = 0}} \frac{1}{2^n} \sum_{\mathbf{k} \in \{-1,1\}^n} \left(\prod_{\ell=1}^n k_{\ell}^{\xi_{\ell}} \right) g(\mathbf{k} \odot \mathbf{x}) \\ &= \frac{1}{2^n} \sum_{\substack{\boldsymbol{\xi} \in \{0,1\}^n \\ \xi_i = 0}} \sum_{\mathbf{k} \in \{-1,1\}^n} \left(\prod_{\ell=1}^n k_{\ell}^{\xi_{\ell}} \right) g(\mathbf{k} \odot \mathbf{x}) \\ &= \frac{1}{2^n} \sum_{\mathbf{k} \in \{-1,1\}^n} g(\mathbf{k} \odot \mathbf{x}) \sum_{\substack{\boldsymbol{\xi} \in \{0,1\}^n \\ \xi_i = 0}} \left(\prod_{\ell=1}^n k_{\ell}^{\xi_{\ell}} \right) \\ &= \frac{1}{2^n} \sum_{\mathbf{k} \in \{-1,1\}^n} g(\mathbf{k} \odot \mathbf{x}) \sum_{\boldsymbol{\xi}_{[n]\setminus i} \in \{0,1\}^{n-1}} \left(\prod_{\ell \neq i} k_{\ell}^{\xi_{\ell}} \right) \\ &= \frac{1}{2^n} \left(\underbrace{2^n g(\mathbf{x})}_{k=1} + \underbrace{2^n g(\mathbf{x}_{[n]\setminus i}, -x_i)}_{k=1-2e_i} \right), \end{aligned}$$

since all the other terms cancel out. It follows directly that $g(\mathbf{x}) = -g(\mathbf{x}_{[n]\setminus i}, -x_i)$. Similarly, this is true for all other $i \in \{1, \dots, n\}$, such that g is odd in all directions. $\square$

3.4 The definition of the spherical ANOVA decomposition

While the decomposition into even and odd parts is straightforward in Euclidean space, the algebraic constraint of the sphere introduces specific dependencies between the variables. Here we discuss how multivariate parity impacts the linear independence of the basis terms and identify which symmetry patterns lead to redundant representations.

The ANOVA operators $\mathcal{A}_{\mathbf{u}}$ from Definition 3.5 produce only odd terms. This means, in the spherical ANOVA decomposition (3.1) only terms of the form $f_{\mathbf{u}, \boldsymbol{\xi}}$ with $\boldsymbol{\xi} = (\mathbf{1}_{\mathbf{u}}, \mathbf{0}_{\mathbf{u}^c})$ are produced by applying $\mathcal{A}_{\mathbf{u}}$ to the function f itself and all even parts of a function f are contained in the last term $\mathcal{A}_{[d+1]}f$.

To include the even terms also in the lower-dimensional terms of the spherical ANOVA decomposition, we propose to define a spherical ANOVA decomposition of the form (3.1) by the following.

Definition 3.9. For a function $f \in L_2(\mathbb{S}^d)$ we use the decomposition into even and odd parts as defined in (2.9) and the projection operator defined in (3.4). With these ingredients we introduce the spherical ANOVA decomposition of the form (3.1) by

1. The constant term:

$$f_{\emptyset, \mathbf{0}} := P_{\emptyset} f = \frac{1}{\omega_d} \int_{\mathbb{S}^d} f(\mathbf{x}) d\mu_d(\mathbf{x}).$$

2. The mixed terms: For all $\mathbf{v} \in \mathcal{D} \setminus \emptyset$ and $\mathbf{u} \supseteq \mathbf{v}$ set $\boldsymbol{\xi} = (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c}) \in \{0, 1\}^{d+1}$. If $d+1 \notin \mathbf{v}$, then $d+1 \notin \mathbf{u}$ and let

$$f_{\mathbf{u}, (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})}(\mathbf{x}_{\mathbf{u}}) := P_{\mathbf{u}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f(\mathbf{x}_{\mathbf{u}})] - \sum_{\mathbf{v} \subseteq \mathbf{a} \subseteq \mathbf{u}} f_{\mathbf{a}, (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})}(\mathbf{x}_{\mathbf{a}}) \quad \text{if } |\mathbf{v}| \leq |\mathbf{u}| \leq d-1,$$

$$f_{[d+1], (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})}(\mathbf{x}) = [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f](\mathbf{x}) - \sum_{\mathbf{v} \subseteq \mathbf{a} \in \mathcal{D} \setminus [d+1]} f_{\mathbf{a}, (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})}(\mathbf{x}_{\mathbf{a}}).$$

3. The pure even terms: For all $\mathbf{u} \subset [d] \setminus \emptyset$ let

$$f_{\mathbf{u}, \mathbf{0}}(\mathbf{x}_{\mathbf{u}}) := P_{\mathbf{u}} [\Xi_{\mathbf{0}} f](\mathbf{x}_{\mathbf{u}}) - \sum_{\mathbf{a} \subsetneq \mathbf{u}} f_{\mathbf{a}, \mathbf{0}}(\mathbf{x}_{\mathbf{a}}).$$

We will discuss in Section 3.5 why we do not include the terms where $d+1 \in \mathbf{u}$ with $\xi_{d+1} = 0$. Our definition of spherical ANOVA terms is of the form (3.1), all the terms $f_{\mathbf{u}, \boldsymbol{\xi}}$ have parity $\boldsymbol{\xi}$ and

$$\Xi_{\boldsymbol{\xi}} f(\mathbf{x}) = \sum_{\substack{\mathbf{a} \in \mathcal{D} \\ \text{supp } \boldsymbol{\xi} \subseteq \mathbf{a}}} f_{\mathbf{a}, \boldsymbol{\xi}}(\mathbf{x}_{\mathbf{a}}).$$

In Example 3.1 we considered the case $d = 2$. A visualization of all the possible indices $\mathbf{u}$ and corresponding $\boldsymbol{\xi}$ for the terms $f_{\mathbf{u}, \boldsymbol{\xi}}$ for $d = 3$ can be found in Figure 2. There we show the totality of all terms $f_{\mathbf{u}, \boldsymbol{\xi}}$. Later, in Section 3.5 we will explain which of these terms we have to omit due to redundancy on the sphere.

The terms in Definition 3.9 can be calculated more easily by the following.

Lemma 3.10. *The Definition 3.9 is equivalent to this definition of spherical ANOVA terms,*

1. The constant term:

$$f_{\emptyset, \mathbf{0}} = P_{\emptyset} \Xi_{\mathbf{0}} f = \frac{1}{\omega_d} \int_{\mathbb{S}^d} f(\mathbf{x}) d\mu_d(\mathbf{x}).$$

2. The mixed terms: For all $\emptyset \neq \mathbf{v} \subseteq [d+1]$ and $\mathbf{u} \supseteq \mathbf{v}$ set $\boldsymbol{\xi} = (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c}) \in \{0, 1\}^{d+1}$. If $d+1 \notin \mathbf{v}$, then $d+1 \notin \mathbf{u}$ and let

$$f_{\mathbf{u}, (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})}(\mathbf{x}_{\mathbf{u}}) = \sum_{\mathbf{v} \subseteq \mathbf{a} \subseteq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{a}|} P_{\mathbf{a}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f]. \quad (3.10)$$

3. The pure even terms: For all $\mathbf{u} \subset [d] \setminus \emptyset$ let

$$f_{\mathbf{u}, \mathbf{0}}(\mathbf{x}_{\mathbf{u}}) = \sum_{\mathbf{a} \subsetneq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{a}|} P_{\mathbf{u}} [\Xi_{\mathbf{0}} f](\mathbf{x}_{\mathbf{u}}).$$

Proof. This proof is similar to the proof of (2.12) for the classical ANOVA decomposition using induction: For $\mathbf{u} = \mathbf{v}$ the sum in (3.10) is just one summand, namely $P_{\mathbf{v}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f]$. For the induction step let (3.10) be fulfilled for all $\mathbf{u} \supseteq \mathbf{v}$ with $|\mathbf{u}| \leq k$ and choose $\mathbf{u} \supseteq \mathbf{v}$ with $|\mathbf{u}| = k+1$. Then, due to the alternating sum of binomial coefficients,

$$\begin{aligned} f_{\mathbf{u}, (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})}(\mathbf{x}_{\mathbf{u}}) &= P_{\mathbf{u}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f(\mathbf{x}_{\mathbf{u}})] - \sum_{\mathbf{v} \subseteq \mathbf{a} \subseteq \mathbf{u}} f_{\mathbf{a}, (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})}(\mathbf{x}_{\mathbf{a}}) \\ &= P_{\mathbf{u}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f(\mathbf{x}_{\mathbf{u}})] - \sum_{\mathbf{v} \subseteq \mathbf{a} \subseteq \mathbf{u}} \sum_{\mathbf{v} \subseteq \mathbf{b} \subseteq \mathbf{a}} (-1)^{|\mathbf{a}| - |\mathbf{b}|} P_{\mathbf{b}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f] \\ &= P_{\mathbf{u}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f(\mathbf{x}_{\mathbf{u}})] - \sum_{\mathbf{v} \subseteq \mathbf{b} \subseteq \mathbf{u}} \sum_{k=0}^{|\mathbf{u} \setminus \mathbf{b}| - 1} \binom{|\mathbf{u} \setminus \mathbf{b}|}{k} (-1)^k P_{\mathbf{b}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f] \\ &= P_{\mathbf{u}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f(\mathbf{x}_{\mathbf{u}})] + \sum_{\mathbf{v} \subseteq \mathbf{b} \subseteq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{b}|} P_{\mathbf{b}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f] \\ &= \sum_{\mathbf{v} \subseteq \mathbf{b} \subseteq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{b}|} P_{\mathbf{b}} [\Xi_{(\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})} f]. \end{aligned}$$

For the pure even case, the proof is the same with $\mathbf{v} = \emptyset$. $\square$

In our spherical ANOVA decomposition we have separated the parities and established the hierarchy. Our proposed definition of a spherical ANOVA decomposition has similarities to the classical ANOVA decomposition, whereby we cannot transfer all properties of the classical ANOVA decomposition to the sphere. See Section 6 for example functions and their decomposition into spherical ANOVA terms.

In our decomposition we have some orthogonality for the spherical ANOVA terms, which we will show in the following section.

3.5 Orthogonality and Redundancy

The decomposition into parity parts belonging to the parity vectors $\boldsymbol{\xi} \in \{0,1\}^{d+1}$, see also (2.9), is an orthogonal decomposition, which also means that for the spherical ANOVA terms defined in Definition 3.9 we have the orthogonality for all $\mathbf{u}, \tilde{\mathbf{u}}$ with $\mathbf{u} \neq \tilde{\mathbf{u}}$,

$$\langle f_{\mathbf{u},\boldsymbol{\xi}}, f_{\tilde{\mathbf{u}},\tilde{\boldsymbol{\xi}}} \rangle_{\mathbb{S}^d} = 0 \quad \text{if } \boldsymbol{\xi} \neq \tilde{\boldsymbol{\xi}}.$$

Unfortunately, we do not have orthogonality for spherical ANOVA terms with the same parity vector $\boldsymbol{\xi}$. Even worse, the equation $\sum_{i=1}^{d+1} x_i^2 = 1$ on the sphere naturally leads to redundancy of the even terms. Every even function $h(x_{d+1})$ can also be written as a $\tilde{h}(x_{d+1}^2) = \tilde{h}(1 - x_1^2 - \dots - x_d^2)$, which is a function depending on the variables x_1 to x_d . Additionally, $\tilde{h}(x_1, \dots, x_d)$ is even in all directions.

For that reason we propose to omit some terms in the Definition 3.9. Namely by the substitution

$$x_{d+1}^2 = 1 - \sum_{i=1}^d x_i^2$$

we propose in Definition 3.9 to omit all terms, which depend on x_{d+1} and where the dependence is even, i.e.,

$$f_{\mathbf{u},\boldsymbol{\xi}} = 0 \quad \text{if } d+1 \in \mathbf{u} \text{ and } \xi_{d+1} = 0.$$

Instead of x_{d+1} one could also choose one other specific variable. For simplicity, we will stick to x_{d+1} in the following. In Figure 2 we illustrate the omitted terms for the case $d = 3$. In total this means, the number of non-zero terms in the decomposition (3.1) reduces to

$$\begin{aligned} 3^{d+1} - 2^d(d+1) - \sum_{i=0}^{d-2} \binom{d}{i} 2^i - 2^d &= 3^{d+1} - 2^d(d+2) - (1+2)^d + \binom{d}{d-1} 2^{d-1} + 2^d \\ &= 2 \cdot 3^d - 2^d(d+2) + d2^{d-1} + 2^d = 2 \cdot 3^d - 2^{d-1}(d+2). \end{aligned} \quad (3.11)$$

3.6 The challenges with the even terms

The aim of this paper is to construct a spherical ANOVA decomposition of the form (3.1), which combines a decomposition into ANOVA terms with the decomposition into parity parts. In Section 3.3 we showed that the terms $\mathcal{A}_{\mathbf{u}}$ should be odd in all directions $i \in \mathbf{u}$. This means that with the operator $\mathcal{A}_{\mathbf{u}}$ we could only represent functions f , which have non-zero spherical ANOVA terms only for $\mathbf{u}$ with $\boldsymbol{\xi} = (\mathbf{1}_{\mathbf{u}}, \mathbf{0}_{\mathbf{u}^c})$. We also include even parts of the function f in our decomposition in Definition 3.9. We analyze these parts here in more detail.

If the function f is a sum of terms, which are odd in all direction, the Definition 3.9 coincides with the decomposition, which we would expect from the classical ANOVA decomposition: Consider for example

$$f(\mathbf{x}) = \underbrace{x_1 x_2^3}_{f_{\{1,2\}, \mathbf{e}_1 + \mathbf{e}_2}} + \underbrace{2x_3 x_4^5}_{f_{\{3,4\}, \mathbf{e}_3 + \mathbf{e}_4}} + \underbrace{0.05x_5}_{f_{\{5\}, \mathbf{e}_5}}$$

for more details about the calculations see A. We marked in Figure 2 all the terms, which are odd in all directions. We do not need to do anything else for these terms.

But as soon as some even term is involved, we run into challenges. We begin with an easy example which shows the problems of the underlying dependence of the variables on the sphere.

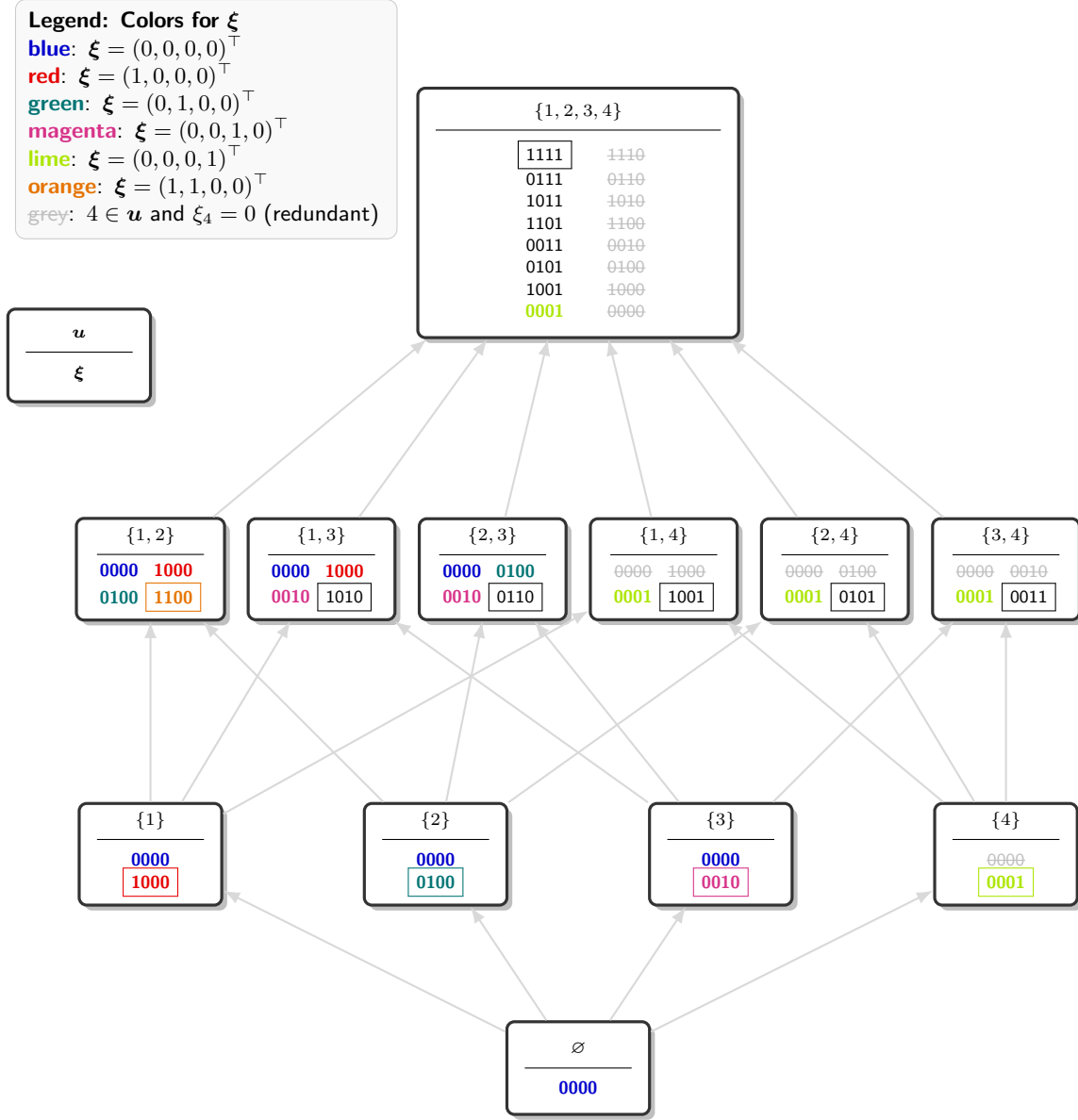


Figure 2: Visualization of all possible indices u and ξ of the terms $f_{u,\xi}$ for the case $d = 3$. As calculated in (3.2), we have in total 49 terms. In Definition 3.9 we defined these terms iteratively. We illustrate this definition by using different colors for the terms belonging to different parity vectors ξ . In Section 3.5 we explain which terms have to be omitted due to redundancies on the sphere. We have grayed out the corresponding vectors ξ . This reduces the decomposition to 34 terms, as calculated in (3.11). We mark all terms where we do not have the challenges in Section 3.6 with a box.

Example 3.11. Let the function be $f(\mathbf{x}) = x_1^2$, which is a pure even function, which depends only on x_1 , but due to the fact that $\sum_{i=1}^d x_i^2 = 1$ on the sphere, the operators $P_{\mathbf{u}}$ applied to this function give

$$\begin{aligned} P_{\emptyset} f &= \frac{1}{d+1} \\ P_1 f &= x_1^2, \\ P_i f &= \frac{1}{d}(1 - x_i^2) & i = \{2, \dots, d+1\}, \\ P_{\{1\} \cup \mathbf{a}} f &= x_1^2 & \text{for } \mathbf{a} \subset [d+1] \setminus \{1\}, \\ P_{\{\mathbf{a}\}} f &= \frac{1}{\omega_{d-|\mathbf{a}|}} (1 - \|\mathbf{x}_{\mathbf{a}}\|^2) \int_{\mathbb{S}^{|\mathbf{a}^c|-1}} x_1^2 d\mu_{|\mathbf{a}^c|-1}(\mathbf{x}_{\mathbf{a}^c}) = \frac{1}{d+1-|\mathbf{a}|} (1 - \|\mathbf{x}_{\mathbf{a}}\|^2) & \text{for } \mathbf{a} \subset [d+1] \setminus \{1\}. \end{aligned}$$

For details about the calculation of the involved integrals, see also Appendix A. Despite the function depending only on x_1 , the application of the operator $P_{\mathbf{u}}$ on the function leads to non-constant terms, even if $1 \notin \mathbf{u}$. In this easy example, we would expect that the spherical ANOVA decomposition is $f_{\emptyset, \mathbf{0}} = \frac{1}{d+1}$, $f_{\{1\}, \mathbf{0}} = x_1^2 - \frac{1}{d+1}$ and all other terms are zero. The Definition 3.9 with the operator $P_{\mathbf{u}}$ leads to a much more complicated decomposition than the expected one, it introduces unnecessarily additional terms. This is due to the natural dependence of the terms $f_{\mathbf{u}, \boldsymbol{\xi}}$ if some entry of $\boldsymbol{\xi}_{\mathbf{u}}$ is zero. Namely, Definition 3.9 leads to

$$\begin{aligned} f_{\emptyset, \mathbf{0}} &= \frac{1}{d+1}, & f_{\{1\}, \mathbf{0}} &= x_1^2 - \frac{1}{d+1}, \\ f_{\{i\}, \mathbf{0}} &= \frac{1}{d}(1 - x_i^2) - \frac{1}{d+1}, & f_{\{1, i\}, \mathbf{0}} &= P_{\{1, i\}} f - f_{\{1\}, \mathbf{0}} - f_{\{i\}, \mathbf{0}} - f_{\emptyset, \mathbf{0}} = -f_{\{i\}, \mathbf{0}}. \end{aligned}$$

And also similar for the higher order terms, which are not zero with this definition we have that

$$f_{\{i\} \cup \mathbf{a}, \mathbf{0}} = -f_{\mathbf{a}, \mathbf{0}} \quad \text{for all } \mathbf{a} \subset \{2, 3, \dots, d+1\}.$$

Our desired function $f = x_1^2$ is completely included in two terms, more precisely, $f = f_{\emptyset, \mathbf{0}} + f_{\{1\}, \mathbf{0}}$. The only non-zero terms are $f_{\mathbf{a}, \boldsymbol{\xi}}$ with $\mathbf{a} \subseteq \mathbf{u} = \{1\}$. All the other terms cancel each other out, such that we can set them to zero.

This example shows how we will modify Definition 3.9. To mitigate the problem of dependencies and make our spherical ANOVA decomposition more similar to the classical decomposition, we show the following.

Lemma 3.12. *Let the function $f(\mathbf{x})$ have the form $f(\mathbf{x}) = g(\mathbf{x}_{\mathbf{u}})$, depend on all variables in $\mathbf{u}$ and have parity $\boldsymbol{\xi}$, i.e. $\Xi_{\boldsymbol{\xi}} f = f$. Let furthermore the parity vector $\boldsymbol{\xi}$ have the form $\boldsymbol{\xi} = (\mathbf{1}_{\mathbf{v}}, \mathbf{0}_{\mathbf{v}^c})^{\top}$ with $\mathbf{v} \subset \mathbf{u}$. Then for all $\emptyset \neq \mathbf{b} \subseteq [d+1] \setminus \mathbf{u}$ we have for the spherical ANOVA terms defined in Definition 3.9 that*

$$0 = \sum_{\mathbf{v} \subseteq \mathbf{a} \subseteq \mathbf{u}} f_{\mathbf{a} \cup \mathbf{b}, \boldsymbol{\xi}}. \quad (3.12)$$

Proof. We illustrate the indices in this proof in Figure 3.

We do the proof by induction over the order $|\mathbf{b}|$. First, let $\mathbf{b} = i$. Then by definition,

$$\begin{aligned} f_{\mathbf{u} \cup i, \boldsymbol{\xi}} &= P_{\mathbf{u} \cup i} [\Xi_{\boldsymbol{\xi}} f] - \sum_{\mathbf{v} \subseteq \mathbf{a} \subset \mathbf{u} \cup i} f_{\mathbf{a}, \boldsymbol{\xi}} = P_{\mathbf{u}} [\Xi_{\boldsymbol{\xi}} f] - \sum_{\mathbf{v} \subseteq \mathbf{a} \subseteq \mathbf{u}} f_{\mathbf{a}, \boldsymbol{\xi}} - \sum_{\mathbf{v} \subseteq \mathbf{a} \subset \mathbf{u}} f_{\mathbf{a} \cup i, \boldsymbol{\xi}} \\ &= - \sum_{\mathbf{v} \subseteq \mathbf{a} \subset \mathbf{u}} f_{\mathbf{a} \cup i, \boldsymbol{\xi}}. \end{aligned}$$

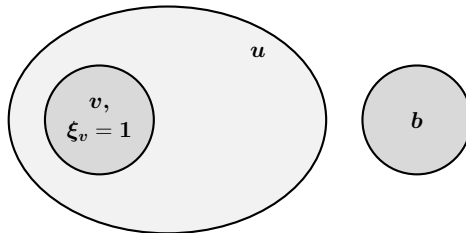


Figure 3: Illustration of the indices $\mathbf{u}, \mathbf{v}, \mathbf{b} \subseteq [d+1]$ in the proof of Lemma 3.12.

Thereby we used that

$$f = \sum_{v \subseteq a \subseteq u} f_{a,\xi}. \quad (3.13)$$

It follows (3.12) for $\mathbf{b} = \{i\}$. For the induction step assume that (3.12) holds for all $\mathbf{b}$ with $|\mathbf{b}| < k$ and choose now $\mathbf{b}$ with $|\mathbf{b}| = k$, then

$$\begin{aligned} f_{u \cup b, \xi} &= P_{u \cup b}[\Xi_\xi f] - \sum_{v \subseteq a \subseteq u \cup b} f_{a, \xi} \\ &= P_u[\Xi_\xi f] - \sum_{v \subseteq a \subseteq u} f_{a, \xi} - \sum_{\emptyset \neq \tilde{\mathbf{b}} \subseteq \mathbf{b}} \sum_{v \subseteq a \subseteq u} f_{a \cup \tilde{\mathbf{b}}, \xi} - \sum_{v \subseteq a \subseteq u} f_{a \cup b, \xi} \\ &= f - \underbrace{\sum_{v \subseteq a \subseteq u} f_{a, \xi}}_{=0, \text{ due to (3.13)}} - \sum_{\emptyset \neq \tilde{\mathbf{b}} \subseteq \mathbf{b}} \underbrace{\sum_{v \subseteq a \subseteq u} f_{a \cup \tilde{\mathbf{b}}, \xi}}_{=0 \text{ due to induction assumption}} - \sum_{v \subseteq a \subseteq u} f_{a \cup b, \xi}. \end{aligned}$$

Hence, (3.12) is also fulfilled for $\mathbf{b}$ with $|\mathbf{u}| = k$. $\square$

For the spherical ANOVA decomposition we want that a function of the form $f(\mathbf{x}) = g(\mathbf{x}_u)$, which depends on all the variables $\mathbf{x}_u$ and not on the variables $\mathbf{x}_{u^c}$, has only non-zero spherical ANOVA terms $f_{\tilde{\mathbf{u}}, \xi}$ if $\tilde{\mathbf{u}} \subseteq \mathbf{u}$. The Definition 3.9 does not fulfill this necessarily. But with the previous lemma, we are able to modify the decomposition by setting $f_{\tilde{\mathbf{u}} \cup \mathbf{a}, \xi} = 0$ for all $\mathbf{a} \subseteq [d+1] \setminus \mathbf{u}$ and $\tilde{\mathbf{u}} \subseteq \mathbf{u}$. Then, we do not have an analytical spherical ANOVA decomposition anymore, but the decomposition is much more similar to the classical ANOVA decomposition.

We will show in the following how to compute the spherical ANOVA decomposition numerically.

4 Choice of basis

For periodic functions one can write every function in $L_2(\mathbb{T}^d)$ with the Fourier series. There is a direct connection between the index $\mathbf{k}$ of the Fourier coefficients and the ANOVA terms $f_{\mathbf{u}}$, i.e. which parts of the Fourier series belong to which ANOVA term, see [26]. Similarly, for orthogonal polynomials with tensor product structure, there is an easy connection between the ANOVA index $\mathbf{u}$ and the frequency indices. See also [35] for an overview about different basis choices for modeling the ANOVA decomposition on tensor product domains.

In our case here, with the sphere as domain, we do not have a tensor product structure of the domain available. However, for every spherical ANOVA term $f_{\mathbf{u}, \xi}$ we construct in the following a basis $\mathcal{B}_{\mathbf{u}, \xi}$.

4.1 The one-dimensional terms

Let $i \in [d+1]$. To describe the space of one-dimensional spherical ANOVA terms $f_{\{i\}, \xi}$ with a basis $\mathcal{B}_{\{i\}, \xi}$, we may choose an orthogonal basis $\eta_k(x_i)$, such that

$$\delta_{k, \ell} = \langle \eta_k(x_i), \eta_\ell(x_i) \rangle_{\mathbb{S}^d} = \int_{-1}^1 \eta_k(x_i) \eta_\ell(x_i) m_i(x_i) dx_i,$$

which means that the basis functions have to be orthogonal on $[-1, 1]$ with respect to $m_i(x_i) \sim \sqrt{1 - x_i^2}^{d-2}$, which are the **Gegenbauer polynomials** with parameter $\alpha = \frac{d-1}{2}$, usually denoted by C_k^α , see [1]. For $k \in \mathbb{N}$ define

$$C_k^\alpha(x) := \frac{1}{\Gamma(\alpha)} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} (-1)^i \frac{\Gamma(\alpha + k - i)}{i! (k - i)!} (2x)^{k-2i}.$$

The first polynomials have the form:

$$\begin{aligned} C_0^{(\alpha)}(x) &= 1, \\ C_1^{(\alpha)}(x) &= 2\alpha x, \\ C_2^{(\alpha)}(x) &= -\alpha + 2\alpha(1 + \alpha)x^2. \end{aligned}$$

For one-dimensional terms, there are only the two possibilities: even and odd functions. The Gegenbauer polynomials have the nice property that for odd k they are odd functions and for even k they are even functions. Since the linear spaces of odd and even function on $[-1, 1]$ are orthogonal with respect to $m_i(x_i) \sim \sqrt{1 - x_i^2}^{d-2}$, we receive an orthogonal basis of the space of odd functions on $[-1, 1]$ for describing the one-dimensional odd spherical ANOVA terms in Definition 3.9 by

$$\mathcal{B}_{i, \mathbf{e}_i} = \left\{ C_k^{(d-1)/2}(x) \mid x \in [-1, 1] \right\}_{k=1,3,5,\dots},$$

which depend on the ambient dimension d . For the even one-dimensional term, naturally a basis can be constructed by the even Gegenbauer polynomials. Let $\mathbf{u} = \{i\}$, then due to our Definition 3.9,

$$P_\emptyset f_{\{i\}, \boldsymbol{\xi}} = P_\emptyset \mathcal{E}_1 P_1 f - P_\emptyset P_\emptyset f = 0,$$

which means that the coefficient for the constant in the even one-dimensional term is zero, such that an even basis for $f_{\{i\}, \mathbf{0}}$ reads

$$\mathcal{B}_{i, \mathbf{0}} = \left\{ C_k^{(d-1)/2}(x) \mid x \in [-1, 1] \right\}_{k=2,4,6,\dots}.$$

4.2 The higher-dimensional terms

For describing the spherical ANOVA terms $f_{\mathbf{u}, \boldsymbol{\xi}}$ on the domain $\mathbb{B}^{|\mathbf{u}|}$, we are searching for an orthogonal basis $\mathcal{B}_{\mathbf{u}, \boldsymbol{\xi}} = \{\eta_k(\mathbf{x}_{\mathbf{u}}) \mid \mathbf{x}_{\mathbf{u}} \in \mathbb{B}^{|\mathbf{u}|}\}_k$ with orthogonality with respect to the weight function $m_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) \sim \sqrt{1 - \|\mathbf{x}_{\mathbf{u}}\|^2}^{d-1-|\mathbf{u}|}$, since then the basis is orthogonal on $\mathbb{S}^d$ with respect to the uniform measure on the sphere μ_d . Changing coordinates to polar coordinates, this weight function becomes

$$m_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) \sim \sqrt{1 - r^2}^{d-1-|\mathbf{u}|},$$

which leads us to the Jacobi polynomials.

Remark 4.1. Jacobi polynomials $P_n^{\alpha, \beta}(x)$ form an orthonormal basis in $L_2([-1, 1])$ with respect to the weight $(1-x)^\alpha (1+x)^\beta$. For polynomials on $[0, 1]$ we use the substitution $t = r^2 - 1$, which means that

$$\begin{aligned} \delta_{k, \ell} &= \int_{-1}^1 P_k^{\alpha, \beta}(t) P_\ell^{\alpha, \beta}(t) (1-t)^\alpha (1+t)^\beta dt \\ &= 4 \int_0^1 P_k^{\alpha, \beta}(2r^2 - 1) P_\ell^{\alpha, \beta}(2r^2 - 1) [2(1-r^2)]^\alpha (2r^2)^\beta r dr \\ &= 4 \cdot 2^{\alpha+\beta} \int_0^1 P_k^{\alpha, \beta}(2r^2 - 1) P_\ell^{\alpha, \beta}(2r^2 - 1) (1-r^2)^\alpha r^{2\beta+1} dr. \end{aligned}$$

With this, an orthogonal basis on $\mathbb{B}^n$, used in the literature [6, Prop. 11.1.13], [7, Prop. 5.2.1] is the following. For $N \in \mathbb{N}$, $0 \leq j \leq \frac{N}{2}$ and $1 \leq k \leq \dim(\mathcal{H}_{N-2j}(\mathbb{S}^{|\mathbf{u}|}))$ define

$$P_{j,k}^N(\mathbf{x}) = P_j^{\frac{d-1-|\mathbf{u}|}{2}, N-2j+\frac{d-2}{2}}(2\|\mathbf{x}\|^2 - 1) Y_{k, N-2j}(\mathbf{x}). \quad (4.1)$$

These form an orthogonal basis of the space of orthogonal polynomials of degree exact N with respect to the inner product

$$\langle f(\mathbf{x}_{\mathbf{u}}), g(\mathbf{x}_{\mathbf{u}}) \rangle_{\mathbb{S}^d} = \langle f(\mathbf{x}_{\mathbf{u}}), g(\mathbf{x}_{\mathbf{u}}) \rangle_{(\mathbb{B}^{|\mathbf{u}|}, m_{\mathbf{u}})} = c \int_{\mathbb{B}^{|\mathbf{u}|}} f(\mathbf{x}_{\mathbf{u}}) g(\mathbf{x}_{\mathbf{u}}) m_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) d\mathbf{x}_{\mathbf{u}},$$

where c is a normalization constant.

4.2.1 The two-dimensional terms

In the two-dimensional case the spherical harmonics are precisely the cosine and sine functions. For the sake of simplicity of the notation, we will construct here a basis for the terms $f_{\{1,2\},\xi}$, but for all other two-dimensional terms the procedure is the same.

In two-dimensional polar coordinates $(x_1, x_2) = (r \cos \theta, r \sin \theta)$, as basis $Y_{k,N-2j}(\mathbf{x})$ we use (2.8). Then, the orthogonal basis in (4.1) for $\mathbb{B}^2$ reads

$$\left\{ P_j^{\frac{d-3}{2}, N-2j} (2r^2 - 1) r^{N-2j} \cos((N-2j)\theta) \right\}_{N \in \mathbb{N}_0, 0 \leq j \leq \frac{N}{2}} \cup \left\{ P_j^{\frac{d-3}{2}, N-2j} (2r^2 - 1) r^{N-2j} \sin((N-2j)\theta) \right\}_{N \in \mathbb{N}_0, 0 \leq j < \frac{N}{2}}. \quad (4.2)$$

In polar coordinates, the parity vector $\xi = (\xi_1, \xi_2, \mathbf{0}_{d-1})^\top$ defines the parity by

$$g(r, \theta) = (-1)^{\xi_1} g(r, -\theta), \quad g(r, \theta) = (-1)^{\xi_2} g(r, \pi - \theta).$$

In (4.2) we have to classify the basis functions to the parity vectors and furthermore, due to the dependencies of the spherical ANOVA terms we have to be careful additionally and have to omit some of the basis functions, which are already included in the terms of lower order 0 or 1.

The selection rule is based on the polynomial degree N and the resulting parity vector $\xi_{\{1,2\}} \in \{0, 1\}^2$ of the term $f_{\mathbf{u},\xi}$. The term $N - 2j$ has the same parity as N , such that a distinction between even and odd N is useful. We exclude all constant ($N = 0$) and linear ($N = 1$) terms. For the purely even parity $\xi = (0, 0)$, we additionally exclude quadratic terms ($N = 2$) that decompose into sums of univariate functions. The parity is determined by N and the angular function type, as detailed in Table 1. There the basis functions for constructing $\mathcal{B}_{\{1,2\},\xi}$ are summarized.

N	j	Type	Form	$\xi_{\{1,2\}}$	Action	Reason
0	0	cos	1	(0, 0)	Omit	Covered by $f_{\emptyset, \mathbf{0}}$
1	0	cos	x_1	(0, 1)	Omit	Covered by $f_{\{1, \mathbf{e}_1\}}$
1	0	sin	x_2	(1, 0)	Omit	Covered by $f_{\{2, \mathbf{e}_2\}}$
2	0	cos	$x_1^2 - x_2^2$	(0, 0)	Omit	Sum of 1d terms
2	1	cos	$x_1^2 + x_2^2$	(0, 0)	Omit	Sum of 1d terms
2	0	sin	$x_1 x_2$	(1, 1)	Keep	True interaction
≥ 3 , odd	$0 \leq j \leq \frac{N-3}{2}$	cos		(1, 0)	Keep	
≥ 3 , odd	$j = \frac{N-1}{2}$	cos		(1, 0)	Omit	Redundancy
≥ 3 , odd	$0 \leq j \leq \frac{N-3}{2}$	sin		(0, 1)	Keep	
≥ 3 , odd	$j = \frac{N-1}{2}$	sin		(0, 1)	Omit	Redundancy
≥ 4 , even	$0 \leq j \leq \frac{N}{2} - 2$	cos		(0, 0)	Keep	
≥ 4 , even	$\frac{N}{2} - 1 \leq j \leq \frac{N}{2}$	cos		(0, 0)	Omit	Redundancy
≥ 4 , even	$0 \leq j < \frac{N}{2}$	sin		(1, 1)	Keep	

Table 1: Classification and selection of two-dimensional basis functions based on degree N and angular frequency $m = N - 2j$. Terms marked as *Omit* are redundant with lower-order ANOVA terms.

The case $N = 2$:

A rationale for discarding the terms with $N = 2$ for $\xi = (0, 0)$ is the specific geometry of the sphere. The radial term $P_1(2r^2 - 1)$ and the angular term $r^2 \cos(2\theta)$ correspond to linear combinations of x_1^2 and x_2^2 . Since the even one-dimensional terms $f_{\{i\}, \mathbf{0}}$ already span the space of x_i^2 (orthogonalized against the constant), including these $N = 2$ terms in the interaction basis would introduce linear dependencies. In contrast, the term $r^2 \sin(2\theta) \sim x_1 x_2$ is kept as it belongs to the parity class $\xi_{\{1,2\}} = (1, 1)$ and represents a pure interaction that cannot be formed by sums of univariate functions.

For $N \geq 3$ only the parity $\xi_{\{1,2\}} = (1, 1)$ is carefree. Concerning the other cases, we have additional redundancy, which leads to more omitted terms. The reason for this is as follows.

The case $\xi_{\{1,2\}} = (0, 0)$:

Let N be even, then the space of homogeneous polynomials in two variables x_1 and x_2 of the form $x_1^{2k}x_2^{2\ell}$ with $2k + 2\ell = N$ has dimension $\frac{N}{2} + 1$, which includes the two one-dimensional terms x_1^N and x_2^N . This means from the possibilities for j , namely $0 \leq j \leq \frac{N}{2}$ we can keep only $\frac{N}{2} - 1$ and have to omit two, which are redundant.

The case $\xi_{\{1,2\}} = (1, 0)$:

Let N be odd, then the space of homogeneous polynomials in two variables x_1 and x_2 of the form $x_1^{2k+1}x_2^{2\ell}$ with $2k + 1 + 2\ell = N$ has dimension $\frac{N+1}{2}$, which includes the two one-dimensional term x_1^N . This means from the possibilities for j , namely $0 \leq j \leq \frac{N-1}{2}$ we can keep only $\frac{N-1}{2}$ and have to omit one, which is redundant. The case $\xi_{\{1,2\}} = (0, 1)$ is similar.

A similar approach can be used for higher-dimensional terms. However, we will limit ourselves here to this explanation for two-dimensional terms. We do reduce the redundancy between the terms of different order. However these terms are not necessarily orthogonal.

5 Low dimensional approximation

In many applications, functions on the sphere depend on a large number of variables but are effectively dominated by terms of low order. The ANOVA decomposition allows us to identify this effective dimension and approximate the function by a sum of terms involving only a few variables. This is particularly valuable for mitigating the **curse of dimensionality** and for developing efficient quadrature or interpolation algorithms. Approaches for low dimensional approximation without the ANOVA decomposition can be found for example in [36].

Similar to the approaches [26, 22] for the classical ANOVA decomposition or [28] for the generalized ANOVA decomposition, we propose to approximate the function f using only low-dimensional terms, i.e.,

$$f \approx f_\emptyset + \sum_{1 \leq |\mathbf{u}| \leq q} \sum_{\xi \in \Upsilon_{\mathbf{u}}} f_{\mathbf{u}, \xi} \quad \text{for some } q \ll d.$$

The next step is to approximate every term by using the specific basis $\mathcal{B}_{\mathbf{u}, \xi} = \{\eta_k^{\mathbf{u}, \xi}\}_{k \in \mathcal{I}_{\mathbf{u}}}$ introduced in Section 4,

$$f_{\mathbf{u}, \xi}(\mathbf{x}_{\mathbf{u}}) \approx \sum_{k \in \mathcal{I}_{\mathbf{u}}} a_k^{\mathbf{u}, \xi} \eta_k^{\mathbf{u}, \xi}(\mathbf{x}_{\mathbf{u}}). \quad (5.1)$$

Substituting the basis expansion (5.1) into the truncated ANOVA decomposition leads to the global approximation ansatz

$$f(\mathbf{x}) \approx a_\emptyset + \sum_{1 \leq |\mathbf{u}| \leq q} \sum_{\xi \in \Upsilon_{\mathbf{u}}} \sum_{k \in \mathcal{I}_{\mathbf{u}}} a_k^{\mathbf{u}, \xi} \eta_k^{\mathbf{u}, \xi}(\mathbf{x}_{\mathbf{u}}).$$

To determine the unknown coefficients $a_k^{\mathbf{u}, \xi}$, we rely on a set of sampling nodes

$$\mathcal{X} = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(M)}\} \subset \mathbb{S}^d$$

and the corresponding function values $\mathbf{f} = (f(\mathbf{x}^{(1)}), \dots, f(\mathbf{x}^{(M)}))^T \in \mathbb{R}^M$. Let N denote the total number of basis functions in the truncated expansion and let $\mathbf{a} \in \mathbb{R}^N$ be the vector stacking all coefficients $a_k^{\mathbf{u}, \xi}$. Evaluating the approximation ansatz at the sampling nodes yields the linear system of equations

$$\mathbf{A}\mathbf{a} \approx \mathbf{y},$$

where the system matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ contains the evaluations of the basis functions. Specifically, the entry $A_{i, \ell}$ corresponds to the ℓ -th basis function evaluated at the i -th sampling node $\mathbf{x}^{(i)}$.

In practical applications, the number of samples M is typically chosen to be larger than the number of unknowns N , resulting in an overdetermined system. Consequently, we seek the coefficient vector $\mathbf{a}$ that minimizes the residual in the Euclidean norm, solving the least squares problem

$$\min_{\mathbf{c} \in \mathbb{R}^N} \|\mathbf{A}\mathbf{c} - \mathbf{y}\|_2. \quad (5.2)$$

While this problem can be solved using direct methods (e.g., via QR decomposition or SVD), the dimensions of $\mathbf{A}$ can become large for high-dimensional approximations. Moreover, as discussed in previous sections, the basis on the sphere may exhibit linear dependencies or ill-conditioning for certain parity combinations. To address these challenges efficiently, we propose to solve (5.2) using the **LSQR algorithm** [25]. LSQR is an iterative Krylov subspace method that is particularly well-suited for large, sparse, or ill-conditioned systems and possesses intrinsic regularization properties when the iteration is terminated early.

Due to the dependence of the terms of different order, the following procedure may be beneficial: We propose to do the approximation first using all terms of order one. Then, we do a second approximation for the residuum using all terms of order 2, which could be done further. In Section 6 and Figure 4 we apply both possibilities, the plain LSQR with all terms at once and the iterative increase of the order of the terms.

5.1 Variances for sensitivity analysis

Global sensitivity analysis has become a cornerstone of computer experiments, providing essential insights into the input-output dependencies of numerical simulation models. Building on the foundational works [32, 31], the field has evolved significantly over the last two decades, yielding a wide array of advanced statistical methodologies and refinements of the original Sobol' indices.

While classical variance-based sensitivity analysis (Sobol' indices) relies on the orthogonality of the ANOVA terms to decompose the total variance into a sum of partial variances, this property assumes independent input variables. On the sphere $\mathbb{S}^d$, the inputs are intrinsically coupled via the constraint $\|\mathbf{x}\| = 1$. Consequently, for even-parity terms, a strict orthogonality cannot be maintained, and the covariance between interaction terms is generally non-zero. Despite this lack of perfect orthogonality, the variances $\sigma^2(f_{\mathbf{u},\boldsymbol{\xi}})$ remain meaningful sensitivity measures for the following reasons:

- Following the framework of sensitivity analysis for dependent variables, see for example [17, 28], the contribution of a variable is always a combination of its marginal effect and its correlation with other variables. On the sphere, these correlations are not artifacts but physical properties of the domain. Our indices capture the combined effect of a variable and its induced changes on the sphere.
- A crucial advantage of our proposed construction is the rigorous exclusion of redundant basis terms. By ensuring that our dictionary of basis functions is linearly independent, we prevent variance inflation. In a redundant system, variance could be arbitrarily shifted between linearly dependent terms (e.g., between f_1 and a redundant part of $f_{1,2}$), leading to unstable and uninterpretable indices. Our restrictive construction ensures that the computed variances are stable and reflect the intrinsic coupling of the sphere rather than numerical artifacts.

Lemma 5.1. *The spherical ANOVA terms defined in Definition 3.9 have mean zero, i.e.,*

$$P_{\emptyset} f_{\mathbf{u},\boldsymbol{\xi}} = 0 \quad \text{for all } \mathbf{u} \in \mathcal{D} \setminus \emptyset \text{ and all } \boldsymbol{\xi} \in \Upsilon_{\mathbf{u}}.$$

Proof. Since the spherical ANOVA terms $f_{\mathbf{u},\boldsymbol{\xi}}$ defined in Definition 3.9 have parity $\boldsymbol{\xi}$, the mean is automatically zero, if the function $f_{\mathbf{u},\boldsymbol{\xi}}$ is odd in at least one direction. The only remaining case is when $\boldsymbol{\xi} = \mathbf{0}$. But in this case we have for $\mathbf{u} \neq \emptyset$,

$$\begin{aligned} P_{\emptyset} f_{\mathbf{u},\mathbf{0}} &= P_{\emptyset} \left(P_{\mathbf{u}}[\Xi_{\mathbf{0}} f](\mathbf{x}_{\mathbf{a}}) - \sum_{\mathbf{a} \subsetneq \mathbf{u}} f_{\mathbf{a},\mathbf{0}}(\mathbf{x}_{\mathbf{a}}) \right) = P_{\emptyset} f - P_{\emptyset} \left(\sum_{\mathbf{a} \subsetneq \mathbf{u}} f_{\mathbf{a},\mathbf{0}}(\mathbf{x}_{\mathbf{a}}) \right) \\ &= P_{\emptyset} \left(\sum_{\emptyset \neq \mathbf{a} \subsetneq \mathbf{u}} f_{\mathbf{a},\mathbf{0}}(\mathbf{x}_{\mathbf{a}}) \right) = \sum_{\emptyset \neq \mathbf{a} \subsetneq \mathbf{u}} P_{\emptyset} f_{\mathbf{a},\mathbf{0}}(\mathbf{x}_{\mathbf{a}}). \end{aligned}$$

By induction this finishes the proof by using the induction assumption. $\square$

Using the previous lemma, the variances of the spherical ANOVA terms can be easily calculated

by

$$\begin{aligned}\sigma^2(f_{\mathbf{u},\boldsymbol{\xi}}) &:= \int_{\mathbb{S}^d} |f_{\mathbf{u},\boldsymbol{\xi}}(\mathbf{x}_{\mathbf{u}}) - \int_{\mathbb{S}^d} f_{\mathbf{u},\boldsymbol{\xi}}(\mathbf{x}_{\mathbf{u}}) d\mu_d(\mathbf{x})|^2 d\mu_d(\mathbf{x}) = \int_{\mathbb{S}^d} |f_{\mathbf{u},\boldsymbol{\xi}}(\mathbf{x}_{\mathbf{u}})|^2 d\mu_d(\mathbf{x}) \\ &\stackrel{(2.3)}{=} \int_{\mathbb{B}^{|\mathbf{u}|}} |f_{\mathbf{u},\boldsymbol{\xi}}(\mathbf{x}_{\mathbf{u}})|^2 m_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) d\mathbf{x}_{\mathbf{u}},\end{aligned}$$

which is a weighted variance of the term $f_{\mathbf{u}}$ on $\mathbb{B}^{|\mathbf{u}|}$. In Section 4 we introduced the basis $\mathcal{B}_{\mathbf{u},\boldsymbol{\xi}}$, which is orthogonal with respect to the weight function $m_{\mathbf{u}}$. Using the approximation (5.1), the variances are approximated by

$$\sigma^2(f_{\mathbf{u},\boldsymbol{\xi}}) \approx \sum_{\mathbf{x} \in \mathcal{X}} |f_{\mathbf{u},\boldsymbol{\xi}}(\mathbf{x})|^2 \approx \sum_{\mathbf{x} \in \mathcal{X}} \left| \sum_{k \in \mathcal{I}_{\mathbf{u}}} a_k^{\mathbf{u},\boldsymbol{\xi}} \eta_k(\mathbf{x}_{\mathbf{u}}) \right|^2.$$

Since terms with different parity vector are orthogonal with respect to $m_{\mathbf{u}}$, we define for $\mathbf{u} \neq \emptyset$ the Sobol indices on the sphere,

$$S_{\mathbf{u}} := \frac{1}{\sigma^2(\mathbf{f})} \sum_{\boldsymbol{\xi} \in \Upsilon_{\mathbf{u}}} \sigma^2(f_{\mathbf{u},\boldsymbol{\xi}}) \approx \frac{1}{\sigma^2(\mathbf{f})} \sum_{\boldsymbol{\xi} \in \Upsilon_{\mathbf{u}}} \sum_{\mathbf{x} \in \mathcal{X}} \left| \sum_{k \in \mathcal{I}_{\mathbf{u}}} a_k^{\mathbf{u},\boldsymbol{\xi}} \eta_k(\mathbf{x}_{\mathbf{u}}) \right|^2 =: \tilde{S}_{\mathbf{u}}, \quad (5.3)$$

where $\sigma^2(\mathbf{f})$ is the variance of the data vector $\mathbf{f}$ and the expression $\tilde{S}_{\mathbf{u}}$ can be calculated easily using parts of the matrix $\mathbf{A}$ and the coefficient vector $\mathbf{a}$. We interpret the normalized ratio $S_{\mathbf{u}}$ and its approximation $\tilde{S}_{\mathbf{u}}$ not strictly as a percentage of disjoint variance, but as a relative importance index reflecting the stable amplitude of the corresponding spherical interaction terms.

For the classical ANOVA decomposition the Sobol indices (or sometimes called global sensitivity indices) also denote the fraction of the variance of the ANOVA term and the total variance of the function. Due to the orthogonality between all terms, they sum up to one.

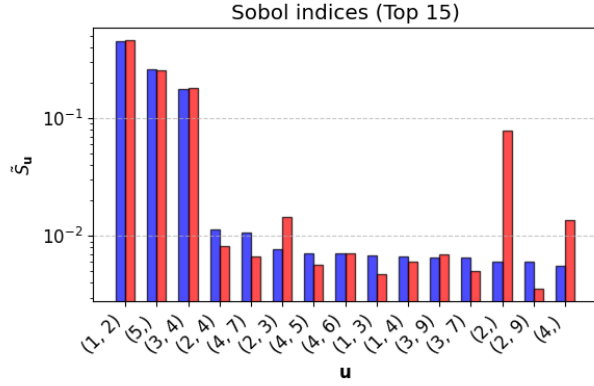
6 Examples and numerical tests

In this chapter, we evaluate the practical applicability and accuracy of the derived spherical ANOVA decomposition through numerical experiments. The primary objective is to demonstrate that the proposed spherical ANOVA decomposition provides a meaningful and interpretable representation of functions on $\mathbb{S}^d$.

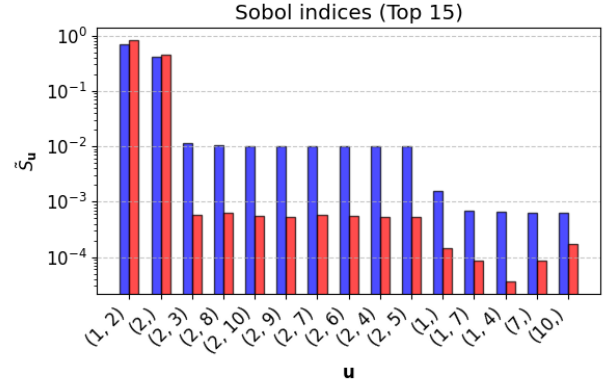
To this end, we consider a set of test functions with known interaction structures. We compute the variances of the individual ANOVA terms $S_{\mathbf{u}}$ from (5.3) numerically using the proposed basis approximation and the LSQR solver in Python.

Let us consider the following test functions $f_A, f_B, f_C: \mathbb{S}^{10} \rightarrow \mathbb{R}$, defined by

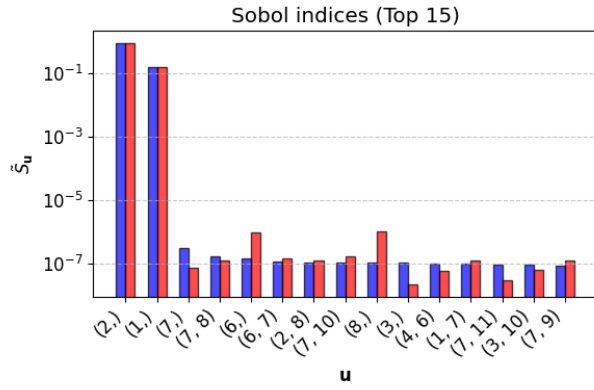
$$\begin{aligned}f_A(\mathbf{x}) &:= x_1 x_2^3 + 2x_3 x_4^5 + 0.05 x_5 \\ f_B(\mathbf{x}) &:= x_2 x_1^2 \\ f_C(\mathbf{x}) &:= x_1^4 + x_2^2 \\ f_D(\mathbf{x}) &:= 5x_1 x_2^2 + x_4 + e^{x_3} + 10 \sin(3\pi x_5) x_2^4 \\ f_E(\mathbf{x}) &:= \sin(x_1) + 7 \sin^2(x_2) + 0.1 x_3^4 \sin(x_1) \\ f_F(\mathbf{x}) &:= x_1^2 x_2^2 x_3^2\end{aligned}$$



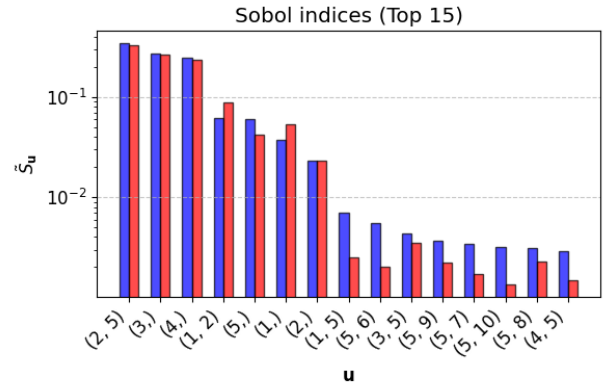
(a) Test function f_A .



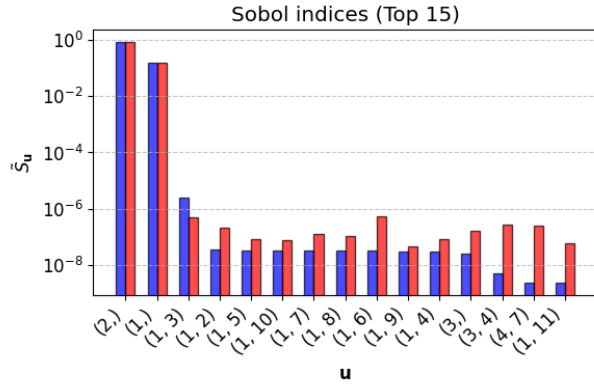
(b) Test function f_B .



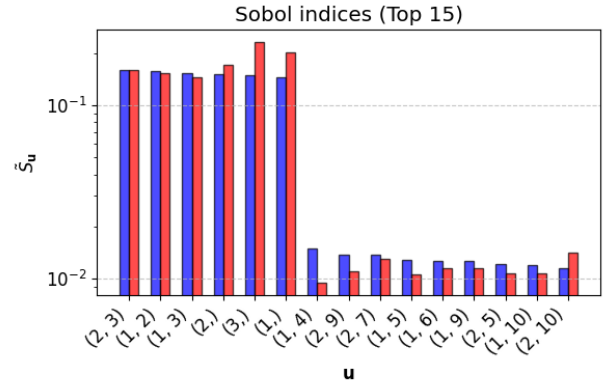
(c) Test function f_C .



(d) Test function f_D .



(e) Test function f_E .



(f) Test function f_F .

Figure 4: Approximated Sobol indices (5.3) for the different test functions. We plot the results from LSQR with all terms up to order $q = 2$ together (red) and the procedure of increasing the order of the used terms iteratively (blue). Note that we use logarithmic scaling.

The spherical ANOVA decomposition of some of these functions is calculated in details in Appendix A. For the numerical tests we draw $M = 10000$ random samples on the sphere $\mathbb{S}^{10}$, which is relatively sparse in this dimension, and use the basis described in Section 4 of order up to $q = 2$ and polynomials of maximal total order $N = 10$. We use (5.3) to approximate the Sobol indices with this basis. In Figure 4 we illustrate the resulting indices.

- If all parts of the function are odd functions, like for f_A , the non-zero terms are detected easily. One can see in the diagram that it can happen that plain LSQR puts some variance in one-dimensional terms, which should be zero.
- The function f_B consists of one term with mixed parity. We clearly see here the effect, that the terms $f_{\{1,i\}}$ for $i > 2$ are not zero, but in the order of magnitude of the Sobol indices $\tilde{S}_{\mathbf{u}}$ one can see that despite the dependence of the terms, our algorithm is able to detect the two important term $f_{\{1,2\},0}$ and $f_{\{2\},0}$.
- The function f_C is a sum of two even functions. Our algorithm clearly detects these two one-dimensional terms. Due to the dependence of the even terms, we get some variance for the terms with $\mathbf{u} = \{2, i\}$ for $i \geq 3$, the reason was described in 3.6. Here, the order of magnitude of the misleading terms is slightly smaller for the case where we use LSQR with all terms of order 0, 1 and 2 together.
- If there are other terms than polynomials involved, like in the function f_D , we still are able to detect the seven non-zero terms in the spherical ANOVA decomposition.
- f_E is the Ishigami function, where we are able to find the three non-zero terms.
- The function f_F is a term of order 3, which means that our approximation with terms of order maximal $q = 2$ could not model the function exactly. However, also here we see clearly the non-zero terms belonging to $\mathbf{u} \subset \{1, 2, 3\}$.

6.1 Summary and outlook

The spherical ANOVA decomposition is a powerful mathematical framework designed to analyze high-dimensional functions defined on the high-dimensional sphere $\mathbb{S}^d$. Unlike the classical ANOVA, this approach respects the intrinsic geometry of the sphere, making it essential for problems involving directional data or isotropic constraints.

Our decomposition allows for the breakdown of a complex function into a sum of orthogonal terms, where each term represents the contribution of specific subsets of variables. This reveals the effective dimension of the function.

By calculating global to Sobol indices, we can quantify how much variance is driven by single variables versus complex interactions, effectively opening the black box of high-dimensional models. We are able to construct sparse approximation algorithms that avoid the curse of dimensionality by identifying that a function is dominated by low-order interactions. It paves the way for designing specialized cubature rules that focus computational effort only on the relevant subspaces identified by the ANOVA terms. In fields dealing with normalized feature vectors (such as text embeddings or gene expression profiles), our spherical ANOVA decomposition can serve as a robust feature selection and noise reduction tool.

Acknowledgment

I would like to thank Daniel Potts for many helpful discussions.

References

- [1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, 1964.

- [2] H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numer.*, 13:147–269, 2004. doi:10.1017/S0962492904000182.
- [3] R. Caffisch, W. Morokoff, and A. Owen. Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension. *J. Comput. Finance*, 1(1):27–46, 1997. doi:10.21314/jcf.1997.005.
- [4] M. Chu, N. Del Buono, L. Lopez, and T. Politi. On the low-rank approximation of data on the unit sphere. *SIAM Journal on Matrix Analysis and Applications*, 27(1):46–60, Jan. 2005. doi:10.1137/s0895479803433295.
- [5] S. da Veiga. Kernel-based ANOVA decomposition and shapley effects – application to global sensitivity analysis, 2021. arXiv:2101.05487, doi:10.48550/ARXIV.2101.05487.
- [6] F. Dai and Y. Xu. *Approximation Theory and Harmonic Analysis on Spheres and Balls*. 01 2013. doi:10.1007/978-1-4614-6660-4.
- [7] C. F. Dunkl and Y. Xu. *Orthogonal Polynomials of Several Variables*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2nd edition, 2014. doi:10.1017/CB09781107786134.
- [8] B. Efron and C. Stein. The Jackknife Estimate of Variance. *Ann. Stat.*, 9(3):586–596, 1981. doi:10.1214/aos/1176345462.
- [9] F. Filbir and W. Themistoclakis. Polynomial approximation on the sphere using scattered data. *Math. Nachr.*, 281:650–668, 2008. doi:10.1002/mana.200710633.
- [10] M. Griebel. Sparse grids and related approximation schemes for higher dimensional problems. In *FOCM 2005*, pages 106–161. Cambridge University Press, 2006.
- [11] M. Griebel and M. Holtz. Dimension-wise integration of high-dimensional functions with applications to finance. *J. Complex.*, 26(5):455–489, 2010. doi:10.1016/j.jco.2010.06.001.
- [12] M. Hegland. Adaptive sparse grids. In *CTAC2001*, volume 44, pages C335–C353, 2003. doi:10.21914/anziamj.v44i0.685.
- [13] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, 19(3):293–325, 1948. doi:10.1214/aoms/1177730196.
- [14] M. Holtz. *Sparse grid quadrature in high dimensions with applications in finance and insurance*, volume 77 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2011. doi:10.1007/978-3-642-16004-2.
- [15] G. Hooker. Generalized functional ANOVA diagnostics for high-dimensional functions of dependent variables. *J. Comput. Graph. Statist.*, 16(3):709–732, 2007. URL: <http://www.jstor.org/stable/27594267>.
- [16] T. Jiang and A. B. Owen. Quasi-regression with shrinkage. *Math. Comput. Simul.*, 62(3):231–241, 2003. 3rd IMACS Seminar on Monte Carlo Methods. doi:10.1016/S0378-4754(02)00253-7.
- [17] S. Kucherenko, S. Tarantola, and P. Annoni. Estimation of global sensitivity indices for models with dependent variables. *Comput. Phys. Commun.*, 183(4):937–946, 2012. doi:10.1016/j.cpc.2011.12.020.
- [18] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski. On decompositions of multi-variate functions. *Math. Comp.*, 79(270):953–966, 2009. doi:10.1090/s0025-5718-09-02319-9.
- [19] C. Labreuche. *Contribution of Subsets of Variables in Global Sensitivity Analysis with Dependent Variables*, pages 233–248. Springer Nature Switzerland, Nov. 2024. doi:10.1007/978-3-031-76235-2_18.
- [20] C. Lemieux and A. B. Owen. Quasi-regression and the relative importance of the ANOVA components of a function. In *MCQMC 2000*, pages 331–344. Springer Berlin Heidelberg, 2002. doi:10.1007/978-3-642-56046-0_22.

- [21] G. Li and H. Rabitz. General formulation of HDMR component functions with independent and correlated variables. *J. Math. Chem.*, 50:99–130, 2012. doi:10.1007/s10910-011-9898-0.
- [22] L. Lippert, D. Potts, and T. Ullrich. Fast hyperbolic wavelet regression meets ANOVA. *Numer. Math.*, 154:155–207, 2023. doi:10.1007/s00211-023-01358-8.
- [23] R. Liu and A. B. Owen. Estimating mean dimensionality of analysis of variance decompositions. *J. Amer. Statist. Assoc.*, 101(474):712–721, 2006. doi:10.1198/016214505000001410.
- [24] E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems Volume I: Linear Information*. Eur. Math. Society, EMS Tracts in Mathematics Vol 6, 2008.
- [25] C. C. Paige and M. A. Saunders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Software*, 8:43–71, 1982. doi:10.1145/355984.355989.
- [26] D. Potts and M. Schmischke. Approximation of high-dimensional periodic functions with Fourier-based methods. *SIAM J. Numer. Anal.*, 59(5):2393–2429, 2021. doi:10.1137/20M1354921.
- [27] D. Potts and M. Schmischke. Learning multivariate functions with low-dimensional structures using polynomial bases. *J. Comput. Appl. Math.*, 403:113821, 2022. doi:10.1016/j.cam.2021.113821.
- [28] D. Potts and L. Weidensager. ANOVA-boosting for random Fourier features. *Appl. Comput. Harmon. Anal.*, 79:101789, 2025. doi:10.1016/j.acha.2025.101789.
- [29] H. Rabitz and Ö. F. Aliş. General foundations of high-dimensional model representations. *J. Math. Chem.*, 25:197–233, 1999. doi:10.1023/A:1019188517934.
- [30] S. Rahman. A generalized ANOVA dimensional decomposition for dependent probability measures. *SIAM-ASA J. Uncertain. Quantif.*, 2(1):670–697, 2014. doi:10.1137/120904378.
- [31] A. Saltelli and I. Sobol. Sensitivity analysis for nonlinear mathematical models: Numerical experience. *Mat. Model.*, 7, 1995.
- [32] I. M. Sobol. On sensitivity estimation for nonlinear mathematical models. *Keldysh Applied Mathematics Institute*, 1:112–118, 1990.
- [33] I. M. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.*, 55(1-3):271–280, 2001. doi:10.1016/s0378-4754(00)00270-6.
- [34] G. Wahba. *Spline Models for Observational Data*. Springer, SIAM, 1990. doi:10.1137/1.9781611970128.
- [35] L. Weidensager. *Efficient High-Dimensional Approximation: ANOVA Decomposition Meets Wavelets and Random Fourier Features*. PhD thesis, Universitätsverlag Technische Universität Chemnitz, 2025. doi:10.51382/978-3-96100-272-6.
- [36] Y. Xie, R. Shi, H. Schaeffer, and R. Ward. SHRIMP: sparser random feature models via iterative magnitude pruning. In *MSML*, volume 190, pages 303–318. PMLR, 2022. URL: <https://proceedings.mlr.press/v190/xie22a.html>.

A Calculation of the spherical ANOVA decompositions in Section 6

In this section we calculate the spherical ANOVA decomposition from Definition 3.9 for some of the test functions. We use in the following these integrals frequently,

$$\begin{aligned} \int_{\mathbb{S}^{d-1}} x_i^2 d\mu_{d-1}(\mathbf{x}) &= \frac{\omega_{d-1}}{d} \quad \text{if } i \in [d+1] \\ \int_{\mathbb{S}^{d-1}} x_i^{2k+1} d\mu_{d-1}(\mathbf{x}) &= 0 \quad \text{if } k \in \mathbb{N}, i \in [d+1]. \end{aligned}$$

- $f_A(\mathbf{x}) = x_1x_2^3 + 2x_3x_4^5 + 0.05x_5$

The parity decomposition (2.9) of this function is

$$\Xi_{\xi}f_A = \begin{cases} x_1x_2^3 & \text{if } \xi = \mathbf{e}_1 + \mathbf{e}_2 \\ 2x_3x_4^5 & \text{if } \xi = \mathbf{e}_3 + \mathbf{e}_4 \\ 0.05x_5 & \text{if } \xi = \mathbf{e}_5 \\ 0 & \text{otherwise .} \end{cases}$$

Applying the operators $P_{\mathbf{u}}$ to this function yields that

$$\begin{aligned} P_{\{1,2\}}f_A &= x_1x_2^3 \\ P_{\{3,4\}}f_A &= 2x_3x_4^5 \\ P_{\{5\}}f_A &= 0.05x_5 \end{aligned}$$

Since all the terms are odd in all directions, the spherical ANOVA decomposition from Definition 3.9 directly gives that

$$\begin{aligned} f_{\emptyset, \mathbf{0}} &= 0, \\ f_{\{1,2\}, \mathbf{e}_1 + \mathbf{e}_2} &= x_1x_2^3 \\ f_{\{3,4\}, \mathbf{e}_3 + \mathbf{e}_4} &= 2x_3x_4^5 \\ f_{\{5\}, \mathbf{e}_5} &= 0.05x_5, \end{aligned}$$

and all other terms are zero.

- $f_B(\mathbf{x}) = x_2x_1^2$

The parity decomposition (2.9) of this function is

$$\Xi_{\xi}f_B = \begin{cases} x_2x_1^2 & \text{if } \xi = \mathbf{e}_2 \\ 0 & \text{otherwise .} \end{cases}$$

Applying the operators $P_{\mathbf{u}}$ with $2 \in \mathbf{u}$ to this function yields that

$$\begin{aligned} P_2f_B &= \frac{1}{d}x_2(1 - x_2^2) \\ P_{\{1,2\}}f_B &= x_2x_1^2 \\ P_{\{2\} \cup \mathbf{a}}f_B &= \frac{1}{\omega_{d-1-|\mathbf{a}|}}x_2 \int_{\mathbb{S}^{d-1-|\mathbf{a}|-1}} (1 - \|\mathbf{x}_{\mathbf{a}}\|^2 - x_2^2)x_1^2 d\mu_{d-|\mathbf{a}|-1}(\mathbf{x}_{\mathbf{a}^c \setminus 2}) \\ &= \frac{x_2}{d - |\mathbf{a}|}(1 - \|\mathbf{x}_{\mathbf{a}}\|^2 - x_2^2) & \text{if } \mathbf{a} \subset [d+1] \setminus \{1, 2\} \\ P_{\{1,2\} \cup \mathbf{a}}f_B &= f_B & \text{if } \mathbf{a} \subset [d+1] \setminus \{1, 2\} \end{aligned}$$

Here we see that the spherical ANOVA decomposition from Definition 3.9 blurs the original function (which depends only on the variables x_1 and x_2 also to the other variables due to the geometric properties on the sphere. But with the cancellations from Section (3.6) we describe the spherical ANOVA decomposition by

$$\begin{aligned} f_{\emptyset, \mathbf{0}} &= 0, \\ f_{2, \mathbf{e}_2} &= \frac{1}{d}x_2 - \frac{1}{d}x_2^3 \\ f_{\{1,2\}, \mathbf{e}_2} &= x_2x_1^2 - \frac{1}{d}x_2 + \frac{1}{d}x_2^3. \end{aligned}$$

All other terms are zero.

- $f_C(\mathbf{x}) = x_1^4 + x_2^2$

This is a pure even function, such that the parity decomposition of this function is

$$\Xi_{\xi}f_C = \begin{cases} x_1^4 + x_2^2 & \text{if } \xi = \mathbf{0} \\ 0 & \text{otherwise .} \end{cases}$$

Applying the operators $P_{\mathbf{u}}$ to this function yields due to the Beta function,

$$\begin{aligned}
P_{\emptyset} x_1^4 &= \frac{1}{\omega_d} \int_{\mathbb{S}^d} f_C \, d\mu_d(\mathbf{x}) = \frac{\omega_{d-1}}{\omega_d} \int_{-1}^1 x_1^4 (1 - x_1^2)^{\frac{d-2}{2}} \, dx_1 \\
&= \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi} \Gamma(\frac{d}{2})} 2 \int_0^1 t^2 (1-t)^{\frac{d-2}{2}} \frac{1}{2\sqrt{t}} \, dt = \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi} \Gamma(\frac{d}{2})} \int_0^1 t^{3/2} (1-t)^{\frac{d-2}{2}} \, dt \\
&= \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi} \Gamma(\frac{d}{2})} \frac{\Gamma(\frac{5}{2}) \Gamma(\frac{d}{2})}{\Gamma(\frac{d+5}{2})} = \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi}} \frac{3\sqrt{\pi}}{4 \Gamma(\frac{d+5}{2})} = \frac{3 \Gamma(\frac{d+1}{2})}{4 \Gamma(\frac{d+5}{2})} = \frac{3}{4 \frac{d+3}{2} \frac{d+1}{2}} = \frac{3}{(d+3)(d+1)}, \\
P_{\emptyset} x_2^2 &= \frac{\omega_{d-1}}{\omega_d} \int_{-1}^1 x_2^2 (1 - x_2^2)^{\frac{d-2}{2}} \, dx_2 \\
&= \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi} \Gamma(\frac{d}{2})} 2 \int_0^1 t (1-t)^{\frac{d-2}{2}} \frac{1}{2\sqrt{t}} \, dt = \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi} \Gamma(\frac{d}{2})} \int_0^1 t^{1/2} (1-t)^{\frac{d-2}{2}} \, dt \\
&= \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi} \Gamma(\frac{d}{2})} \frac{\Gamma(\frac{3}{2}) \Gamma(\frac{d}{2})}{\Gamma(\frac{d+3}{2})} = \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi}} \frac{\frac{1}{2} \sqrt{\pi}}{\Gamma(\frac{d+3}{2})} = \frac{\Gamma(\frac{d+1}{2})}{2 \Gamma(\frac{d+3}{2})} = \frac{1}{d+1}, \\
P_{\{1\}} f_C &= x_1^4 + \frac{1}{d} (1 - x_1^2), \\
P_{\{i\}} f_C &= \frac{3}{d(d+2)} (1 - x_2^2)^2, \quad i \in \{2, 3, \dots, d+1\}, \\
P_{\{2\}} f_C &= \frac{3}{d(d+2)} (1 - x_2^2)^2 + x_2^2, \\
P_{\{i\}} f_C &= \frac{3}{d(d+2)} (1 - x_i^2)^2 + \frac{1}{d} (1 - x_i^2) \quad i \in \{3, 4, 5, \dots, d+1\} \\
P_{\{1,2\}} f_C &= x_1^4 + x_2^2.
\end{aligned}$$

The spherical ANOVA decomposition of this function, with regard to the cancellations in Section 3.6 then is

$$\begin{aligned}
f_{\emptyset, \mathbf{0}} &= \frac{d+9}{(d+3)(d+1)}, \\
f_{1, \mathbf{e}_1} &= 0, & f_{1, \mathbf{0}} &= x_1^4 - \frac{3}{(d+3)(d+1)}, \\
f_{2, \mathbf{e}_2} &= 0, & f_{2, \mathbf{0}} &= x_2^2 - \frac{1}{d+1},
\end{aligned}$$

and all other terms are zero.