Master's Thesis

Multi-Direction Analysis in Ambisonics

Thomas Deppisch
01431769

Supervisor: Ass.Prof. Dr.rer.nat. Franz Zotter

Master's Degree Programme
Electrical Engineering and Audio Engineering (F 066 413)
University of Music and Performing Arts Graz
and Graz University of Technology

Graz, November 4, 2020

Abstract

Ambisonics is known to provide a generalised framework for recording, analysis, and playback of spatial audio signals. While recording and playback in high spatial resolution has, due to the availability of multichannel microphone and loudspeaker arrays, already been adopted by the industry, the development of signal processing algorithms for the analysis of such recordings is still a matter of ongoing research. Especially subspace-based algorithms recently proved to effectively detect the directions of multiple simultaneous sounds. Consequently, multiple direction-of-arrival detection in this thesis considers such subspace-based algorithms as a basis to introduce two new algorithms: (i) a real-valued variant of the vector-based eigenbeam ESPRIT and (ii) a variant of the spherical harmonics MUSIC algorithm that uses Newton's method and recurrence relations of spherical harmonics to avoid an expensive grid search. Additionally, two specific applications of the algorithms are shown. The first application is an extension of the Ambisonic spatial decomposition method (ASDM) to higher Ambisonic orders and enables an enhancement of the directional resolution of Ambisonic room impulse responses by higher-order re-encoding of sound instances (acoustic reflections or diffraction) from multiple directions per times instant. The second application considers the detection of peaks and the two tangential curvatures per peak to analyse and resynthesise directivity patterns of sound sources.

Kurzfassung

Ambisonics bietet ein generelles Rahmenkonzept zur Aufnahme, Analyse und Wiedergabe räumlicher Audiosignale. Während die Aufnahme und Wiedergabe in hoher räumlicher Auflösung durch entsprechende, mittlerweile etablierte, mehrkanalige Mikrofon- und Lautsprechersysteme den Sprung von der Forschung in die Industrie bereits geschafft hat, ist die Entwicklung von Signalverarbeitungsalgorithmen zur Analyse der entsprechenden Aufnahmen noch lebendiger Gegenstand der aktuellen Forschung. Hier haben sich zuletzt besonders signalunterraumbasierte Algorithmen aufgrund ihrer hohen Genauigkeit als vielversprechend erwiesen. Diese Arbeit beschäftigt sich mit solchen unterraumbasierten Algorithmen zur gleichzeitigen Schätzung mehrerer Quellrichtungen. Es werden zwei neue Algorithmen vorgestellt, (i) eine reellwertige Variante des vektorbasierten Eigenbeam ESPRIT und (ii) eine Variante des kugelflächenfunktionsbasierten MUSIC Algorithmus, die Newtons Methode und Rekurrenzbeziehungen der Kugelflächenfunktionen benutzt, um eine aufwändige Rastersuche zu vermeiden. Anschließend werden zwei spezifische Anwendungen der Algorithmen aufgezeigt. Die erste Anwendung ist eine Erweiterung der ambisonischen räumlichen Zerlegungsmethode (engl. *Spatial Decomposition Method*) auf höhere ambisonische Ordnungen und erlaubt eine Erhöhung der Richtungsauflösung ambisonischer Raumimpulsantworten basierend auf der gleichzeitigen Richtungsschätzung mehrerer Schallereignisse (Reflektionen oder Diffraktion). Die zweite Anwendung beinhaltet die Analyse und Resynthese von Richtwirkungsmustern von Schallquellen basierend auf den beiden tangentialen Krümmungen an den Maxima.

Contents

1 Introduction **5**

2 Subspace-Based Algorithms for Multiple DOA Estimation **11**

2.1 Decomposition into Signal and Noise Subspace 11

2.2 MUSIC 14

2.3 ESPRIT 16

2.4 Real-Valued Extended Vector-Based EB-ESPRIT (REVEB-ESPRIT) . . . 18

2.4.1 Extended Vector-Based EB-ESPRIT (EVEB-ESPRIT) 18

2.4.2 Multiplication Recurrences for Real-Valued SHs 20

2.4.3 Joint Triangularisation using Rayleigh Quotient Iterations 22

2.4.4 Algorithm Summary and Complexity 25

2.4.5 Evaluation 25

2.5 New MUSIC 29

2.5.1 Gradient Addition Recurrences 29

2.5.2 Newton's Method on the Surface of the Sphere 30

2.5.3 Summary of the New MUSIC Algorithm for DOA estimation . . 35

2.5.4 Maximum Number of Extrema in SH Patterns 35

2.6 Performance Comparison of REVEB-ESPRIT and New MUSIC 39

2.6.1 Evaluation Setup 39

2.6.2 Results 43

3 Applications **49**

3.1 Higher-Order Ambisonic Spatial Decomposition Method (HO-ASDM) . . 49

3.1.1 Algorithm 50

3.1.2 Listening Experiment 52

3.2 Analysis and Synthesis of Directivity Patterns in Peak Value and Tangential Curvature 55
3.2.1 Analysis of Directivity Patterns 55
3.2.2 Synthesis of Hessian-Constrained Directivity Patterns 57
3.2.3 Exemplary Resynthesis of a Random Directivity Pattern 61

4 Conclusion and Outlook **65**

A Derivation of SH Recurrence Relations **67**
A.1 Derivation of Gradient Addition Theorems for Real-Valued SHs 67
A.2 Derivation of Multiplication Theorems for Real-Valued SHs 69
A.3 Derivation of Extending Multiplication Theorems for Real-Valued SHs . . 71
A.4 Code Listings for the Derived Recurrence Relations 73

B Extended Pseudo-Intensity Vector from Real Recurrences **79**

C Why eigenvalue n-tuples of joint eigendecomposition are displayed as consistent diagonal-entry tuples in joint Schur decomposition **81**

D Additions to the Listening Experiment **83**
D.1 Compared ASDM Equalisation Methods 83
D.2 Extensive Results 84

Nomenclature

$(\cdot)^{\mathrm{H}}$	Conjugate transpose.
$(\cdot)^{\mathrm{T}}$	Transpose.
$(\cdot)^{*}$	Complex conjugate.
$(\cdot)^{\dagger}$	Moore-Penrose pseudoinverse.
$\boldsymbol{\Theta}$	Diagonal matrix containing direction-vector components.
$\boldsymbol{\theta}$	Direction vector.
$\boldsymbol{E}_{\mathrm{n}}$	Matrix containing orthogonal noise eigenvectors.
$\boldsymbol{E}_{\mathrm{s}}$	Matrix containing orthogonal signal eigenvectors.
$\boldsymbol{g}$	Gradient vector.
$\boldsymbol{G}_{\{x,y,z\}}$	Matrices containing coefficients of real-valued gradient addition recurrences.
$\boldsymbol{H}$	Hessian.
$\boldsymbol{T}$	Transformation matrix relating array steering matrix and signal subspace.
$\boldsymbol{U}_{\mathrm{n}}$	Matrix containing orthogonal spherical-harmonics-domain noise eigenvectors.
$\boldsymbol{U}_{\mathrm{s}}$	Matrix containing orthogonal spherical-harmonics-domain signal eigenvectors.
$\boldsymbol{y}_N$	Spherical harmonics vector holding coefficients up to order N.
δ	Kronecker delta.
$\mathcal{O}$	Number of observations.
φ	Azimuth angle.
ϑ	Zenith angle.
i	Imaginary unit $i = \sqrt{-1}$.
J, j	Number of trials and trial index.
L, l	Number of array sensors (microphones) and sensor index.
N, n	Maximum spherical harmonics order and order index.
Q, q	Number of sources and source index.
Y_n^m	Complex-valued spherical harmonic of order n and degree m.
$Y_{n,m}$	Real-valued spherical harmonic of order n and degree m.

Chapter 1

Introduction

This thesis addresses subspace-based algorithms for direction-of-arrival (DOA) estimation of sound sources and applications built upon. Specifically, it covers algorithms working in the spherical harmonics domain (SHD), i.e. the input signals are expected to be expressed via spherical harmonics (SHs). Expansion over SHs is used throughout many fields, e.g. for representation of electromagnetic fields, analysis of the cosmic microwave background, description of atomic orbitals and efficient simulation of light sources in computer graphics. In acoustics, SHs are of particular interest as they are solutions to the wave equation in spherical coordinates. In context of surround sound loudspeaker reproduction systems, SHs were first used in the 1970s by Michael Gerzon [Ger73], building upon work by Cooper and Shiga [CS71]. Soon, a corresponding recording technique was developed [Ger75], the Sound Field microphone was built [CG77][Far79] and the name Ambisonics was established [Ger77][Ger85]. With the first higher-order coincident microphone array, the Eigenmike em32 [ME02], the achievable spatial resolution was considerably increased and a need for algorithms processing signals from spherical microphone arrays arose. Nowadays, a multitude of higher-order spherical microphone arrays is available and the development of signal processing algorithms to retrieve information from such recordings is a highly-researched topic.

Spherical Harmonics. The complex-valued spherical harmonics [Wil99] are composed of a normalisation term, the associated Legendre function P_n^m and a complex exponential,

$$Y_n^m(\varphi, \vartheta) = \sqrt{\frac{2n+1}{4\pi}\frac{(n-m)!}{(n+m)!}}\, P_n^m(\cos\vartheta)\, e^{im\varphi}\,, \tag{1.1}$$

where n is the order of the spherical harmonic, m is the degree [1], φ and ϑ are azimuth and zenith angle, and $i = \sqrt{-1}$ is the imaginary unit. In many cases it is sufficient and has a

1. Note that in many subjects other than acoustics, the denotation of order and degree would be switched.

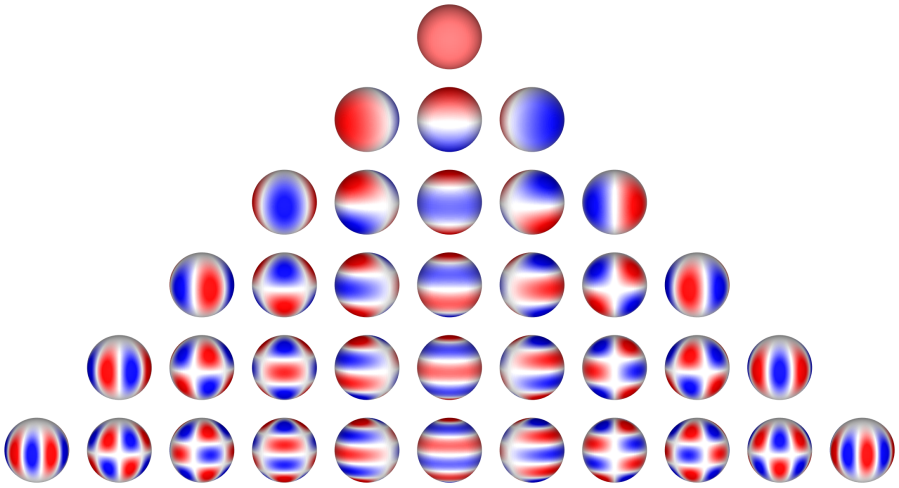


Figure 1.1 – Real-valued spherical harmonics $Y_{n,m}(\varphi, \vartheta)$ from order $n = 0$ (top) to $n = 5$ (bottom) and for degree $m = -n$ (left) to $m = n$ (right), evaluated on the surface of the unit sphere. Red colour depicts positive values, blue colour negative values and colour intensity depicts magnitude.

computational advantage to use real-valued spherical harmonics,

$$Y_{n,m}(\varphi, \vartheta) = \sqrt{\frac{(2n+1)(2-\delta_m)}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} \, P_n^{|m|}(\cos\vartheta) \begin{cases} \cos(m\varphi) & \text{for } m \geq 0 \\ \sin(|m|\varphi) & \text{for } m < 0 \end{cases}, \tag{1.2}$$

where

$$\delta_m = \begin{cases} 1, & \text{for } m = 0 \\ 0, & \text{else} \end{cases}, \tag{1.3}$$

is the Kronecker delta. Figure 1.1 shows real-valued spherical harmonics up to fifth order. The spherical harmonics form a complete set of orthonormal functions on the unit sphere $\mathbb{S}^2$, such that any square-integrable function on the unit sphere $f(\varphi, \vartheta)$ can be expanded over the spherical harmonics $Y_n^m(\varphi, \vartheta)$ [Wil99],

$$f_{nm} = \int_{\Omega \in \mathbb{S}^2} f(\varphi, \vartheta) Y_n^{m^*}(\varphi, \vartheta) \, d\Omega := \mathcal{SHT}\{f(\varphi, \vartheta)\}, \tag{1.4}$$

denoted as spherical harmonics transform (SHT) or spherical Fourier transform. The inverse operation, the inverse spherical harmonics transform (ISHT), is defined as

$$f(\varphi, \vartheta) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{nm} Y_n^m(\varphi, \vartheta) := \mathcal{ISHT}\{f_{nm}\}. \tag{1.5}$$

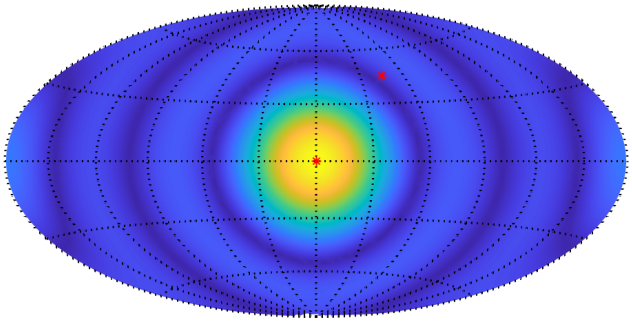
It is often helpful to describe SHD operations in vector notation. Hence, a $(N+1)^2 \times 1$ vector $\boldsymbol{y}_N$ is defined to contain the SHs sorted by their increasing Ambisonic Channel Number (ACN) [CMZ+09],

$$\boldsymbol{y}_N = [Y_0^0, Y_1^{-1}, Y_1^0, Y_1^1, \ldots, Y_N^{-N}, \ldots, Y_N^0, \ldots, Y_N^N]^{\mathrm{T}}, \tag{1.6}$$

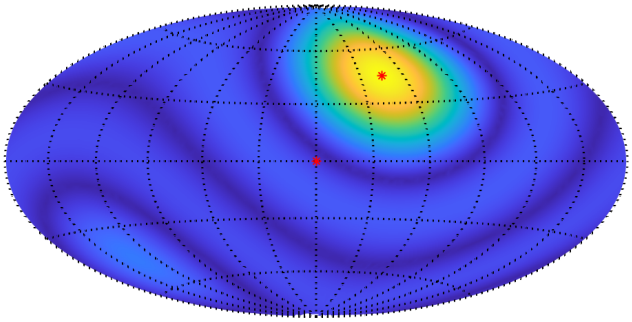
up to the maximum order N.

Motivation for subspace-based methods. Consider a two-source direction-of-arrival (DOA) estimation scenario, where two source signals interfere as shown in figure 1.2. The magnitude information of the steered response power in fig. 1.2(c) typically indicates the interference of the signals rather than their original source directions. However, if several independent observations of the interfered signal are available, signal and noise subspace can be obtained using the method shown in section 2.1. In the signal and noise subspace magnitudes, depicted in figure 1.3, the two interference-free DOAs appear as peaks and roots, respectively. These can be retrieved using the methods described in chapter 2.

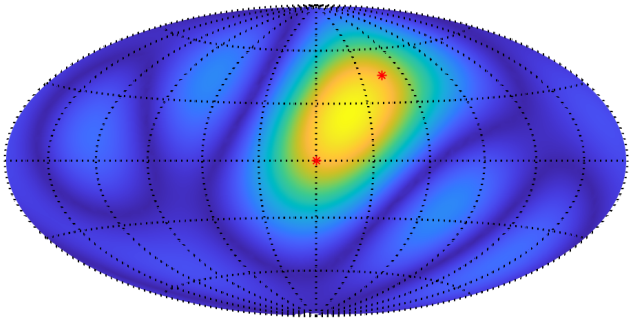
Structure of this thesis. Chapter 2 begins by defining the decomposition of a signal into a signal and a noise subspace in section 2.1. Then, MUSIC and ESPRIT as the two most common subspace methods for multiple DOA estimation in sensor array problems are summarised and their application in the SHD is shown in sections 2.2 and 2.3. Sections 2.4 and 2.5 develop and propose two new algorithms, namely a real-valued version of the extended vector-based EB-ESPRIT with an iterative joint-triangularisation method, and the New MUSIC algorithm, a variant of SH-MUSIC, that avoids the computationally-intensive grid search by using Newton's method and SH gradient recurrences. The performance of both new algorithms is evaluated and compared in section 2.6. Chapter 3 considers two applications of the proposed algorithms. Section 3.1 utilises the real-valued vector-based EB-ESPRIT for estimation of sound-event directions (reflections and diffractions) in measured Ambisonic room impulse responses (ARIRs) and subsequent upmixing to an arbitrary higher Ambisonics order. The algorithm is evaluated via a listening experiment regarding created artefacts and preservation of timbre. Section 3.2 shows a new way of describing and resynthesising measured directivity patterns by evaluation of the Hessian at the directions of the maxima. Both, the locations of the maxima and their corresponding Hessians are found using the New MUSIC algorithm. In chapter 4 the outcomes of this thesis are concluded and an outlook for possible further work is given.



(a) Source 1 at direction $\varphi = 0$, $\vartheta = \pi/2$.

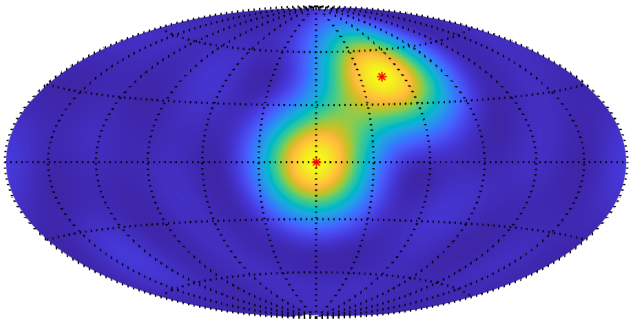


(b) Source 2 at direction $\varphi = \pi/4$, $\vartheta = \pi/4$.

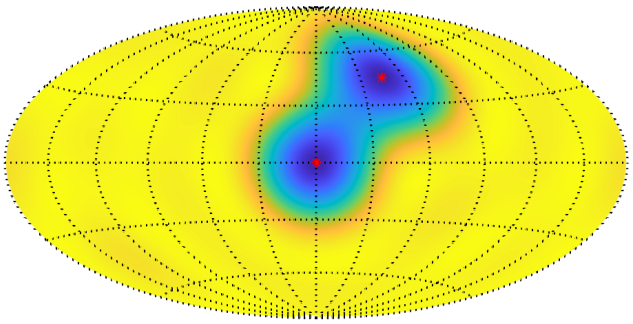


(c) Interference of the two source signals.

Figure 1.2 – Directional magnitude of two signals and the corresponding interference pattern observed in the steered response power and projected onto a world map. Both original source directions are marked by red asterisks. The peak(s) of the steered response power strongly depends on the correlation of both signals and often deviates from the original directions.



(a) Signal subspace.



(b) Noise subspace.

Figure 1.3 – The directional magnitude of signal and noise subspace consistently displays the original source directions (marked by red asterisks), despite the specific signal interference sampled by the independent observations.

Chapter 2

Subspace-Based Algorithms for Multiple DOA Estimation

2.1 Decomposition into Signal and Noise Subspace

Decomposition into a signal and a noise subspace, that either contains all signals or (diffuse) noise, is a helpful way to reduce the dimensionality of a data set without losing the desired information of the measurement. An introduction to the topic is given in [RK89] and [Teu07, chapter 4]. When applied to sensor array problems, a good estimation of signal parameters is achieved by finding intersections (or closest distances) between the observed signal subspace and the array manifold. The array manifold contains the array response (also called array steering vectors) for every direction of a noise-free signal. Array steering vectors intersecting with the observed signal subspace indicate a signal direction, and equivalently steering vectors orthogonal to the observed noise subspace do so. Alternatively, array steering vectors orthogonal to the observed noise subspace allow for similar deductions. When a particular set of array steering vectors is found, the signal parameters are extracted from a known unique mapping.

For estimation of the signal subspace, the data model

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{\nu}(t) \tag{2.1}$$

is introduced, where $\boldsymbol{x}(t)$ is the measurement vector, the columns of $\boldsymbol{A}$ contain the array steering vectors for signals of every direction $\boldsymbol{\theta}$, $\boldsymbol{s}(t)$ are the sensor signals, and $\boldsymbol{\nu}(t)$ is additive white Gaussian noise (AWGN) with variance σ_ν^2. By observing that the columns of $\boldsymbol{A}$ span the signal subspace S_X, span$\{\boldsymbol{A}\} = S_X$, an estimation of the Q-dimensional signal subspace S_X is found via the Q orthogonal eigenvectors $\boldsymbol{e}_i$, corresponding to the Q largest eigenvalues of the normal, positive-definite measurement covariance matrix [RK89],

$$\boldsymbol{R}_{XX} = \mathcal{E}\{\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\} = \boldsymbol{A}\boldsymbol{R}_{SS}\boldsymbol{A}^{\mathrm{H}} + \sigma_\nu^2\boldsymbol{I}\,, \tag{2.2}$$

such that $S_X = \text{span}\{\boldsymbol{E}_s\} = \text{span}\{[\boldsymbol{e}_1, \ldots, \boldsymbol{e}_Q]\}$. Similarly, the eigenvectors correspond-

ing to the remaining eigenvalues of lower magnitude span the noise subspace, $S_N = \text{span}\{\boldsymbol{E}_n\} = \text{span}\{[\boldsymbol{e}_{Q+1}, \ldots, \boldsymbol{e}_{\mathcal{O}}]\}$. In practical scenarios, $\boldsymbol{R}_{XX}$ cannot be obtained directly but is estimated via the sample covariance

$$\hat{\boldsymbol{R}}_{XX} = \frac{1}{\mathcal{O}} \boldsymbol{X}\boldsymbol{X}^{\mathrm{H}}\,, \tag{2.3}$$

where $\boldsymbol{X}$ is the data matrix that collects $\mathcal{O}$ observations of $\boldsymbol{x}(t)$ in its columns. Alternatively, $\hat{\boldsymbol{R}}_{XX}$ can be estimated directly from the data matrix $\boldsymbol{X}$ via singular value decomposition (SVD) $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{H}}$,

$$\hat{\boldsymbol{R}}_{XX} = \frac{1}{\mathcal{O}} \boldsymbol{U}\boldsymbol{\Sigma}^{\mathrm{H}}\boldsymbol{\Sigma}\boldsymbol{U}^{\mathrm{H}} = \boldsymbol{E}\boldsymbol{\Lambda}\boldsymbol{E}^{\mathrm{H}}\,, \tag{2.4}$$

where the left singular vectors $\boldsymbol{U}$ are equal to the (orthogonal) eigenvectors $\boldsymbol{E}$ of $\hat{\boldsymbol{R}}_{XX}$, avoiding numerical problems arising from squaring $\boldsymbol{X}$ [RK89]. If an estimation of the eigenvalues of the (sample) covariance matrix is needed (as below for the SORTE algorithm), the eigenvalues can hence be calculated from the singular values as

$$\boldsymbol{\Lambda} = \frac{1}{\mathcal{O}} \boldsymbol{\Sigma}^{\mathrm{H}}\boldsymbol{\Sigma}\,. \tag{2.5}$$

Estimation of the signal subspace dimension. Signal and noise eigenvectors are distinguished by defining a threshold dividing the magnitude-sorted eigenvalues of the covariance matrix into eigenvalues corresponding to the signal eigenvectors and eigenvalues corresponding to the noise eigenvectors. More precisely, if $\mathcal{T}$ eigenvalues are obtained, the eigenvectors corresponding to the Q largest eigenvalues are assigned to the signal subspace and collected in $\boldsymbol{E}_{\mathrm{s}}$, and the remaining $\mathcal{T} - Q$ eigenvectors are spanning the noise subspace and are collected in $\boldsymbol{E}_{\mathrm{n}}$. Following the data model in eq. 2.1, the eigenvalues of the covariance $\boldsymbol{R}_{XX}$ corresponding to the noise eigenvectors would be equal to the noise variance σ_ν^2 in an ideal measurement system. Hence an algorithm considering the index in the magnitude-sorted eigenvalues where the variance of the eigenvalues stays approximately constant can effectively define such a threshold. SORTE [HCXC10] compares the variances of subsequences of eigenvalue differences,

$$\text{var}\left(\{\nabla\lambda_i\}_{i=k_1}^{k_2}\right) = \frac{1}{k_2 - k_1 + 1} \sum_{i=k_1}^{k_2} \left(\nabla\lambda_i - \frac{1}{k_2 - k_1 + 1} \sum_{j=k_1}^{k_2} \nabla\lambda_j \right)^2\,, \tag{2.6}$$

with $\nabla\lambda_i = \lambda_i - \lambda_{i+1}$. More precisely, SORTE analyses the ratio of variances of the adjacent subsequences of eigenvalue differences starting at index k and $k+1$,

$$o(k) = \begin{cases} \frac{\text{var}\left(\{\nabla\lambda_i\}_{i=k+1}^{\mathcal{T}-1}\right)}{\text{var}\left(\{\nabla\lambda_i\}_{i=k}^{\mathcal{T}-1}\right)}\,, & \text{for var}\left(\{\nabla\lambda_i\}_{i=k+1}^{\mathcal{T}-1}\right) \neq 0 \\ \infty\,, & \text{for var}\left(\{\nabla\lambda_i\}_{i=k+1}^{\mathcal{T}-1}\right) = 0 \end{cases}\,, \tag{2.7}$$

where $k \in [1, \mathcal{T} - 2]$. The number of signal subspace eigenvectors is then found as the index at which the ratio of the variances of subsequence eigenvalue differences is the smallest,

$$Q = \arg\min_k o(k)\,. \tag{2.8}$$

In non-ideal measurement systems, $\text{var}\left(\{\nabla\lambda_i\}_{i=k+1}^{\mathcal{T}-1}\right)$ will generally be larger than 0. Hence the authors of [HCXC10] propose to only consider eigenvalues containing a large percentage of the cumulative energy,

$$o(k) = \begin{cases} \frac{\text{var}\left(\{\nabla\lambda_i\}_{i=k+1}^{\mathcal{T}-1}\right)}{\text{var}\left(\{\nabla\lambda_i\}_{i=k}^{\mathcal{T}-1}\right)}\,, & \text{for var}\left(\{\nabla\lambda_i\}_{i=k+1}^{\mathcal{T}-1}\right) \neq 0 \\ \infty\,, & \text{for var}\left(\{\nabla\lambda_i\}_{i=k+1}^{\mathcal{T}-1}\right) = 0 \\ \infty\,, & \text{for } k > \mathcal{J} \end{cases}, \tag{2.9}$$

where

$$\mathcal{J} = \arg\min_j \left(\sum_{i=1}^{j} \lambda_i \,/\, \sum_{i=1}^{\mathcal{T}} \lambda_j > \epsilon_1 \right), \tag{2.10}$$

for e.g. $\epsilon_1 = 0.99$.

As the last entry $o(\mathcal{T}-2)$ is zero by design (due to $\text{var}\left(\{\nabla\lambda_i\}_{i=\mathcal{T}-1}^{\mathcal{T}-1}\right) = 0$), SORTE produces a maximum number of $\mathcal{T}-3$ non-zero entries. Hence, a different criterion has to be used if the number of available eigenvalues is small. In SHD problems, such special cases occur for first-order Ambisonics material, where only $\mathcal{T}=4$ eigenvalues are available. Hence, in this work only the variances of subsequences of eigenvalue differences are analysed in such cases,

$$\hat{o}(k) = \begin{cases} \infty\,, & \text{for var}\left(\{\nabla\lambda_i\}_{i=k}^{\mathcal{T}-1}\right) / \sum_{j=1}^{\mathcal{T}-1} \text{var}\left(\{\nabla\lambda_i\}_{i=j}^{\mathcal{T}-1}\right) < \epsilon_2 \\ \text{var}\left(\{\nabla\lambda_i\}_{i=k}^{\mathcal{T}-1}\right)\,, & \text{else} \end{cases}, \tag{2.11}$$

i.e. if the variance of a subsequence if smaller than a percentage of the total variance, e.g. $\epsilon_2 = 0.1$, such entries are neglected. When assuming $\mathcal{T}=4$ eigenvalues as in first-order SHD problems, $\mathcal{T}-1=3$ eigenvalue differences $\nabla\lambda_i$ and $\mathcal{T}-2=2$ corresponding variances are obtained, which is the maximum number of DOAs extended EB-ESPRIT algorithms (cf. section 2.4.1) can estimate. If the variation of this difference from the second eigenvalue onward is smaller than 10% of the total variation (starting from the first eigenvalue and second eigenvalue, respectively), variations starting from the second eigenvalue are considered so small, that the corresponding subspace components are assigned to the noise subspace, resulting in a number of $Q=1$ signal subspace components.

2.2 MUSIC

The Multiple Signal Classification (MUSIC) algorithm was proposed by Schmidt [Sch86] and independently by Bienvenue [BK79] for multiple-emitter sensor array problems with arbitrary arrays and applications including DOA estimation and multiple frequency estimation. MUSIC assumes a superposition of Q plane-waves impinging on the L array sensors. This is described by the data model

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{\nu}(t)\,, \tag{2.12}$$

where $\boldsymbol{x}(t)$ are the individual sensor signals $\boldsymbol{x}(t) = [x_1(t), \ldots, x_L(t)]$, $\boldsymbol{A} = [\boldsymbol{a}(\boldsymbol{\theta}_1), \ldots, \boldsymbol{a}(\boldsymbol{\theta}_Q)]$ contains the array steering vectors for all source directions $\boldsymbol{\theta}_q$, $\boldsymbol{s}(t) = [s_1(t), \ldots, s_Q(t)]$ are the corresponding plane-wave signals and $\boldsymbol{\nu}(t)$ denotes additive noise. The steering vectors describe the wave propagation from source to sensor using the corresponding Green's function (including possible direction-dependent sensitivities)

$$\boldsymbol{a}(\boldsymbol{\theta}_q) = [a_1(\boldsymbol{\theta}_q)e^{-j\omega_0\tau_1(\boldsymbol{\theta}_q)}, \ldots, a_L(\boldsymbol{\theta}_q)e^{-j\omega_0\tau_L(\boldsymbol{\theta}_q)}]^{\mathrm{T}}\,. \tag{2.13}$$

Generally, MUSIC involves a decomposition into signal and noise subspace, e.g. via eigendecomposition of the covariance matrix, followed by a grid search of roots in the steered, squared noise subspace, the MUSIC cost function, or by a search of peaks in the inverse function, the MUSIC pseudo-spectrum,

$$P_{\mathrm{MUSIC}}(\boldsymbol{\theta}) = \frac{1}{\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{\theta})\boldsymbol{E}_{\mathrm{n}}\boldsymbol{E}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{a}(\boldsymbol{\theta})}\,. \tag{2.14}$$

Here, the matrix $\boldsymbol{E}_{\mathrm{n}}$ holds the orthogonal noise eigenvectors and $\boldsymbol{a}(\boldsymbol{\theta})$ is the array steering vector for an arbitrary search-grid direction $\boldsymbol{\theta}$. As $\boldsymbol{a}^{\mathrm{H}}\boldsymbol{E}_{\mathrm{n}}\boldsymbol{E}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{a}^{\mathrm{H}}(\boldsymbol{I} - \boldsymbol{E}_{\mathrm{s}}\boldsymbol{E}_{\mathrm{s}}^{\mathrm{H}})\boldsymbol{a} = \boldsymbol{a}^{\mathrm{H}}\boldsymbol{a} - \boldsymbol{a}^{\mathrm{H}}\boldsymbol{E}_{\mathrm{s}}\boldsymbol{E}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}$, the MUSIC cost function exhibits roots at directions in the signal subspace independent of the norm of the steering vectors.

Schmidt also provides a geometrical interpretation of the algorithm: When disregarding noise for a moment, the array output $\boldsymbol{x}(t)$ lies in the Q-dimensional signal subspace (e.g. one-dimensional for one signal) of the L-dimensional column space span$\{\boldsymbol{A}\}$. The steering vectors $\boldsymbol{a}(\boldsymbol{\theta})$ for all possible directions $\boldsymbol{\theta}$, which can be obtained by physical measurements or, in some cases, by analytical expressions, stem from a set which was later denoted as the array manifold [RK89]. After obtaining Q independent observations $\boldsymbol{x}(t)$, the signal subspace is known and the searched parameters can be determined by a mapping between signal subspace and parameters. When considering noise, MUSIC searches for the solutions of the observed signal subspace which are closest to the array manifold, or equivalently, by searching directions in the array manifold which are (nearly) orthogonal to the noise subspace.

For spherical sensor arrays, MUSIC is reformulated in the SH domain, resulting in SH-MUSIC [LYMH11][KR12]. DOAs are then found as the peaks of the SHD pseudo-spectrum $P_{\mathrm{SH-MUSIC}}$, which evaluates (the inverse of) the squared amplitude of the SHD noise subspace $\boldsymbol{U}_{\mathrm{n}}$ at the direction $\boldsymbol{\theta}$,

$$P_{\mathrm{SH-MUSIC}}(\boldsymbol{\theta}) = \frac{1}{\boldsymbol{y}_N^{\mathrm{H}}(\boldsymbol{\theta})\boldsymbol{U}_{\mathrm{n}}\boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{y}_N(\boldsymbol{\theta})}\,, \tag{2.15}$$

where $\boldsymbol{y}_N$ is the steering vector containing SHs up to order N. For uniform, linear arrays (or generally, in cases where the array manifold can be expressed as Vandermonde matrix), the MUSIC cost function $\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{\theta})\boldsymbol{E}_{\mathrm{n}}\boldsymbol{E}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{a}(\boldsymbol{\theta})$ can be reformulated as polynomial such that the roots of the polynomial are the angles of arrival (AOAs) [Bar83]. This enables super-resolution AOA estimation without discretisation of the MUSIC cost function and is called Root-MUSIC. SH-Root-MUSIC [KBH16] formulates the Root-MUSIC polynomial in the SH-domain for azimuth-only AOA estimation. In [ZDWC18] and [CKR09], manifold separation techniques (MST) are used to create bivariate polynomials which are solved for the DOAs in azimuth and elevation. To reduce the computational complexity needed for bivariate polynomial solving, a one-dimensional MUSIC-type approach (ODMUSIC) for spherical arrays is proposed in [HC20], where, after a one-dimensional spectral search, two independent polynomials for azimuth and elevation are solved iteratively for joint DOA estimation.

Building upon SH-MUSIC, the New MUSIC algorithm is proposed in section 2.5. New MUSIC efficiently minimises the SH-MUSIC cost function $\boldsymbol{y}_N^{\mathrm{H}}(\boldsymbol{\theta})\boldsymbol{U}_{\mathrm{n}}\boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{y}_N(\boldsymbol{\theta})$ using Newton's method. The gradient and Hessian are expressed analytically via SH recurrence relations. New MUSIC achieves super-resolution DOA estimation without discretisation of the MUSIC cost function.

2.3 ESPRIT

Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) was proposed by Roy and Kailath [RK89] and avoids a computationally expensive grid search. What is more, ESPRIT does not require knowledge of the array manifold. ESPRIT employs the same data model as MUSIC (cf. eq. 2.12) but exploits a translational invariance to relate the signal subspaces of subarray responses. This translational invariance is created via a translational displacement, i.e. the array needs to consist of matched sensor pairs, where one element in each pair is displaced by the same displacement vector in the dimension in which parameters are estimated. For a simultaneous estimation of a multitude of parameters (e.g. azimuth and elevation), multiple matched sensors are grouped and displaced by multiple displacement vectors. When assuming two subarrays with output signals $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$, only one searched parameter φ (e.g. azimuthal AOA), and a displacement vector $\boldsymbol{\Delta}$ of magnitude Δ, the data model is extended as

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{\nu}_x(t)\,, \tag{2.16}$$

$$\boldsymbol{y}(t) = \boldsymbol{A}\boldsymbol{\Phi}\boldsymbol{s}(t) + \boldsymbol{\nu}_y(t)\,. \tag{2.17}$$

The array manifold rotation operator $\boldsymbol{\Phi}$ is a diagonal matrix containing phase delays between the sensor pairs for each of the Q signals,

$$\boldsymbol{\Phi} = \mathrm{diag}\left(\left[e^{i\omega_0/c\,\Delta\sin(\varphi_1)}, \ldots, e^{i\omega_0/c\,\Delta\sin(\varphi_Q)}\right]\right)\,, \tag{2.18}$$

rotating the phase of the array responses corresponding to a translation of the entire array and hence explaining the acronym ESPRIT. It can further be shown that the signal subspace of the second subarray is a rotated version of the signal subspace of the first subarray and both subspaces are related to the array steering matrix $\boldsymbol{A}$ by a unique transformation $\boldsymbol{T}$,

$$\boldsymbol{E}_{s_1} = \boldsymbol{A}\boldsymbol{T}\,, \tag{2.19}$$

$$\boldsymbol{E}_{s_2} = \boldsymbol{A}\boldsymbol{\Phi}\boldsymbol{T}\,, \tag{2.20}$$

such that $\mathrm{span}(\boldsymbol{E}_{s_1}) = \mathrm{span}(\boldsymbol{E}_{s_2}) = \mathrm{span}(\boldsymbol{A})$ and the subspace $\boldsymbol{E}_{s_2}$ is related to $\boldsymbol{E}_{s_1}$ by a new operator $\boldsymbol{\Psi}$,

$$\boldsymbol{E}_{s_2} = \boldsymbol{E}_{s_1}\boldsymbol{T}^{-1}\boldsymbol{\Phi}\boldsymbol{T} := \boldsymbol{E}_{s_1}\boldsymbol{\Psi}\,. \tag{2.21}$$

Hence, the array manifold rotation operator $\boldsymbol{\Phi}$ can be calculated from the observable matrix $\boldsymbol{\Psi}$ that relates both subspaces via eigendecomposition defining the eigenvectors contained in $\boldsymbol{T}$, and the phases of the eigenvalues λ_q of $\boldsymbol{\Psi}$ reveal the angles $\varphi_q = \arcsin(\angle(\lambda_q)/(\omega_0/c\,\Delta))$. In real physical systems, where measurement noise cannot be neglected, $\boldsymbol{\Psi}$ needs to be estimated. In [RK89] a least-squares (LS) and a total least-squares (TLS) solution are derived.

This technique was adapted for use with uniform circular arrays by Mathews and Zoltowski [MZ94] and for spherical arrays by Teutsch and Kellermann [TK08]. The deduced algorithms, UCA-ESPRIT and EB-ESPRIT (also called SH-ESPRIT), are expressed in circular or spherical harmonics (also called beamspace, eigenspace or

modal domain). Instead of translational invariances of subarrays, these algorithms use recurrence relations that relate subsets of circular or spherical harmonics to each other after multiplication by the direction parameters. As before, the subspace $\boldsymbol{U}_\mathrm{s}$ and the SH coefficient matrix $\boldsymbol{Y}$ share the same column space, $\mathrm{span}\{\boldsymbol{U}_\mathrm{s}\} = \mathrm{span}\{\boldsymbol{Y}\}$, such that they are related by a transformation $\boldsymbol{U}_\mathrm{s} = \boldsymbol{Y}\boldsymbol{T}$. For EB-ESPRIT, the diagonal matrix $\boldsymbol{\Phi}$ contains the directional parameters ϑ_q and φ_q in form of the trigonometric function $\tan(\vartheta_q)e^{\pm j\varphi_q}$ and hence involves an angular ambiguity regarding $\pm\varphi_q$ and a singularity for $\vartheta_q \to \pm\pi/2$. With EB-ESPRIT $\lfloor N^2/2 \rfloor$ DOAs can be estimated simultaneously. Several beamspace ESPRIT variants were proposed after that, with variations in the specific multiplication recurrence relations employed: Nonsingular ESPRIT [JC18] utilises the trigonometric function $\sin(\vartheta_q)e^{i\varphi_q}$ as directional parameter and therefore solves the singularity problem but exhibits an ambiguity due to the symmetry of the sine function. It enables simultaneous estimation of $\lfloor N^2 + N/2 \rfloor$ DOAs. Two-Step SH-ESPRIT [HZF18] employs two SH recurrence relations and estimates azimuth and zenith angles separately, avoiding the ambiguity issue but requiring an additional matching of the corresponding parameter pairs. With Two-Step SH-ESPRIT, $(N-1)^2$ DOAs can be estimated simultaneously. Vector-Based EB-ESPRIT (VEB-ESPRIT) was developed independently by Jo and Choi [JC19] and by Herzog and Habets [HH19a]. As it uses all three DOA-vector-related direction parameters and the corresponding multiplication recurrences, VEB-ESPRIT necessitates a joint diagonalisation procedure to accomplish simultaneous estimation of up to N^2 DOAs. VEB-ESPRIT avoids singularities and ambiguities and was recently extended by Jo, Zotter and Choi [JZC20] by two additional relations only for the highest order N, enabling simultaneous estimation of up to $N^2 + \lfloor 4/3N \rfloor$ DOAs. Extended Vector-Based EB-ESPRIT (EVEB-ESPRIT) is revised in section 2.4.1 and reformulated using real-valued recurrences in section 2.4.2, in particular because the three DOA parameters in literature are still complex-valued and their multiplication recurrences most often still refer to complex-valued spherical harmonics.

2.4 Real-Valued Extended Vector-Based EB-ESPRIT (REVEB-ESPRIT)

The proposed Real-Valued Extended Vector-Based EB-ESPRIT (REVEB-ESPRIT) uses the same techniques as the complex-valued EVEB-ESPRIT [JZC20], as introduced in section 2.3 above and formally outlined below. REVEB-ESPRIT however employs recurrences re-formulated fully to real-valued direction vectors and real-valued spherical harmonics.

2.4.1 Extended Vector-Based EB-ESPRIT (EVEB-ESPRIT)

This section presents VEB-ESPRIT [JC19, HH19a] and EVEB-ESPRIT as a complex-valued starting point towards REVEB-ESPRIT. VEB-ESPRIT uses three recurrence relations,

$$\theta_{\mathrm{xy}}^{*} Y_n^{m^*} = w_n^{-m} Y_{n-1}^{m+1^*} - w_{n+1}^{m+1} Y_{n+1}^{m+1^*} \,, \tag{2.22}$$

$$\theta_{\mathrm{xy}} Y_n^{m^*} = -w_n^{m} Y_{n-1}^{m-1^*} + w_{n+1}^{-m+1} Y_{n+1}^{m-1^*} \,, \tag{2.23}$$

$$\theta_{\mathrm{z}} Y_n^{m^*} = v_n^{m} Y_{n-1}^{m^*} + v_{n+1}^{m} Y_{n+1}^{m^*} \,, \tag{2.24}$$

relating a complex-valued spherical harmonic coefficient Y_n^m multiplied by the directional parameters $\theta_{\mathrm{xy}}^{*} = \sin\vartheta e^{-i\varphi}$, $\theta_{\mathrm{xy}} = \sin\vartheta e^{i\varphi}$ and $\theta_{\mathrm{z}} = \cos\vartheta$, to a linear combination of shifted coefficients, weighted by recurrence coefficients

$$w_n^m = \sqrt{\frac{(n+m-1)(n+m)}{(2n-1)(2n+1)}} \,, \tag{2.25}$$

$$v_n^m = \sqrt{\frac{(n-m)(n+m)}{(2n-1)(2n+1)}} \,. \tag{2.26}$$

Next, the recurrence coefficients and shift operations are stacked into the matrices $\boldsymbol{D}_{xy^*}$, $\boldsymbol{D}_{xy}$ and $\boldsymbol{D}_{z}$, and the relation $\boldsymbol{U}_{\mathrm{s}} = \boldsymbol{Y}\boldsymbol{T}$ is used, yielding the three independent matrix equations

$$\boldsymbol{M}\boldsymbol{U}_{\mathrm{s}}\boldsymbol{T}^{-1}\boldsymbol{\Theta}_{\{xy^*,xy,z\}}\boldsymbol{T} = \boldsymbol{D}_{\{xy^*,xy,z\}}\boldsymbol{U}_{\mathrm{s}} \,, \tag{2.27}$$

where $\boldsymbol{\Theta}_{\{xy^*,xy,z\}}$ are three diagonal matrices holding the directional parameters $\{\theta_{xy,q}^{*}, \theta_{xy,q}, \theta_{z,q}\}$, and the matrix $\boldsymbol{M} = \begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\end{bmatrix}$ extracts orders $[0, N-1]$ from $\boldsymbol{U}_{\mathrm{s}}$. Hence, without explicit knowledge of $\boldsymbol{T}$, the diagonal matrix with the parameters $\boldsymbol{\Theta}_{\{xy^*,xy,z\}}$ can be estimated via pseudo-inversion of $\boldsymbol{M}\boldsymbol{U}_{\mathrm{s}}$ from the left, followed by a joint diagonalisation,

$$\hat{\boldsymbol{\Theta}}_{\{xy^*,xy,z\}} = \boldsymbol{T}\left(\boldsymbol{M}\boldsymbol{U}_{\mathrm{s}}\right)^{\dagger}\boldsymbol{D}_{\{xy^*,xy,z\}}\boldsymbol{U}_{\mathrm{s}}\boldsymbol{T}^{-1} \,. \tag{2.28}$$

The number of simultaneously detectable DOAs is determined by the rank of $\boldsymbol{M}$ [JZC20], which is $Q_{\max} = \mathrm{rank}\{\boldsymbol{M}\} = N^2$.

In EVEB-ESPRIT [JZC20] the recurrences in eqs. 2.22-2.24 are recombined to eliminate higher-order terms Y_{n+1} for the highest order N, yielding

$$\theta^*_{\mathrm{xy}}\eta_n^{-m}Y_n^{m^*} + \theta_{\mathrm{z}}\eta_n^{m+1}Y_n^{m+1^*} = \eta_{n-1}^{-m}Y_{n-1}^{m+1^*} \qquad \text{for } -n \leq m \leq n-1\,, \tag{2.29}$$

$$\theta_{\mathrm{xy}}\eta_n^{m}Y_n^{m^*} + \theta_{\mathrm{z}}\eta_n^{-m+1}Y_n^{m-1^*} = -\eta_{n-1}^{m}Y_{n-1}^{m-1^*} \qquad \text{for } -n+1 \leq m \leq n\,, \tag{2.30}$$

with the coefficients

$$\eta_n^m = \sqrt{\frac{n+m}{2n+1}}\,. \tag{2.31}$$

After reformulation of these relations in matrix form, with the new coefficients and shift operations in matrices $\boldsymbol{A}$, $\bar{\boldsymbol{A}}$, $\boldsymbol{C}$ and $\bar{\boldsymbol{C}}$, the three equations from 2.27 are not independent anymore,

$$\begin{bmatrix} \boldsymbol{M} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{M} \\ \bar{\boldsymbol{A}} & \boldsymbol{0} & \boldsymbol{A} \\ \boldsymbol{0} & \boldsymbol{A} & -\bar{\boldsymbol{A}} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{\mathrm{s}} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_{\mathrm{s}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{U}_{\mathrm{s}} \end{bmatrix} \begin{bmatrix} \boldsymbol{T}^{-1}\boldsymbol{\Theta}_{xy^*}\boldsymbol{T} \\ \boldsymbol{T}^{-1}\boldsymbol{\Theta}_{xy}\boldsymbol{T} \\ \boldsymbol{T}^{-1}\boldsymbol{\Theta}_{z}\boldsymbol{T} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}_{xy^*} \\ \boldsymbol{D}_{xy} \\ \boldsymbol{D}_{z} \\ \bar{\boldsymbol{C}} \\ -\boldsymbol{C} \end{bmatrix} \boldsymbol{U}_{\mathrm{s}}\,. \tag{2.32}$$

Due to the extension via the matrices $\boldsymbol{A}$ and $\bar{\boldsymbol{A}}$, the rank of the leftmost matrix was extended from $3N^2$ to $3N^2+4N$, enabling a simultaneous estimation of up to $Q_{\max} = N^2 + \lfloor 4/3N \rfloor$ DOAs, again by using the pseudoinverse and subsequent joint-diagonalisation.

Instead of a dedicated joint-diagonalisation algorithm, in practice an ad-hoc method has shown to be effective [HH19a, JZC20]. In the ad-hoc method first all three matrices $\boldsymbol{\Psi}_{\{xy^*,xy,z\}} := \boldsymbol{T}^{-1}\boldsymbol{\Theta}_{\{xy^*,xy,z\}}\boldsymbol{T}$ are diagonalised individually, and then the eigenvector matrix $\hat{\boldsymbol{T}}$ creating the lowest off-diagonal Frobenius norm is chosen for diagonalisation of all three matrices,

$$\hat{\boldsymbol{T}} = \arg\min_{\boldsymbol{T}} \sum_{\mu\in\{xy^*,xy,z\}} \left\| \mathrm{zdiag}(\boldsymbol{T}\boldsymbol{\Psi}_\mu\boldsymbol{T}^{-1}) \right\|_{\mathrm{F}}^2\,, \tag{2.33}$$

where $\mathrm{zdiag}(\cdot)$ extracts off-diagonal elements (sets diagonal elements to zero). As first proposal of this thesis, instead of an eigenvalue decomposition, a Schur decomposition $\boldsymbol{Q}\boldsymbol{\Psi}_\mu\boldsymbol{Q}^{\mathrm{T}}$ is employed which might yield similar performance at a lower computational expense. With the operator $\mathrm{tril}(\cdot)$ extracting the strictly-lower triangular matrix (setting elements of the upper triangular to zero), a Schur-decomposition-based ad-hoc method is formulated as

$$\hat{\boldsymbol{Q}} = \arg\min_{\boldsymbol{Q}} \sum_{\mu\in\{xy^*,xy,z\}} \left\| \mathrm{tril}(\boldsymbol{Q}\boldsymbol{\Psi}_\mu\boldsymbol{Q}^{\mathrm{T}}) \right\|_{\mathrm{F}}^2\,, \tag{2.34}$$

and compared to the eigendecomposition-based ad-hoc method in section 2.4.5. It is shown in appendix C that the Schur decomposition does not jumble the order of eigenvalue tuples and hence might serve as an effective alternative to joint triangularisation.

2.4.2 Multiplication Recurrences for Real-Valued SHs

Real-valued multiplication recurrence relations are deducted from the complex-valued ones in eqs. 2.22-2.24 by evaluating real and imaginary parts of the relations separately, as shown in the appendices A.2 and A.3. From appendix A.2 we get the following real-valued recurrence relations for $m > 0$,

$$\begin{aligned}\frac{(-1)^m}{\sqrt{2-\delta_m}}\theta_{\mathrm{x}}Y_{nm} = \tfrac{1}{2}\Big[&- \tfrac{(-1)^{m-1}w_n^m}{\sqrt{2-\delta_{m-1}}}Y_{n-1,m-1} + \tfrac{(-1)^{m-1}w_{n+1}^{-m+1}}{\sqrt{2-\delta_{m-1}}}Y_{n+1,m-1} \\ &+ \tfrac{(-1)^{m+1}w_n^{-m}}{\sqrt{2-\delta_{m+1}}}Y_{n-1,m+1} - \tfrac{(-1)^{m+1}w_{n+1}^{m+1}}{\sqrt{2-\delta_{m+1}}}Y_{n+1,m+1}\Big], \qquad (2.35)\end{aligned}$$

$$\begin{aligned}\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)\theta_{\mathrm{y}}Y_{n,-m} = \tfrac{1}{2}\Big[&- \tfrac{(-1)^{m-1}w_n^m}{\sqrt{2-\delta_{m-1}}}Y_{n-1,m-1} + \tfrac{(-1)^{m-1}w_{n+1}^{-m+1}}{\sqrt{2-\delta_{m-1}}}Y_{n+1,m-1} \\ &- \tfrac{(-1)^{m+1}w_n^{-m}}{\sqrt{2-\delta_{m+1}}}Y_{n-1,m+1} + \tfrac{(-1)^{m+1}w_{n+1}^{m+1}}{\sqrt{2-\delta_{m+1}}}Y_{n+1,m+1}\Big], \qquad (2.36)\end{aligned}$$

$$\begin{aligned}\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)\theta_{\mathrm{x}}Y_{n,-m} = \tfrac{1}{2}\Big[&- \tfrac{(-1)^{m-1}w_n^m}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n-1,-m+1} \\ &+ \tfrac{(-1)^{m-1}w_{n+1}^{-m+1}}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n+1,-m+1} + \tfrac{(-1)^{m+1}w_n^{-m}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n-1,-m-1} \\ &- \tfrac{(-1)^{m+1}w_{n+1}^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n+1,-m-1}\Big], \qquad (2.37)\end{aligned}$$

$$\begin{aligned}\frac{(-1)^m}{\sqrt{2-\delta_m}}\theta_{\mathrm{y}}Y_{n,m} = \tfrac{1}{2}\Big[&\tfrac{(-1)^{m-1}w_n^m}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n-1,-m+1} \\ &- \tfrac{(-1)^{m-1}w_{n+1}^{-m+1}}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n+1,-m+1} + \tfrac{(-1)^{m+1}w_n^{-m}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n-1,-m-1} \\ &- \tfrac{(-1)^{m+1}w_{n+1}^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n+1,-m-1}\Big], \qquad (2.38)\end{aligned}$$

for $m = 0$,

$$\theta_{\mathrm{x}}Y_{n0} = \frac{-1}{\sqrt{2}}w_n^0 Y_{n-1,1} + \frac{1}{\sqrt{2}}w_{n+1}^1 Y_{n+1,1}, \qquad (2.39)$$

$$\theta_{\mathrm{y}}Y_{n0} = \frac{-1}{\sqrt{2}}w_n^0 Y_{n-1,-1} + \frac{1}{\sqrt{2}}w_{n+1}^1 Y_{n+1,-1}\,, \qquad (2.40)$$

and for θ_{z},

$$\theta_{\mathrm{z}}Y_{nm} = v_n^{|m|}Y_{n-1,m} + v_{n+1}^{|m|}Y_{n+1,m}\,, \qquad (2.41)$$

re-using the weights w_n^m and v_n^m from eqs. 2.25, 2.26 and some additional factors/terms. From appendix A.3 we get the extending recurrences, yielding $4N-2$ independent equations valid for $m > 0$,

$$\begin{aligned}&\theta_{\mathrm{x}}\eta_n^{-m}\frac{(-1)^m}{\sqrt{2-\delta_m}}Y_{n,m} - \theta_{\mathrm{y}}\eta_n^{-m}\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-m} + \theta_{\mathrm{z}}\eta_n^{m+1}\frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}Y_{n,m+1} \\ &\quad = \eta_{n-1}^{-m}\frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}Y_{n-1,m+1}\,, \qquad (2.42)\end{aligned}$$

$$\theta_\mathrm{x}\eta_n^m\frac{(-1)^m}{\sqrt{2-\delta_m}}Y_{n,m}+\theta_\mathrm{y}\eta_n^m\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-m}-\theta_\mathrm{z}\eta_n^{-m+1}\frac{(-1)^{-m+1}}{\sqrt{2-\delta_{-m+1}}}Y_{n,|-m+1|}$$
$$=-\eta_{n-1}^m\frac{(-1)^{m-1}}{\sqrt{2-\delta_{m-1}}}Y_{n-1,m-1}\,, \tag{2.43}$$

$$\theta_\mathrm{x}\eta_n^{-m}\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-m}+\theta_\mathrm{y}\eta_n^{-m}\frac{(-1)^m}{\sqrt{2-\delta_m}}Y_{n,m}$$
$$+\theta_\mathrm{z}\eta_n^{m+1}\frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n,-|m+1|}$$
$$=\eta_{n-1}^{-m}\frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n-1,-|m+1|}\,, \tag{2.44}$$

$$\theta_\mathrm{x}\eta_n^m\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-m}-\theta_\mathrm{y}\eta_n^m\frac{(-1)^m}{\sqrt{2-\delta_m}}Y_{n,m}$$
$$-\theta_\mathrm{z}\eta_n^{-m+1}\frac{(-1)^{-m+1}}{\sqrt{2-\delta_{-m+1}}}(1-\delta_{-m+1})Y_{n,-|m+1|}$$
$$=-\eta_{n-1}^m\frac{(-1)^{m-1}}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n-1,-|m-1|}\,, \tag{2.45}$$

and additional 2 independent equations for $m=0$,

$$\theta_\mathrm{x}\eta_n^0Y_{n,0}-\theta_\mathrm{z}\eta_n^1\frac{1}{\sqrt{2}}Y_{n,1}=-\eta_{n-1}^0\frac{1}{\sqrt{2}}Y_{n-1,1}\,, \tag{2.46}$$

$$-\theta_\mathrm{y}\eta_n^0Y_{n,0}+\theta_\mathrm{z}\eta_n^1\frac{1}{\sqrt{2}}Y_{n,-1}=\eta_{n-1}^0\frac{1}{\sqrt{2}}Y_{n-1,-1}\,, \tag{2.47}$$

also re-using the constant η_n^m introduced for the complex-valued recurrences (eq. 2.31). Here, the real-valued directional Cartesian parameters $\{\theta_\mathrm{x},\theta_\mathrm{y},\theta_\mathrm{z}\}$ are directly contained in the recurrences, such that DOA angles are obtained by a conventional conversion from Cartesian to spherical coordinates. As implemented in the corresponding MATLAB code (provided online [1] and as listing in appendix A.4), these recurrences are stacked into matrices $\boldsymbol{M}_x$, $\boldsymbol{M}_y$, $\boldsymbol{M}_z$, $\boldsymbol{M}_L$ and $\boldsymbol{C}_L$, and arranged similarly to eqs. 2.28 and 2.32 to obtain a REVEB-ESPRIT formulation,

$$\begin{bmatrix}\boldsymbol{T}^{-1}\hat{\boldsymbol{\Theta}}_x\boldsymbol{T}\\ \boldsymbol{T}^{-1}\hat{\boldsymbol{\Theta}}_y\boldsymbol{T}\\ \boldsymbol{T}^{-1}\hat{\boldsymbol{\Theta}}_z\boldsymbol{T}\end{bmatrix}=\left(\begin{bmatrix}\boldsymbol{M}_T\\ \boldsymbol{M}_L\end{bmatrix}\begin{bmatrix}\boldsymbol{U}_\mathrm{s}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{U}_\mathrm{s}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{U}_\mathrm{s}\end{bmatrix}\right)^\dagger\begin{bmatrix}\boldsymbol{M}_x\\ \boldsymbol{M}_y\\ \boldsymbol{M}_z\\ \boldsymbol{C}_L\end{bmatrix}\boldsymbol{U}_\mathrm{s}\,, \tag{2.48}$$

where

$$\boldsymbol{M}_T=\begin{bmatrix}\boldsymbol{M}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{M}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{M}\end{bmatrix}\,. \tag{2.49}$$

These recurrences also allow for a simple expression of the extended pseudo-intensity vector as proposed in [HH19b]. A derivation is provided in appendix B.

1. `https://git.iem.at/thomasdeppisch/real-sh-recurrence-relations`

2.4.3 Joint Triangularisation using Rayleigh Quotient Iterations

Power iterations. A simultaneous Schur decomposition of an ensemble of matrices $\mathcal{A}$ can be achieved via power or Rayleigh quotient iterations and subsequent deflation. First, the power iterations algorithm [TB97, chapter 5] for a single matrix $\boldsymbol{A}$ is revised. The recurrence relation

$$\boldsymbol{v}^{(k+1)} = \frac{\boldsymbol{A}\boldsymbol{v}^{(k)}}{\|\boldsymbol{A}\boldsymbol{v}^{(k)}\|} \tag{2.50}$$

converges to the eigenvector $\boldsymbol{v}_1$ corresponding to the largest eigenvalue λ_1 of $\boldsymbol{A}$, if λ_1 is significantly larger in magnitude than the other eigenvalues. Once the eigenvector is found (up to chosen numerical precision), the corresponding eigenvalue is calculated as

$$\lambda_1 = \boldsymbol{v}_1^{\mathrm{T}} \boldsymbol{A} \boldsymbol{v}_1 \,. \tag{2.51}$$

Note that the power iteration algorithm reduces the error by a linear factor $\approx |\lambda_2/\lambda_1|$ in each iteration. If the two largest eigenvalues λ_1, λ_2 have similar magnitude, the power iterations algorithm only converges slowly.

Rayleigh quotient iterations. Cubic convergence is reached by combining inverse iterations [TB97, chapter 5]

$$\boldsymbol{v}^{(k+1)} = \frac{(\boldsymbol{A} - \mu\boldsymbol{I})^{-1}\boldsymbol{v}^{(k)}}{\|(\boldsymbol{A} - \mu\boldsymbol{I})^{-1}\boldsymbol{v}^{(k)}\|} \,, \tag{2.52}$$

where $(\boldsymbol{A} - \mu\boldsymbol{I})^{-1}$ has the same eigenvectors as $\boldsymbol{A}$ and eigenvalues $(\lambda_i - \mu)^{-1}$, making $(\boldsymbol{A} - \mu\boldsymbol{I})^{-1}\boldsymbol{v}^{(k)}$ converge fast to $\boldsymbol{v}_i$ if μ is close to λ_i, and the Rayleigh quotient [TB97, chapter 5] for the estimation of an eigenvalue

$$\mu = \lambda^{(k)} = \left(\boldsymbol{v}^{(k)}\right)^{\mathrm{T}} \boldsymbol{A}\boldsymbol{v}^{(k)} \,, \tag{2.53}$$

not converging to the largest eigenvalue λ_1 but to the eigenvalue closest to the initial guess $\lambda^{(0)}$. The power iterations and the Rayleigh quotient iterations [TB97, chapter 5] both only yield an estimate for one eigenvector, exhibiting the need for a deflation step to find the other eigenvectors and eigenvalues.

Convergence criterion. For both methods, convergence is reached, if the iteration does not change the direction of the eigenvector anymore. A useful convergence criterion is hence found by projecting $\boldsymbol{v}^{(k+1)}$ onto $\boldsymbol{v}^{(k)}$ and calculating the norm,

$$\left\| \boldsymbol{v}^{(k+1)} - \left(\boldsymbol{v}^{(k+1)^{\mathrm{T}}} \boldsymbol{v}^{(k)} \right) \boldsymbol{v}^{(k)} \right\| < \epsilon \,, \tag{2.54}$$

where ϵ is a small positive constant, e.g. $\epsilon = 10^{-6}$ and $\boldsymbol{v}^{(k)}$, $\boldsymbol{v}^{(k+1)}$ are unit vectors.

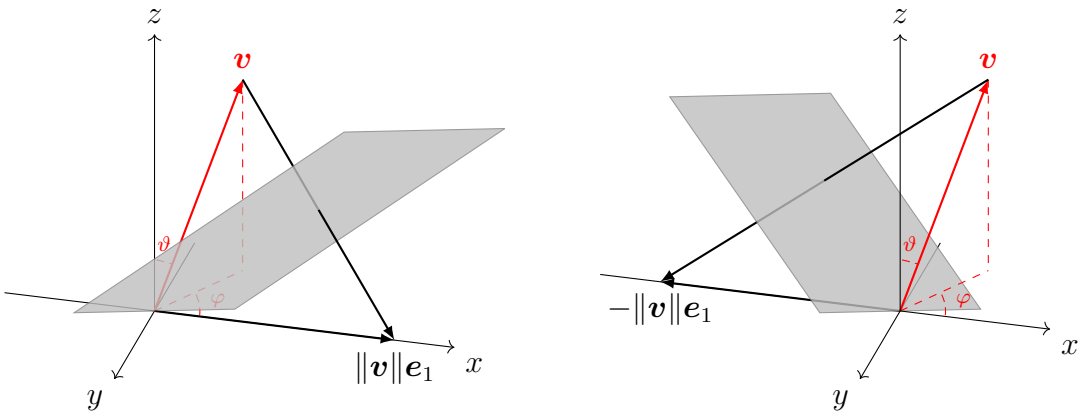


Figure 2.1 – The two possible Householder reflections that align the vector $\boldsymbol{v}$ with the unit vector of the first line $\boldsymbol{e}_1$.

Householder deflation. The idea of the deflation step is to remove the found eigenvector from the matrix $\boldsymbol{A}$ via projection to a subspace orthogonal to $\boldsymbol{v}$, e.g. $\boldsymbol{I} - \frac{\boldsymbol{v}\boldsymbol{v}^\mathrm{T}}{\boldsymbol{v}^\mathrm{T}\boldsymbol{v}}$. A numerically more stable deflation method is found by using (orthogonal) Householder reflectors $\boldsymbol{P}$ to find the similar matrix $\boldsymbol{B} = \boldsymbol{P}^\mathrm{T}\boldsymbol{A}\boldsymbol{P}$, where the first column of $\boldsymbol{B}$ holds the first eigenvalue λ_1 followed by zeros,

$$\boldsymbol{B} = \begin{bmatrix} \lambda_1 & \boldsymbol{b}^\mathrm{T} \\ \boldsymbol{0} & \boldsymbol{B}_1 \end{bmatrix} . \tag{2.55}$$

After matrix $\boldsymbol{B}$ is found, an eigenvalue of $\boldsymbol{B}_1$ is calculated via one of the iterative algorithms above and deflation can continue. A Householder reflector $\boldsymbol{P}$ is expressed as [TB97, chapter 2]

$$\boldsymbol{P} = \boldsymbol{I} - 2\frac{\boldsymbol{u}\boldsymbol{u}^\mathrm{T}}{\boldsymbol{u}^\mathrm{T}\boldsymbol{u}} , \tag{2.56}$$

with

$$\boldsymbol{u} = \mathrm{sign}(v_1)\|\boldsymbol{v}\|\boldsymbol{e}_1 + \boldsymbol{v} , \tag{2.57}$$

such that

$$\boldsymbol{P}\boldsymbol{v} = \|\boldsymbol{v}\|\boldsymbol{e}_1 . \tag{2.58}$$

As shown in figure 2.1, the idea of Householder reflections can be expressed geometrically: As a first idea to remove the eigenvector $\boldsymbol{v}$ from $\boldsymbol{A}$ one could project $\boldsymbol{A}$ onto the hyperplane orthogonal to $\boldsymbol{v}$, $\boldsymbol{H} = \boldsymbol{I} - \frac{\boldsymbol{v}\boldsymbol{v}^\mathrm{T}}{\boldsymbol{v}^\mathrm{T}\boldsymbol{v}}$. In case of Householder reflections, the reflected matrix does not have reduced rank but has a unit vector (scaled by the first eigenvalue) as first column and hence is also the first step of triangularisation. When reflecting $\boldsymbol{A}$ across the orthogonal hyperplane, going twice as far $\boldsymbol{P} = \boldsymbol{I} - 2\frac{\boldsymbol{u}\boldsymbol{u}^\mathrm{T}}{\boldsymbol{u}^\mathrm{T}\boldsymbol{u}}$, a scaled version of the Cartesian coordinate vector $\boldsymbol{e}_1$, $\boldsymbol{P}\boldsymbol{v} = \pm\|\boldsymbol{v}\|\boldsymbol{e}_1$ is supposed to be reached. This requirement is expressed by $(\boldsymbol{I} - 2\frac{\boldsymbol{u}\boldsymbol{u}^\mathrm{T}}{\boldsymbol{u}^\mathrm{T}\boldsymbol{u}})\boldsymbol{v} = \pm\|\boldsymbol{v}\|\boldsymbol{e}_1$ which is fulfilled by $\boldsymbol{u} = \boldsymbol{v} \pm \|\boldsymbol{v}\|\boldsymbol{e}_1$. For numerical stability, the sign creating the larger reflection is chosen, yielding equation 2.57. As the first column of $\boldsymbol{P}$ holds $\pm\boldsymbol{v}$, the first column of $\boldsymbol{P}^\mathrm{T}\boldsymbol{A}\boldsymbol{P}$ holds $\boldsymbol{e}_1\lambda_1$.

Simultaneous iterations. The initial idea for joint Schur decomposition in this thesis was a power iteration sending an eigenvector candidate through the set of matrices that is to be jointly decomposed. However, the power iteration algorithm exhibits linear convergence and only converges if $\boldsymbol{A}$ has a dominant eigenvalue (and the initial guess $\boldsymbol{v}^{(0)}$ has a non-zero component in direction of the corresponding eigenvector). For a rapid convergence, we only consider the Rayleigh quotient iterations algorithm below. Only for initialisation of the eigenvalues for the Rayleigh procedure, a single power iteration can be beneficial. Following the procedure above, we extend the algorithm for simultaneous Schur decomposition of J matrices in the ensemble $\mathcal{A} = \{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_J\}$. In each iteration, the eigenvector $\boldsymbol{v}_j^{(k+1)}$ is approximated by applying Rayleigh quotient iterations to each of the matrices,

$$\boldsymbol{v}_j^{(k+1)} = \frac{(\boldsymbol{A}_j - \lambda_j^{(k)}\boldsymbol{I})^{-1}\boldsymbol{v}^{(k)}}{\left\|(\boldsymbol{A}_j - \lambda_j^{(k)}\boldsymbol{I})^{-1}\boldsymbol{v}^{(k)}\right\|} \,. \tag{2.59}$$

Then, the eigenvector approximation for the next iteration $\boldsymbol{v}^{(k+1)}$ is calculated by constructively summing up the individual updated eigenvector estimates $\boldsymbol{v}_j^{(k+1)}$ and their renormalisation,

$$\boldsymbol{v}^{(k+1)} = \frac{\sum_{j=1}^{J} \boldsymbol{v}_j^{(k+1)} z_j^{(k+1)}}{\left\|\sum_{j=1}^{J} \boldsymbol{v}_j^{(k+1)} z_j^{(k+1)}\right\|} \,, \tag{2.60}$$

where the sign $z^{(k+1)}$ ensures constructive addition with the eigenvector estimate from the first matrix in the set

$$z_j^{(k+1)} = \operatorname{sign}\left(\left(\boldsymbol{v}_j^{(k+1)}\right)^{\mathrm{T}} \boldsymbol{v}_1^{(k+1)}\right) \,. \tag{2.61}$$

The corresponding eigenvalue estimates are updated as for a single-matrix problem,

$$\lambda_j^{(k)} = \left(\boldsymbol{v}^{(k)}\right)^{\mathrm{T}} \boldsymbol{A}_j \boldsymbol{v}^{(k)} \,. \tag{2.62}$$

If utilised in VEB-ESPRIT algorithms, where $\sum_j \lambda_j^2 \stackrel{!}{=} 1$, the eigenvalue estimates should moreover be normalised at the end of each iteration,

$$\lambda_j^{(k)} \leftarrow \frac{\lambda_j^{(k)}}{\sqrt{\sum_{i=1}^{J} \left(\lambda_i^{(k)}\right)^2}} \,. \tag{2.63}$$

Numerical simulations suggest an additional weighting of the eigenvector estimates $\boldsymbol{v}_j^{(k+1)}$ that is inversely proportional to the length, such that directions that could not converge yet and thus get shorter in the iteration are boosted for the subsequent iteration. Correspondingly, we propose the use of an inverse quadratic weighting, that modifies eq. 2.60

$$\boldsymbol{v}^{(k+1)} = \frac{\sum_{j=1}^{J} \boldsymbol{v}_j^{(k+1)} z_j^{(k+1)}}{\left\|\sum_{j=1}^{J} \boldsymbol{v}_j^{(k+1)} z_j^{(k+1)}\right\|^2} \,. \tag{2.64}$$

2.4.4 Algorithm Summary and Complexity

The proposed REVEB-ESPRIT algorithm is a reformulated version of EVEB-ESPRIT [JZC20], but uses real-valued SHs and corresponding recurrence relations with real-valued DOA-vector entries $\{\theta_\mathrm{x}, \theta_\mathrm{y}, \theta_\mathrm{z}\}$ as its directional parameters

$$\begin{bmatrix} \boldsymbol{T}^{-1}\hat{\boldsymbol{\Theta}}_x\boldsymbol{T} \\ \boldsymbol{T}^{-1}\hat{\boldsymbol{\Theta}}_y\boldsymbol{T} \\ \boldsymbol{T}^{-1}\hat{\boldsymbol{\Theta}}_z\boldsymbol{T} \end{bmatrix} = \left(\begin{bmatrix} \boldsymbol{M}_T \\ \boldsymbol{M}_L \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_\mathrm{s} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_\mathrm{s} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{U}_\mathrm{s} \end{bmatrix} \right)^\dagger \begin{bmatrix} \boldsymbol{M}_x \\ \boldsymbol{M}_y \\ \boldsymbol{M}_z \\ \boldsymbol{C}_L \end{bmatrix} \boldsymbol{U}_\mathrm{s}\,, \tag{2.65}$$

with the definitions in section 2.4.2. Three eigenvalue-revealing approaches were shown: Two ad-hoc approaches, either diagonalising (eq. 2.33) or triangularising (eq. 2.34) the three matrices individually and then choosing the eigen- or Schur vectors creating the lowest error in a common transform, and one iterative joint triangularisation approach using Rayleigh quotient iterations (section 2.4.3). As shown in [JZC20], EVEB-ESPRIT involves a complexity of $4O(N^6)$ operations for the subspace identification, $27\times4O(N^2Q^2)$ operations for the pseudo-inversion (by QR decomposition) of the EB-ESPRIT matrix, and $3 \times 4O(Q^3)$ operations for the ad-hoc diagonalisation. As REVEB-ESPRIT only involves real-valued operations, the complexity of all three steps is reduced by a factor of 4, i.e. subspace identification requires $O(N^6)$ operations, pseudoinversion $27 \times O(N^2Q^2)$ operations and ad-hoc diagonalisation (or triangularisation) $3 \times O(Q^3)$ operations. The Rayleigh quotient algorithm is of complexity $O(Q^3)$ [TB97] and hence the joint variant from section 2.4.3 is of complexity $3 \times O(Q^3)$ as it involves 3 matrix inversions at each iteration. The complexity of EVEB-ESPRIT, and also REVEB-ESPRIT, can be further reduced by subspace tracking techniques, such as PASTd [Yan95], and by two-step inversion [JZC20].

2.4.5 Evaluation

Evaluation setup. For comparison, REVEB-ESPRIT is evaluated using three different types of eigenvalue-revealing methods with the complex-valued EVEB-ESPRIT using the ad-hoc method for joint diagonalisation (cf. section 2.4.1, eq. 2.33) as a reference. For REVEB-ESPRIT, the same ad-hoc eigendecomposition is compared to an ad-hoc Schur decomposition (cf. section 2.4.1, eq. 2.34) and the iterative approach from section 2.4.3 using $\epsilon = 10^{-6}$ as a stopping criterion. The methods are evaluated using a simple free-field scenario: A free-field, aliasing-free SHD signal $s_n^m[\tau]$ of order $N = 3$ is simulated by creating $\mathcal{O} = 20$ samples of a varying number of $Q \in \{2, 6, 13\}$ plane-wave, unit-variance white-noise source signals $s_q[\tau]$, where τ denotes the discrete time index. The source signals are distributed in randomly-drawn directions $\boldsymbol{\theta}_q$ of a spherical 48-point 9-design [HS96]. Additive white Gaussian noise (AWGN) $\nu[\tau]$ is added directly in the SHD, creating varying signal-to-noise ratios (SNRs) in 10 dB steps from 10 dB to 80 dB,

$$s_n^m[\tau] = \left(\sum_{q=1}^{Q} Y_n^{m^*}(\boldsymbol{\theta}_q)\, s_q[\tau] \right) + \nu[\tau]\,. \tag{2.66}$$

Each combination of the different number of sources Q and varying SNR-step is repeated for a number of $J = 500$ trials. RMSEs of the great-circle distance

$$\Delta\boldsymbol{\theta}_q^{(j)} = \arccos\left\{\boldsymbol{\theta}_q^{(j)\mathrm{T}}\hat{\boldsymbol{\theta}}_q^{(j)}\right\}, \tag{2.67}$$

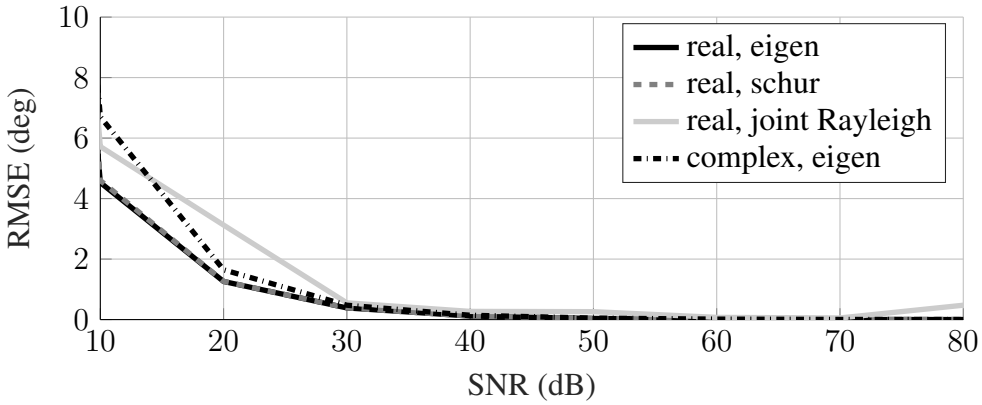
for each source q and each trial j, where $\boldsymbol{\theta}_q^{(j)}$ is the found DOA and $\hat{\boldsymbol{\theta}}_q^{(j)}$ is the true DOA, are calculated for each SNR step,

$$\mathrm{RMSE} = \sqrt{\sum_{j=1}^{J}\sum_{q=1}^{Q}\left|\Delta\boldsymbol{\theta}_q^{(j)}\right|^2/(JQ)}\,. \tag{2.68}$$

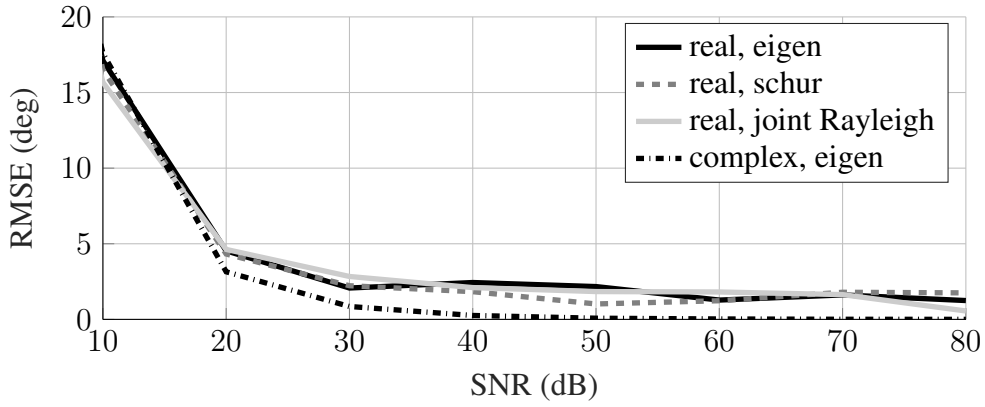
Results. Figure 2.2 shows the evaluation results for simultaneous estimation of $Q \in \{2, 6, 13\}$ DOAs. In most cases, the complex-valued ad-hoc diagonalisation (*complex, eigen*) outperforms the other algorithms in accuracy as indicated by its low RMSEs. Only the low SNR and $Q = 2$ conditions are exceptional and the real-valued ad-hoc variants (*real, eigen* and *real, schur*) yield slightly lower RMSEs. Moreover for the many-sources case with $Q = 13$ simultaneously estimated DOAs, the joint Rayleigh iterations approach can compete with the complex ad-hoc algorithm by delivering similarly accurate results.

A closer look on the inferior accuracy in the results of the real-valued ad-hoc algorithms (for $Q > 2$ cases) reveals the occurrence of eigenvalues of algebraic multiplicity as a cause. As those algorithms do not find a decomposition of all three matrices simultaneously, eigenvectors/Schur vectors of the joint problem can lack definition whenever their eigenvalue is of algebraic multiplicity. For instance, if multiple DOAs lie on the horizon, their multiple trivial eigenvalues in the z-coordinate problem cannot resolve the x and y components of the corresponding DOA vectors, even if the ad-hoc decomposition in z delivers the best approximation of diagonal/upper triangular matrices for the joint decomposition problem in x, y, and z. The complex-valued ad-hoc algorithm has the advantage that multiple eigenvalues are less likely as the eigenvalues in x and y can only vanish at nadir and zenith, and the problem in z is mainly useful to resolve coincidences in azimuth with elevations that are symmetric with regard to the horizon. Figure 2.3 shows RMSEs evaluated in the same way as before, but with a small jitter applied to the directional test layout, what could be considered as a more natural test scenario free of strict symmetries, etc. and avoiding algebraic multiplicities of the eigenvalues. The jitter was applied as Gaussian noise with variance $\sigma_\mathrm{j}^2 \approx 0.57°$. For this somewhat less artificial DOA test set, the real-values ad-hoc variants are able to compete slightly better with their complex-valued counterpart.

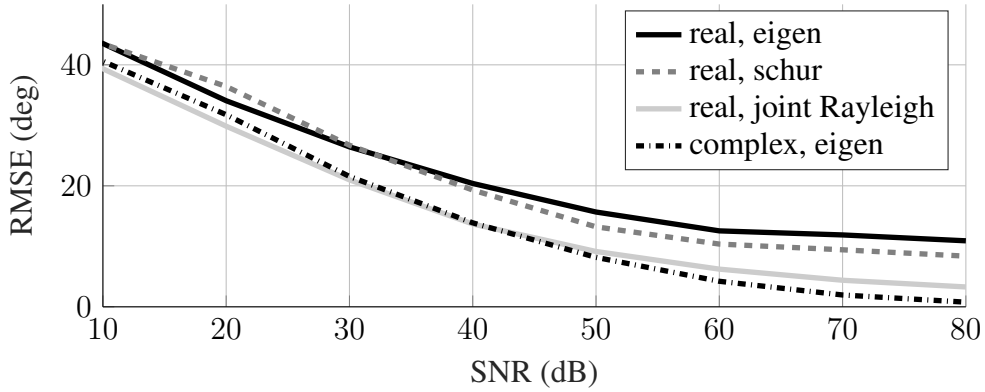
The larger RMSEs at high SNRs of the joint Rayleigh iterations approach (e.g. in figure 2.2(a) for an SNR of 80 dB) are possibly due to spontaneous errors occasionally caused by disadvantageous random initialisation of a Schur vector.



(a) $Q = 2$.



(b) $Q = 6$.



(c) $Q = 13$.

Figure 2.2 – RMSE as a function of SNR for complex-valued EVEB-ESPRIT and three real-valued EVEB-ESPRIT versions using different eigenvalue-revealing decompositions. The three plots show results for different numbers of sources Q (simultaneously estimated DOAs). The maximum SH order is set to $N = 3$.

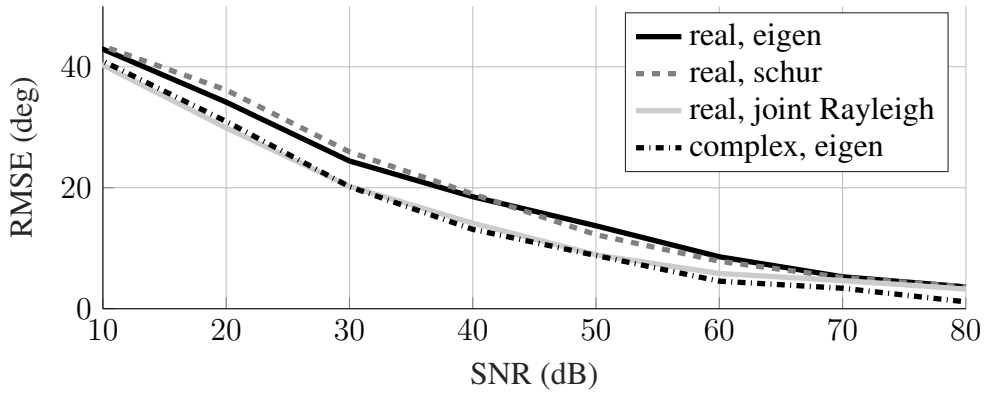


Figure 2.3 – RMSEs for simultaneous estimation of $Q = 13$ DOAs, where a jitter was applied to the source positions to avoid a multiplicity of eigenvalues.

2.5 New MUSIC

This section introduces the Newton-based Multiple Signal Classification algorithm (New MUSIC) as another contribution of this work. New MUSIC iteratively retrieves the direction vectors of peaks of the MUSIC spectrum (cf. section 2.2). To enable iterations via Newton's method, New MUSIC applies gradient recurrences of real-valued SHs for efficient calculation of gradient and Hessian. The employed recurrences allow for differentiation of the direction vector components $\partial/\partial x$, $\partial/\partial y$ and $\partial/\partial z$ and are defined on the whole surface of the unit sphere $\mathbb{S}^2$. Corresponding recurrences for complex-valued SHs were developed by Chew [Che92] and e.g. applied in an acoustic centreing algorithm by Deboy and Zotter [DZ11]. Gumerov and Duraiswami [GD01] developed recurrence relations for complex-valued directional components $\partial/\partial x + i\partial/\partial y$, $\partial/\partial x - i\partial/\partial y$ and $\partial/\partial z$. Gräf and Potts [GP11] developed a similar Newton-based algorithm for the computation of spherical designs based on recurrences of complex-valued SHs for differentiation with respect to the spherical coordinates $\partial/\partial\varphi$ and $\partial/\partial\vartheta$. Differentiation with respect to azimuth angle φ is however not well-defined at zenith ($\vartheta = 0$) and nadir ($\vartheta = \pi$).

2.5.1 Gradient Addition Recurrences

From appendix A.1, recurrence relations for the calculation of the gradient of real-valued SHs,

$$\boldsymbol{Y}_{nm}(\varphi,\mu) = N_{nm} P_n^m(\mu)\boldsymbol{\Phi}_m(\varphi)\,, \tag{2.69}$$

with sine and cosine harmonics stacked as

$$\boldsymbol{\Phi}_m(\varphi) = \begin{pmatrix}\cos(m\varphi)\\ \sin(m\varphi)\end{pmatrix}\,, \tag{2.70}$$

are obtained as,

$$\begin{aligned}\partial/\partial x\boldsymbol{Y}_{nm} = &-\frac{(n+1)(n+m-1)(n+m)N_{nm}}{2(2n+1)N_{n-1,m-1}}\boldsymbol{Y}_{n-1,m-1} \\ &-\frac{n(n-m+1)(n-m+2)N_{nm}}{2(2n+1)N_{n+1,m-1}}\boldsymbol{Y}_{n+1,m-1} \\ &+\frac{(n+1)N_{nm}}{2(2n+1)N_{n-1,m+1}}\boldsymbol{Y}_{n-1,m+1}+\frac{nN_{nm}}{2(2n+1)N_{n+1,m+1}}\boldsymbol{Y}_{n+1,m+1}\,,\end{aligned} \tag{2.71}$$

$$\begin{aligned}\partial/\partial y\boldsymbol{Y}_{nm} = &-\frac{(n+1)(n+m-1)(n+m)N_{nm}}{2(2n+1)N_{n-1,m-1}}\boldsymbol{L}\boldsymbol{Y}_{n-1,m-1} \\ &-\frac{n(n-m+1)(n-m+2)N_{nm}}{2(2n+1)N_{n+1,m-1}}\boldsymbol{L}\boldsymbol{Y}_{n+1,m-1} \\ &-\frac{(n+1)N_{nm}}{2(2n+1)N_{n-1,m+1}}\boldsymbol{L}\boldsymbol{Y}_{n-1,m+1}-\frac{nN_{nm}}{2(2n+1)N_{n+1,m+1}}\boldsymbol{L}\boldsymbol{Y}_{n+1,m+1}\,,\end{aligned} \tag{2.72}$$

$$\partial/\partial z\boldsymbol{Y}_{nm} = \frac{(n+1)(n+m)N_{nm}}{(2n+1)N_{n-1,m}}\boldsymbol{Y}_{n-1,m}-\frac{n(n-m+1)N_{nm}}{(2n+1)N_{n+1,m}}\boldsymbol{Y}_{n+1,m}\,, \tag{2.73}$$

where $\boldsymbol{L}$ changes cosine to minus sine and sine to cosine for the derivative w.r.t. y

$$\boldsymbol{L} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} , \tag{2.74}$$

and for $m = 0$ we get the exception

$$\partial/\partial x Y_{n,0}^{(\cos)} = \frac{(n+1)N_{n,0}}{(2n+1)N_{n-1,1}} Y_{n-1,1}^{(\cos)} + \frac{nN_{n,0}}{(2n+1)N_{n+1,1}} Y_{n+1,1}^{(\cos)} , \tag{2.75}$$

$$\partial/\partial y Y_{n,0}^{(\cos)} = \frac{(n+1)N_{n,0}}{(2n+1)N_{n-1,1}} Y_{n-1,1}^{(\sin)} + \frac{nN_{n,0}}{(2n+1)N_{n+1,1}} Y_{n+1,1}^{(\sin)} . \tag{2.76}$$

The recurrence relations are stacked into the constant matrices $\boldsymbol{G}_x$, $\boldsymbol{G}_y$ and $\boldsymbol{G}_z$, such that a tangential gradient w.r.t. x, y and z of an order N SH pattern $\boldsymbol{\gamma}$, evaluated at the variable direction $\hat{\boldsymbol{\theta}}$, is obtained as

$$\nabla_{\boldsymbol{\theta}} \boldsymbol{\gamma}^{\mathrm{T}} \boldsymbol{y}_N(\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} := \boldsymbol{g} = \begin{bmatrix} \boldsymbol{\gamma}^{\mathrm{T}} \boldsymbol{G}_x \boldsymbol{y}_{N+1}(\hat{\boldsymbol{\theta}}) \\ \boldsymbol{\gamma}^{\mathrm{T}} \boldsymbol{G}_y \boldsymbol{y}_{N+1}(\hat{\boldsymbol{\theta}}) \\ \boldsymbol{\gamma}^{\mathrm{T}} \boldsymbol{G}_z \boldsymbol{y}_{N+1}(\hat{\boldsymbol{\theta}}) \end{bmatrix} , \tag{2.77}$$

where $\boldsymbol{y}_{N+1}(\hat{\boldsymbol{\theta}})$ holds SHs up to the order $N+1$, evaluated at $\hat{\boldsymbol{\theta}}$. MATLAB code to obtain the matrices $\boldsymbol{G}_x$, $\boldsymbol{G}_y$ and $\boldsymbol{G}_z$ is provided online [2] and as listing in appendix A.4.

2.5.2 Newton's Method on the Surface of the Sphere

In the following, the gradient recurrences of the previous section are used to formulate Newton's method in the SHD. The gradient matrices are employed as a tool: They contain what is required to evaluate the gradient and Hessian, which both enable to set up Newton's method for (i) finding the zeros (roots) of an SH expansion, (ii) finding its extrema, and finding extrema (iii) or zeros (iv) of the squared SH expansion; the latter one is applied in the New MUSIC algorithm.

Roots of a linear SH expansion. Evaluation of an SH pattern $\boldsymbol{\gamma}$ at the direction $\boldsymbol{\theta}$ is expressed as

$$f(\boldsymbol{\theta}) = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta}) \boldsymbol{\gamma} . \tag{2.78}$$

Now Newton's method is employed to iteratively estimate its zeros $f(\boldsymbol{\theta}) = 0$ using the re-fining update

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{j}_f^{\mathrm{T}} \frac{\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta}_k)\boldsymbol{\gamma}}{\boldsymbol{j}_f \boldsymbol{j}_f^{\mathrm{T}}} , \tag{2.79}$$

where $\boldsymbol{j}_f$ is the Jacobian that for $f : \mathbb{R}^3 \to \mathbb{R}$ degenerates to the transpose of the gradient vector $\boldsymbol{g}$,

$$\boldsymbol{j}_f = \boldsymbol{g}_f^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{\gamma}^{\mathrm{T}} \boldsymbol{G}_x \boldsymbol{y}_{N+1} & \boldsymbol{\gamma}^{\mathrm{T}} \boldsymbol{G}_y \boldsymbol{y}_{N+1} & \boldsymbol{\gamma}^{\mathrm{T}} \boldsymbol{G}_z \boldsymbol{y}_{N+1} \end{bmatrix} . \tag{2.80}$$

2. `https://git.iem.at/thomasdeppisch/real-sh-recurrence-relations`

Extrema of a linear SH expansion. Newton's method is expanded to the search of extrema by replacing function and gradient by gradient and Hessian, and searching for the roots of the gradient,

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{H}_f(\boldsymbol{\theta}_k)^{-1}\boldsymbol{g}_f(\boldsymbol{\theta}_k)\,, \tag{2.81}$$

using the Hessian

$$\boldsymbol{H}_f = \begin{bmatrix} \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_x\boldsymbol{G}_x\boldsymbol{y}_{N+2} & \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_x\boldsymbol{G}_y\boldsymbol{y}_{N+2} & \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_x\boldsymbol{G}_z\boldsymbol{y}_{N+2} \\ \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_y\boldsymbol{G}_x\boldsymbol{y}_{N+2} & \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_y\boldsymbol{G}_y\boldsymbol{y}_{N+2} & \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_y\boldsymbol{G}_z\boldsymbol{y}_{N+2} \\ \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_z\boldsymbol{G}_x\boldsymbol{y}_{N+2} & \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_z\boldsymbol{G}_y\boldsymbol{y}_{N+2} & \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_z\boldsymbol{G}_z\boldsymbol{y}_{N+2} \end{bmatrix}\,, \tag{2.82}$$

which is also known as *Newton's method in optimisation*.

Roots of a squared SH expansion. Similarly, the method can be applied to a squared SH-expanded function,

$$f^2(\boldsymbol{\theta}) = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{y}(\boldsymbol{\theta})\,, \tag{2.83}$$

to iteratively retrieve its roots by the update

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \frac{1}{\boldsymbol{g}_{f^2}^{\mathrm{T}}\boldsymbol{g}_{f^2}}\boldsymbol{g}_{f^2}\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{y}(\boldsymbol{\theta})\,, \tag{2.84}$$

and the gradient is obtained by applying the product rule,

$$\boldsymbol{g}_{f^2} = \begin{bmatrix} \boldsymbol{y}_{N+1}^{\mathrm{T}}\boldsymbol{G}_x^{\mathrm{T}}\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{y}_N + \boldsymbol{y}_N^{\mathrm{T}}\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_x\boldsymbol{y}_{N+1} \\ \boldsymbol{y}_{N+1}^{\mathrm{T}}\boldsymbol{G}_y^{\mathrm{T}}\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{y}_N + \boldsymbol{y}_N^{\mathrm{T}}\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_y\boldsymbol{y}_{N+1} \\ \boldsymbol{y}_{N+1}^{\mathrm{T}}\boldsymbol{G}_z^{\mathrm{T}}\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{y}_N + \boldsymbol{y}_N^{\mathrm{T}}\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_z\boldsymbol{y}_{N+1} \end{bmatrix}\,. \tag{2.85}$$

Extrema of a squared SH expansion. Furthermore, extrema of squared SH-expanded patterns are found via the update equation

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{H}_{f^2}(\boldsymbol{\theta}_k)^{-1}\boldsymbol{g}_{f^2}(\boldsymbol{\theta}_k)\,, \tag{2.86}$$

with the Hessian

$$\boldsymbol{H}_{f^2} = \begin{bmatrix} H_{xx} & H_{yx} & H_{zx} \\ H_{xy} & H_{yy} & H_{zy} \\ H_{xz} & H_{yz} & H_{zz} \end{bmatrix}\,, \tag{2.87}$$

where

$$\begin{aligned} H_{\alpha\beta} = {} & \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_\beta\boldsymbol{G}_\alpha\boldsymbol{y}_{N+2}\boldsymbol{y}_N^{\mathrm{T}}\boldsymbol{\gamma} + \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_\alpha\boldsymbol{y}_{N+1}\boldsymbol{y}_{N+1}^{\mathrm{T}}\boldsymbol{G}_\beta^{\mathrm{T}}\boldsymbol{\gamma} \\ & + \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{G}_\beta\boldsymbol{y}_{N+1}\boldsymbol{y}_{N+1}^{\mathrm{T}}\boldsymbol{G}_\alpha^{\mathrm{T}}\boldsymbol{\gamma} + \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{y}_{N+1}\boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_\beta^{\mathrm{T}}\boldsymbol{G}_\alpha^{\mathrm{T}}\boldsymbol{\gamma}\,. \end{aligned} \tag{2.88}$$

Notice that this Hessian is generally non-symmetric as $\boldsymbol{G}_\alpha\boldsymbol{G}_\beta \neq \boldsymbol{G}_\beta\boldsymbol{G}_\alpha$.

Towards New MUSIC: Analysis of squared SH expansions via signal and noise subspaces. For squared SH-expanded functions we can analyse signal subspace $\boldsymbol{U}_\mathrm{s}$ and noise subspace $\boldsymbol{U}_\mathrm{n}$ separately using eigendecomposition, $\boldsymbol{\gamma}\boldsymbol{\gamma}^\mathrm{T} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\mathrm{T}$ and $\boldsymbol{Q} = \begin{bmatrix}\boldsymbol{U}_\mathrm{s} & \boldsymbol{U}_\mathrm{n}\end{bmatrix}$, yielding

$$f^2(\boldsymbol{\theta}) = \boldsymbol{y}^\mathrm{T}(\boldsymbol{\theta})\boldsymbol{\gamma}\boldsymbol{\gamma}^\mathrm{T}\boldsymbol{y}(\boldsymbol{\theta}) = \boldsymbol{y}^\mathrm{T}(\boldsymbol{\theta})\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\mathrm{T}\boldsymbol{y}(\boldsymbol{\theta})\,. \tag{2.89}$$

The MUSIC cost function consists of eigenvectors of the signal covariance (cf. section 2.2) and either maximises the signal function

$$f_\mathrm{s}^2(\boldsymbol{\theta}) = \boldsymbol{y}^\mathrm{T}(\boldsymbol{\theta})\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{y}(\boldsymbol{\theta})\,, \tag{2.90}$$

or minimises the noise function

$$f_\mathrm{n}^2(\boldsymbol{\theta}) = \boldsymbol{y}^\mathrm{T}(\boldsymbol{\theta})\boldsymbol{U}_\mathrm{n}\boldsymbol{U}_\mathrm{n}^\mathrm{T}\boldsymbol{y}(\boldsymbol{\theta})\,, \tag{2.91}$$

at DOAs $\boldsymbol{\theta}$ of the signal. In particular, the signal and noise functions have roots at directions of the noise subspace and the signals subspace, respectively, as $\boldsymbol{y}(\boldsymbol{\theta})$ is orthogonal to the vectors in $\boldsymbol{U}_\mathrm{s}$ and $\boldsymbol{U}_\mathrm{n}$ for directions in the noise and signal subspace. Notice that searching for maxima in the signal space is equivalent to searching for minima in the noise space and vice versa, as $\boldsymbol{Q}\boldsymbol{Q}^\mathrm{T} = \boldsymbol{I} = \boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T} + \boldsymbol{U}_\mathrm{n}\boldsymbol{U}_\mathrm{n}^\mathrm{T}$ and hence

$$\boldsymbol{y}^\mathrm{T}\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{y} = \boldsymbol{y}^\mathrm{T}(\boldsymbol{I} - \boldsymbol{U}_\mathrm{n}\boldsymbol{U}_\mathrm{n}^\mathrm{T})\boldsymbol{y} = \frac{(N+1)^2}{4\pi} - \boldsymbol{y}^\mathrm{T}\boldsymbol{U}_\mathrm{n}\boldsymbol{U}_\mathrm{n}^\mathrm{T}\boldsymbol{y}\,. \tag{2.92}$$

Now four objectives can be analysed: A root search of the signal subspace yields minima of the quadratic function,

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{g}_{f_\mathrm{s}^2}\frac{\boldsymbol{y}(\boldsymbol{\theta}_k)^\mathrm{T}\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{y}(\boldsymbol{\theta}_k)}{\boldsymbol{g}_{f_\mathrm{s}^2}^\mathrm{T}\boldsymbol{g}_{f_\mathrm{s}^2}}\,, \tag{2.93}$$

and a root search of the noise subspace yields peaks of the quadratic function,

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{g}_{f_\mathrm{n}^2}\frac{\boldsymbol{y}(\boldsymbol{\theta}_k)^\mathrm{T}\boldsymbol{U}_\mathrm{n}\boldsymbol{U}_\mathrm{n}^\mathrm{T}\boldsymbol{y}(\boldsymbol{\theta}_k)}{\boldsymbol{g}_{f_\mathrm{n}^2}^\mathrm{T}\boldsymbol{g}_{f_\mathrm{n}^2}}\,. \tag{2.94}$$

Extrema of the signal subspace are found via

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{H}_{f_\mathrm{s}^2}(\boldsymbol{\theta}_k)^{-1}\boldsymbol{g}_{f_\mathrm{s}^2}(\boldsymbol{\theta}_k)\,, \tag{2.95}$$

and extrema of the noise subspace are found via

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{H}_{f_\mathrm{n}^2}(\boldsymbol{\theta}_k)^{-1}\boldsymbol{g}_{f_\mathrm{n}^2}(\boldsymbol{\theta}_k)\,. \tag{2.96}$$

Here, the gradient of the signal function is described as

$$\boldsymbol{g}_{f_\mathrm{s}^2} = \begin{bmatrix}\boldsymbol{y}_{N+1}^\mathrm{T}\boldsymbol{G}_x^\mathrm{T}\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{y}_N + \boldsymbol{y}_N^\mathrm{T}\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{G}_x\boldsymbol{y}_{N+1} \\ \boldsymbol{y}_{N+1}^\mathrm{T}\boldsymbol{G}_y^\mathrm{T}\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{y}_N + \boldsymbol{y}_N^\mathrm{T}\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{G}_y\boldsymbol{y}_{N+1} \\ \boldsymbol{y}_{N+1}^\mathrm{T}\boldsymbol{G}_z^\mathrm{T}\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{y}_N + \boldsymbol{y}_N^\mathrm{T}\boldsymbol{U}_\mathrm{s}\boldsymbol{U}_\mathrm{s}^\mathrm{T}\boldsymbol{G}_z\boldsymbol{y}_{N+1}\end{bmatrix}\,, \tag{2.97}$$

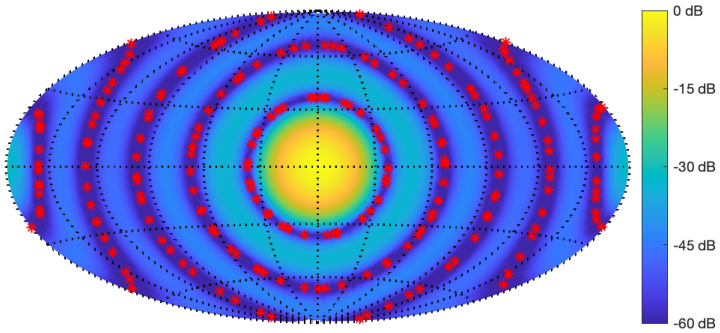
and the Hessian of the signal function is

$$\boldsymbol{H}_{f_\mathrm{s}^2} = \begin{bmatrix} H_{\mathrm{s},xx} & H_{\mathrm{s},xy} & H_{\mathrm{s},xz} \\ H_{\mathrm{s},yx} & H_{\mathrm{s},yy} & H_{\mathrm{s},yz} \\ H_{\mathrm{s},zx} & H_{\mathrm{s},zy} & H_{\mathrm{s},zz} \end{bmatrix} , \tag{2.98}$$

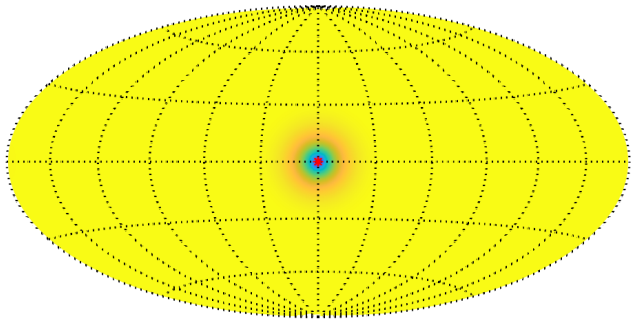
where

$$\begin{aligned} H_{\mathrm{s},\alpha\beta} = {} & \boldsymbol{y}_{N+2}^\mathrm{T} \boldsymbol{G}_\beta^\mathrm{T} \boldsymbol{G}_\alpha^\mathrm{T} \boldsymbol{U}_\mathrm{s} \boldsymbol{U}_\mathrm{s}^\mathrm{T} \boldsymbol{y}_N + \boldsymbol{y}_{N+1}^\mathrm{T} \boldsymbol{G}_\alpha^\mathrm{T} \boldsymbol{U}_\mathrm{s} \boldsymbol{U}_\mathrm{s}^\mathrm{T} \boldsymbol{G}_\beta \boldsymbol{y}_{N+1} \\ & + \boldsymbol{y}_{N+1}^\mathrm{T} \boldsymbol{G}_\beta^\mathrm{T} \boldsymbol{U}_\mathrm{s} \boldsymbol{U}_\mathrm{s}^\mathrm{T} \boldsymbol{G}_\alpha \boldsymbol{y}_{N+1} + \boldsymbol{y}_N^\mathrm{T} \boldsymbol{U}_\mathrm{s} \boldsymbol{U}_\mathrm{s}^\mathrm{T} \boldsymbol{G}_\alpha \boldsymbol{G}_\beta \boldsymbol{y}_{N+2} \,. \end{aligned} \tag{2.99}$$

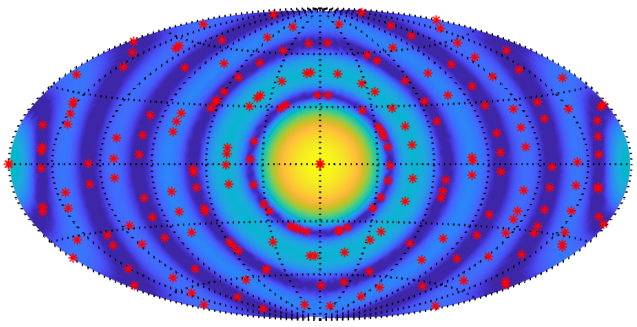
The matrices for the noise subspace are defined by replacing $\boldsymbol{U}_\mathrm{s}$ by $\boldsymbol{U}_\mathrm{n}$. Note that the Hessian can only be of rank two at most, as its derivatives are all tangential and exclude any radial component. For its inversion, section 3.2.1 shows how to project the Hessian to its two relevant tangential dimensions before inversion. Figure 2.4 shows the magnitude of the signal and noise subspaces of a fifth-order white-noise signal encoded at $\varphi = 0$, $\vartheta = \pi/2$ and converged New MUSIC optimisation points to either find the roots of the signal subspace, roots of the noise subspace, extrema of the signal subspace, or extrema of the noise subspace (eqs. 2.93-2.96).



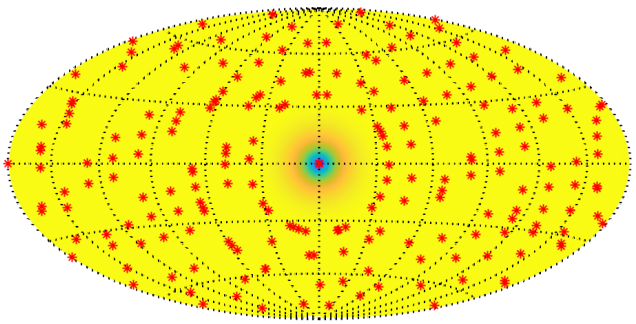
(a) Objective 1: Roots of the signal subspace.



(b) Objective 2: Root of the noise subspace.



(c) Objective 3: Extrema of the signal subspace.



(d) Objective 4: Extrema of the noise subspace.

Figure 2.4 – Directional magnitude of signal and noise subspaces of a white-noise signal encoded at $\varphi = 0$, $\vartheta = \pi/2$, depicted via a map projection. Red asterisks show converged New MUSIC optimisation points for the four different objectives (eqs. 2.93-2.96).

2.5.3 Summary of the New MUSIC Algorithm for DOA estimation

The search of roots in the noise subspace (objective 2) exhibits a low computational cost as no Hessian needs to be computed. Moreover, it is particularly effective for DOA estimation as it avoids optimisation towards the minima of the function and hence is less sensitive to local extrema than the other objectives (cf. figure 2.4). As described in section 2.5.2, the Newton steps for this objective are performed as

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \mu\, \boldsymbol{g}_{f_\mathrm{n}^2} \frac{\boldsymbol{y}(\boldsymbol{\theta}_k)^\mathrm{T} \boldsymbol{U}_\mathrm{n} \boldsymbol{U}_\mathrm{n}^\mathrm{T} \boldsymbol{y}(\boldsymbol{\theta}_k)}{\boldsymbol{g}_{f_\mathrm{n}^2}^\mathrm{T} \boldsymbol{g}_{f_\mathrm{n}^2}}\,, \tag{2.100}$$

with

$$\boldsymbol{g}_{f_\mathrm{n}^2} = \begin{bmatrix} \boldsymbol{y}_{N+1}^\mathrm{T} \boldsymbol{G}_x^\mathrm{T} \boldsymbol{U}_\mathrm{n} \boldsymbol{U}_\mathrm{n}^\mathrm{T} \boldsymbol{y}_N + \boldsymbol{y}_N^\mathrm{T} \boldsymbol{U}_\mathrm{n} \boldsymbol{U}_\mathrm{n}^\mathrm{T} \boldsymbol{G}_x \boldsymbol{y}_{N+1} \\ \boldsymbol{y}_{N+1}^\mathrm{T} \boldsymbol{G}_y^\mathrm{T} \boldsymbol{U}_\mathrm{n} \boldsymbol{U}_\mathrm{n}^\mathrm{T} \boldsymbol{y}_N + \boldsymbol{y}_N^\mathrm{T} \boldsymbol{U}_\mathrm{n} \boldsymbol{U}_\mathrm{n}^\mathrm{T} \boldsymbol{G}_y \boldsymbol{y}_{N+1} \\ \boldsymbol{y}_{N+1}^\mathrm{T} \boldsymbol{G}_z^\mathrm{T} \boldsymbol{U}_\mathrm{n} \boldsymbol{U}_\mathrm{n}^\mathrm{T} \boldsymbol{y}_N + \boldsymbol{y}_N^\mathrm{T} \boldsymbol{U}_\mathrm{n} \boldsymbol{U}_\mathrm{n}^\mathrm{T} \boldsymbol{G}_z \boldsymbol{y}_{N+1} \end{bmatrix}. \tag{2.101}$$

In addition to eq. 2.94, a step size parameter $\mu \in [0, 1]$ is introduced into the equation to control the convergence speed.

Similar to the case of REVEB-ESPRIT (cf. section 2.4.4), the most complex part of New MUSIC is the eigen- or singular value decomposition with a general complexity of $O(N^6)$. This complexity can be reduced by using subspace tracking techniques, such as PASTd [Yan95]. As the matrix multiplications of $\boldsymbol{U}_\mathrm{n}\boldsymbol{U}_\mathrm{n}^\mathrm{T}$ during the Newton step and of $\boldsymbol{G}_{\{x,y,z\}}^\mathrm{T}\boldsymbol{U}_\mathrm{n}\boldsymbol{U}_\mathrm{n}^\mathrm{T}$ in the gradient expression do not change while iterating, they can be precomputed, and the evaluation of SHs can be implemented efficiently using look-up tables or recurrence relations. This leaves the matrix-vector product $\left(\boldsymbol{G}_{\{x,y,z\}}^\mathrm{T}\boldsymbol{U}_\mathrm{n}\boldsymbol{U}_\mathrm{n}^\mathrm{T}\right)\boldsymbol{y}_N$ as the computationally most complex operation performed in every iteration. This product consists of a multiplication of a $(N+2)^2 \times (N+1)^2$ matrix by a $(N+1)^2 \times 1$ vector and hence generally involves $O(N^4)$ multiplication operations.

2.5.4 Maximum Number of Extrema in SH Patterns

When using New MUSIC to retrieve extrema in order N SH patterns, the question arises, how many extrema such a pattern can exhibit. In the following an upper bound for this maximum is found by first describing the width of the main lobe of an SH basic beam of order N and then investigating how close two main lobes can come together without collapsing into a single maximum.

Beam width of a SH basic beam. Rafaely [Raf04] uses the SH addition theorem,

$$\sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}_1) Y_n^{m^*}(\boldsymbol{\theta}_2) = \frac{2n+1}{4\pi} P_n(\cos\Theta)\,, \tag{2.102}$$

where $\boldsymbol{\theta} = \{\varphi, \vartheta\}$ and Θ is the angle between $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, and a recurrence relation of Legendre polynomials P_n, to describe the spherical function $f_N(\boldsymbol{\theta})$ of maximum order N

that peaks at $\boldsymbol{\theta}_0$,

$$f_N(\boldsymbol{\theta}) = \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}) Y_n^{m^*}(\boldsymbol{\theta}_0) \tag{2.103}$$

$$= \sum_{n=0}^{N} \frac{2n+1}{4\pi} P_n(\cos\Theta) \tag{2.104}$$

$$= \frac{N+1}{4\pi(\cos\Theta - 1)} \left(P_{N+1}(\cos\Theta) - P_N(\cos\Theta)\right) . \tag{2.105}$$

Its main-lobe width is found as twice the distance of the first zero Θ_0. Note that even though one might expect a zero at $\Theta = 0$ as $P_{N+1}(1) - P_N(1) = 1 - 1 = 0$, this zero is canceled by the pole $1/(\cos(\Theta) - 1)$. Rafaely [Raf04] provides analytic expressions for the emerging polynomials for orders 1 to 5 and $z = \cos\Theta$,

$$f_0(z) = \frac{1}{4\pi} , \tag{2.106}$$

$$f_1(z) = \frac{1}{4\pi}(3z+1) , \tag{2.107}$$

$$f_2(z) = \frac{1}{4\pi}\frac{3}{2}(5z^2+2z-1) , \tag{2.108}$$

$$f_3(z) = \frac{1}{4\pi}\frac{1}{2}(35z^3+15z^2-15z-3) , \tag{2.109}$$

$$f_4(z) = \frac{1}{4\pi}\frac{5}{8}(63z^4+28z^3-42z^2-12z+3) , \tag{2.110}$$

$$f_5(z) = \frac{1}{4\pi}\frac{6}{16}(231z^5+105z^4-210z^3-70z^2+35z+5) , \tag{2.111}$$

which allows for the calculation of the first zero or beam width. Rafaely also proposes an approximation for the distance of the first zero, which creates an error of less than two degree for orders $N \in [4, 40]$,

$$\Theta_0 \approx \frac{\pi}{N} . \tag{2.112}$$

Minimum distance of two maxima. As we are interested in the maximum number of peaks an SH pattern can have, it is not sufficient two distribute SH beams regularly on the sphere, overlapping at the border of the main lobe. Hence, we now calculate the minimum angular spacing that two plane-wave sources (SH basic beam peaks) need to have, such that their combined peaks do not collapse into a single peak. Two additively superimposed SH beams are denoted as

$$f_N(\boldsymbol{\theta}) = \sum_{n=0}^{N} \sum_{m=-n}^{n} (Y_n^{m^*}(\boldsymbol{\theta}_1) + Y_n^{m^*}(\boldsymbol{\theta}_2)) Y_n^m(\boldsymbol{\theta})$$

$$= \sum_{n=0}^{N} \frac{2n+1}{4\pi} (P_n(\cos(\Theta_1)) + P_n(\cos(\Theta_2))) , \tag{2.113}$$

where Θ_1 and Θ_2 are the angles between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_1$ and between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_2$, respectively. The two peaks collapse into one peak if $f_N(\Theta_\mathrm{m})$, evaluated in the middle between both peaks, $\Theta_1 = -\Theta_2 = \Theta_\mathrm{m}$, is larger than $f_N(\boldsymbol{\theta}_1)$ evaluated at one of the peaks. We describe the peak locations by their angular translation $\Delta\Theta$ from their common centre, such that the centre of the peaks Θ_m has the angular distance $\Delta\Theta$ to both peaks and original plane-wave source locations $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ have distance $2\Delta\Theta$ to each other. Now, we again utilise the Legendre recurrence relation and get,

$$\begin{aligned}\frac{N+1}{2\pi(\cos(\Delta\Theta)-1)}&(P_{N+1}(\cos(\Delta\Theta)) - P_N(\cos(\Delta\Theta))) \\ &> \frac{(N+1)^2}{4\pi} + \frac{(N+1)(P_{N+1}(\cos(2\Delta\Theta) - P_N(\cos(2\Delta\Theta))))}{4\pi(\cos(2\Delta\Theta)-1)}\,,\end{aligned} \tag{2.114}$$

where the first part of the equation evaluates the function in the middle of both peaks and the second expression evaluates the function at the direction $\boldsymbol{\theta}_1$, such that the distance to peak one is zero (yielding $(N+1)^2/(4\pi)$) and the distance to the second peak is $2\Delta\Theta$. The transition where two peaks collapse to a single peak is found by using an equality relation instead of the inequality above, and simplifying,

$$\begin{aligned}\frac{2}{\cos(\Delta\Theta)-1}&(P_{N+1}(\cos(\Delta\Theta)) - P_N(\cos(\Delta\Theta))) - N - 1 \\ &- \frac{1}{\cos(2\Delta\Theta)-1}(P_{N+1}(\cos(2\Delta\Theta) - P_N(\cos(2\Delta\Theta)))) = 0\,.\end{aligned} \tag{2.115}$$

We numerically find a relation for $\Delta\Theta_\mathrm{min}$ by polynomial fitting,

$$\Delta\Theta_\mathrm{min} \approx \frac{-21}{N^2} + \frac{109}{N} + 1.6\,, \tag{2.116}$$

which creates an absolute error of less than $1.4°$ for $N \in [1, 20]$ and an error less than $0.7°$ for $N \in [1, 20] \setminus 2$. The minimum distance at which two SH basic beam peaks collapse into one, is calculated accordingly as $d_\mathrm{min} = 2\Delta\Theta_\mathrm{min}$, exhibiting twice the error margins.

Maximum number of extrema. In practical applications we need to consider not only maxima but extrema in general, as phase-shifted, "negative" peaks (minima) lead to intensity peaks just as the maxima. This complicates the derivations as it is not sufficient anymore to investigate the collapsing of the main lobes, while neglecting side lobes. However, we can describe an upper bound of extrema in an order N SH pattern, if we describe the number of main lobes that fit on the surface on the sphere, distributed regularly, with a minimum distance of d_min. For that, we calculate the surface segment of the unit sphere covered by a spherical segment with radius $\Delta\Theta_\mathrm{min}$,

$$\int_{-\Delta\Theta_\mathrm{min}}^{\Delta\Theta_\mathrm{min}} \int_{\pi/2-\Delta\Theta_\mathrm{min}}^{\pi/2+\Delta\Theta_\mathrm{min}} \sin\vartheta \; d\vartheta d\varphi = 4\Delta\Theta_\mathrm{min}\sin(\Delta\Theta_\mathrm{min})\,, \tag{2.117}$$

and then take the ratio of the whole surface of the unit sphere to the segment,

$$\frac{4\pi}{4\Delta\Theta_\mathrm{min}\sin(\Delta\Theta_\mathrm{min})} \approx \frac{\pi}{\left(\frac{-21}{N^2} + \frac{109}{N} + 1.6\right)\sin(\frac{-21}{N^2} + \frac{109}{N} + 1.6)}\,. \tag{2.118}$$

Order N	Upper Bound b	Monte-Carlo
1	2	2
2	4	4
3	8	8
4	13	13
5	19	18
6	26	23
7	34	31
8	44	38
9	54	46
10	66	56

Table 2.1 – Calculated upper bound b of the number of extrema found in a Monte-Carlo simulation using uniformly-distributed SH coefficients and $100\,(N+1)^2$ runs.

This ratio (the surface of the unit sphere divided by the surface covered by the spherical segment) describes, after rounding down to the next integer, an upper bound b for the number of extrema a pattern on the sphere created with SHs of order N can have,

$$b = \left\lfloor \frac{\pi}{\Delta\Theta_{\min} \sin(\Delta\Theta_{\min})} \right\rfloor . \tag{2.119}$$

Table 2.1 shows the upper bound b for orders $N \in [1, 10]$. Note that several assumptions make this result an upper bound and not an approximation of the real maximum number of extrema in an SH pattern: i) The sphere surface is divided by the spherical surface segment, neglecting the actual shape of the (circular) peaks and possible regular distributions of multiple circular peaks. ii) Side lobes of the maximum directivity SH beam patterns are neglected, assuming that a distribution of such beam peaks is a valid upper bound for the number of peaks. However, a $100\,(N+1)^2$-run Monte-Carlo simulation with random directivity, where SH coefficients are distributed uniformly between -1 and 1 and extrema are found by the New MUSIC algorithm, appears to obey the upper bound, cf. table 2.1.

2.6 Performance Comparison of REVEB-ESPRIT and New MUSIC

2.6.1 Evaluation Setup

Four scenarios are considered to compare the New MUSIC and REVEB-ESPRIT algorithms. Scenarios (i)-(iii) evaluate the performance for DOA estimation of broadband acoustic sources while scenario (iv) evaluates the performance for DOA estimation of individual reflections in a room impulse response. The first three scenarios contain (i) a simple free-field scenario, similar to the one in section 2.4.5, (ii) a free-field scenario including spatial aliasing, and (iii), a reverberant scenario employing a room simulation via the image method [AB79]. In all scenarios the maximum Ambisonics order is set to $N = 3$ and the SNR is varied via additive white Gaussian noise (AWGN) simulating the microphone self noise. Keeping the source distance to the array constant at 2 m, the source directions were drawn randomly from a 48-point 9-design [HS96]. The evaluation for the scenarios (i)-(iii) is carried out twice, using either two or six simultaneously estimated source directions. In each trial, the SNR is varied in 10 dB steps from 0 to 50 dB, where each SNR step is repeated $J = 100$ times using different AWGN instances. In scenarios (ii)-(iv) that include a microphone array simulation, the array is simulated as rigid-sphere array with a radius of $r = 4.2$ cm and the microphone layout of a 32-point 7-design [HS96], similar to the Eigenmike em32. In all scenarios, RMSEs of the great-circle distance

$$\Delta\boldsymbol{\theta}_q^{(j)} = \arccos\left\{\boldsymbol{\theta}_q^{(j)\mathrm{T}}\hat{\boldsymbol{\theta}}_q^{(j)}\right\} \tag{2.120}$$

are calculated for each source q and each trial j,

$$\mathrm{RMSE} = \sqrt{\sum_{j=1}^{J}\sum_{q=1}^{Q}\left|\Delta\boldsymbol{\theta}_q^{(j)}\right|^2/(JQ)}\,, \tag{2.121}$$

where $\boldsymbol{\theta}_q^{(j)}$ is the found DOA vector and $\hat{\boldsymbol{\theta}}_q^{(j)}$ is the true DOA vector.

For simulation of room impulse responses in the scenarios (iii) and (iv), the Spherical Microphone Impulse Response (SMIR) generator [JHTN12] was used. To be able to retrieve true DOAs, the SMIR generator was extended to additionally output reflection directions and corresponding timestamps. This extended version is made available online. [3]

(i) Simple free-field scenario. For the simple free-field scenario, Q unit-variance, WGN source signals $s_q[\tau]$ are transformed to the spherical harmonics domain as plane waves from direction $\boldsymbol{\theta}_q$, where AWGN noise $\nu[\tau]$ is added with a trial-dependent SNR,

$$s_n^m[\tau] = \left(\sum_{q=1}^{Q} Y_n^{m^*}(\boldsymbol{\theta}_q)\, s_q[\tau]\right) + \nu[\tau]\,. \tag{2.122}$$

3. `https://github.com/thomasdeppisch/SMIR-Generator`

5120 samples of $s_n^m[\tau]$ are created and directly fed into the REVEB-ESPRIT and New MUSIC algorithms.

(ii) Free-field scenario including spatial aliasing. Scenarios (ii) and (iii) are carried out in the frequency domain using block-processing and hence include a common pre-processing step, where all source signals are short-time Fourier transformed using a frame length of 512 samples, 50% overlap, Hann windowing and an FFT size of 1024 samples, yielding the following formalism for windowing of one frame of one source signal $s_q[\tau]$,

$$\tilde{s}_q[\tau] = w[\tau]\, s_q[\tau]\,. \tag{2.123}$$

After zero padding of $\tilde{s}_q[\tau]$ to $T = 1024$ samples,

$$\hat{s}_q[\tau] = \begin{cases} \tilde{s}_q[\tau], & \text{for } 0 < \tau < 512\,, \\ 0 & \text{for } 512 < \tau < 1024\,, \end{cases} \tag{2.124}$$

the signal is transformed to the frequency domain,

$$S_q[k] = \frac{1}{\sum_\tau w[\tau]} \sum_{\tau=0}^{T-1} \hat{s}_q[\tau] e^{-i\frac{2\pi k}{T}\tau}\,. \tag{2.125}$$

For further processing, we only consider the single-sided spectrum $\tilde{S}_q[k]$ (with $0 < k < 513$) and hence apply a factor of 2,

$$\tilde{S}_q[k] = \frac{2}{\sum_\tau w[\tau]} \sum_{\tau=0}^{T-1} \hat{s}_q[\tau] e^{-i\frac{2\pi k}{T}\tau}\,. \tag{2.126}$$

For the free-field scenario with spatial aliasing, the frequency-domain signals $\tilde{S}_q[k]$ are evaluated at bin κ corresponding to $kr = 2$ (and to $f = 2600$ Hz when assuming $c = 343$ m/s and $r = 4.2$ cm), before being transformed to the SHD. Here, scattering of a rigid-sphere microphone array of radius $r = 4.2$ cm is simulated via multiplication by $b_n(kr) = 4\pi i^{n+1}/((kr)^2 h_n^{(2)'}(kr))$ [ZF19, chapter 6],

$$\tilde{S}_n^m[\kappa] = b_n(kr) \sum_{q=1}^{Q} Y_n^m(\boldsymbol{\theta}_q)^* \tilde{S}_q[\kappa]\,. \tag{2.127}$$

Note that at this stage, the spherical harmonics are truncated at a high simulation order of $N_\mathrm{s} = 30$ to be able to account for spatial aliasing in a later step. Next, microphone signals $S_l[\kappa]$ are obtained via the inverse Spherical Fourier Transform (SFT),

$$S_l[\kappa] = \left(\sum_{n=0}^{N_s} \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}_l) \tilde{S}_n^m[\kappa] \right) + \tilde{\nu}[\kappa]\,, \tag{2.128}$$

where the $L = 32$ microphones are located at directions $\boldsymbol{\theta}_l$ distributed according to a 32-point 7-design [HS96] and AWGN $\tilde{\nu}[\kappa]$ is added after transforming it to the frequency domain (to account for the correct scaling) using eq. 2.126. Finally, the single-bin,

frequency-domain microphone signals $S_l[\kappa]$ are transformed back to the SHD at evaluation order $N = 3$, where Tikhonov-regularised mode-strength compensation [MDB06] is applied,

$$S_n^m[\kappa] = \frac{b_n^*(kr)}{|b_n(kr)|^2 + 10^{-2}} \frac{4\pi}{L} \sum_{l=0}^{L} Y_n^m(\boldsymbol{\theta}_l)^* \, S_l[\kappa] \,. \tag{2.129}$$

This procedure is performed for 10 frames which are then fed into the REVEB-ESPRIT and the New MUSIC algorithms for evaluation.

(iii) Reverberant scenario. For the reverberant scenario, spherical microphone array room impulse responses of 2048 samples length (at $f_\mathrm{s} = 48$ kHz) are generated with the Spherical Microphone Impulse Response (SMIR) generator [JHTN12], which accounts for rigid-sphere scattering and spatial aliasing (by using a high simulation order $N_\mathrm{s} = 30$ as in the previous scenario) and uses the image method [AB79]. The microphone array is placed at $[4.103 \text{ m}, 3.471 \text{ m}, 2.912 \text{ m}]$ in a $8 \times 7 \times 6$ m^3 room with reverberation time $T_{60} = 0.6$ s. The retrieved time-domain impulse responses for each of the $L = 32$ microphones are convolved with their corresponding source signals and AWGN is added,

$$s_l[\tau] = \left(\sum_{q=0}^{Q} h_{q,l}[\tau] * s_q[\tau] \right) + \nu[\tau] \,. \tag{2.130}$$

The time-domain microphone signals are then transformed to the SHD,

$$s_n^m[\tau] = \frac{4\pi}{L} \sum_{l=0}^{L} Y_n^m(\boldsymbol{\theta}_l)^* \, s_l[\tau] \,, \tag{2.131}$$

before frequency-domain block-processing is applied for all $(N+1)^2$ channels as in the previous scenario. Here, the signals are not only evaluated at $kr = 2$ but at a range of frequencies between 1300 - 5200 Hz ($kr \in [1, 4]$), where the lower bound is chosen to include highest-order SH components, and the upper bound is chosen to avoid spatial aliasing. The zero-padded and windowed SHD signals $\hat{s}_n^m[\tau]$ are transformed to the frequency-domain and are mode-strength compensated using Thikonov-regularised filters,

$$S_n^m[k] = \frac{b_n^*(kr)}{|b_n(kr)|^2 + 10^{-2}} \frac{2}{\sum_\tau w[\tau]} \sum_{\tau=0}^{T-1} \hat{s}_n^m[\tau] e^{-i\frac{2\pi k}{T}\tau} \,. \tag{2.132}$$

The procedure is again performed for 10 frames, after which the signals $S_n^m[k]$ for all frames and all evaluated frequencies are fed into the REVEB-ESPRIT and New MUSIC algorithms.

(iv) DOA estimation with ARIRs. The processing in this scenario is similar to the processing for the reverberant scenario, with the difference that the block processing is performed in the time domain and SNRs are evaluated in 10 dB steps from 30 to 80 dB.

As this scenario analyses room impulse responses, no source signals are convolved with the simulated RIR,

$$s_l[\tau] = h_l[\tau] + \nu[\tau]\,. \tag{2.133}$$

The impulse responses for all microphones are again transformed to the SHD,

$$\tilde{s}_n^m[\tau] = \frac{4\pi}{L} \sum_{l=0}^{L} Y_n^m(\boldsymbol{\theta}_l)^*\, s_l[\tau]\,, \tag{2.134}$$

where the 2048-sample RIRs are first zero-padded to $T = 4096$ samples, and then transformed to the frequency domain and mode-strength compensated,

$$S_n^m[k] = \frac{b_n^*(kr)}{|b_n(kr)|^2 + 10^{-2}} \sum_{\tau=0}^{T-1} \hat{s}_n^m[\tau] e^{-i\frac{2\pi k}{T}\tau}\,. \tag{2.135}$$

After full-bandwidth mode-strength compensation, the RIRs are transformed back to the time domain, where they are band limited by front and reverse direction filtering to create zero phase distortion. For this, a second-order bandpass filter $\mathcal{B}_{\mathrm{bp}}\{\cdot\}$ with cut-off frequencies $f_{\mathrm{hp}} = 1300$ Hz ($kr = 1$) and $f_{\mathrm{lp}} = 5200$ Hz ($kr = 4$) is used,

$$s_n^m[\tau] = \mathcal{B}_{\mathrm{bp}} \left\{ \frac{1}{T} \sum_{k=0}^{T-1} S_n^m[k] e^{i\frac{2\pi k}{T}\tau} \right\}\,. \tag{2.136}$$

The time-domain signal $s_n^m[\tau]$ is then fed into the REVEB-ESPRIT and New MUSIC algorithms in frames of 21 samples with 20 samples overlap. A DOA estimation is only performed if a true reflection is present in the current frame. RMSEs are divided by the number of found reflections in a trial and frame $Q^{(j,p)}$, the number of active frames P (where a DOA estimation was carried out) and the number of trials J,

$$\mathrm{RMSE} = \sqrt{\sum_{j=1}^{J} \sum_{q=1}^{Q} \sum_{p=1}^{P} \left| \Delta\boldsymbol{\theta}_q^{(j,p)} \right|^2 / (JQ^{(j,p)}P)}\,. \tag{2.137}$$

Note that in contrast to the other scenarios, where the number of sources Q is static and fed into the algorithms, in this scenario, the number of reflections (defining the subspace dimension) is calculated for each frame using the SORTE algorithm (cf. section 2.1). However, for calculation of the RMSE, only $\min(Q^{(j,p)}, \hat{Q}^{(j,p)})$ DOAs are considered, where $\hat{Q}^{(j,p)}$ are the true DOAs. This means that errors resulting from a wrong estimation of the number of subspace components are neglected, e.g. if the algorithm finds 5 DOAs in a frame which only has 3 true DOAs, the RMSE will be calculated only for the 3 found DOAs with respect to the 3 true DOAs closest to them.

In view of DOA estimation of reflections in an Ambisonic room impulse response (ARIR), as done by the higher-order Ambisonic Spatial Decomposition Method (HO-ASDM, introduced in section 3.1), a performance comparison of the New MUSIC and REVEB-ESPRIT algorithms to the pseudo-intensity vector (PIV) method employed in the Ambisonic spatial decomposition method (ASDM, see the introduction of

section 3.1, [TPKL13] for the original SDM and [FZ16] for the use in an Ambisonic context) is of particular interest. Hence, scenario (iv) is repeated for a maximum SH order of $N = 1$, where the lower frequency bound of the bandpass filter is set to $f_{\mathrm{hp}} = 100$ Hz and the PIV method is added to the comparison. Here, two cases are investigated: One where the subspace based algorithms (New MUSIC and REVEB-ESPRIT) are simultaneously estimating up to $Q = 2$ DOAs, and one where the number of estimated DOAs is fixed to $Q = 1$, which is the maximum number of DOAs the PIV method can estimate.

Parametrisation. In all four scenarios the New MUSIC algorithm is set up identically. It is initialised using $20 \times Q$ points distributed according to a spherical Fibonacchi grid. The maximum number of iterations is set to $k_{\max} = 40$ and the step size is calculated adaptively as $1 - ((k-1)/k_{\max})$. All scenarios are using the objective for finding roots in the noise subspace, i.e. the Newton update is calculated as

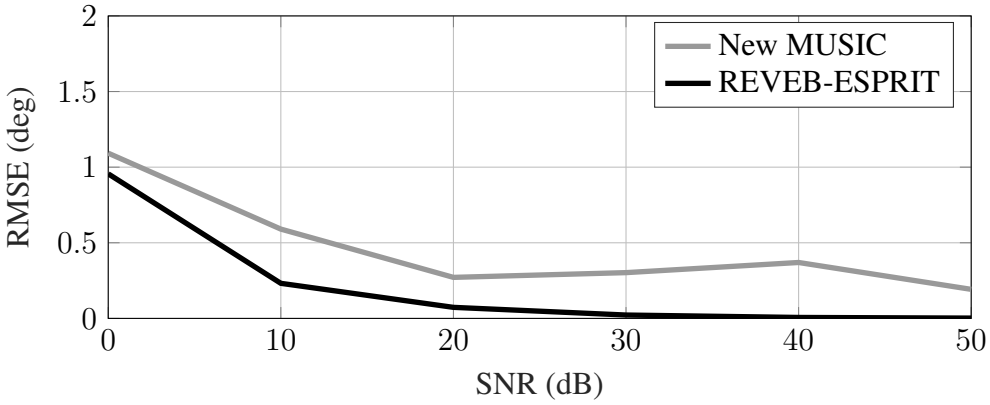
$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \left(1 - \frac{k-1}{k_{\max}}\right) \boldsymbol{g}_{f_{\mathrm{n}}^2} \frac{\boldsymbol{y}(\boldsymbol{\theta}_k)^{\mathrm{T}} \boldsymbol{U}_{\mathrm{n}} \boldsymbol{U}_{\mathrm{n}}^{\mathrm{T}} \boldsymbol{y}(\boldsymbol{\theta}_k)}{\boldsymbol{g}_{f_{\mathrm{n}}^2}^{\mathrm{T}} \boldsymbol{g}_{f_{\mathrm{n}}^2}} \,. \tag{2.138}$$

A convergence criterion terminates the algorithm if the mean of all Newton steps of the current iteration is smaller than 10^{-3}, or if it is larger than 1.5 times the mean of all Newton steps of the previous iteration while being smaller than 10^{-1}. After the algorithm converged or the maximum number of iterations is reached, points that did not converge (their corresponding last Newton step being larger than 2 times the median Newton step of all points while being smaller larger than 10^{-2}) are excluded from the result. Furthermore, points at which the Hessian is not positive-definite are also excluded. Finally, the remaining points are divided into Q clusters using the k-means algorithm and the means of the points in each cluster are the resulting Q DOAs. These rather sophisticated convergence criteria are chosen due to the versatile test setup, under which the algorithm is exposed to a multitude of different testing conditions. In most applications, static convergence thresholds, set after an estimation of the SNR and with regard to the specific environment, should be sufficient.

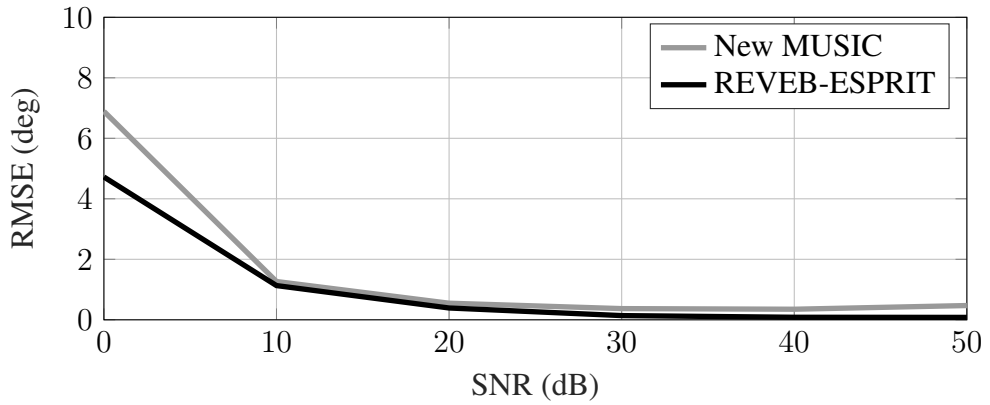
The REVEB-ESPRIT algorithm uses the ad-hoc Schur decomposition (cf. section 2.4.1) for simultaneous triangularisation. In scenario (iv), the SORTE algorithm was used with $\epsilon_1 = 0.99$ and $\epsilon_2 = 0.1$ (cf. section 2.1).

2.6.2 Results

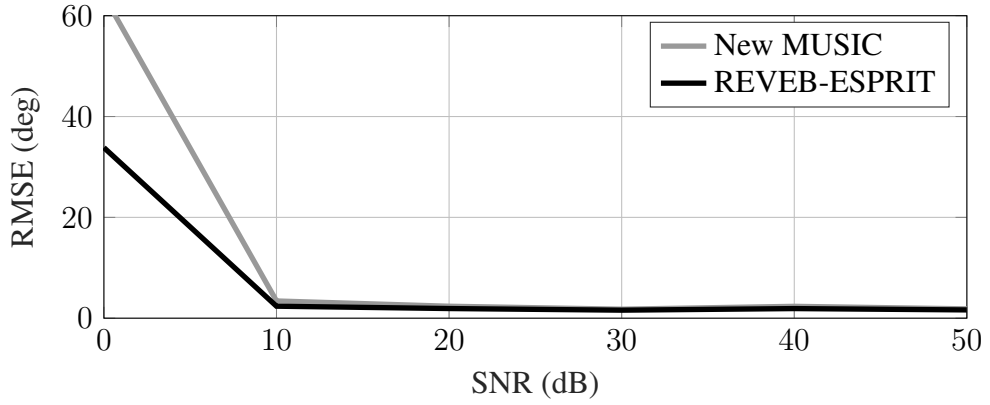
Figure 2.5 shows RMSEs for scenarios (i)-(iii) and $Q = 2$ simultaneously estimated DOAs. In most of the cases that were analysed, New MUSIC and REVEB-ESPRIT perform similar. In the simple free-field scenario, REVEB-ESPRIT achieves between $0.2^\circ - 0.4^\circ$ lower RMSEs than New MUSIC. The only larger differences occur for a low SNR of 0 dB in scenarios (ii) and (iii), where the RMSEs of REVEB-ESPRIT are lower than the RMSEs of New MUSIC by 2° and 15°, respectively. From 10 dB SNR upwards, both algorithms perform very similar again, with REVEB-ESPRIT having an advantage of up to 0.4° (for scenario (ii), 50 dB SNR).



(a) Scenario (i): Simple free-field scenario.



(b) Scenario (ii): Free-field scenario including spatial aliasing.

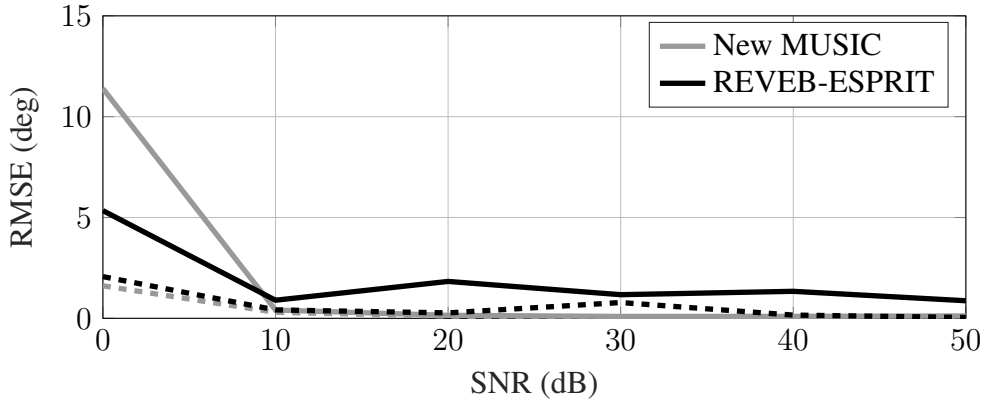


(c) Scenario (iii): Reverberant scenario.

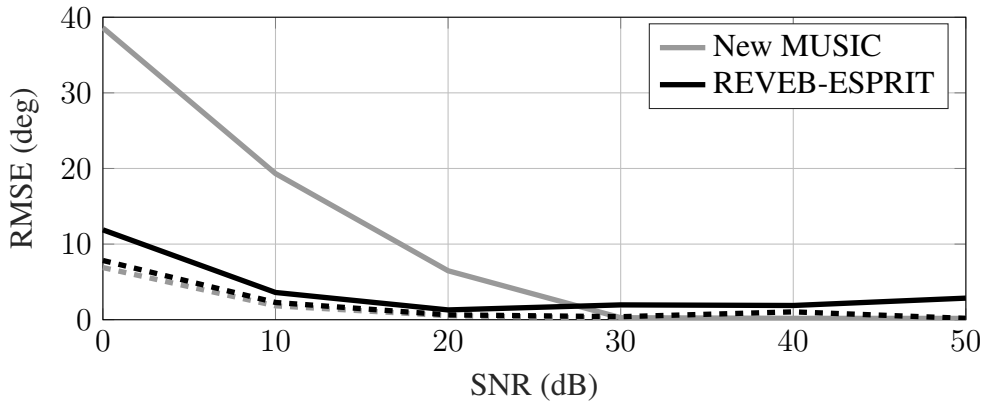
Figure 2.5 – RMSEs for simultaneous DOA estimation of $Q = 2$ sources for the New MUSIC and REVEB-ESPRIT algorithms.

The results for simultaneous estimation of $Q = 6$ DOAs in figure 2.6 are quite different. Although REVEB-ESPRIT again performs better for low SNRs in scenarios (i) and (ii), New MUSIC achieves lower RMSEs for SNRs from 10 dB upwards in scenario (i) (by up to $1.7°$) and from 30 dB upwards in scenario (ii) (by up to $2.7°$). In the reverberant condition, scenario (iii), the result is flipped: Here, New MUSIC achieves a lower RMSE by $5°$ for a low SNR of 0 dB, while for higher SNRs REVEB-ESPRIT performs clearly better with lower SNRs by between $14° - 24°$. In addition to RMSEs, the figure shows dashed-line RMSEs for both algorithms that retrieve the average of all trials from the median of the errors of all six simultaneously estimated source directions. In all tested scenarios, and in scenario (iii) that exhibits large RMSEs in particular, the median-based RMSEs are distinctly lower (by up to $31°$ for New MUSIC and up to $16°$ for REVEB-ESPRIT) than their mean-based counterparts. Hence, the large occurring RMSEs are due to few or single estimated directions that miss the ground truth by far, while most of the estimations are rather accurate.

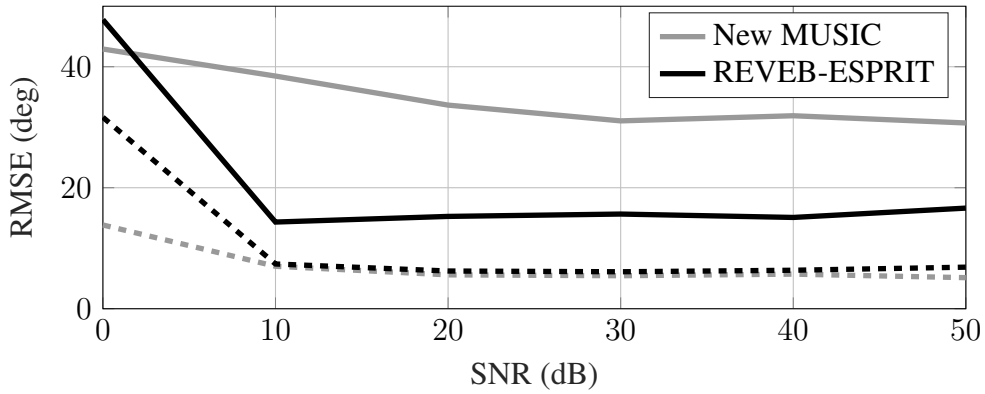
Figure 2.7 shows the results for scenario (iv) that evaluates the estimation accuracy of DOAs of individual reflections in a simulated ARIR. In this task, New MUSIC and REVEB-ESPRIT perform similarly, which is not surprising as in most trials a maximum of 3 simultaneously arriving reflections were found by the SORTE algorithm and hence the scenario is comparable to scenario (iii) for $Q = 2$, figure 2.5(c). In figure 2.7(b) and 2.7(c) the PIV method is added to the comparison and the maximum SH order is set to $N = 1$. Regardless of whether the number of DOAs to be estimated for REVEB-ESPRIT and New MUSIC is fixed to $Q = 1$ (figure 2.7(c)), or is allowed vary between one and two depending on the output of the SORTE algorithm (figure 2.7(b)), both algorithms outperform the PIV method (which in both cases can only estimated a single DOA) for SNRs larger than 50 dB, by up to $4.5°$ and up to $2.9°$, respectively. For SNRs of less than 50 dB, the PIV method has the advantage of up to $5.7°$ for an SNR of 30 dB.



(a) Scenario (i): Simple free-field scenario.

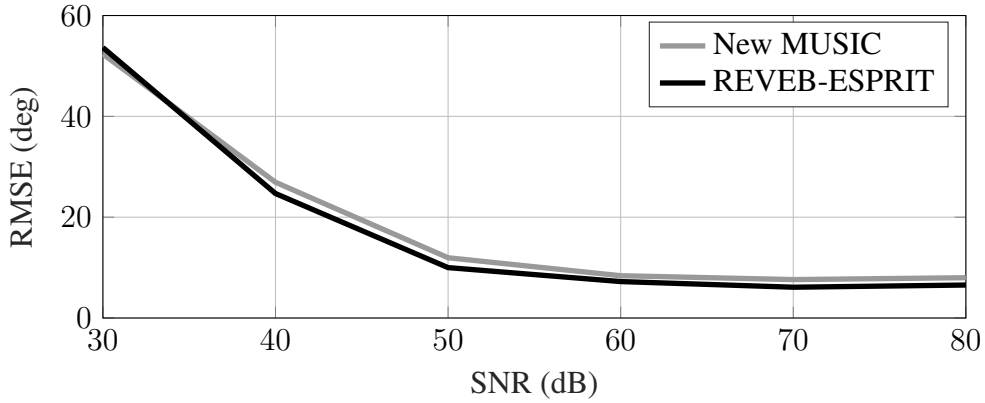


(b) Scenario (ii): Free-field scenario including spatial aliasing.

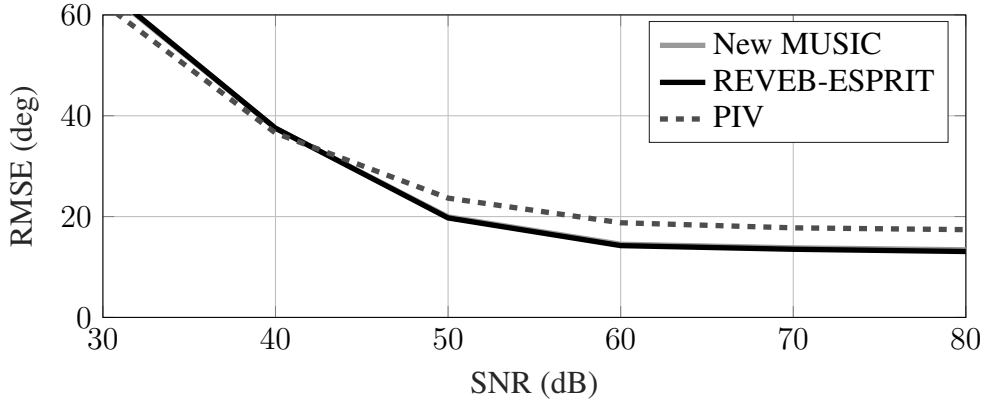


(c) Scenario (iii): Reverberant scenario.

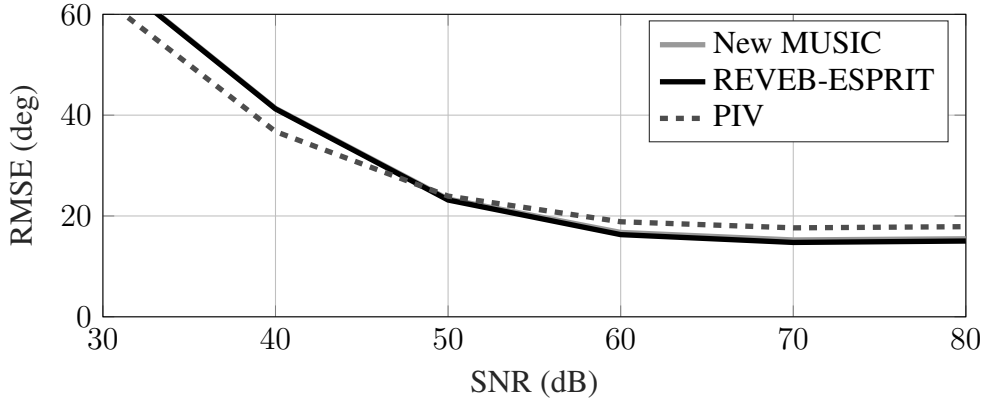
Figure 2.6 – Solid lines show RMSEs for simultaneous DOA estimation of $Q = 6$ sources for the New MUSIC and REVEB-ESPRIT algorithms. Dashed lines show RMSEs based on the median error of all six simultaneously estimated directions for both algorithms.



(a) Maximum SH order $N = 3$.



(b) Maximum SH order $N = 1$.



(c) Maximum SH order $N = 1$, number of DOAs for simultaneous estimation fixed to $Q = 1$.

Figure 2.7 – Scenario (iv): RMSEs for simultaneous estimation of reflection directions in an ARIR for the New MUSIC and REVEB-ESPRIT algorithms. In (b) and (c) the maximum SH order is set to $N = 1$ and the PIV method is added to the comparison.

Chapter 3

Applications

3.1 Higher-Order Ambisonic Spatial Decomposition Method (HO-ASDM)

Convolution of a signal with a measured room impulse response is a widely applied technique for reproduction of the acoustic characteristics of a room. This concept was extended to multichannel loudspeaker arrays and directional impulse responses by Merimaa and Pulkki with the spatial impulse response rendering (SIRR) method [MP05, PM06]. In each frequency band, the method splits the input RIR into a direct part and a diffuse part and reproduces the direct part directionally sharpened via vector-base amplitude panning (VBAP) [Pul97] after DOA estimation via the pseudo-intensity vector (PIV). The diffuse part is reproduced after a decorrelation process. The spatial decomposition method (SDM) is a similar method, but in contrast to SIRR applies broadband processing and does not explicitly split the signal into a direct and a diffuse part [TPKL13]. Although originally proposed for reproduction via VBAP, the method was also found useful in the SHD framework, called Ambisonic SDM (ASDM), see e.g. [FZ16][ZF19, chapter 5]. In the SHD of higher-order Ambisonics, the method can be seen as a directional upmixing procedure that renders higher-order RIRs from measured first- and zeroth-order RIRs. So far, the procedure was based on estimating a single direction by PIV or magnitude sensor response. The directional enhancement of first-order Ambisonic RIRs by the 2+2 directional signal estimator (2DSE2) [GZ20] is a similar method that however manages to simultaneously estimate two directions in first-order Ambisonic RIRs and is based on the high angular resolution plane-wave expansion (HARPEX) method [BB10, BB11]. Recently, SIRR was extended to higher-order SH input, called higher-order SIRR (HO-SIRR), by applying sector-based processing [MPP$^+$20] that is also capable of detecting multiple simultaneous directions. Clearly, HO-SIRR is the first method to exploit higher-order SH information in ARIRs. The aim of this section is to introduce HO-ASDM as a novel counterpart to HO-SIRR.

Similarly to HO-SIRR for SIRR, the higher-order ASDM (HO-ASDM) extends the ASDM to higher-order SH input. HO-ASDM uses a VEB-ESPRIT algorithm for

multiple DOA estimation, which is able to simultaneously estimate $N^2 + \lfloor 4/3N \rfloor$ directions (in an order N input ARIR) and has shown to yield superior performance when compared to the PIV method (cf. section 2.6). The subspace-based DOA estimation procedure controls a multi-direction beamformer that extracts directional signals for the subsequent higher-order re-encoding (upmixing). As directional signals with stable directions do not include diffuse sounds, a residuum stemming from the noise subspace is added to the upmixed higher-order ARIR. The resulting ARIR is equalised based on the squared spectral envelope. Coding and multidirectional parameterisation of Ambisonic sound scenes (COMPASS) [PTP18] is a similar algorithm for analysis and synthesis of Ambisonic sound scene recordings that applies sub-band processing. COMPASS uses MUSIC for the DOA estimation, which has been designed for SHD signals and but which has not been tested within the context of HO-SIRR to upmix ARIRs. Another technique for multiple-direction retrieval is compressed sensing, which was used in an SHD upmixing method by Wabnitz et al. [WEMJ11]. Similar to the non-parametric MUSIC algorithm that applies a grid-search to estimate directions, the method employs a large plane-wave dictionary in the optimisation process.

3.1.1 Algorithm

Subspace decomposition and smoothing of the number of subspace components. As shown in section 2.1, an estimation of the signal and noise eigenvectors in matrices $\boldsymbol{U}_\mathrm{s}$ and $\boldsymbol{U}_\mathrm{n}$ is accomplished either by eigendecomposition of the sample covariance or by singular value decomposition of the data matrix. The dimension of the signal subspace, i.e. the number of directional components to be extracted from the original ARIR is determined using the SORTE algorithm (cf. section 2.1) as it is also suggested in COMPASS [PTP18]. Strong fluctuations in the number of observable directional components Q in $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_Q]$ are found to create noticeable switching artefacts that strongly distort the upmixed ARIR. Hence, a temporal moving-median smoothing of the number of directional components Q is applied before the actual DOA estimation and upmixing process.

DOA estimation and upmixing. A frame-wise multiple-DOA estimation is achieved by using the (extended) VEB-ESPRIT algorithm (cf. section 2.4.1) with a pre-specified number of directional components Q derived from the preceding smoothing process. Upmixing to an arbitrary Ambisonic order $\tilde{N}$ is achieved by a multi-direction beamformer that assumes a (mode-strength compensated) Ambisonic RIR $\boldsymbol{h}_N(t)$. Simplistically, a maximum-directivity beamformer would just evaluate the input SH pattern $\boldsymbol{h}_N(t)$ for each input sample at the found VEB-ESPRIT DOAs $\boldsymbol{\Theta}(t)$, and upmixing re-encodes the resulting directional signals in an arbitrary order $\tilde{N}$,

$$\boldsymbol{h}_{\tilde{N}}(t) = \boldsymbol{Y}_{\tilde{N}}(\boldsymbol{\Theta}(t))\, \boldsymbol{Y}_N^T(\boldsymbol{\Theta}(t))\, \boldsymbol{h}_N(t)\,.$$

Signal cross-talk across the outputs of the maximum-directivity beamformers $\boldsymbol{Y}_N^T(\boldsymbol{\Theta}(t))$ is removed by a crosstalk canceller $\left(\boldsymbol{Y}_N^T(\boldsymbol{\Theta}(t))\boldsymbol{Y}_N(\boldsymbol{\Theta}(t))\right)^{-1}$, yielding the desired direc-

tional upmixing procedure

$$\boldsymbol{h}_{\tilde{N}}(t) = \boldsymbol{Y}_{\tilde{N}}(\boldsymbol{\Theta}(t)) \left(\boldsymbol{Y}_N^T(\boldsymbol{\Theta}(t))\boldsymbol{Y}_N(\boldsymbol{\Theta}(t))\right)^{-1} \boldsymbol{Y}_N^T(\boldsymbol{\Theta}(t))\, \boldsymbol{h}_N(t)\,. \tag{3.1}$$

In cases when some of the DOAs get close to each other, the cross-talk cancelling inverse $\left(\boldsymbol{Y}_N^T(\boldsymbol{\Theta}(t))\boldsymbol{Y}_N(\boldsymbol{\Theta}(t))\right)^{-1}$ is ill-conditioned. Hence, a regularised pseudoinverse is employed by discarding small singular values $\sigma_i < \epsilon$ and their corresponding singular vectors, denoted by the diminished matrices $\tilde{\boldsymbol{U}}$, $\tilde{\boldsymbol{S}}$ and $\tilde{\boldsymbol{V}}$,

$$\boldsymbol{Y}_N = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}} \approx \tilde{\boldsymbol{U}}\tilde{\boldsymbol{S}}\tilde{\boldsymbol{V}}^{\mathrm{T}}\,, \tag{3.2}$$

where ϵ is a small positive constant, e.g. $\epsilon = 10^{-6}$. To also capture the spectral characteristics of the ARIR that are not directional or diffuse, a residual signal $\boldsymbol{r}_N(t)$ is calculated by projecting the ARIR $\boldsymbol{h}_N(t)$ onto the noise subspace (i.e. the column space of the matrix $\boldsymbol{U}_{\mathrm{n}}$ holding the noise eigenvectors),

$$\boldsymbol{r}_N(t) = \boldsymbol{U}_{\mathrm{n}}\boldsymbol{U}_{\mathrm{n}}^{\mathrm{T}}\boldsymbol{h}_N(t) = (\boldsymbol{I} - \boldsymbol{U}_{\mathrm{s}}\boldsymbol{U}_{\mathrm{s}}^{\mathrm{T}})\boldsymbol{h}_N(t)\,, \tag{3.3}$$

where the noise eigenvectors are obtained by QR-decomposition of the encoding matrix,

$$\boldsymbol{Y}_N = \boldsymbol{Q}\boldsymbol{R} = \begin{bmatrix}\boldsymbol{U}_{\mathrm{s}} & \boldsymbol{U}_{\mathrm{n}}\end{bmatrix} \boldsymbol{R}\,. \tag{3.4}$$

Equivalently, the residuum can be obtained by using the pseudoinverse as done in [PTP18],

$$\boldsymbol{r}_N(t) = (\boldsymbol{I} - \boldsymbol{Y}_N\boldsymbol{Y}_N^{\dagger})\boldsymbol{h}_N(t)\,. \tag{3.5}$$

By renormalisation of the directional components with their new Ambisonics order $\tilde{N}$ using a suitable factor, they can be matched in level to the unmodified residual, and the upmixed ARIR is expressed as sum

$$\boldsymbol{h}_{\tilde{N}}(t) = \sqrt{\frac{(N+1)^2}{(\tilde{N}+1)^2}}\, \boldsymbol{Y}_{\tilde{N}}(\boldsymbol{\Theta}(t))\, \tilde{\boldsymbol{V}}\tilde{\boldsymbol{S}}^{-1}\tilde{\boldsymbol{U}}^{\mathrm{T}}\, \boldsymbol{h}_N(t) + \boldsymbol{r}_N(t)\,. \tag{3.6}$$

Spectral equalisation of the squared signal envelope. To equalise increased spectral brightness that is caused by directional fluctuations of the estimated DOAs [TPKL13, FZ16] that amplitude modulate the typical longer reverberation of lower frequencies to higher frequencies, the third-octave envelope of the upmixed ARIR $\tilde{h}_n^m(t,b)$ is equalised with regard to the third-octave envelope of the original ARIR $h_{n'}^{m'}(t,b)$,

$$\tilde{h}_{n,\mathrm{eq}}^m(t,b) = \tilde{h}_n^m(t,b) \sqrt{\frac{1/(N+1)^2 \sum_{n'=0}^{N} \sum_{m'=-n'}^{n'} \mathcal{E}\{|h_{n'}^{m'}(t,b)|^2\}}{1/(\tilde{N}+1)^2 \sum_{n=0}^{\tilde{N}} \sum_{m=-n}^{n} \mathcal{E}\{|\tilde{h}_n^m(t,b)|^2\}}}\,, \tag{3.7}$$

where $\mathcal{E}\{|h_n^m(t,b)|^2\}$ denotes a short-term estimator of the squared signal envelope at the time t and frequency band b. A similar approach was found to be effective in context of the ASDM, also denoted as recovery of the spectral decay [ZF19, chapter 5].

3.1.2 Listening Experiment

Experiment setup. In some cases, despite the spectral equalisation described above, the upmixed ARIRs exhibit audible artefacts, especially if the input ARIR is only a first-order response. The listening experiment presented below systematically investigates the perceived similarity of six different upmixed ARIRs (denoted as conditions) to a reference in six trials with different RIRs. The reference condition in every trial is a fourth-order ARIR recorded with an Eigenmike EM32 microphone array. The conditions under test contain the same ARIR in first order and several upmixed ARIRs using HO-ASDM and ASDM, see table 3.1 for a description and appendix D for a description of the equalisation methods. On all input ARIRs, radial filtering is done using the responses described in [LZ15][ZF19, chapter 6] with 15 dB noise gain and without max-$\boldsymbol{r}_\mathrm{E}$ side-lobe suppression [1]. The ARIRs are decoded to a 7+4+1 loudspeaker array at the IEM production studio (*Produktionsstudio*) using the AllRADecoder plug-in from the IEM plug-in suite [2] and are played back directly (without convolution) to allow for a detailed evaluation of artefacts. Due to the short response of the directly played-back ARIRs, this experiment does not facilitate an evaluation of the diffuseness or envelopment created by the respective upmixing technique by design.

HO-ASDM used a 21-sample, Hann-windowed block-wise processing for DOA estimation to facilitate the maximum number of simultaneously detectable DOAs and a 49-sample (≈ 1 ms) moving median filtering for smoothing of the number of subspace components, while both ASDM conditions used 14-sample Hann-windowed block-processing. For the HO-ASDM, the SORTE algorithm estimates the number of subspace components using $\epsilon_1 = 0.95$ and $\epsilon_2 = 0.01$, cf. section 2.1. The input RIRs are zero-phase bandpass-filtered with the input-order-dependent lower cut-off frequency set to the cut-on frequency of the corresponding highest-order radial filter $f_\mathrm{l} = \{150, 950, 2000, 3150\}$ Hz to allow for full-order DOA estimation and an upper cut-off frequency of $f_\mathrm{u} = 5.2$ kHz (corresponding to $kr = 4$ for the EM32) to avoid spatial aliasing.

The experiment was conducted using an open-source MUSHRA tool [3] as OSC control for Reaper [4] in the IEM production studio. All conditions and trials were presented in randomised order and the participants were allowed to listen to the conditions as often as desired. The test design is based on the multiple stimulus comparison with hidden reference and anchor (MUSHRA) paradigm. It is similar to the MUSHRA paradigm in the sense that multiple conditions are compared to a reference by multiple subjects in a repeated (within-subjects) manner, but differs in the way that it neither employs a hidden reference nor a dedicated anchor. However, it is assumed that the `o3-o4` condition serves as a quasi-reference and the `ASDM-PO` condition as a quasi-anchor, due to the strong similarity to the reference and the strong artefacts, respectively.

1. Responses including max-$\boldsymbol{r}_\mathrm{E}$ weights are available here: `https://phaidra.kug.ac.at/detail/o:69296`
2. `https://plugins.iem.at/`
3. `https://git.iem.at/rudrich/mushra`, available after login.
4. `http://reaper.fm/`
5. `https://phaidra.kug.ac.at/view/o:104385`
6. `https://zenodo.org/record/3477602`

trials	
Ligeti13	Measured at Ligeti hall ($T_{60} = 1.4$ s), Graz, with an EM32, $0°$ AOA, loudspeaker pointing away from receiver, available online [5].
Ligeti06	Measured at Ligeti hall ($T_{60} = 1.4$ s), Graz, with an EM32, $0°$ AOA, loudspeaker pointing towards the receiver, available online[5].
M	Measured at St. Paul's concert hall ($T_{60} = 2.1$ s), Huddersfield, with an EM32, $0°$ AOA, taken from the 3D-MARCo library [6].
Ligeti13 +70°	Same as in Ligeti13, rotated by 70°.
Ligeti06 +70°	Same as in Ligeti06, rotated by 70°.
M +70°	Same as in M, rotated by 70°.
conditions	
`reference`	Fourth-order ARIR, no upmixing.
`o1`	First-order ARIR, no upmixing.
`o1-o4`	Upmix from first to fourth order using HO-ASDM.
`o2-o4`	Upmix from second to fourth order using HO-ASDM.
`o3-o4`	Upmix from third to fourth order using HO-ASDM.
`ASDM-PO`	Upmix from first to fourth order using ASDM, per-order equalisation (cf. appendix D).
`ASDM-M`	Upmix from first to fourth order using ASDM, equalisation of the mean RMS value (cf. appendix D).

Table 3.1 – Description of the ARIRs for all trials and conditions used in the experiment.

Evaluation and results. 11 trained listeners with previous experience in listening experiments between 24 and 47 years with an average of 32 years took part in the experiment. On average the experiment took 13.5 minutes. Due to inherent violations in MUSHRA-like paradigms of several assumptions required for ANOVA-type methods, such as interval scale, normality, equal variances and independence (see [MDM18] for a discussion), which are also not mitigated by the specific design of this experiment, the non-parametric Friedman test is chosen to test if there are significant differences between the tested conditions. An independent Friedman test was conducted for each trial and for each trial the null hypothesis (H_0: *The medians of the data are equal for all conditions*) was rejected with a significance level of $\alpha = 0.05$. Hence, a pairwise Wilcoxon signed-rank test was conducted as post-hoc measure to test for significant differences between each pair of conditions. To account for the inflation of the type I error, the resulting p-values were corrected by the Bonferroni-Holm method [Hol79]. To verify whether the results are meaningful, effect sizes were calculated using the matched-pairs rank biserial correlation coefficient r_c [KRM18, chapter 22]. An extensive overview over the results of all pairwise tests are listed in appendix D. The pairwise Wilcoxon signed-rank test yielded significant differences between all tested pairs except for the pair `o1, o1-o4` in three of the six trials and the pair `o2-o4, o3-o4` in the second trial (Ligeti06). The matched-pairs rank biserial correlation coefficient is larger than 0.9 in all but one significant cases (which has $r_c = 0.77$), showing a large effect size. Figure 3.1 shows the medians and 95% confidence intervals for all conditions and trials.

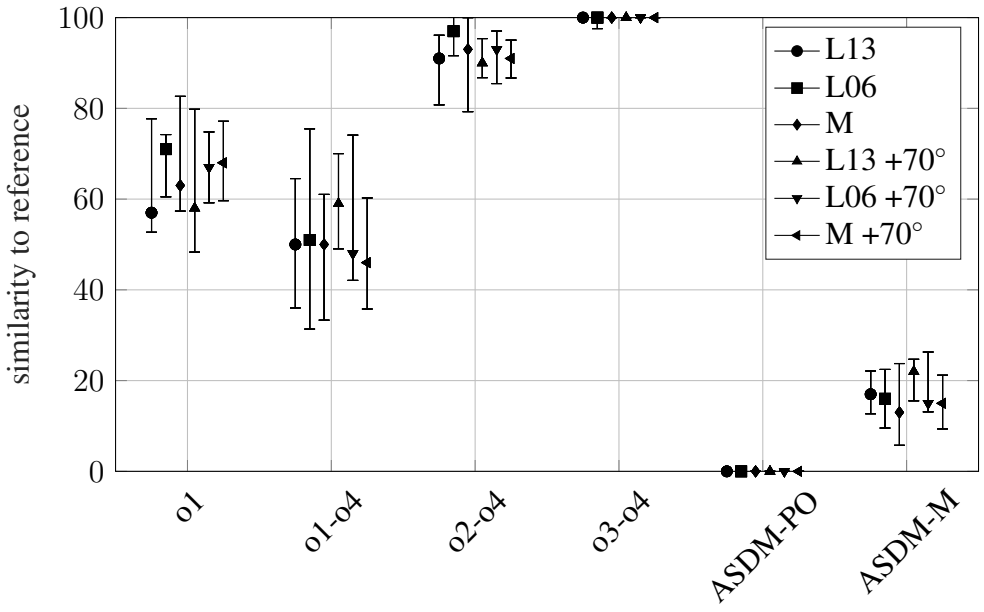


Figure 3.1 – Medians and 95% confidence intervals for all conditions and trials.

Assessment of the results. The above results resemble subjective experiences of the author and of individual participants (gathered via an informal interview directly after the listening experiment) closely: For stimuli of type `o2-o4` and `o3-o4` the difference to the reference is perceived as a subtle difference in timbre, while artefects are hardly ever noticeable. The subtle timbre difference (as a main difference) only turned out significant in three of six cases. In the environment of the test, the first-order impulse response in stimulus `o1` only differed from the reference in timbre. The timbre difference is presumably caused by the order-dependent radial filtering. The upmix from first to fourth order, `o1-o4`, has a similar timbre as `o1` but exhibits slight artefacts. Again, the slight artefacts were not significant in all trials. It is concluded that both ASDM-based stimuli are rated inferior mainly because of their strong artefacts that dominate over any perception of timbre. Despite the clearly inferior ratings of both ASDM conditions, overall spectral equalisation (`ASDM-M`) proved to be superior to the per-order equalisation (`ASDM-PO`) in avoiding artefacts when using ASDM.

The listening experiment revealed advantages of the HO-ASDM over the ASDM possibly regarding preservation of sound colour and reduction of artefacts for all tested input orders. Despite the fact that listeners could also have based their answers on properties like spatial envelopment or diffuseness, it is found from informal interviews after the tasks that their main focus lied on artefacts and changes in timbre. In the given setting, timbral changes appeared to yield more pronounced distinction than spatial attributes. Hence, future investigations should consider attributes like diffuseness, envelopment and size of the sweet area to rate the actual improvement made by the upmixing procedure, maybe in a larger listening environment. Especially for the pair `o1, o1-o4` significant differences would be expected in such a test design.

3.2 Analysis and Synthesis of Directivity Patterns in Peak Value and Tangential Curvature

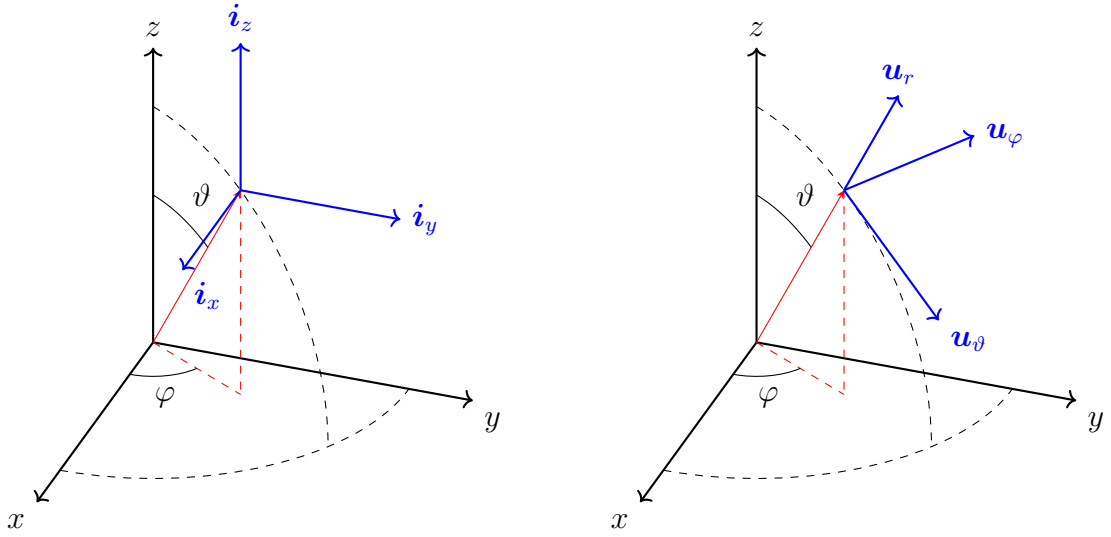
When measuring directivity patterns of sound sources such as musical instruments by means of a surrounding microphone array, an ongoing problem in research remains in the acoustic centreing of the sound source [DZ10, BPVR11, SV15, Pol15, Zag19]. Acoustic centreing of the source is a non-trivial task as the acoustic centre of a sound source is frequency dependent and not generally known. When the measured sound source is not properly centred, the resulting directivity pattern exhibits an energy shift towards higher-order patterns, resulting in a more complex directivity pattern than the one of the originally measured source. As the acoustic centreing problem is essentially a phase-retrieval problem, common solutions lie in finding an optimal phase to shift the energy of the directivity pattern back towards lower orders [Zag19] or in evaluating the measured directivity patterns only in magnitude and adding a linear or minimum phase term [AB20, BMKF20]. Recently, the Directivity Sample Combination algorithm (DISCO) [GP19] was proposed as a simple and efficient parametric alternative. DISCO allows the rendering of a simplified directivity patterns via VBAP after sampling it at a few regularly-distributed, pattern-independent points.

By exploiting the algorithms introduced in this work, another parametric approach is developed that is able to describe simplified directivity patterns after analysis of its peak locations, peak values and corresponding tangential curvatures. The New MUSIC algorithm is employed to not only find peaks of the measured directivity pattern but also to evaluate the Hessian at each peak. From this information, complexity-reduced directivity patterns exhibiting the original's principal peaks and also following their shape are synthesised. The introduced approach is not only applicable to source centreing problems, but can be seen as a hands-on tool to reduce the complexity of directivity patterns while preserving their main directional characteristics, which reduces the computational effort for simulation of directional sound sources, e.g. in extended reality applications.

3.2.1 Analysis of Directivity Patterns

Analysis of the Hessian. The Hessian contains information like definiteness and curvature at extrema locations, which is beneficial for analysis and synthesis of directivity patterns. However, interpretation of the 3×3 Hessian with tangential derivatives in x, y and z (as derived in section 2.5.2) is inconvenient and a rotation to tangential azimuthal, zenithal and radial directions as shown in figure 3.2 is beneficial. Such a mapping on the tangential plane is accomplished by rotation to unit vectors in azimuthal, zenithal and radial directions, $\boldsymbol{u}_\varphi$, $\boldsymbol{u}_\vartheta$ and $\boldsymbol{u}_r$,

$$\begin{aligned}\begin{bmatrix}\boldsymbol{u}_\varphi & \boldsymbol{u}_\vartheta & \boldsymbol{u}_r\end{bmatrix}^{\mathrm{T}} &= \boldsymbol{R}^{\mathrm{T}} \begin{bmatrix}\boldsymbol{i}_x & \boldsymbol{i}_y & \boldsymbol{i}_z\end{bmatrix} \\ &= \begin{bmatrix}\boldsymbol{u}_\varphi & \boldsymbol{u}_\vartheta & \boldsymbol{u}_r\end{bmatrix}^{\mathrm{T}} \boldsymbol{I}\end{aligned}$$



(a) Before rotation.

(b) After rotation.

Figure 3.2 – Rotation of the Cartesian coordinate system using spherical unit vectors $\boldsymbol{u}_\varphi$, $\boldsymbol{u}_\vartheta$, $\boldsymbol{u}_r$ at a point on the unit sphere.

$$= \begin{bmatrix} -\sin(\varphi) & \cos(\varphi)\cos(\vartheta) & \cos(\varphi)\sin(\vartheta) \\ \cos(\varphi) & \sin(\varphi)\cos(\vartheta) & \sin(\varphi)\sin(\vartheta) \\ 0 & -\sin(\vartheta) & \cos(\vartheta) \end{bmatrix}^{\mathrm{T}} . \tag{3.8}$$

For calculation of the Hessian in spherical coordinates, the coordinate system is rotated from spherical to Cartesian coordinates, where the Hessian is taken and rotated back,

$$\tilde{\boldsymbol{H}} = \begin{bmatrix} \tilde{H}_{\varphi\varphi} & \tilde{H}_{\varphi\vartheta} & \tilde{H}_{\varphi r} \\ \tilde{H}_{\vartheta\varphi} & \tilde{H}_{\vartheta\vartheta} & \tilde{H}_{\vartheta r} \\ \tilde{H}_{r\varphi} & \tilde{H}_{r\vartheta} & \tilde{H}_{rr} \end{bmatrix} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{R} . \tag{3.9}$$

A useful 2×2 Hessian is found as

$$\boldsymbol{M}^{\mathrm{T}}\tilde{\boldsymbol{H}}\boldsymbol{M} = \begin{bmatrix} \tilde{H}_{\varphi\varphi} & \tilde{H}_{\varphi\vartheta} \\ \tilde{H}_{\vartheta\varphi} & \tilde{H}_{\vartheta\vartheta} \end{bmatrix} = \boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{R}\boldsymbol{M} , \tag{3.10}$$

with the projector $\boldsymbol{M}$ that projects $\tilde{\boldsymbol{H}}$ onto azimuthal and zenithal unit vectors and neglects the trivial radial components,

$$\boldsymbol{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} , \tag{3.11}$$

allowing for the analysis of the definiteness of the Hessian at extrema locations via eigenvalue decomposition. Following this scheme, the Newton steps are defined via the Hessian in spherical coordinates, where regularisation of the Hessian is avoided as the 2×2 Hessian in spherical coordinates has full rank,

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \boldsymbol{R}\boldsymbol{M}\left(\boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{H}_{f^2}\boldsymbol{R}\boldsymbol{M}\right)^{-1}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{g}_{f^2} . \tag{3.12}$$

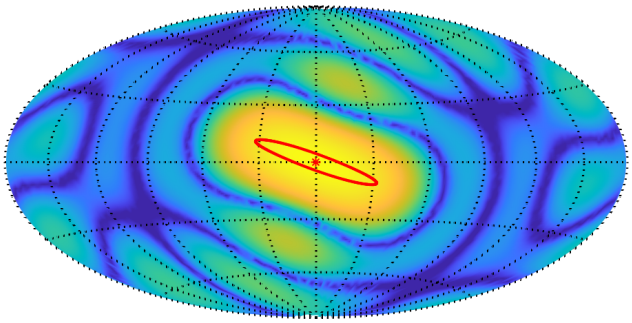


Figure 3.3 – Directional magnitude of an order $N = 5$ anisotropic beam pattern and an ellipse characterising the relative tangential flatness.

Note that these calculations could also be performed directly using recurrence relations in spherical coordinates. However as explained in section 2.5, the derivative w.r.t. azimuth angle is not defined a zenith and nadir.

Depicting relative tangential flatness to measure the peak extent in 2D. The ellipsoidal shape of the extrema is analysed by calculation of major and minor principal semi-axes via singular value decomposition,

$$\tilde{\boldsymbol{H}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}} \,. \tag{3.13}$$

Here, the left-singular vectors in $\boldsymbol{U}^{\mathrm{T}}$ rotate $\tilde{\boldsymbol{H}}$ to be aligned with the axes of the coordinate system spanned by the tangential azimuthal and zenithal unit vectors at the maximum location $\hat{\boldsymbol{\theta}}$, so an ellipse characterising the curvature at the maximum can be drawn by rotating and stretching the unit circle accordingly, and then rotating back from the local tangential azimuth and zenith directions to world coordinates,

$$\begin{bmatrix} \xi_x(t) \\ \xi_y(t) \\ \xi_z(t) \end{bmatrix} = \boldsymbol{R}\boldsymbol{M}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\|\boldsymbol{\gamma}\| \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} + \hat{\boldsymbol{\theta}} \,, \tag{3.14}$$

where $t \in [0, 2\pi]$. Here, the ellipse's axes are weighted with the inverse of the singular values, to plot flatness of the maximum instead of curvature, and normalised by the norm of the pattern $\boldsymbol{\gamma}$ the Hessian is evaluated at. Figure 3.3 shows the directional magnitude of an order $N = 5$ anisotropic beam pattern and an ellipse characterising its relative tangential flatness using the described method.

3.2.2 Synthesis of Hessian-Constrained Directivity Patterns

To resynthesise simplified directivity patterns, a minimum-norm SH pattern $\hat{\boldsymbol{\gamma}}$ subject to a Hessian constraint is searched,

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|^2 \tag{3.15}$$

$$\text{s.t. } \tilde{\boldsymbol{h}}_\gamma = \tilde{\boldsymbol{h}}\,, \tag{3.16}$$

where

$$\tilde{\boldsymbol{h}}_\gamma = \begin{bmatrix} \tilde{h}_{\varphi\varphi} \\ \tilde{h}_{\varphi\vartheta} \\ \tilde{h}_{\vartheta\vartheta} \end{bmatrix} = \boldsymbol{A}\boldsymbol{\gamma}\,, \tag{3.17}$$

contains the entries of the 2×2 Hessian of the pattern $\boldsymbol{\gamma}$, which is calculated as before via projection of a 3×3 Hessian in Cartesian coordinates $\boldsymbol{H}_\gamma$ onto tangential azimuth and zenith angles,

$$\tilde{\boldsymbol{H}}_\gamma = \begin{bmatrix} \tilde{h}_{\varphi\varphi} & \tilde{h}_{\varphi\vartheta} \\ \tilde{h}_{\varphi\vartheta} & \tilde{h}_{\vartheta\vartheta} \end{bmatrix} = \boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{H}_\gamma\boldsymbol{R}\boldsymbol{M}\,, \tag{3.18}$$

and $\tilde{\boldsymbol{h}}$ contains the entries of a target Hessian, both evaluated at location $\boldsymbol{\theta}$. With the same projection matrix $\boldsymbol{M}$ as in the previous section (eq. 3.11), the contents of $\boldsymbol{R}$ from eq. 3.8 described by variables r for mathematical brevity,

$$\boldsymbol{R} = \begin{bmatrix} -\sin(\varphi) & \cos(\varphi)\cos(\vartheta) & \cos(\varphi)\sin(\vartheta) \\ \cos(\varphi) & \sin(\varphi)\cos(\vartheta) & \sin(\varphi)\sin(\vartheta) \\ 0 & -\sin(\vartheta) & \cos(\vartheta) \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}\,, \tag{3.19}$$

and the Hessian expressed via SH recurrences as in section 2.5.2,

$$\boldsymbol{H}_\gamma = \begin{bmatrix} \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_x^{\mathrm{T}}\boldsymbol{G}_x^{\mathrm{T}}\boldsymbol{\gamma} & \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_x^{\mathrm{T}}\boldsymbol{G}_y^{\mathrm{T}}\boldsymbol{\gamma} & \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_x^{\mathrm{T}}\boldsymbol{G}_z^{\mathrm{T}}\boldsymbol{\gamma} \\ \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_y^{\mathrm{T}}\boldsymbol{G}_x^{\mathrm{T}}\boldsymbol{\gamma} & \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_y^{\mathrm{T}}\boldsymbol{G}_y^{\mathrm{T}}\boldsymbol{\gamma} & \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_y^{\mathrm{T}}\boldsymbol{G}_z^{\mathrm{T}}\boldsymbol{\gamma} \\ \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_z^{\mathrm{T}}\boldsymbol{G}_x^{\mathrm{T}}\boldsymbol{\gamma} & \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_z^{\mathrm{T}}\boldsymbol{G}_y^{\mathrm{T}}\boldsymbol{\gamma} & \boldsymbol{y}_{N+2}^{\mathrm{T}}\boldsymbol{G}_z^{\mathrm{T}}\boldsymbol{G}_z^{\mathrm{T}}\boldsymbol{\gamma} \end{bmatrix}\,, \tag{3.20}$$

or in vectorised form,

$$\mathrm{vec}(\boldsymbol{H}_\gamma) = \begin{bmatrix} \boldsymbol{h}_{11}^{\mathrm{T}} \\ \boldsymbol{h}_{21}^{\mathrm{T}} \\ \boldsymbol{h}_{31}^{\mathrm{T}} \\ \boldsymbol{h}_{12}^{\mathrm{T}} \\ \boldsymbol{h}_{22}^{\mathrm{T}} \\ \boldsymbol{h}_{32}^{\mathrm{T}} \\ \boldsymbol{h}_{13}^{\mathrm{T}} \\ \boldsymbol{h}_{23}^{\mathrm{T}} \\ \boldsymbol{h}_{33}^{\mathrm{T}} \end{bmatrix} \boldsymbol{\gamma}\,, \tag{3.21}$$

the evaluated Hessian in half-vectorised form (vectorisation of the upper triangular matrix assuming a symmetric Hessian at location $\boldsymbol{\theta}$) is found after rotation,

$$\begin{aligned} \boldsymbol{A}\boldsymbol{\gamma} &= \mathrm{vech}(\tilde{\boldsymbol{H}}_\gamma) \\ &= \begin{bmatrix} r_{11}\left(\boldsymbol{h}_{11}^{\mathrm{T}} r_{11} + \boldsymbol{h}_{21}^{\mathrm{T}} r_{21} + \boldsymbol{h}_{31}^{\mathrm{T}} r_{31}\right) + r_{21}\left(\boldsymbol{h}_{12}^{\mathrm{T}} r_{11} + \boldsymbol{h}_{22}^{\mathrm{T}} r_{21} + \boldsymbol{h}_{32}^{\mathrm{T}} r_{31}\right) + r_{31}\left(\boldsymbol{h}_{13}^{\mathrm{T}} r_{11} + \boldsymbol{h}_{23}^{\mathrm{T}} r_{21} + \boldsymbol{h}_{33}^{\mathrm{T}} r_{31}\right) \\ r_{12}\left(\boldsymbol{h}_{11}^{\mathrm{T}} r_{11} + \boldsymbol{h}_{21}^{\mathrm{T}} r_{21} + \boldsymbol{h}_{31}^{\mathrm{T}} r_{31}\right) + r_{22}\left(\boldsymbol{h}_{12}^{\mathrm{T}} r_{11} + \boldsymbol{h}_{22}^{\mathrm{T}} r_{21} + \boldsymbol{h}_{32}^{\mathrm{T}} r_{31}\right) + r_{32}\left(\boldsymbol{h}_{13}^{\mathrm{T}} r_{11} + \boldsymbol{h}_{23}^{\mathrm{T}} r_{21} + \boldsymbol{h}_{33}^{\mathrm{T}} r_{31}\right) \\ r_{12}\left(\boldsymbol{h}_{11}^{\mathrm{T}} r_{12} + \boldsymbol{h}_{21}^{\mathrm{T}} r_{22} + \boldsymbol{h}_{31}^{\mathrm{T}} r_{32}\right) + r_{22}\left(\boldsymbol{h}_{12}^{\mathrm{T}} r_{12} + \boldsymbol{h}_{22}^{\mathrm{T}} r_{22} + \boldsymbol{h}_{32}^{\mathrm{T}} r_{32}\right) + r_{32}\left(\boldsymbol{h}_{13}^{\mathrm{T}} r_{12} + \boldsymbol{h}_{23}^{\mathrm{T}} r_{22} + \boldsymbol{h}_{33}^{\mathrm{T}} r_{32}\right) \end{bmatrix} \boldsymbol{\gamma}\,. \end{aligned} \tag{3.22}$$

We solve using a Lagrange multiplier $\boldsymbol{\lambda}$,

$$J(\boldsymbol{\gamma}, \boldsymbol{\lambda}) = \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{\gamma} + (\boldsymbol{A}\boldsymbol{\gamma} - \boldsymbol{h})^{\mathrm{T}}\boldsymbol{\lambda}\,, \tag{3.23}$$

by zeroing the gradients

$$\frac{\partial}{\partial \boldsymbol{\gamma}} J = 2\boldsymbol{\gamma} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{0}\,, \tag{3.24}$$

$$\frac{\partial}{\partial \boldsymbol{\lambda}} J = \boldsymbol{A}\boldsymbol{\gamma} - \boldsymbol{h} = \mathbf{0}\,, \tag{3.25}$$

and solving for

$$\hat{\boldsymbol{\gamma}} = -\frac{1}{2}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}\,. \tag{3.26}$$

Inserting into the constraint yields the Lagrange multiplier

$$\boldsymbol{\lambda} = -2(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\tilde{\boldsymbol{h}}\,, \tag{3.27}$$

and inserting back into the equation for $\hat{\boldsymbol{\gamma}}$ yields the optimum SH pattern

$$\hat{\boldsymbol{\gamma}} = \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\tilde{\boldsymbol{h}}\,. \tag{3.28}$$

Adding an amplitude constraint. We can add an amplitude constraint, forcing the amplitude at the target location $\boldsymbol{\theta}$ to 1,

$$\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})\boldsymbol{\gamma} = 1\,, \tag{3.29}$$

by augmenting the constraint (eq. 3.16) to

$$\begin{bmatrix} \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta}) \\ \boldsymbol{A} \end{bmatrix} \boldsymbol{\gamma} - \begin{bmatrix} 1 \\ \tilde{\boldsymbol{h}} \end{bmatrix} = \mathbf{0}\,. \tag{3.30}$$

If the amplitude a_{t} at the location of the target Hessian is known, the amplitude constraint can be adapted to perfectly meet it,

$$\begin{bmatrix} 1 \\ \tilde{\boldsymbol{h}} \end{bmatrix} \leftarrow \begin{bmatrix} a_{\mathrm{t}} \\ \tilde{\boldsymbol{h}} \end{bmatrix}\,, \tag{3.31}$$

or to create a beam with amplitude 1 that has the same Hessian as the input Hessian scaled to an amplitude of 1 at the target location,

$$\begin{bmatrix} 1 \\ \tilde{\boldsymbol{h}} \end{bmatrix} \leftarrow \begin{bmatrix} 1 \\ \tilde{\boldsymbol{h}}\frac{1}{a_{\mathrm{t}}} \end{bmatrix}\,, \tag{3.32}$$

such that the directivity pattern will be reconstructed with correct relation of Hessian to amplitude.

Adding inverse weights. To favor energy in lower orders, inverse weights $\boldsymbol{w}$, such as max-$\boldsymbol{r}_{\mathrm{E}}$ or in-phase weights [Dan01, ZF12], can be added to the cost function,

$$J(\boldsymbol{\gamma}, \boldsymbol{\lambda}) = \boldsymbol{\gamma}^{\mathrm{T}}\mathrm{diag}(\boldsymbol{w})^{-1}\boldsymbol{\gamma} + (\boldsymbol{A}\boldsymbol{\gamma} - \boldsymbol{h})^{\mathrm{T}}\boldsymbol{\lambda}\,, \tag{3.33}$$

yielding

$$\hat{\boldsymbol{\gamma}} = -\frac{1}{2}\mathrm{diag}(\boldsymbol{w})\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}\,, \tag{3.34}$$

$$\boldsymbol{\lambda} = -2(\boldsymbol{A}\,\mathrm{diag}(\boldsymbol{w})\boldsymbol{A}^{\mathrm{T}})^{-1}\tilde{\boldsymbol{h}}\,, \tag{3.35}$$

and

$$\hat{\boldsymbol{\gamma}} = \mathrm{diag}(\boldsymbol{w})\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\,\mathrm{diag}(\boldsymbol{w})\boldsymbol{A}^{\mathrm{T}})^{-1}\tilde{\boldsymbol{h}}\,. \tag{3.36}$$

Relation of curvature to amplitude of a SH basic beam. We use a helpful relation of Legendre polynomials to analyse the beam curvature of an SH pattern at zenith, where all SHs of degree $m \neq 0$ vanish,

$$\frac{d^2}{d\vartheta^2}P_n(z) = -z\frac{d}{dz}P_n(z) + (1-z^2)\frac{d^2}{dz^2}P_n(z)\,, \tag{3.37}$$

yielding at zenith $z = \cos\vartheta = 1$,

$$\frac{d^2}{d\vartheta^2}P_n(1) = -\frac{d}{dz}P_n(1)\,. \tag{3.38}$$

Via the Legendre differential equation

$$(1-z^2)\frac{d^2}{dz^2}P_n(z) - 2z\frac{d}{dz}P_n(z) + n(n+1)P_n(z) = 0\,, \tag{3.39}$$

we find at zenith,

$$\frac{d}{dz}P_n(1) = \frac{n(n+1)}{2}P_n(1)\,. \tag{3.40}$$

Inserting this in eq. 3.38 yields the curvature at zenith,

$$\frac{d^2}{d\vartheta^2}P_n(1) = -\frac{n(n+1)}{2}P_n(1) = -\frac{n(n+1)}{2}\,. \tag{3.41}$$

Now, the curvature of a basic SH beam (at zenith, summing over $m = 0$ SHs is sufficient) is derived as,

$$\begin{aligned}
\frac{\partial^2}{\partial\vartheta^2}\sum_{n=0}^{N} Y_n^0(0)Y_n^0(\vartheta)\bigg|_{\vartheta=0} &= \sum_{n=0}^{N}(N_n^0)^2 P_n(1)\ \frac{\partial^2}{\partial\vartheta^2}P_n(\cos(\vartheta))\bigg|_{\vartheta=0} \\
&= \frac{1}{4\pi}\sum_{n=0}^{N}(2n+1)\ \frac{\partial^2}{\partial\vartheta^2}P_n(\cos(\vartheta))\bigg|_{\vartheta=0} \\
&= -\frac{1}{2}\frac{1}{4\pi}\sum_{n=0}^{N}(2n+1)(n(n+1)) \\
&= -\frac{1}{4}\frac{1}{4\pi}N(N+1)^2(N+2)\,,
\end{aligned} \tag{3.42}$$

where the series is evaluated using Gauss' sum formula (or Faulhaber's formula for higher-order terms) and fully-normalised SHs ($N_n^0 = \sqrt{\frac{2n+1}{4\pi}}$) are used. We normalise by the amplitude $\boldsymbol{y}_\mathrm{b}^\mathrm{T}\boldsymbol{y}_\mathrm{b} = \frac{(N+1)^2}{4\pi}$ of a basic beam, and finally get the ratio of curvature to amplitude of a basic SH beam,

$$\frac{\frac{\partial^2}{\partial\vartheta^2}\sum_{n=0}^{N}\sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}_0)Y_n^m(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}}{\sum_{n=0}^{N}\sum_{m=-n}^{n}(Y_n^m(\boldsymbol{\theta}_0))^2} = -\frac{1}{4}N(N+2)\,. \tag{3.43}$$

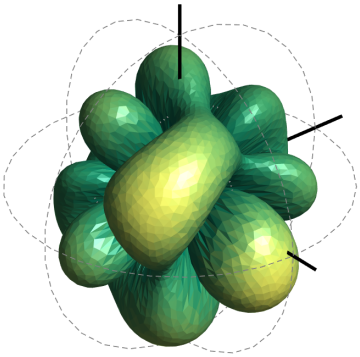
This information can be used to manipulate the target Hessian,

$$\begin{bmatrix}1\\ \tilde{\boldsymbol{h}}\end{bmatrix} \leftarrow \begin{bmatrix}1\\ \frac{\tilde{\boldsymbol{h}}}{\max(\tilde{\boldsymbol{h}})}\frac{1}{4}N(N+2)\end{bmatrix}\,, \tag{3.44}$$

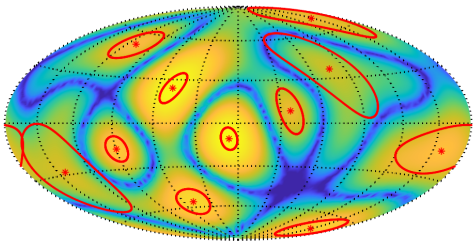
such that the Hessian of a basic beam will be recreated perfectly (scaled to have amplitude 1) without knowledge of the original amplitude.

3.2.3 Exemplary Resynthesis of a Random Directivity Pattern

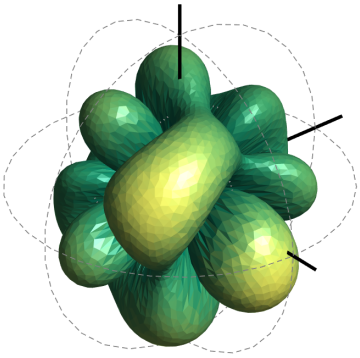
To demonstrate the proposed approach, a random directivity pattern of maximum SH order $N = 5$ is analysed and resynthesised in different ways. Figure 3.4 shows (a) the original random directivity pattern, (b) a map projection and found peaks using New MUSIC, (c) a resynthesised pattern using a Hessian constraint and an amplitude constraint with the original amplitudes (eq. 3.32), and (d) a resynthesised pattern with the Hessian and amplitude constraints, using only the two dominant peaks. In (c) all 11 peaks of the original pattern are used and the original pattern is reconstructed perfectly. In (d) a simplified pattern is obtained, where the two dominant peaks from (a) are still clearly visible but all other peaks are strongly reduced. Further simplification of the pattern is achieved by reducing the reconstruction order, as shown in figure 3.5, where the maximum SH order is set to $N = 3$. In cases, where a lower reconstruction order is used, the original relations of amplitude to curvature might not be reproducible anymore. Hence, in figure 3.5, the amplitude constraint using the original amplitudes (eq. 3.32) is not used anymore, but rather only the Hessian constraint in (a), the Hessian constraint and inverse max-$\boldsymbol{r}_{\mathrm{E}}$ weights (eq. 3.36) in (b), and the Hessian constraint and inverse in-phase weights in (c). As expected, the addition of inverse weights decreases side-lobes at the cost of increasing main-lobe widths.



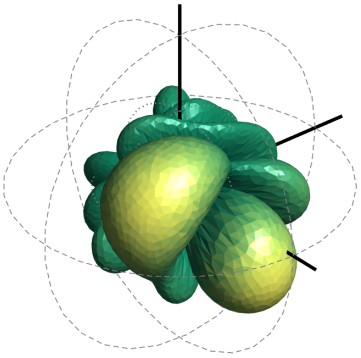
(a) Balloon plot of the original pattern.



(b) Map projection showing 11 peak locations and tangential peak curvatures of the original pattern.

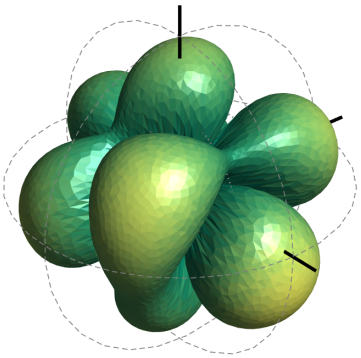


(c) Resynthesised pattern using all $Q = 11$ found peaks.

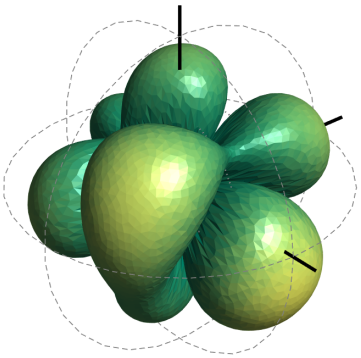


(d) Resynthesised pattern using the $Q = 2$ dominant peaks.

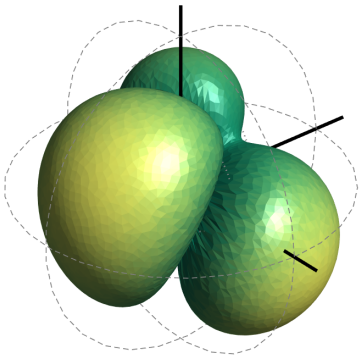
Figure 3.4 – Figures (a) and (b) show a balloon plot and a map projection of a random directivity pattern of maximum SH order $N = 5$. In the map projection red asterisks and ellipses mark peak locations and tangential curvature at the peaks as found by New MUSIC. Figures (c) and (d) show resynthesised patterns using a Hessian constraint and an amplitude constraint, a maximum SH order $N = 5$ and the Q dominant peaks.



(a) Resynthesised pattern using a Hessian constraint only.



(b) Resynthesised pattern using a Hessian constraint and inverse max-$\boldsymbol{r}_{\mathrm{E}}$ weights.



(c) Resynthesised pattern using a Hessian constraint and inverse in-phase weights.

Figure 3.5 – Resynthesised patterns based on the higher-order original in figure 3.4(a), using a maximum SH order $N = 3$, the $Q = 2$ dominant peaks and different constraints.

Chapter 4

Conclusion and Outlook

This thesis developed and introduced two novel, real-valued subspace-based simultaneous DOA estimation algorithms: REVEB-ESPRIT and New MUSIC (cf. chapter 2). Both algorithms perform similar to each other and to related algorithms in several tested scenarios. While the RMSEs generated by REVEB-ESPRIT are a bit higher than the RMSEs of the complex-valued counterpart in many cases, REVEB-ESPRIT has the advantage of being generally four times more efficient and easier to implement, since all computations are done using real numbers only. New MUSIC performs similar to REVEB-ESPRIT, and there are cases in which either New MUSIC or REVEB-ESPRIT brings advantages. While in black-box scenarios, REVEB-ESPRIT has the advantage that (apart from the thresholds ϵ_1 and ϵ_2 for the subspace identification with SORTE) no parameters have to be set, in a highly specific but constant scenario, New MUSIC has an advantage as it can be adapted to the specific environment by setting convergence criteria and step sizes, and analysing the convergence behaviour. Moreover, New MUSIC could easily be integrated into a statistical framework including Bayesian analysis of the convergence points or time-variant tracking via a Kalman filter, for instance, making use of information from previous time steps to provide a well-controlled tracking behaviour for the detected DOAs. To support reproducible research, all recurrence relations of the spherical harmonics required to express tangential derivatives and multiplication by unit-vector components are made available online [1] (and in the listings of appendix A.4) along with the required numerical procedures for verification.

Chapter 3 outlined two possible applications of the algorithms. REVEB-ESPRIT was demonstrated to enable a powerful, novel upmixing algorithm for ARIRs, called HO-ASDM. By combination of simultaneous DOA estimation, beamforming, and addition of a residual signal, a higher-order ARIR is created from a lower-order ARIR of arbitrary order. By using REVEB-ESPRIT, multiple reflection directions are estimated simultaneously at higher accuracy compared to the conventional pseudo-intensity vector approach (assuming a large measurement SNR for the RIR measurement). HO-ASDM outperformed the conventional ASDM in a listening experiment, in particular concerning ASDM's transient-response and preservation of timbre. However, HO-ASDM still needs

1. `https://git.iem.at/thomasdeppisch/real-sh-recurrence-relations`

to be regarded as a work in progress because evaluation of envelopment and diffuseness still remains to be investigated in greater detail in the future. A new way of analysing and synthesising directivity patterns using the curvature at the peaks was introduced using New MUSIC. Here, New MUSIC can help to better understand the specific characteristics of directivity patterns and to resynthesise simplified directivity patterns as an alternative to source centreing and for computationally efficient synthesis of directive sound sources in virtual environments. In particular the anisotropic curvatures it delivers seem to offer an appropriate characterisation in analysis and re-synthesis of directivity.

Appendix A

Derivation of SH Recurrence Relations

A.1 Derivation of Gradient Addition Theorems for Real-Valued SHs

Following a technical note from Franz Zotter, in this section gradient addition theorems for real-valued SHs are derived. For derivation of a tangential gradient in Cartesian coordinates, the tangential gradient operator (neglecting radial derivatives) is defined by using the chain rule,

$$\begin{aligned}\nabla_{\boldsymbol{r},\text{tangential}} &= \begin{pmatrix} \partial\varphi/\partial x & \partial\mu/\partial x \\ \partial\varphi/\partial y & \partial\mu/\partial y \\ \partial\varphi/\partial z & \partial\mu/\partial z \end{pmatrix} \begin{pmatrix} \partial/\partial\varphi \\ \partial/\partial\mu \end{pmatrix} \\ &= \begin{pmatrix} -\frac{\sin(\varphi)}{\sqrt{1-\mu^2}} & -\mu\cos(\varphi)\sqrt{1-\mu^2} \\ \frac{\cos(\varphi)}{\sqrt{1-\mu^2}} & -\mu\sin(\varphi)\sqrt{1-\mu^2} \\ 0 & 1-\mu^2 \end{pmatrix} \begin{pmatrix} \partial/\partial\varphi \\ \partial/\partial\mu \end{pmatrix}, \end{aligned} \tag{A.1}$$

with

$$\varphi = \arctan(y/x)\,, \tag{A.2}$$

$$\mu = \cos(\vartheta) = z/r = z/\sqrt{x^2+y^2+z^2}\,. \tag{A.3}$$

For mathematical brevity, real-valued SHs are defined as

$$\boldsymbol{Y}_{nm}(\varphi,\mu) = N_{nm}P_n^m(\mu)\boldsymbol{\Phi}_m(\varphi)\,, \tag{A.4}$$

with

$$\boldsymbol{\Phi}_m(\varphi) = \begin{pmatrix} \cos(m\varphi) \\ \sin(m\varphi) \end{pmatrix} \tag{A.5}$$

and

$$\boldsymbol{L} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{A.6}$$

In the derivation, the trigonometric identities

$$\begin{aligned}\cos(\varphi)\boldsymbol{\Phi}_m(\varphi) &= \frac{1}{2}\begin{pmatrix}\cos((m-1)\varphi)+\cos((m+1)\varphi)\\ \sin((m-1)\varphi)+\sin((m+1)\varphi)\end{pmatrix}\\ &= \frac{1}{2}\left(\boldsymbol{\Phi}_{m-1}(\varphi)+\boldsymbol{\Phi}_{m+1}(\varphi)\right),\end{aligned} \tag{A.7}$$

and

$$\begin{aligned}\sin(\varphi)\boldsymbol{\Phi}_m(\varphi) &= \frac{1}{2}\begin{pmatrix}-\sin((m-1)\varphi)+\sin((m+1)\varphi)\\ \cos((m-1)\varphi)-\cos((m+1)\varphi)\end{pmatrix}\\ &= \frac{1}{2}\boldsymbol{L}\left(\boldsymbol{\Phi}_{m-1}(\varphi)-\boldsymbol{\Phi}_{m+1}(\varphi)\right),\end{aligned} \tag{A.8}$$

the partial derivative in azimuth,

$$\partial/\partial\varphi\,\boldsymbol{\Phi}_m(\varphi) = m\boldsymbol{L}\boldsymbol{\Phi}_m(\varphi), \tag{A.9}$$

and the following recurrence relations of associated Legendre functions from [GD04, chapter 2],

$$(1-\mu^2)\partial/\partial\mu P_n^m = \frac{(n+1)(n+m)}{2n+1}P_{n-1}^m - \frac{n(n-m+1)}{2n+1}P_{n+1}^m, \tag{A.10}$$

$$\mu\sqrt{1-\mu^2}\partial/\partial\mu P_n^m + \frac{mP_n^m}{\sqrt{1-\mu^2}} = -\frac{n+1}{2n+1}P_{n-1}^{m+1} - \frac{n}{2n+1}P_{n+1}^{m+1}, \tag{A.11}$$

$$\begin{aligned}\mu\sqrt{1-\mu^2}\partial/\partial\mu P_n^m - \frac{mP_n^m}{\sqrt{1-\mu^2}} =\\ \frac{(n+1)(n+m-1)(n+m)}{2n+1}P_{n-1}^{m-1} + \frac{n(n-m+1)(n-m+2)}{2n+1}P_{n+1}^{m-1},\end{aligned} \tag{A.12}$$

will be useful. Applying the tangential gradient yields

$$\nabla_{\boldsymbol{r},\text{tangential}}\boldsymbol{Y}_{nm} = \begin{pmatrix}\partial/\partial x\\ \partial/\partial y\\ \partial/\partial z\end{pmatrix}\boldsymbol{Y}_{nm} \tag{A.13}$$

$$= \begin{pmatrix}-\frac{\sin(\varphi)}{\sqrt{1-\mu^2}}N_{nm}P_n^m\partial/\partial\varphi\boldsymbol{\Phi}_m - \mu\cos(\varphi)\sqrt{1-\mu^2}N_{nm}\boldsymbol{\Phi}_m\partial/\partial\mu P_n^m\\ \frac{\cos(\varphi)}{\sqrt{1-\mu^2}}N_{nm}P_n^m\partial/\partial\varphi\boldsymbol{\Phi}_m - \mu\sin(\varphi)\sqrt{1-\mu^2}N_{nm}\boldsymbol{\Phi}_m\partial/\partial\mu P_n^m\\ (1-\mu^2)N_{nm}\boldsymbol{\Phi}_m\partial/\partial\mu P_n^m\end{pmatrix}.$$

By applying equations A.9, A.7 and A.8 the partial derivative w.r.t x is reformulated as

$$\frac{\partial}{\partial x}\boldsymbol{Y}_{nm} = \frac{N_{nm}}{2\sqrt{1-\mu^2}}P_n^m m(\boldsymbol{\Phi}_{m-1}-\boldsymbol{\Phi}_{m+1}) - \frac{N_{nm}}{2}\mu\sqrt{1-\mu^2}(\boldsymbol{\Phi}_{m-1}+\boldsymbol{\Phi}_{m+1})\partial/\partial\mu P_n^m, \tag{A.14}$$

which is expanded and simplified using the recurrences in eq. A.11 and A.12, yielding

$$\begin{aligned}\partial/\partial x \boldsymbol{Y}_{nm} = &-\frac{(n+1)(n+m-1)(n+m)N_{nm}}{2(2n+1)N_{n-1,m-1}}\boldsymbol{Y}_{n-1,m-1} \\ &- \frac{n(n-m+1)(n-m+2)N_{nm}}{2(2n+1)N_{n+1,m-1}}\boldsymbol{Y}_{n+1,m-1} \\ &+ \frac{(n+1)N_{nm}}{2(2n+1)N_{n-1,m+1}}\boldsymbol{Y}_{n-1,m+1} + \frac{nN_{nm}}{2(2n+1)N_{n+1,m+1}}\boldsymbol{Y}_{n+1,m+1}\,.\end{aligned} \tag{A.15}$$

Similarly, the result for the partial derivative w.r.t. y is found as

$$\begin{aligned}\partial/\partial y \boldsymbol{Y}_{nm} = &-\frac{(n+1)(n+m-1)(n+m)N_{nm}}{2(2n+1)N_{n-1,m-1}}\boldsymbol{L}\boldsymbol{Y}_{n-1,m-1} \\ &- \frac{n(n-m+1)(n-m+2)N_{nm}}{2(2n+1)N_{n+1,m-1}}\boldsymbol{L}\boldsymbol{Y}_{n+1,m-1} \\ &- \frac{(n+1)N_{nm}}{2(2n+1)N_{n-1,m+1}}\boldsymbol{L}\boldsymbol{Y}_{n-1,m+1} - \frac{nN_{nm}}{2(2n+1)N_{n+1,m+1}}\boldsymbol{L}\boldsymbol{Y}_{n+1,m+1}\,.\end{aligned} \tag{A.16}$$

The partial derivative w.r.t. z is obtained after applying the recurrence in eq. A.10,

$$\partial/\partial z \boldsymbol{Y}_{nm} = \frac{(n+1)(n+m)N_{nm}}{(2n+1)N_{n-1,m}}\boldsymbol{Y}_{n-1,m} - \frac{n(n-m+1)N_{nm}}{(2n+1)N_{n+1,m}}\boldsymbol{Y}_{n+1,m}\,. \tag{A.17}$$

For $m = 0$ the expressions for the derivatives w.r.t. x and y are derived explicitly,

$$\partial/\partial x Y^{(\cos)}_{n,0} = \frac{(n+1)N_{n,0}}{(2n+1)N_{n-1,1}}Y^{(\cos)}_{n-1,1} + \frac{nN_{n,0}}{(2n+1)N_{n+1,1}}Y^{(\cos)}_{n+1,1}\,, \tag{A.18}$$

$$\partial/\partial y Y^{(\cos)}_{n,0} = \frac{(n+1)N_{n,0}}{(2n+1)N_{n-1,1}}Y^{(\sin)}_{n-1,1} + \frac{nN_{n,0}}{(2n+1)N_{n+1,1}}Y^{(\sin)}_{n+1,1}\,, \tag{A.19}$$

where $Y^{(\cos)}_{n,m}$ and $Y^{(\sin)}_{n,m}$ hold the upper and lower entry of $\boldsymbol{Y}_{nm}$, respectively.

A.2 Derivation of Multiplication Theorems for Real-Valued SHs

The following derivations of multiplication theorems for real-valued SHs and corresponding extension relations (cf. section A.3) are based on technical notes from Franz Zotter and Byeongho Jo, developed during their work on [JZC20]. From [JZC20] the following recurrences for the complex-valued spherical harmonics Y^m_n are known, for multiplication with $\theta_{\text{xy}} = \sin\vartheta\, e^{i\varphi}$, its conjugate θ^*_{xy}, and $\theta_{\text{z}} = \cos\vartheta$, with the coefficients $w^m_n = \sqrt{\frac{(n+m-1)(n+m)}{(2n-1)(2n+1)}}$, and $v^m_n = \sqrt{\frac{(n-m)(n+m)}{(2n-1)(2n+1)}}$,

$$\theta_{\text{xy}}Y^m_n = w^{-m}_n Y^{m+1}_{n-1} - w^{m+1}_{n+1}Y^{m+1}_{n+1}, \tag{A.20}$$

$$\theta^*_{\text{xy}}Y^m_n = -w^m_n Y^{m-1}_{n-1} + w^{-m+1}_{n+1}Y^{m-1}_{n+1}, \tag{A.21}$$

$$\theta_{\text{z}}Y^m_n = v^m_n Y^m_{n-1} + v^m_{n+1}Y^m_{n+1}. \tag{A.22}$$

The conjugated parameters are $\theta_{\mathrm{xy}} = \theta_{\mathrm{x}} + i\theta_{\mathrm{y}}$ and $\theta^*_{\mathrm{xy}} = \theta_{\mathrm{x}} - i\theta_{\mathrm{y}}$, and the complex harmonics are related to the real ones by [JHN17, chapter 3]

$$Y_n^m = \begin{cases} \frac{(-1)^m}{\sqrt{2-\delta_m}}[Y_{nm} + i(1-\delta_m)Y_{n,-m}], & m \geq 0 \\ \frac{1}{\sqrt{2-\delta_m}}[Y_{n|m|} - i(1-\delta_m)Y_{n,-|m|}], & m < 0\,. \end{cases} \tag{A.23}$$

We first take the real part of eqs. A.20 and A.21 for $m > 0$

$$\frac{(-1)^m}{\sqrt{2-\delta_m}}\left(\theta_{\mathrm{x}}Y_{nm} - (1-\delta_m)\theta_{\mathrm{y}}Y_{n,-m}\right) = \frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}(w_n^{-m}Y_{n-1,m+1} - w_{n+1}^{m+1}Y_{n+1,m+1}), \tag{A.24}$$

$$\frac{(-1)^m}{\sqrt{2-\delta_m}}\left(\theta_{\mathrm{x}}Y_{nm} + (1-\delta_m)\theta_{\mathrm{y}}Y_{n,-m}\right) = \frac{(-1)^{m-1}}{\sqrt{2-\delta_{m-1}}}(-w_n^{m}Y_{n-1,m-1} + w_{n+1}^{-m+1}Y_{n+1,m-1}), \tag{A.25}$$

and we see that sum and difference of eqs. A.24 and A.25 divided by two yields for $m > 0$

$$\begin{aligned} \frac{(-1)^m}{\sqrt{2-\delta_m}}\theta_{\mathrm{x}}Y_{nm} = \tfrac{1}{2}\Big[&- \frac{(-1)^{m-1}w_n^m}{\sqrt{2-\delta_{m-1}}}Y_{n-1,m-1} + \frac{(-1)^{m-1}w_{n+1}^{-m+1}}{\sqrt{2-\delta_{m-1}}}Y_{n+1,m-1} \\ &+ \frac{(-1)^{m+1}w_n^{-m}}{\sqrt{2-\delta_{m+1}}}Y_{n-1,m+1} - \frac{(-1)^{m+1}w_{n+1}^{m+1}}{\sqrt{2-\delta_{m+1}}}Y_{n+1,m+1}\Big], \end{aligned} \tag{A.26}$$

$$\begin{aligned} \frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)\theta_{\mathrm{y}}Y_{n,-m} = \tfrac{1}{2}\Big[&- \frac{(-1)^{m-1}w_n^m}{\sqrt{2-\delta_{m-1}}}Y_{n-1,m-1} + \frac{(-1)^{m-1}w_{n+1}^{-m+1}}{\sqrt{2-\delta_{m-1}}}Y_{n+1,m-1} \\ &- \frac{(-1)^{m+1}w_n^{-m}}{\sqrt{2-\delta_{m+1}}}Y_{n-1,m+1} + \frac{(-1)^{m+1}w_{n+1}^{m+1}}{\sqrt{2-\delta_{m+1}}}Y_{n+1,m+1}\Big]. \end{aligned} \tag{A.27}$$

We repeat the process for the imaginary parts of eqs. A.20 and A.21,

$$\begin{aligned} &\frac{(-1)^m}{\sqrt{2-\delta_m}}(\theta_{\mathrm{x}}(1-\delta_m)Y_{n,-m} + \theta_{\mathrm{y}}Y_{nm}) \\ &\qquad = \frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})[w_n^{-m}Y_{n-1,-m-1} - w_{n+1}^{m+1}Y_{n+1,-m-1}], \end{aligned} \tag{A.28}$$

$$\begin{aligned} &\frac{(-1)^m}{\sqrt{2-\delta_m}}(\theta_{\mathrm{x}}(1-\delta_m)Y_{n,-m} - \theta_{\mathrm{y}}Y_{nm}) \\ &\qquad = \frac{(-1)^{m-1}}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})[-w_n^{m}Y_{n-1,-m+1} + w_{n+1}^{-m+1}Y_{n+1,-m+1}]. \end{aligned} \tag{A.29}$$

As before, we calculate the halved sums and differences thereof,

$$\begin{aligned} \frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)\theta_{\mathrm{x}}Y_{n,-m} = \tfrac{1}{2}\Big[&- \frac{(-1)^{m-1}w_n^m}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n-1,-m+1} \\ &+ \frac{(-1)^{m-1}w_{n+1}^{-m+1}}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n+1,-m+1} + \frac{(-1)^{m+1}w_n^{-m}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n-1,-m-1} \\ &- \frac{(-1)^{m+1}w_{n+1}^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n+1,-m-1}\Big], \end{aligned} \tag{A.30}$$

$$\frac{(-1)^m}{\sqrt{2-\delta_m}}\theta_{\mathrm{y}}Y_{n,m} = \tfrac{1}{2}\Big[\frac{(-1)^{m-1}w_n^m}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n-1,-m+1} - \frac{(-1)^{m-1}w_{n+1}^{-m+1}}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n+1,-m+1} + \frac{(-1)^{m+1}w_n^{-m}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n-1,-m-1} - \frac{(-1)^{m+1}w_{n+1}^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n+1,-m-1}\Big]\,. \tag{A.31}$$

We denote the first recurrence explicitly for $m = 0$,

$$\theta_{\mathrm{xy}}Y_n^0 = w_n^0 Y_{n-1}^1 - w_{n+1}^1 Y_{n+1}^1, \tag{A.32}$$

and thereof the real and imaginary parts are

$$\theta_{\mathrm{x}}Y_{n0} = \frac{-1}{\sqrt{2}}w_n^0 Y_{n-1,1} + \frac{1}{\sqrt{2}}w_{n+1}^1 Y_{n+1,1}, \tag{A.33}$$

$$\theta_{\mathrm{y}}Y_{n0} = \frac{-1}{\sqrt{2}}w_n^0 Y_{n-1,-1} + \frac{1}{\sqrt{2}}w_{n+1}^1 Y_{n+1,-1}\,. \tag{A.34}$$

For θ_{z} there is no shift in degree m, therefore the recurrence is obtained directly,

$$\theta_{\mathrm{z}}Y_{nm} = v_n^{|m|}Y_{n-1,m} + v_{n+1}^{|m|}Y_{n+1,m}. \tag{A.35}$$

A.3 Derivation of Extending Multiplication Theorems for Real-Valued SHs

From [JZC20] the following extending recurrence relations for the complex-valued spherical harmonics Y_n^m are known, for multiplication with $\theta_{\mathrm{xy}} = \sin\vartheta\, e^{i\varphi}$, its conjugate θ_{xy}^*, and $\theta_{\mathrm{z}} = \cos\vartheta$, with $\eta_n^m = \sqrt{\frac{n+m}{2n+1}}$,

$$\theta_{\mathrm{xy}}\eta_n^{-m}Y_n^m + \theta_{\mathrm{z}}\eta_n^{m+1}Y_n^{m+1} = \eta_{n-1}^{-m}Y_{n-1}^{m+1}, \qquad \text{for } -N \le m \le N-1 \tag{A.36}$$

$$\theta_{\mathrm{xy}}^*\eta_n^m Y_n^m - \theta_{\mathrm{z}}\eta_n^{-m+1}Y_n^{-m+1} = -\eta_{n-1}^m Y_{n-1}^{m-1} \qquad \text{for } -N+1 \le m \le N\,. \tag{A.37}$$

We first take the real part of the above equations for $m > 0$,

$$\theta_{\mathrm{x}}\eta_n^{-m}\frac{(-1)^m}{\sqrt{2-\delta_m}}Y_{n,m} - \theta_{\mathrm{y}}\eta_n^{-m}\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-m} + \theta_{\mathrm{z}}\eta_n^{m+1}\frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}Y_{n,m+1} = \eta_{n-1}^{-m}\frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}Y_{n-1,m+1}\,, \tag{A.38}$$

$$\theta_{\mathrm{x}}\eta_n^{m}\frac{(-1)^m}{\sqrt{2-\delta_m}}Y_{n,m} + \theta_{\mathrm{y}}\eta_n^{m}\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-m} - \theta_{\mathrm{z}}\eta_n^{-m+1}\frac{(-1)^{-m+1}}{\sqrt{2-\delta_{-m+1}}}Y_{n,|-m+1|} = -\eta_{n-1}^{m}\frac{(-1)^{m-1}}{\sqrt{2-\delta_{m-1}}}Y_{n-1,m-1}\,, \tag{A.39}$$

then we take the imaginary part, $m > 0$,

$$\begin{aligned}
&\theta_{\mathrm{x}}\eta_n^{-m}\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-m} + \theta_{\mathrm{y}}\eta_n^{-m}\frac{(-1)^m}{\sqrt{2-\delta_m}}Y_{n,m} \\
&\quad + \theta_{\mathrm{z}}\eta_n^{m+1}\frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n,-|m+1|} \\
&\quad = \eta_{n-1}^{-m}\frac{(-1)^{m+1}}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n-1,-|m+1|}\,, \qquad \text{(A.40)} \\
&\theta_{\mathrm{x}}\eta_n^{m}\frac{(-1)^m}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-m} - \theta_{\mathrm{y}}\eta_n^{m}\frac{(-1)^m}{\sqrt{2-\delta_m}}Y_{n,m} \\
&\quad - \theta_{\mathrm{z}}\eta_n^{-m+1}\frac{(-1)^{-m+1}}{\sqrt{2-\delta_{-m+1}}}(1-\delta_{-m+1})Y_{n,-|m+1|} \\
&\quad = -\eta_{n-1}^{m}\frac{(-1)^{m-1}}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n-1,-|m-1|}\,, \qquad \text{(A.41)}
\end{aligned}$$

then we take the real part of the equations for $m < 0$,

$$\begin{aligned}
&\theta_{\mathrm{x}}\eta_n^{-m}\frac{1}{\sqrt{2-\delta_m}}Y_{n,|m|} + \theta_{\mathrm{y}}\eta_n^{-m}\frac{1}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-|m|} + \theta_{\mathrm{z}}\eta_n^{m+1}\frac{1}{\sqrt{2-\delta_{m+1}}}Y_{n,|m+1|} \\
&\qquad = \eta_{n-1}^{-m}\frac{1}{\sqrt{2-\delta_{m+1}}}Y_{n-1,|m+1|}\,, \qquad \text{(A.42)} \\
&\theta_{\mathrm{x}}\eta_n^{m}\frac{1}{\sqrt{2-\delta_m}}Y_{n,|m|} - \theta_{\mathrm{y}}\eta_n^{m}\frac{1}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-|m|} - \theta_{\mathrm{z}}\eta_n^{-m+1}\frac{1}{\sqrt{2-\delta_{-m+1}}}Y_{n,-|m+1|} \\
&\qquad = -\eta_{n-1}^{m}\frac{1}{\sqrt{2-\delta_{m-1}}}Y_{n-1,|m-1|}\,, \qquad \text{(A.43)}
\end{aligned}$$

and again the imaginary part, $m < 0$,

$$\begin{aligned}
&-\theta_{\mathrm{x}}\eta_n^{-m}\frac{1}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-|m|} + \theta_{\mathrm{y}}\eta_n^{-m}\frac{1}{\sqrt{2-\delta_m}}Y_{n,|m|} \\
&\quad - \theta_{\mathrm{z}}\eta_n^{m+1}\frac{1}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n,-|m+1|} \\
&\quad = -\eta_{n-1}^{-m}\frac{1}{\sqrt{2-\delta_{m+1}}}(1-\delta_{m+1})Y_{n-1,-|m+1|}\,, \qquad \text{(A.44)} \\
&-\theta_{\mathrm{x}}\eta_n^{m}\frac{1}{\sqrt{2-\delta_m}}(1-\delta_m)Y_{n,-|m|} - \theta_{\mathrm{y}}\eta_n^{m}\frac{1}{\sqrt{2-\delta_m}}Y_{n,|m|} \\
&\quad + \theta_{\mathrm{z}}\eta_n^{-m+1}\frac{1}{\sqrt{2-\delta_{-m+1}}}(1-\delta_{-m+1})Y_{n,-|m+1|} \\
&\quad = \eta_{n-1}^{m}\frac{1}{\sqrt{2-\delta_{m-1}}}(1-\delta_{m-1})Y_{n-1,-|m-1|}\,. \qquad \text{(A.45)}
\end{aligned}$$

We notice that the equations from the first recurrence for $m > 0$ and $m < 0$ are linearly dependent with respect to the equations from the second recurrence for $m < 0$ and $m > 0$, leaving us with $4N - 2$ independent equations so far. Only the recurrences for $m = 0$ are missing. We denote the real part of the first recurrence explicitly for $m = 0$

$$\theta_x \eta_n^0 Y_{n,0} - \theta_z \eta_n^1 \frac{1}{\sqrt{2}} Y_{n,1} = -\eta_{n-1}^0 \frac{1}{\sqrt{2}} Y_{n-1,1}\,, \tag{A.46}$$

and do the same for the imaginary part,

$$-\theta_y \eta_n^0 Y_{n,0} + \theta_z \eta_n^1 \frac{1}{\sqrt{2}} Y_{n,-1} = \eta_{n-1}^0 \frac{1}{\sqrt{2}} Y_{n-1,-1}\,. \tag{A.47}$$

Note that the real and imaginary parts of the second recurrence yield the same results for $m = 0$ and can therefore be left out. Finally we found the additional $4N$ equations for the REVEB-ESPRIT.

A.4 Code Listings for the Derived Recurrence Relations

This section holds MATLAB code listings to obtain matrices containing the derived recurrence relations. Corresponding files are provided online [1].

```
function [Gx,Gy,Gz] = tg_gradient_matrix_real(N)
% tangential gradient addition theorems for real-valued SHs of
    maximum order N
% Franz Zotter and Thomas Deppisch, 2020

Nnm = @(n_,m_) (-1)^m_ * sqrt(((2*n_+1)*factorial(n_-abs(m_))*(2-(
    m_==0))) ./ (4*pi*factorial(n_+abs(m_))));
nm2acn = @(n_,m_) n_.^2 + n_ + m_ + 1;

Gx=zeros((N+1)^2,(N+2)^2);
Gy=zeros((N+1)^2,(N+2)^2);
Gz=zeros((N+1)^2,(N+2)^2);

for n=0:N
    for m=0:n
        if abs(m-1)<=n-1 && n-1>=0
            a=-(n+1)*(n+m-1)*(n+m)*Nnm(n,m)/((2*n+1)*Nnm(n-1,m-1))
    /2;
            if m==1
                Gx(nm2acn(n,m),nm2acn(n-1,abs(m-1)))=a;
                Gy(nm2acn(n,-m),nm2acn(n-1,abs(m-1)))=a;
            elseif m~=0
                Gx(nm2acn(n,m),nm2acn(n-1,abs(m-1)))=a;
                Gx(nm2acn(n,-m),nm2acn(n-1,-abs((m-1))))=a;
                Gy(nm2acn(n,-m),nm2acn(n-1,abs(m-1)))=a;
```

1. https://git.iem.at/thomasdeppisch/real-sh-recurrence-relations

```
                Gy(nm2acn(n,m),nm2acn(n-1,-abs((m-1))))=-a;
            end
        end

        if abs(m-1)<=n+1 && n+1<=N+1
            a=-n*(n-m+1)*(n-m+2)*Nnm(n,m)/((2*n+1)*Nnm(n+1,m-1))/2;
            if m==1
                Gx(nm2acn(n,m),nm2acn(n+1,abs(m-1)))=a;
                Gy(nm2acn(n,-m),nm2acn(n+1,abs(m-1)))=a;
            elseif m~=0
                Gx(nm2acn(n,m),nm2acn(n+1,abs(m-1)))=a;
                Gx(nm2acn(n,-m),nm2acn(n+1,-abs(m-1)))=a;
                Gy(nm2acn(n,-m),nm2acn(n+1,abs(m-1)))=a;
                Gy(nm2acn(n,m),nm2acn(n+1,-abs(m-1)))=-a;
            end
        end

        if abs(m+1)<=n-1 && n-1>=0
            a=(n+1)*Nnm(n,m)/((2*n+1)*Nnm(n-1,m+1))/2;
            if m==0
                Gx(nm2acn(n,m),nm2acn(n-1,m+1))=2*a;
                Gy(nm2acn(n,m),nm2acn(n-1,-(m+1)))=2*a;
            else
                Gx(nm2acn(n,m),nm2acn(n-1,m+1))=a;
                Gx(nm2acn(n,-m),nm2acn(n-1,-(m+1)))=a;
                Gy(nm2acn(n,-m),nm2acn(n-1,m+1))=-a;
                Gy(nm2acn(n,m),nm2acn(n-1,-(m+1)))=a;
            end
        end

        if abs(m+1)<=n+1 && n+1<=N+1
            a=n*Nnm(n,m)/((2*n+1)*Nnm(n+1,m+1))/2;
            if m==0
                Gx(nm2acn(n,m),nm2acn(n+1,m+1))=2*a;
                Gy(nm2acn(n,m),nm2acn(n+1,-(m+1)))=2*a;
            else
                Gx(nm2acn(n,m),nm2acn(n+1,m+1))=a;
                Gx(nm2acn(n,-m),nm2acn(n+1,-(m+1)))=a;
                Gy(nm2acn(n,-m),nm2acn(n+1,m+1))=-a;
                Gy(nm2acn(n,m),nm2acn(n+1,-(m+1)))=a;
            end
        end

        if abs(m)<=n-1 && n-1>=0
            a=(n+1)*(n+m)*Nnm(n,m)/((2*n+1)*Nnm(n-1,m));
            if m==0
                Gz(nm2acn(n,m),nm2acn(n-1,m))=a;
            else
                Gz(nm2acn(n,m),nm2acn(n-1,m))=a;
                Gz(nm2acn(n,-m),nm2acn(n-1,-m))=a;
            end
        end
        if abs(m)<=n+1 && n+1<=N+1
            a=-n*(n-m+1)*Nnm(n,m)/((2*n+1)*Nnm(n+1,m));
```

```
            if m==0
                Gz(nm2acn(n,m),nm2acn(n+1,m))=a;
            else
                Gz(nm2acn(n,m),nm2acn(n+1,m))=a;
                Gz(nm2acn(n,-m),nm2acn(n+1,-m))=a;
            end
        end
    end
end
```

```
function [Mx,My,Mz] = multiplication_theorems_real(N)
% multiplication theorems for real-valued SHs of maximum order N
% Franz Zotter and Thomas Deppisch, 2020

    Mx=zeros((N)^2,(N+1)^2);
    My=zeros((N)^2,(N+1)^2);
    Mz=zeros((N)^2,(N+1)^2);

    wnm = @(n_,m_) sqrt((n_+m_-1)*(n_+m_)/((2*n_-1)*(2*n_+1)));
    vnm = @(n_,m_) sqrt((n_-m_)*(n_+m_)/((2*n_-1)*(2*n_+1)));
    sm = @(m1_,m2_) sqrt(2-(m1_==0))/sqrt(2-(m2_==0));
    nm = @(n_,m_) n_.^2+n_+m_+1;

    for n=0:N-1
        for m=0:n
            %% Mx
            % n-1,m-1
            if (n-1>=0)&&(abs(m-1)<=n-1)
                Mx(nm(n,m),nm(n-1,abs(m-1))) = -wnm(n,m)/2*sm(m,m
-1) * (-1)^(m-1)/(-1)^m;
                if (m>1)
                    Mx(nm(n,-m),nm(n-1,-abs(m-1))) = -wnm(n,m)/2*sm
(m,m-1) * (-1)^(m-1)/(-1)^m;
                end
            end
            % n-1,m+1
            if (m+1<=n-1)&&(n-1>=0)
                Mx(nm(n,m),nm(n-1,(m+1))) = +wnm(n,-m)/2*sm(m,m+1)
*(1+(m==0)) * (-1)^(m+1)/(-1)^m;
                if m>0
                    Mx(nm(n,-m),nm(n-1,-(m+1))) = +wnm(n,-m)/2*sm(m
,m+1) * (-1)^(m+1)/(-1)^m;
                end
            end
            % n+1,m-1
            Mx(nm(n,m),nm(n+1,abs(m-1))) = +wnm(n+1,-m+1)/2*sm(m,m
-1) * (-1)^(m-1)/(-1)^m;
            if m>1
                Mx(nm(n,-m),nm(n+1,-abs(m-1))) = +wnm(n+1,-m+1)/2*
sm(m,m-1) * (-1)^(m-1)/(-1)^m;
            end
            % n+1,m+1
            Mx(nm(n,m),nm(n+1,(m+1))) = -wnm(n+1,m+1)/2*sm(m,m+1)
*(1+(m==0)) * (-1)^(m+1)/(-1)^m;
```

```
38              if m>0
39                  Mx(nm(n,-m),nm(n+1,-(m+1))) = -wnm(n+1,m+1)/2*sm(m,
    m+1) * (-1)^(m+1)/(-1)^m;
40              end
41              %% My
42              % n-1,m-1
43              if (n-1>=0)&&(abs(m-1)<=n-1)
44                  if (m~=1)
45                      My(nm(n,m),nm(n-1,-abs(m-1))) = wnm(n,m)/2*sm(m
    ,m-1) * (-1)^(m-1)/(-1)^m;
46                  end
47                  if (m>0)
48                      My(nm(n,-m),nm(n-1,abs(m-1))) = -wnm(n,m)/2*sm(
    m,m-1) * (-1)^(m-1)/(-1)^m;
49                  end
50              end
51              % n-1,m+1
52              if (m+1<=n-1)&&(n-1>=0)
53                  My(nm(n,m),nm(n-1,-(m+1))) = wnm(n,-m)/2*sm(m,m+1)
    *(1+(m==0)) * (-1)^(m+1)/(-1)^m;
54                  if (m>0)
55                      My(nm(n,-m),nm(n-1,(m+1))) = -wnm(n,-m)/2*sm(m,
    m+1) * (-1)^(m+1)/(-1)^m;
56                  end
57              end
58              % n+1,m-1
59              if (m>1)
60                  My(nm(n,m),nm(n+1,-abs(m-1))) = -wnm(n+1,-m+1)/2*sm
    (m,m-1) * (-1)^(m-1)/(-1)^m;
61              end
62              if (m>0)
63                  My(nm(n,-m),nm(n+1,abs(m-1))) = +wnm(n+1,-m+1)/2*sm
    (m,m-1) * (-1)^(m-1)/(-1)^m;
64              end
65              % n+1,m+1
66              My(nm(n,m),nm(n+1,-(m+1))) = -wnm(n+1,m+1)/2*sm(m,m+1)
    *(1+(m==0)) * (-1)^(m+1)/(-1)^m;
67              if (m>0)
68                  My(nm(n,-m),nm(n+1,(m+1))) = wnm(n+1,m+1)/2*sm(m,m
    +1) * (-1)^(m+1)/(-1)^m;
69              end
70              %% Mz
71              if (n>0)&&(m<=n-1)
72                  Mz(nm(n,m),nm(n-1,m)) = vnm(n,m);
73                  if m>0
74                      Mz(nm(n,-m),nm(n-1,-m)) = vnm(n,m);
75                  end
76              end
77              Mz(nm(n,m),nm(n+1,m)) = vnm(n+1,m);
78              if m>0
79                  Mz(nm(n,-m),nm(n+1,-m)) = vnm(n+1,m);
80              end
81
82          end
```

```
    end
end
```

```
function [ML,CL] = extended_multiplication_theorems_real(N)
% extension theorems for real-valued SHs of maximum order N
% Franz Zotter and Thomas Deppisch, 2020

    h_nm = @(n_,m_) sqrt((n_+m_) ./ (2*n_+1));
    nm2acn = @(n_,m_) n_.^2 + n_ + m_ + 1;
    cspsqrt = @(m_) (-1)^m_ / sqrt(2 - (m_==0));
    omdm = @(m_) (1 - (m_==0));

    L = (N+1)^2;
    % first equation
    A = zeros(2*N, L); % theta_x
    B = A; % theta_y
    C = A; % theta_z
    D = A; % right side of eq
    % second equation
    A2 = A; % theta_x
    B2 = A; % theta_y
    C2 = A; % theta_z
    D2 = A; % right side of eq

    % first recurrence
    idx = 1;
    for m = 0:N-1
        if (m==0)
            A(idx, nm2acn(N,0)) = h_nm(N,0);
            C(idx, nm2acn(N,1)) = -h_nm(N,1) / sqrt(2);
            if (N>1)
                D(idx, nm2acn(N-1,1)) = -h_nm(N-1,0) / sqrt(2);
            end

            B(idx+1, nm2acn(N,0)) = -h_nm(N,0);
            C(idx+1, nm2acn(N,-1)) = h_nm(N,1) / sqrt(2);
            if (N>1)
                D(idx+1, nm2acn(N-1,-1)) = h_nm(N-1,0) / sqrt(2);
            end
        else
            % real, m>0
            A(idx, nm2acn(N,m)) = h_nm(N,-m) * cspsqrt(m);
            B(idx, nm2acn(N,-m)) = -h_nm(N,-m) * cspsqrt(m) * omdm(
    m);
            C(idx, nm2acn(N,m+1)) = h_nm(N,m+1) * cspsqrt(m+1);
            if (m<N-1)
                D(idx, nm2acn(N-1,m+1)) = h_nm(N-1,-m) * cspsqrt(m
    +1);
            end

            % imaginary, m>0
            A(idx+1, nm2acn(N,-m)) = h_nm(N,-m) * cspsqrt(m) * omdm
    (m);
            B(idx+1, nm2acn(N,m)) = h_nm(N,-m) * cspsqrt(m);
```

```
            C(idx+1, nm2acn(N,-(m+1))) = h_nm(N,m+1) * cspsqrt(m+1) * omdm(m+1);
            if (m<N-1)
                D(idx+1, nm2acn(N-1,-(m+1))) = h_nm(N-1,-m) * cspsqrt(m+1) * omdm(m+1);
            end
        end

        idx = idx + 2;
    end

    % second recurrence
    idx = 1;
    for m = 1:N
        % real, m>0
        A2(idx, nm2acn(N,m)) = h_nm(N,m) * cspsqrt(m);
        B2(idx, nm2acn(N,-m)) = h_nm(N,m) * cspsqrt(m) * omdm(m);
        C2(idx, nm2acn(N,abs(-m+1))) = -h_nm(N,-m+1) * cspsqrt(-m+1);
        D2(idx, nm2acn(N-1,m-1)) = -h_nm(N-1,m) * cspsqrt(m-1);

        % imaginary, m>0
        A2(idx+1, nm2acn(N,-m)) = h_nm(N,m) * cspsqrt(m) * omdm(m);
        B2(idx+1, nm2acn(N,m)) = -h_nm(N,m) * cspsqrt(m);
        C2(idx+1, nm2acn(N,-abs(-m+1))) = -h_nm(N,-m+1) * cspsqrt(-m+1) * omdm(-m+1);
        D2(idx+1, nm2acn(N-1,-(m-1))) = -h_nm(N-1,m) * cspsqrt(m-1) * omdm(m-1);

        idx = idx + 2;
    end

    ML = [A, B, C;
          A2, B2, C2];

    CL = [D; D2];

end
```

Appendix B

Extended Pseudo-Intensity Vector from Real Recurrences

If the real-valued VEB-ESPRIT is reformulated for one source and one observation $\boldsymbol{y}_N$ as done for the complex-valued VEB-ESPRIT in [HH19b], a simple formulation for the extended pseudo-intensity vector (PIV) is obtained as the least squares solutions

$$x = \frac{\boldsymbol{y}_N^{\mathrm{T}} \boldsymbol{M}^{\mathrm{T}} \boldsymbol{M}_x \boldsymbol{y}_N}{||\boldsymbol{M}\boldsymbol{y}_N||^2} \tag{B.1}$$

$$y = \frac{\boldsymbol{y}_N^{\mathrm{T}} \boldsymbol{M}^{\mathrm{T}} \boldsymbol{M}_y \boldsymbol{y}_N}{||\boldsymbol{M}\boldsymbol{y}_N||^2} \tag{B.2}$$

$$z = \frac{\boldsymbol{y}_N^{\mathrm{T}} \boldsymbol{M}^{\mathrm{T}} \boldsymbol{M}_z \boldsymbol{y}_N}{||\boldsymbol{M}\boldsymbol{y}_N||^2} \tag{B.3}$$

of

$$\boldsymbol{M}\boldsymbol{y}_N x = \boldsymbol{M}_x \boldsymbol{y}_N \,, \tag{B.4}$$

$$\boldsymbol{M}\boldsymbol{y}_N y = \boldsymbol{M}_y \boldsymbol{y}_N \,, \tag{B.5}$$

$$\boldsymbol{M}\boldsymbol{y}_N z = \boldsymbol{M}_z \boldsymbol{y}_N \,, \tag{B.6}$$

where $\boldsymbol{M} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix}$ reduces $\boldsymbol{y}_N$ by one order and $\boldsymbol{M}_x$, $\boldsymbol{M}_y$, $\boldsymbol{M}_z$ hold multiplication recurrences of real-valued spherical harmonics (cf. section 2.4.2).

Appendix C

Why eigenvalue n-tuples of joint eigendecomposition are displayed as consistent diagonal-entry tuples in joint Schur decomposition

The following shows that a simultaneous Schur decomposition $\boldsymbol{U}_k = \boldsymbol{Q}^{\mathrm{H}} \boldsymbol{A}_k \boldsymbol{Q}$ manages to preserve the order of the eigenvalues of the simultaneously diagonalisable matrices $\boldsymbol{A}_k$ and hence is effective for simultaneous DOA estimation via matrix decomposition, e.g. in an EB-ESPRIT context. Simultaneous diagonalisation of matrices $\boldsymbol{A}_k$ is achieved via a similarity transform using the common eigenvectors in matrix $\boldsymbol{X}$,

$$\boldsymbol{A}_k \boldsymbol{X} = \boldsymbol{X} \operatorname{diag}(\boldsymbol{\lambda}_k) \,. \tag{C.1}$$

Any real, square matrix $\boldsymbol{A}$ may be decomposed into an orthogonal matrix $\boldsymbol{Q}$ and an upper triangular matrix $\boldsymbol{R}$ by means of QR decomposition. We apply the QR decomposition to the eigenvector matrix $\boldsymbol{X}$ and get

$$\boldsymbol{A}_k \boldsymbol{Q} \boldsymbol{R} = \boldsymbol{Q} \boldsymbol{R} \operatorname{diag}(\boldsymbol{\lambda}_k) \,, \tag{C.2}$$

or equivalently

$$\boldsymbol{Q}^{\mathrm{H}} \boldsymbol{A}_k \boldsymbol{Q} = \boldsymbol{R} \operatorname{diag}(\boldsymbol{\lambda}_k) \boldsymbol{R}^{-1} \,, \tag{C.3}$$

where $\boldsymbol{R}$ is invertible if $\boldsymbol{A}_k$ is diagonalisable. The upper triangular matrix $\boldsymbol{R}$ can be expressed as the sum of a diagonal matrix $\operatorname{diag}(\boldsymbol{r})$ and a strictly-upper triangular and hence nilpotent matrix $\boldsymbol{N}$. The inverse of upper triangular matrices is still upper triangular and holds the inverted eigenvalues on its main diagonal. Hence we can also express the inverse as a sum of diagonal and strictly-upper matrices, $\boldsymbol{R}^{-1} := \operatorname{diag}(\boldsymbol{r})^{-1} + \tilde{\boldsymbol{N}}$, and rearrange the similarity transform as

$$\boldsymbol{Q}^{\mathrm{H}} \boldsymbol{A}_k \boldsymbol{Q} = (\operatorname{diag}(\boldsymbol{r}) + \boldsymbol{N}) \operatorname{diag}(\boldsymbol{\lambda}_k) (\operatorname{diag}(\boldsymbol{r})^{-1} + \tilde{\boldsymbol{N}}) \,. \tag{C.4}$$

After expansion of the product we get

$$\begin{aligned}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{Q} = \mathrm{diag}(\boldsymbol{r})\mathrm{diag}(\boldsymbol{\lambda}_k)\mathrm{diag}(\boldsymbol{r})^{-1} + \mathrm{diag}(\boldsymbol{r})\mathrm{diag}(\boldsymbol{\lambda}_k)\tilde{\boldsymbol{N}} \\ + \boldsymbol{N}\mathrm{diag}(\boldsymbol{\lambda}_k)\mathrm{diag}(\boldsymbol{r})^{-1} + \boldsymbol{N}\mathrm{diag}(\boldsymbol{\lambda}_k)\tilde{\boldsymbol{N}}\,. \end{aligned} \tag{C.5}$$

On the left side we recognise the Schur decomposition $\boldsymbol{U}_k = \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{Q}$ and after simplification on right side we get

$$\boldsymbol{U}_k = \mathrm{diag}(\boldsymbol{\lambda}_k) + \mathrm{diag}(\boldsymbol{r})\mathrm{diag}(\boldsymbol{\lambda}_k)\tilde{\boldsymbol{N}} + \boldsymbol{N}\mathrm{diag}(\boldsymbol{\lambda}_k)\mathrm{diag}(\boldsymbol{r})^{-1} + \boldsymbol{N}\mathrm{diag}(\boldsymbol{\lambda}_k)\tilde{\boldsymbol{N}}\,. \tag{C.6}$$

As the product of (strictly) upper triangular matrices is also (strictly) upper triangular, the terms containing the strictly-upper triangular, nilpotent matrix $\boldsymbol{N}$ or $\tilde{\boldsymbol{N}}$ do not contribute any non-zero values to the main diagonal such that we find the eigenvalues on the main diagonal of $\boldsymbol{U}_k$ as

$$\mathrm{diag}(\boldsymbol{U}_k) = \boldsymbol{\lambda}_k\,, \tag{C.7}$$

which proves the effectiveness of the simultaneous Schur decomposition.

Appendix D

Additions to the Listening Experiment

D.1 Compared ASDM Equalisation Methods

The compared methods of spectral equalisation for the ASDM are based on the RMS equalisation in [ZF19, chaper 5]. The condition `ASDM-PO` is equalised per-order,

$$\tilde{h}^m_{n,\mathrm{eq}}(t,b) = \tilde{h}^m_n(t,b)\sqrt{\frac{(2n+1)\left(1/2\,\mathcal{E}\{|h^0_0(t,b)|^2\} + 1/6\sum_{m'=-1}^{1}\mathcal{E}\{|h^{m'}_1(t,b)|^2\}\right)}{\sum_{m=-n}^{n}\mathcal{E}\{|\tilde{h}^m_n(t,b)|^2\}}}\,, \tag{D.1}$$

while the condition `ASDM-M` is equalised via the mean RMS value, as done for the HO-ASDM (eq. 3.7),

$$\tilde{h}^m_{n,\mathrm{eq}}(t,b) = \tilde{h}^m_n(t,b)\sqrt{\frac{1/(N+1)^2\sum_{n'=0}^{N}\sum_{m'=-n'}^{n'}\mathcal{E}\{|h^{m'}_{n'}(t,b)|^2\}}{1/(\tilde{N}+1)^2\sum_{n=0}^{\tilde{N}}\sum_{m=-n}^{n}\mathcal{E}\{|\tilde{h}^m_n(t,b)|^2\}}}\,. \tag{D.2}$$

D.2 Extensive Results

condition pair	L13		L06		M		L13 +70°		L06 +70°		M +70°	
	p_{BH}	r_{c}	p_{BH}	r_{c}	p_{BH}	r_{c}	p_{BH}	r_{c}	p_{BH}	r_{c}	p_{BH}	r_{c}
o1, o1-o4	0.0205*	0.7727	0.3477	0.4848	0.0059*	0.9545	0.3311	0.3485	0.1982	0.4545	0.0049*	0.9091
o1, o2-o4	0.0146*	1.0000	0.0146*	1.0000	0.0078*	0.9697	0.0146*	1.0000	0.0146*	1.0000	0.0146*	1.0000
o1, o3-o4	0.0146*	1.0000	0.0146*	1.0000	0.0078*	0.9697	0.0146*	1.0000	0.0146*	1.0000	0.0146*	1.0000
o1, ASDM-PO	0.0127*	1.0000	0.0127*	1.0000	0.0137*	1.0000	0.0127*	1.0000	0.0127*	1.0000	0.0127*	1.0000
o1, ASDM-M	0.0117*	1.0000	0.0059*	0.9697	0.0127*	1.0000	0.0117*	1.0000	0.0117*	1.0000	0.0117*	1.0000
o1-o4, o2-o4	0.0107*	1.0000	0.0117*	1.0000	0.0117*	1.0000	0.0107*	1.0000	0.0107*	1.0000	0.0107*	1.0000
o1-o4, o3-o4	0.0098*	1.0000	0.0107*	1.0000	0.0107*	1.0000	0.0098*	1.0000	0.0098*	1.0000	0.0098*	1.0000
o1-o4, ASDM-PO	0.0088*	1.0000	0.0098*	1.0000	0.0098*	1.0000	0.0088*	1.0000	0.0088*	1.0000	0.0088*	1.0000
o1-o4, ASDM-M	0.0078*	1.0000	0.0088*	1.0000	0.0088*	1.0000	0.0078*	1.0000	0.0078*	1.0000	0.0078*	1.0000
o2-o4, o3-o4	0.0039*	0.9848	0.2188	0.5455	0.0078*	0.9091	0.0039*	0.9848	0.0078*	0.9545	0.0078*	0.9242
o2-o4, ASDM-PO	0.0068*	1.0000	0.0078*	1.0000	0.0078*	1.0000	0.0068*	1.0000	0.0068*	1.0000	0.0068*	1.0000
o2-o4, ASDM-M	0.0059*	1.0000	0.0068*	1.0000	0.0068*	1.0000	0.0059*	1.0000	0.0059*	1.0000	0.0059*	1.0000
o3-o4, ASDM-PO	0.0049*	1.0000	0.0059*	1.0000	0.0059*	1.0000	0.0049*	1.0000	0.0049*	1.0000	0.0049*	1.0000
o3-o4, ASDM-M	0.0039*	1.0000	0.0049*	1.0000	0.0049*	1.0000	0.0039*	1.0000	0.0039*	1.0000	0.0039*	1.0000
ASDM-PO, ASDM-M	0.0029*	1.0000	0.0039*	1.0000	0.0059*	0.9848	0.0029*	1.0000	0.0029*	1.0000	0.0029*	1.0000

Table D.1 – Bonferroni-Holm corrected p-values p_{BH} of the pairwise Wilcoxon signed-rank test and corresponding matched-pairs rank biserial correlation coefficient r_{c} for each trial and each condition pair. Statistical significance at $\alpha = 0.05$ is marked by an asterisk.

Bibliography

[AB79] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979. [Online]. Available: http://asa.scitation.org/doi/10.1121/1.382599

[AB20] J. Ahrens and S. Bilbao, "Computation of Spherical Harmonic Representations of Source Directivity Based on the Finite-Distance Signature," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. PP, 2020.

[Bar83] A. J. Barabell, "Improving the Resolution Performance of Eigenstructure-Based Direction-Finding Algorithms." *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 1, pp. 336–339, 1983.

[BB10] S. Berge and N. Barrett, "High Angular Resolution Planewave Expansion," in *Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, Paris, 2010.

[BB11] ——, "A new method for B-format to binaural Transcoding," *Proceedings of the AES International Conference*, pp. 1–10, 2011.

[BK79] G. Bienvenu and L. Kopp, "Principe de la goniométrie passive adaptative," in *7. Colloque sur le traitement du signal et des images*. Nice, France: GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 1979, pp. 106/1–106/10.

[BMKF20] M. Brandner, N. Meyer-Kahlen, and M. Frank, "Directivity pattern measurement of a grand piano for augmented acoustic reality," *Proc. DAGA*, pp. 138–141, 2020.

[BPVR11] I. Ben Hagai, M. Pollow, M. Vorländer, and B. Rafaely, "Acoustic centering of sources measured by surrounding spherical microphone arrays," *The Journal of the Acoustical Society of America*, vol. 130, no. 4, pp. 2003–2015, 2011.

[CG77] P. G. Craven and M. A. Gerzon, "Coincident microphone simulation covering three dimensional space and yielding various directional outputs," 1977.

[Che92] W. C. Chew, "Recurrence Relations for Three-Dimensional Scalar Addition Theorem," *Journal of Electromagnetic Waves and Applications*, vol. 6, no. 2, pp. 133–142, 1992.

[CKR09] M. Costa, V. Koivunen, and A. Richter, "Low complexity azimuth and elevation estimation for arbitrary array configurations," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, no. July, pp. 2185–2188, 2009.

[CMZ+09] M. Chapman, T. Musil, J. Zmölnig, H. Pomberger, F. Zotter, A. Sontacchi, W. Ritsch, T. Musil, and A. Sontacchi, "A STANDARD FOR INTERCHANGE OF AMBISONIC SIGNAL SETS Including a file standard with metadata," *International Symposium on Ambisonics and Spherical Acoustics*, vol. 1996, pp. 1–6, 2009. [Online]. Available: http://iem.kug.ac.at/fileadmin/media/iem/projects/2009/ambixchange09.pdf

[CS71] D. H. Cooper and T. Shiga, "Discrete-Matrix Multichannel Stereo," in *Proc. of the 41st AES Conv.*, New York, 1971, pp. 346–360.

[Dan01] J. Daniel, "Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia," Ph.D. dissertation, Université de Paris, 2001.

[DZ10] D. Deboy and F. Zotter, "Acoustic center and orientation analysis of sound-radiation recorded with a surrounding spherical microphone array," in *2nd International Symposium on Ambisonics and Spherical Acoustics*, Paris, 2010.

[DZ11] ——, "Tangential Intensity Algorithm for Acoustic Centering," in *Fortschritte der Akustik, DAGA*, 2011.

[Far79] K. Farrar, "Soundfield microphone," pp. 48–50, 1979.

[FZ16] M. Frank and F. Zotter, "Spatial impression and directional resolution in the reproduction of reverberation," in *Proc. DAGA*, 2016, pp. 1304–1307.

[GD01] N. A. Gumerov and R. Duraiswami, "Fast, exact, and stable computation of multipole translation and rotation coefficients for the 3-d helmholtz equation," *Institute for Advanced Computer Studies Technical Report*, vol. 44, p. 45, 2001. [Online]. Available: http://drum.lib.umd.edu/handle/1903/1141

[GD04] ——, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Amsterdam: Elsevier Science, 2004.

[Ger73] M. A. Gerzon, "Periphohy: With-Height Sound Reproduction." *AES: Journal of the Audio Engineering Society*, vol. 21, no. 1, pp. 2–10, 1973.

[Ger75] ——, "The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound," *J. Audio Eng. Soc.*, vol. 23, pp. 402–404, 1975.

[Ger77] ——, "Design of Ambisonic Decoders for Multispeaker Surround Sound," in *Proc. of the 58th AES Convention*, 1977.

[Ger85] ——, "Ambisonics in multichannel broadcasting and video," *Journal of the Audio Engineering Society*, vol. 33, no. 11, pp. 859–871, 1985.

[GP11] M. Gräf and D. Potts, "On the computation of spherical designs by a new optimization approach based on fast spherical Fourier transforms," *Numerische Mathematik*, vol. 119, no. 4, pp. 699–724, 2011.

[GP19] G. Götz and V. Pulkki, "Simplified Source Directivity Rendering in Acoustic Virtual Reality using the Directivity Sample Combination," *AES 147th Convention Paper*, pp. 1–11, 2019.

[GZ20] L. Gölles and F. Zotter, "Directional Enhancement of First-Order Ambisonic Room Impulse Responses By The 2 + 2 Directional Signal Estimator," in *Proceedings of the 15th International Conference on Audio Mostly*. Graz: Association for Computing Machinery, 2020.

[HC20] Q. Huang and T. Chen, "One-Dimensional MUSIC-Type Algorithm for Spherical Microphone Arrays," *IEEE Access*, vol. 8, pp. 28 178–28 187, 2020.

[HCXC10] Z. He, A. Cichocki, S. Xie, and K. Choi, "Detecting the number of clusters in n-way probabilistic clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 2006–2021, 2010.

[HH19a] A. Herzog and E. A. P. Habets, "Eigenbeam-ESPRIT for DOA-Vector estimation," *IEEE Signal Processing Letters*, vol. 26, no. 4, pp. 572–576, 2019.

[HH19b] ——, "On the Relation Between DOA-Vector Eigenbeam ESPRIT and Subspace Pseudointensity-Vector," *Eusipco*, no. 1, pp. 1–5, 2019.

[Hol79] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian journal of statistics*, vol. 6, no. 2, pp. 65–70, 1979.

[HS96] R. H. Hardin and N. J. Sloane, "McLaren's improved snub cube and other new spherical designs in three dimensions," *Discrete and Computational Geometry*, vol. 15, no. 4, pp. 429–441, 1996.

[HZF18] Q. Huang, L. Zhang, and Y. Fang, "Two-Step Spherical Harmonics ESPRIT-Type Algorithms and Performance Analysis," *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 26, no. 9, pp. 1684–1697, 2018.

[JC18] B. Jo and J.-W. Choi, "Direction of arrival estimation using nonsingular spherical ESPRIT," *The Journal of the Acoustical Society of America*, vol. 143, no. 3, pp. EL181–EL187, 2018. [Online]. Available: http://dx.doi.org/10.1121/1.5026122

[JC19] ——, "Parametric direction-of-arrival estimation with three recurrence relations of spherical harmonics," *The Journal of the Acoustical Society of America*, vol. 145, no. 1, pp. 480–488, 2019. [Online]. Available: http://dx.doi.org/10.1121/1.5087698

[JHN17] D. P. Jarrett, E. A. P. Habets, and P. A. Naylor, *Theory and Applications of Spherical Microphone Array Processing*. Springer, 2017.

[JHTN12] D. P. Jarrett, E. A. P. Habets, M. R. P. Thomas, and P. A. Naylor, "Rigid sphere room impulse response simulation: Algorithm and applications," *The Journal of the Acoustical Society of America*, vol. 132, no. 3, pp. 1462–1472, 2012.

[JZC20] B. Jo, F. Zotter, and J.-W. Choi, "Extended Vector-Based EB-ESPRIT Method," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1692–1705, 2020.

[KBH16] L. Kumar, G. Bi, and R. M. Hegde, "The spherical harmonics root-music," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2016-May, no. March, pp. 3046–3050, 2016.

[KR12] D. Khaykin and B. Rafaely, "Acoustic analysis by spherical microphone array processing of room impulse responses," *The Journal of the Acoustical Society of America*, vol. 132, no. 1, pp. 261–270, 2012.

[KRM18] B. M. King, P. J. Rosopa, and E. W. Minium, *Statistical Reasoning in the Behavioral Sciences*, 7th ed. Wiley, 2018.

[LYMH11] X. Li, S. Yan, X. Ma, and C. Hou, "Spherical harmonics MUSIC versus conventional MUSIC," *Applied Acoustics*, vol. 72, no. 9, pp. 646–652, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.apacoust.2011.02.010

[LZ15] S. Lösler and F. Zotter, "Comprehensive Radial Filter Design for Practical higher-order Ambisonic Recording," in *Fortschritte der Akustik, DAGA*, no. 1, Nürnberg, 2015, pp. 452–455.

[MDB06] S. Moreau, J. Daniel, and S. Bertet, "3D Sound Field Recording with Higher Order Ambisonics - Objective Measurements and Validation of a 4th Order Spherical Microphone," in *Proc. of the 120th Convention of the Audio Eng. Soc.*, 2006.

[MDM18] C. Mendonça and S. Delikaris-Manias, "Statistical tests with MUSHRA data," in *144th Audio Engineering Society Convention 2018*, 2018.

[ME02] J. Meyer and G. W. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 2002, pp. II–1781–II–1784.

[MP05] J. Merimaa and V. Pulkki, "Spatial impulse response rendering I: Analysis and synthesis," *AES: Journal of the Audio Engineering Society*, vol. 53, no. 12, pp. 1115–1127, 2005.

[MPP+20] L. McCormack, V. Pulkki, A. Politis, O. Scheuregger, and M. Marschall, "Higher-Order Spatial Impulse Response Rendering: Investigating the Perceived Effects of Spherical Order, Dedicated Diffuse Rendering, and Frequency Resolution," *Journal of the Audio Engineering Society*, vol. 68, no. 5, pp. 338–354, 2020.

[MZ94] C. P. Mathews and M. D. Zoltowski, "Eigenstructure Techniques for 2-D Angle Estimation with Uniform Circular Arrays," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2395–2407, 1994.

[PM06] V. Pulkki and J. Merimaa, "Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests," *AES: Journal of the Audio Engineering Society*, vol. 54, no. 1-2, pp. 3–20, 2006.

[Pol15] M. Pollow, "Directivity Patterns for Room Acoustical Measurements and Simulations," Ph.D. dissertation, RWTH Aachen, 2015.

[PTP18] A. Politis, S. Tervo, and V. Pulkki, "COMPASS: Coding and multidirectional parameterization of ambisonic sound scenes," *ICASSP, IEEE International*

Conference on Acoustics, Speech and Signal Processing - Proceedings, vol. 2018-April, pp. 6802–6806, 2018.

[Pul97] V. Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, 1997.

[Raf04] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 2149–2157, 2004.

[RK89] R. Roy and T. Kailath, "ESPRIT - Estimation of Signal Parameters Via Rotational Invariance Techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[Sch86] R. O. Schmidt, "Multiple emitter location and parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. AP-34, no. March, pp. 276–280, 1986.

[SV15] N. R. Shabtai and M. Vorländer, "Acoustic centering of sources with high-order radiation patterns," *The Journal of the Acoustical Society of America*, vol. 137, no. 4, pp. 1947–1961, 2015.

[TB97] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, 1997.

[Teu07] H. Teutsch, *Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition*. Springer, 2007.

[TK08] H. Teutsch and W. Kellermann, "Detection and localization of multiple wideband acoustic sources based on wavefield decomposition using spherical apertures," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, no. 3, 2008, pp. 5276–5279.

[TPKL13] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial decomposition method for room impulse responses," *J. Audio Eng. Soc*, vol. 61, no. 1/2, pp. 17–28, 2013. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=16664

[WEMJ11] A. Wabnitz, N. Epain, A. McEwan, and C. T. Jin, "Upscaling Ambisonic sound scenes using compressed sensing techniques," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, no. June, 2011, pp. 1–4.

[Wil99] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. New York: Academic Press, 1999.

[Yan95] B. Yang, "Projection Approximation Subspace Tracking," *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 95–107, 1995.

[Zag19] F. Zagala, "Optimum-phase primal signal and radiation-filter modelling of musical instruments," Master's Thesis, University of Music and Performing Arts, Graz, 2019.

[ZDWC18] J. Zhuang, C. Duan, W. Wang, and Z. Chen, "Joint estimation of azimuth and elevation via manifold separation for arbitrary array structures," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 5585–5596, 2018.

[ZF12] F. Zotter and M. Frank, “All-round ambisonic panning and decoding,” *AES: Journal of the Audio Engineering Society*, vol. 60, no. 10, pp. 807–820, 2012.

[ZF19] ——, *Ambisonics, A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, 1st ed. Springer, 2019. [Online]. Available: https://www.springer.com/gp/book/9783030172060