

Toningenieur-Projekt

**Richtungsschärfung in ambisonischen Raumimpuls-
antworten erster Ordnung durch den 2+2
Richtungssignalschätzer**

Lukas Göllés

Betreuung: Dr. Franz Zotter

Graz, 12. Mai 2020



institut für elektronische musik und akustik



Zusammenfassung

Zur Virtualisierung eines Lautsprechers im Raum kann die Messung mit einem Tetraedermikrofon sehr hilfreich sein. Ein solches Tetraedermikrofonarray liefert üblicherweise nur Ambisonics 1. Ordnung und besitzt folglich nur eine sehr geringe Richtungsauflösung. In der Praxis haben sich Verfahren zur Verbesserung der Richtungsauflösung von ambisonischen Raumimpulsantworten (engl. ARIR) etabliert, in denen die Richtungsschätzung ein zentraler Aspekt ist. Der bekannte Algorithmus *Ambisonic Spatial Decomposition Method* (ASDM) führt eine Richtungszerlegung anhand des Pseudointensitätsvektors als Richtungsschätzer durch. Dieser Vektor detektiert eine einzige Richtung pro betrachtetem Zeitabschnitt/Abtastzeitpunkt. Ihm zu Grunde liegt die Annahme, dass zum Betrachtungszeitpunkt nur eine Richtung präsent sein darf. Für den Direktschall und die frühen Anteile einer Raumimpulsantwort mag das in der Regel genügen, jedoch nicht für den immer dichter werden und dann schließlich diffusen Nachhall. Um speziell den Übergang zwischen Direktschall- und Diffusanteil zu verbessern, schlägt diese Arbeit einen Doppelrichtungsschätzer vor. Die Technik wurde von der *High Angular Plane Wave Expansion* (HARPEX) motiviert und verwendet einen vereinfachten und auf Raumimpulsantworten zugeschnittenen Doppelrichtungsdetektor. Ein überarbeiteter Beamformer extrahiert die beiden Signale vervollständigt mit zwei Komplementärrichtungen eines optimalen, unregelmäßigen Tetraeders. In der Arbeit werden sämtliche Schritte zuerst technisch und dann in einem vergleichenden Hörversuch anhand gemessener Raumimpulsantworten überprüft.

Abstract

Tetrahedral microphone arrays can be used for measurements to virtualize a loudspeaker in a room. These arrays only deliver a first-order Ambisonic output and therefore their directional resolution is rather limited. In practice there are several methods to improve the resolution of Ambisonic Room Impulse Responses (ARIR), for which estimation of direction is a central aspect. In *Ambisonic Spatial Decomposition Method* (ASDM), an established algorithm, a directional decomposition is done based on the pseudo intensity vector as direction estimator. This vector detects a unique direction per sampling interval. It is based on the assumption that only one direction may be present at the time of observation. For direct sound and early reflections of a room impulse response this is usually sufficient, but not for the later parts of reverberations, which become dense and finally diffuse. To improve the transition between direct sound and diffuse sound, a dual-direction estimator is proposed. The technique was motivated by the *High Angular Plane Wave Expansion* (HARPEX) and uses a simplified dual-direction estimator tailored for room impulse responses. A revised beamformer extracts these two signals which are complemented by the signals of two more directions of an optimal, irregular tetrahedron. The paper presents a technical analysis of all steps and a comparative listening test based on measured room impulse responses.

Inhaltsverzeichnis

1	Einleitung	5
2	Theoretischer Überblick	7
2.1	Doppelrichtungsschätzer	7
2.2	Kurzzeitmittelung	9
2.3	Vervollständigen auf ein Tetraeder	10
2.4	Optimaler Öffnungswinkel	11
2.5	Rotation für beliebige F-L-O-Systeme	12
2.6	Beamforming und Reenkodieren	13
2.7	Spektrale Korrektur	14
2.8	Quellaufweitung für diffusen Nachhall	15
3	Technische Analyse	16
3.1	Richtungsdetektionsquote bei einer aktiven Wiedergaberichtung	16
3.2	Ergebnis für eine aktive Wiedergaberichtung	17
3.3	Ergebnis für zwei aktive Wiedergaberichtungen	19
3.4	Ergebnis für drei aktive Wiedergaberichtungen	22
3.5	Winkelabhängiger Fehler	23
4	Hörversuch	26
4.1	Vorbereitung	26
4.2	Durchführung	27
4.2.1	György-Ligeti-Saal	27
4.2.2	Dom im Berg	27
4.3	Ergebnisse	27
5	Schlussfolgerung und Ausblick	30

<i>L. Gölles: Richtungsschärfung durch den 2+2 Richtungssignalschätzer</i>	4
A Soundfield ST450: Signalamplituden nullter und erster Ordnung	31
B Konditionszahl der Tetraederenkodierungsmatrix und optimaler Öffnungswinkel ψ	32
C Singulärwertzerlegung der quadrierten Enkodierungsmatrix	34

1 Einleitung

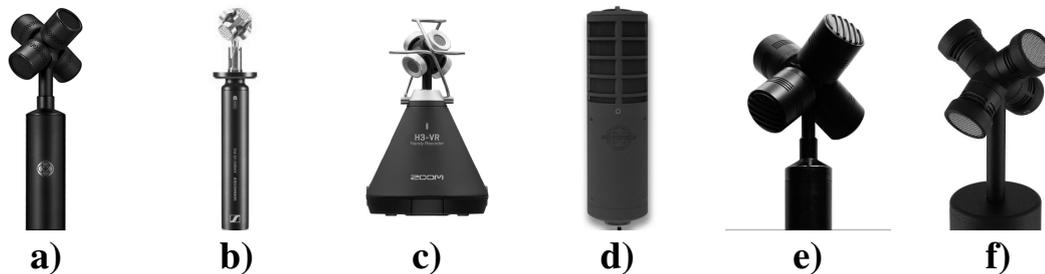


Abbildung 1 – Diverse Tetraedermikrofone: a) Røde NT-SF1 [Rod] b) Sennheiser Ambeo VR Mic [Sen] c) Zoom H3-VR [Zoo] d) Soundfield ST450 [SouB] e) Soundfield SPS200 [SouA] und f) Octava MK-4012 [Oct].

Mit Einführung der virtuellen Realität, die es bereits ermöglicht, Räume mit 3D-Brillen virtuell darzustellen, steigt auch dementsprechend das Interesse Lautsprecher in Räumen akustisch zu virtualisieren. Tetraedermikrofone, siehe Abbildung 1, als akustisches Pendant zur 3D-Kamera, erlauben es quell- und empfängerdirektionale ambisonische Raumpulsantworten (engl.: *Ambisonic Room Impulse Response*, ARIR) zu messen, um später einen Raum virtuell akustisch darstellen zu können. Jedoch ist die Richtungsauflösung dieser Raumpulsantworten erster Ordnung begrenzt. In der Praxis verwendet man daher unterschiedliche Algorithmen zur Richtungsauflösungsverbesserung.

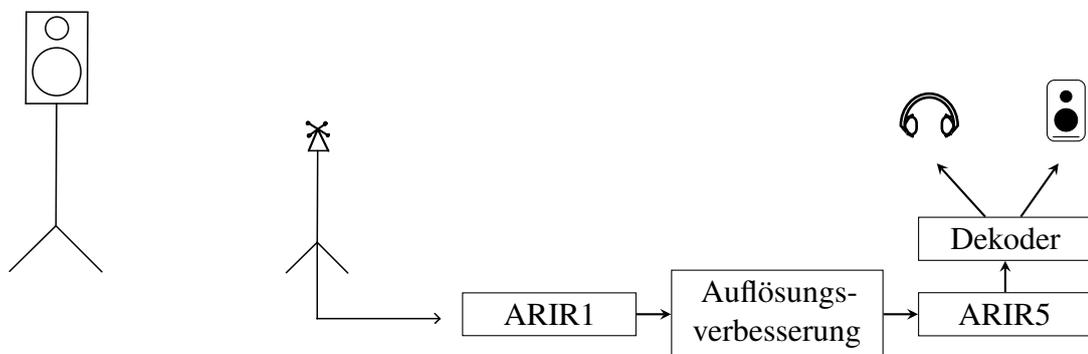


Abbildung 2 – Schematische Darstellung: Messung einer ambisonischen Raumpulsantwort erster Ordnung (ARIR1) mit anschließender Richtungsauflösungsverbesserung und Dekodierung der beispielsweise ambisonischen Raumpulsantwort fünfter Ordnung (ARIR5) auf Kopfhörer oder Lautsprecher.

Ein sehr verbreiteter und einfacher Algorithmus ist die *Ambisonic Spatial Decomposition Method* (ASDM), bei der versucht wird, mit dem Pseudointensitätsvektor die Richtung zu schätzen. Eine Multiplikation der Kugelflächenfunktionen, an den detektierten Richtungen ausgewertet, mit der omnidirektionalen Signalkomponente erlaubt die Berechnung einer ARIR beliebiger Ordnung [TPKL13], [Ahr19], [ZF19], [ZFZ20]. Eine andere Methode stellt *Spatial Impulse Response Rendering* (SIRR) [MP05], [PM06] dar, die jedoch

mit höherem Berechnungsaufwand verbunden ist und im Gegensatz zur ASDM im Frequenzbereich arbeitet.

Die Annahme, dass nur eine Richtung im Signal präsent ist, kann für den Direktschallanteil der Raumimpulsantwort angenommen werden. Jedoch im späten diffusen Teil kann eine einzige Richtung nicht mehr bestimmt werden, weil sich im idealen Diffusfeld der Schall in alle Richtungen gleichmäßig ohne Vorzugsrichtung ausbreitet [Deu06]. Um den Übergang von Direktschallanteil und Diffusanteil der Impulsantwort besser darstellen zu können, müssen Algorithmen ausgewählt werden, die mehr als zwei Richtungen schätzen.

Die Methode *High Angular Plane Wave Expansion* (HARPEX) bietet die Möglichkeit zwei Richtungen zu detektieren. In [BB10a] und [BB10b] wird vorgeschlagen, diese beiden mit zwei weiteren zu kombinieren, sodass ein unregelmäßiges Tetraeder entsteht. Angelehnt an diese Methode wird in dieser Arbeit ein Doppelrichtungsschätzer (engl.: *Dual-Direction Estimator*, DDE) präsentiert. Ein Beamformer extrahiert die Signale an den beiden geschätzten Richtungen, die mit zwei Komplementärrichtungen zu einem optimalen, unregelmäßigen Tetraeder vervollständigt werden, um durch die anschließende Reenkodierung Impulsantworten höherer ambisonischer Ordnung zu erhalten. Die Vorgehensweise wird unter dem Begriff *Richtungsschärfung durch den 2+2 Richtungssignalschätzer* (engl.: *Directional Sharpening By The 2+2 Directional Signal Estimator*, 2DSE2) zusammengefasst.

Diese Projektarbeit präsentiert zunächst den Doppelrichtungsschätzer, sowie die Vervollständigung auf ein unregelmäßiges Tetraeder mit Betracht auf optimaler Kondition der Tetraederenkodierungsmatrix. Es steht anschließend ein Vergleich der Detektion mit dem PIV im Vordergrund, sowie ein Test mit unterschiedlicher Richtungsanzahl im Signal. Ein vergleichender Hörversuch soll die theoretischen Ergebnisse unterstreichen.

2 Theoretischer Überblick

Folgender Abschnitt beleuchtet die Einführung eines Doppelrichtungsschätzers, angelehnt an HARPEX [BB10a], [BB10b], mit einer stark vereinfachten Berechnung und einer neuen Kurzzeitmittelungsmethode. Anschließend werden Berechnungsvorschriften für zwei weitere Richtungen, sowie deren optimaler Öffnungswinkel eingeführt. Nach dem Beamforming und der Reenkodierung erfolgt die spektrale Korrektur, die notwendig ist, um die typischer Weise auftretenden unnatürlichen Höhenanhebungen zu kompensieren.

2.1 Doppelrichtungsschätzer

Es wird zunächst die im B-Format aufgenommene Impulsantwort erster Ordnung \mathbf{b} durch den omidirektionalen Kanal W normiert, sodass die Information in den richtungsabhängigen Kanälen X, Y, Z steckt und vom omidirektionalen Kanal W unabhängig ist,

$$\frac{\mathbf{b}}{w} = \frac{1}{w} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \boldsymbol{\beta} \end{bmatrix}. \quad (1)$$

Ab hier ist für eine erfolgreiche Detektion zwingend erforderlich, dass die Signale in nullter und erster Ordnung zueinander korrekt angesteuert sind. Im Falle einer Aufnahme mit dem *Soundfield ST450* (Abbildung 1d)) muss der omidirektionale Kanal W mit $\sqrt{3}$ multipliziert werden (siehe Anhang A). Um eine stabile Richtungsdetektion gewährleisten zu können, wird die Anwendung eines Bandpassfilter mit den Grenzfrequenzen 200 Hz und 4000 Hz empfohlen [ZF19].

Nun wird das Vorhandensein zweier Signale s_1 und s_2 aus den Richtungen $\boldsymbol{\theta}_1$ und $\boldsymbol{\theta}_2$, mit $\|\boldsymbol{\theta}_1\| = \|\boldsymbol{\theta}_2\| = 1$, angenommen. Die normierte Impulsantwort $\frac{\mathbf{b}}{w}$ wird, wie folgt, aus den beiden Signalen ausgedrückt,

$$\frac{\mathbf{b}}{w} = \begin{bmatrix} 1 \\ \boldsymbol{\beta} \end{bmatrix} = \frac{s_1}{s_1 + s_2} \begin{bmatrix} 1 \\ \boldsymbol{\theta}_1 \end{bmatrix} + \frac{s_2}{s_1 + s_2} \begin{bmatrix} 1 \\ \boldsymbol{\theta}_2 \end{bmatrix}. \quad (2)$$

Die Substitution mit Hilfe eines Parameters γ , $\frac{s_1}{s_1 + s_2} = \gamma$ bzw. $\frac{s_2}{s_1 + s_2} = 1 - \gamma$ wird zur Erfüllung der Bedingung in der erste Zeile $1 = \frac{s_1}{s_1 + s_2} + \frac{s_2}{s_1 + s_2}$ verwendet. Sie reduziert das Problem auf,

$$\boldsymbol{\beta} = \gamma \boldsymbol{\theta}_1 + (1 - \gamma) \boldsymbol{\theta}_2. \quad (3)$$

Somit kann festgehalten werden, dass $\boldsymbol{\beta}$ Werte, die auf einer Linie $\boldsymbol{\beta}_g$, welche durch $\boldsymbol{\theta}_1$ und $\boldsymbol{\theta}_2$ festgelegt wird, annehmen kann. Diese beiden Richtungen $\boldsymbol{\theta}_1$ und $\boldsymbol{\theta}_2$ können als Schnittpunkt dieser Linie mit der Einheitskugel gesehen werden, $\boldsymbol{\theta}_{1,2} = \{\boldsymbol{\beta}_g : \|\boldsymbol{\beta}_g\|^2 = 1\}$. Die Linie wird dabei durch zwei unabhängige Beobachtungen $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$ festgelegt. Durch Differenzbildung $\Delta \boldsymbol{\beta} = \boldsymbol{\beta}_2 - \boldsymbol{\beta}_1$ mit anschließender Normierung $\mathbf{v} = \frac{\Delta \boldsymbol{\beta}}{\|\Delta \boldsymbol{\beta}\|}$ entsteht ein Einheitsrichtungsvektor, welcher zur Bildung einer Geradengleichung in Parameterform verwendet wird,

$$\boldsymbol{\beta}_g = \boldsymbol{\beta}_1 + g \mathbf{v} \quad \text{mit } g \in \mathbb{R}. \quad (4)$$

Zur Berechnung des Punktes mit kleinstem Radius auf der Geraden wird der optimale Parameter g der Gleichung $\beta_g = \beta_1 + g \mathbf{v}$ als Minimum der quadrierten euklidischen Norm

$$\|\beta_g\|^2 = \beta_g^T \beta_g = \beta_1^T \beta_1 + 2g \mathbf{v}^T \beta_1 + g^2 \mathbf{v}^T \mathbf{v} \quad (5)$$

gesucht. Dies wird durch Nullsetzung der Ableitung der Gleichung 5 nach g erreicht,

$$\frac{d\|\beta_g\|^2}{dg} = 2 \mathbf{v}^T \beta_1 + 2g \mathbf{v}^T \mathbf{v} \stackrel{!}{=} 0. \quad (6)$$

Da zuvor festgelegt wurde, dass $\|\mathbf{v}\| = 1$, vereinfacht sich die Gleichung zu

$$2 \mathbf{v}^T \beta_1 + 2g \mathbf{v}^T \mathbf{v} = 2 \mathbf{v}^T \beta_1 + 2g \stackrel{!}{=} 0, \quad (7)$$

wodurch sich für den Parameter g die Lösung

$$g = -\mathbf{v}^T \beta_1 \quad (8)$$

ergibt. Eingesetzt in Gleichung 4, sieht man, dass es sich hierbei um eine Orthogonalisierung nach Gram-Schmidt handelt [Rai],

$$\beta_{min} = \beta_1 - (\mathbf{v}^T \beta_1) \mathbf{v}. \quad (9)$$

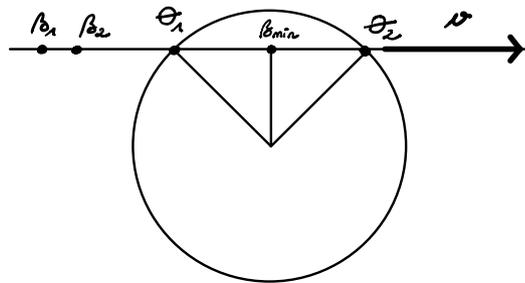


Abbildung 3 – Gram-Schmidt Orthogonalisierungsverfahren.

Die Richtungen θ_1 und θ_2 können schließlich mit Hilfe der Gleichung

$$\beta_g = \beta_{min} + h \mathbf{v} \quad \text{mit } h \in \mathbb{R} \quad (10)$$

gefunden werden. Zur Berechnung des Parameters h der Gleichung 10, wird zunächst die quadrierte euklidische Norm

$$\|\beta_g\|^2 = \beta_g^T \beta_g = \|\beta_{min}\|^2 + 2h \mathbf{v}^T \beta_1 + h^2 \mathbf{v}^T \mathbf{v} \quad (11)$$

benötigt. Da die Richtungsvektoren θ_1 und θ_2 die Schnittpunkte der Gerade mit der Einheitskugel sind, muss $\|\beta_g\|^2 = 1$ gelten. Weiters ergibt sich auf Grund des Orthogonalisierungsschrittes $\mathbf{v}^T \beta_{min} = 0$ und letztendlich

$$\|\beta_{min}\|^2 + h^2 = 1 \implies h = \pm \sqrt{1 - \|\beta_{min}\|^2} \quad (12)$$

Daher folgt die Berechnungsvorschrift für die Richtungen

$$\theta_{1,2} = \beta_{min} \pm \sqrt{1 - \|\beta_{min}\|^2} \mathbf{v}. \quad (13)$$

2.2 Kurzzeitmittelung

Um eine stabile Lösung für das Richtungspaar $\theta_{1,2}$ unter zeitlich variierenden Signalen zu gewährleisten, ist es notwendig Mittelwerte, mit dem quadrierten omnidirektionalen Signal gewichtet, zu berechnen. Damit entsteht einerseits ein stabiler Vektor β , andererseits ein stabiler Richtungsvektor für die Gerade $\Delta\beta$, der aus der Differenz $\Delta\beta = \beta[n] - \beta[n-1]$ errechnet wird, wobei n für den Zeitindex steht. Innerhalb eines Zeitfensters der Länge N kann die Mittelung mit Hilfe einer beliebigen Fensterfunktion $a[n]$ erfolgen,

$$\bar{\beta} = \frac{\sum_{n=-\frac{N}{2}}^{\frac{N}{2}} a[n] w^2[n] \beta[n]}{\sum_{n=-\frac{N}{2}}^{\frac{N}{2}} a[n] w^2[n]}, \quad (14)$$

$$\overline{\Delta\beta} = \frac{\sum_{n=-\frac{N}{2}}^{\frac{N}{2}} a[n] w^2[n] \Delta\beta[n] \text{sign}\{\Delta\beta_{ref}^T \Delta\beta[n]\}}{\sum_{n=-\frac{N}{2}}^{\frac{N}{2}} a[n] w^2[n]}. \quad (15)$$

Hier ist $\Delta\beta_{ref}$ ein Referenzrichtungsvektor, für den der Ausdruck $a[n] w^2[n] \|\Delta\beta[n]\|$ innerhalb des Zeitfensters maximal ist. Der Ausdruck $\text{sign}\{\Delta\beta_{ref}^T \Delta\beta[n]\}$ verhindert die destruktive Interferenz, die auftritt, falls die zwei Beobachtungen im Vorzeichen gegeneinander gerichtete Vektoren liefern. β_1 und $\Delta\beta$ aus dem vorigen Abschnitt werden durch $\bar{\beta}$ und $\overline{\Delta\beta}$ ersetzt, wodurch sich für den Gram-Schmidt'schen Orthogonalisierungsschritt ($\mathbf{v} = \frac{\overline{\Delta\beta}}{\|\overline{\Delta\beta}\|}$)

$$\beta_{min} = \bar{\beta} - (\mathbf{v}^T \bar{\beta}) \mathbf{v} \quad (16)$$

ergibt, der schließlich zum kurzzeitgemittelten Richtungspaar

$$\theta_{1,2} = \beta_{min} \pm \sqrt{(1 - \|\beta_{min}\|^2)} \mathbf{v} \quad (17)$$

führt. Für $h = \Re\{\sqrt{1 - \|\beta_{min}\|^2}\} = 0$, d.h. dass entweder kein oder nur ein Schnittpunkt mit der Einheitskugel existiert, ist es sinnvoll den Doppelrichtungsschätzer auf die Richtung von $\bar{\beta}$ überzublenzen. Zweckdienlich erweist sich die Einführung der Formel

$$\theta_{1,2} = \frac{d}{1+d} (\beta_{min} \pm h \mathbf{v}) + \frac{1}{1+d} \bar{\beta}, \quad (18)$$

mit dem Überblendparameter $d = \frac{0.01}{h}$.

2.3 Vervollständigen auf ein Tetraeder

Aus dem stabilen Richtungspaar $\theta_{1,2}$ wird in [BB10a] ein komplementäres Paar $\theta_{3,4}$ berechnet, das den diffusen Anteil behandelt. Diese vier Richtungen spannen ein (unregelmäßiges) Tetraeder auf.

Zur Berechnung des komplementären Richtungspaares wird zunächst von einem regulären Tetraeder mit den Richtungen frontal-links-oben (**FLO**), frontal-rechts-unten (**FRU**), hinten-rechts-oben (**HRO**) und hinten-links-unten (**HLU**) ausgegangen,

$$\begin{bmatrix} \mathbf{FLO} \\ \mathbf{FRU} \\ \mathbf{HRO} \\ \mathbf{HLU} \end{bmatrix} = \begin{bmatrix} \theta_1^T \\ \theta_2^T \\ \theta_3^T \\ \theta_4^T \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} +1 & +1 & +1 \\ +1 & -1 & -1 \\ -1 & -1 & +1 \\ -1 & +1 & -1 \end{bmatrix}. \quad (19)$$

Zur Vereinfachung der nun folgenden Betrachtungen ist es wünschenswert, dass die Richtung von **F** der x-Achse $\theta_x^T = [1 \ 0 \ 0]^T$ entspricht, sodass zwei Richtungen vorne links bzw. rechts und zwei Richtungen hinten oben bzw. unten existieren. Zur Erfüllung dieser Kriterien muss das Tetraeder aus Gleichung 19 um die x-Achse um -45° gedreht werden, siehe Abbildung 4.

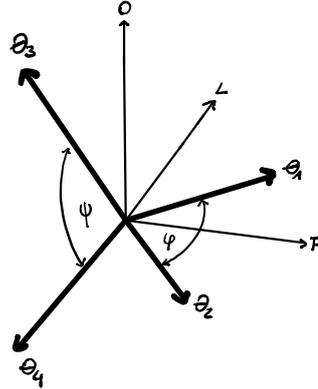


Abbildung 4 – Um -45° in x-Richtung rotierte Richtungsvektoren θ_1 , θ_2 , θ_3 und θ_4 .

$$\begin{bmatrix} \mathbf{FL} \\ \mathbf{FR} \\ \mathbf{HO} \\ \mathbf{HU} \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} +1 & +1 & +1 \\ +1 & -1 & -1 \\ -1 & -1 & +1 \\ -1 & +1 & -1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & +\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & +\frac{1}{\sqrt{2}} & +\frac{1}{\sqrt{2}} \end{bmatrix}}_{R_x^T(-45^\circ)} = \frac{1}{\sqrt{3}} \begin{bmatrix} +1 & +\sqrt{2} & 0 \\ +1 & -\sqrt{2} & 0 \\ -1 & 0 & +\sqrt{2} \\ -1 & 0 & -\sqrt{2} \end{bmatrix} \quad (20)$$

Die Eleganz an der gezeigten Betrachtung in der gewählten Hauptlage ist, dass die Berechnung von **F**, **O**, $\theta_{3,4}$ nur mehr von θ_1 und θ_2 abhängt. Nach Ersetzen von **FL** und **FR** durch die im dazwischen aufgespannten Winkel freien zwei Richtungen θ_1 und θ_2 kann zunächst der Richtungsvektor der Frontalen (zeigt nach vorne)

$$\mathbf{F} = \frac{\theta_1 + \theta_2}{\|\theta_1 + \theta_2\|}, \quad (21)$$

sowie der Richtungsvektor der Vertikalen (zeigt nach oben)

$$\mathbf{O} = \frac{\boldsymbol{\theta}_1 \times \boldsymbol{\theta}_2}{\|\boldsymbol{\theta}_1 \times \boldsymbol{\theta}_2\|}, \quad (22)$$

berechnet werden. Ausgehend von Gleichung 20 kann man durch Kombination von \mathbf{HO} und \mathbf{HU} ein Gleichungssystem aufstellen, welches von den bekannten Richtungsvektoren \mathbf{F} und \mathbf{O} abhängt,

$$\begin{aligned} I: \mathbf{HO} - \mathbf{HU} &= \frac{1}{\sqrt{3}} [0 \ 0 \ 2\sqrt{2}] = 2 \frac{\sqrt{2}}{\sqrt{3}} \mathbf{O} \\ II: -\mathbf{HO} - \mathbf{HU} &= \frac{1}{\sqrt{3}} [2 \ 0 \ 0] = \frac{2}{\sqrt{3}} \mathbf{F} \end{aligned}$$

$$\begin{aligned} I - II: 2\mathbf{HO} &= 2 \frac{\sqrt{2}}{\sqrt{3}} \mathbf{O} - \frac{2}{\sqrt{3}} \mathbf{F} \\ \implies \mathbf{HO} &= \frac{\sqrt{2}}{\sqrt{3}} \mathbf{O} - \frac{1}{\sqrt{3}} \mathbf{F} \\ I + II: -2\mathbf{HU} &= 2 \frac{\sqrt{2}}{\sqrt{3}} \mathbf{O} + \frac{2}{\sqrt{3}} \mathbf{F} \\ \implies \mathbf{HU} &= -\frac{\sqrt{2}}{\sqrt{3}} \mathbf{O} - \frac{1}{\sqrt{3}} \mathbf{F} \end{aligned} \quad (23)$$

Durch Ersetzen von \mathbf{HO} und \mathbf{HU} lassen sich Berechnungsvorschriften für $\boldsymbol{\theta}_3$ und $\boldsymbol{\theta}_4$ finden,

$$\boldsymbol{\theta}_{3,4} = -\frac{1}{\sqrt{3}} \mathbf{F} \pm \frac{\sqrt{2}}{\sqrt{3}} \mathbf{O} = -\frac{1}{\sqrt{3}} \frac{\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2}{\|\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2\|} \pm \frac{\sqrt{2}}{\sqrt{3}} \frac{\boldsymbol{\theta}_1 \times \boldsymbol{\theta}_2}{\|\boldsymbol{\theta}_1 \times \boldsymbol{\theta}_2\|}. \quad (24)$$

2.4 Optimaler Öffnungswinkel

Damit für spätere Beamforminganwendungen, die Matrix der Kugelflächenfunktionen, ausgewertet an den Positionen $\boldsymbol{\theta}_{1,2,3,4}$ gut konditioniert ist, bedarf es einer Betrachtung des Richtungsraars $\boldsymbol{\theta}_{3,4}$. Bei gegebenem Öffnungswinkel $\varphi = \angle(\boldsymbol{\theta}_1; \boldsymbol{\theta}_2)$ soll der Öffnungswinkel $\psi = \angle(\boldsymbol{\theta}_3; \boldsymbol{\theta}_4)$ entsprechend angepasst werden, siehe Abbildung 4.

Um den Zusammenhang der Winkel φ und ψ auf die Konditionszahl zu studieren, ist es zunächst notwendig eine Enkodierungsmatrix einzuführen, die von den beiden Öffnungswinkel abhängt,

$$\tilde{\mathbf{Y}}_1^T = \begin{bmatrix} \frac{1}{\sqrt{3}} & +\cos \frac{\varphi}{2} & +\sin \frac{\varphi}{2} & 0 \\ \frac{1}{\sqrt{3}} & +\cos \frac{\varphi}{2} & -\sin \frac{\varphi}{2} & 0 \\ \frac{1}{\sqrt{3}} & -\cos \frac{\psi}{2} & 0 & +\sin \frac{\psi}{2} \\ \frac{1}{\sqrt{3}} & -\cos \frac{\psi}{2} & 0 & -\sin \frac{\psi}{2} \end{bmatrix}. \quad (25)$$

Durch numerisches Berechnen der Konditionszahl (analytische Herleitung siehe Anhang B) der Matrix $\tilde{\mathbf{Y}}_1$ bei gegebenem Öffnungswinkel $\varphi \in [0; \pi]$ für alle Öffnungswinkel

$\psi \in [0; \pi]$ ergibt sich jeweils ein Winkel ψ mit minimaler Konditionszahl. Der genaue Verlauf für ψ ist in Abbildung 5 dargestellt. Da in Anhang B analytisch festgestellt werden konnte, dass für $\varphi \leq 110^\circ$ der optimale Öffnungswinkel ψ durch

$$\psi_{opt} = 2 \arccos \left(\frac{1}{3 \cos \frac{\varphi}{2}} \right) \quad (26)$$

gegeben ist, kann mit einer Approximation für größere Öffnungswinkel eine Berechnungsvorschrift für den optimalen Öffnungswinkel ψ angegeben werden,

$$\psi_{approx} = \begin{cases} 2 \arccos \left(\frac{1}{3 \cos \frac{\varphi}{2}} \right) & \text{wenn } \varphi < 110^\circ \\ 26 \arctan \left(\frac{109^\circ - \varphi}{26^\circ} \right) + 109^\circ & \text{sonst.} \end{cases} \quad (27)$$

Anhang B gibt die analytischen Ausdrücke der 4 Singulärwerte an, aus denen die Konditionszahl und der optimale Öffnungswinkel ψ berechnet werden können.

Für große Öffnungswinkel φ ist die Approximation ausreichend, weil sich die Konditionszahl in beiden Fällen (Optimum und Approximation) nur gering unterscheidet.

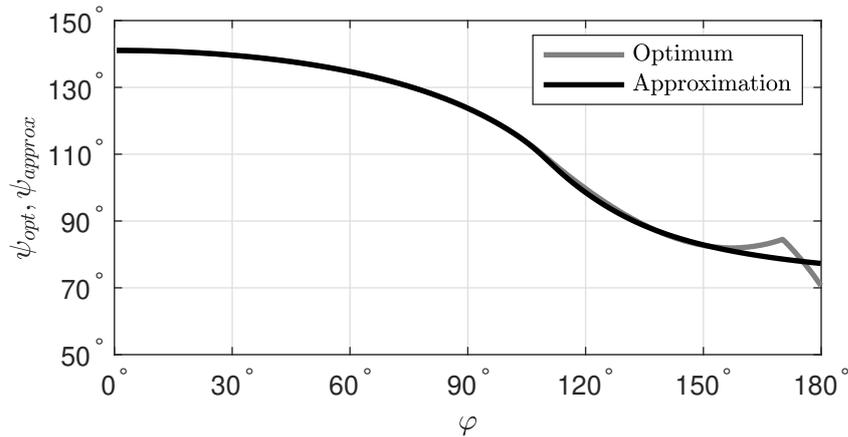


Abbildung 5 – Optimaler Öffnungswinkel ψ bei gegebenem Öffnungswinkel φ in Betracht auf optimaler Konditionszahl der Matrix $\tilde{\mathbf{Y}}_1$.

2.5 Rotation für beliebige F-L-O-Systeme

Für ein beliebig im Raum liegendes F - L - O -System können zunächst alle vier Richtungsvektoren in Hauptlage, siehe Abbildung 6 angeschrieben werden,

$$\mathbf{T} = [\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\theta}}_3, \hat{\boldsymbol{\theta}}_4] = \begin{bmatrix} \cos \frac{\varphi}{2} & \cos \frac{\varphi}{2} & -\cos \frac{\psi}{2} & -\cos \frac{\psi}{2} \\ \sin \frac{\varphi}{2} & -\sin \frac{\varphi}{2} & 0 & 0 \\ 0 & 0 & \sin \frac{\psi}{2} & -\sin \frac{\psi}{2} \end{bmatrix}. \quad (28)$$

Für ψ wird der approximierte Wert aus Gleichung 27 verwendet. Als praktische Rotationsmatrix, damit die Richtungen im F - L - O -System vorliegen, zeigt sich

$$\mathbf{R} = [\mathbf{F}, \mathbf{L}, \mathbf{O}] = \begin{bmatrix} F_x & L_x & O_x \\ F_y & L_y & O_y \\ F_z & L_z & O_z \end{bmatrix}. \quad (29)$$

Somit ergeben sich die rotierten Richtungen für das optimale, unregelmäßige Tetraeder aus

$$\mathbf{T} = [\theta_1, \theta_2, \tilde{\theta}_3, \tilde{\theta}_4] = \mathbf{R} \cdot \mathbf{T}. \quad (30)$$

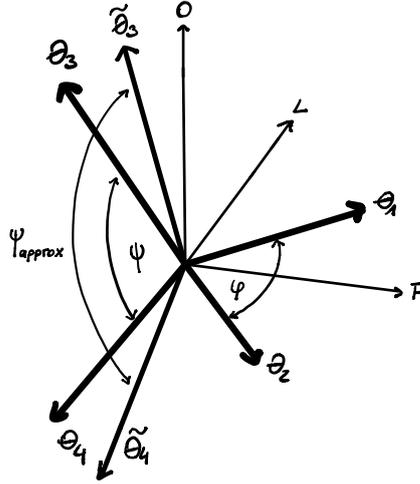


Abbildung 6 – Unregelmäßiges Tetraeder in Hauptlage mit den Richtungsvektoren $\theta_{1,2,3,4}$ bzw. optimales unregelmäßiges Tetraeder mit den Richtungsvektoren $\theta_{1,2}$ und $\tilde{\theta}_{3,4}$.

2.6 Beamforming und Reenkodieren

Zur Reenkodierung des Signals, ist es notwendig, die Signalanteile an den berechneten Richtungen $\theta_{1,2}$ und $\tilde{\theta}_{3,4}$ zu kennen.

Eine relative einfache Lösung des Problems bietet der sogenannte *Least-Square-Beamformer*, bei dem versucht wird, die quadrierte Norm des Fehlersignals

$$\|e\|^2 = e^T e = (\mathbf{b} - \mathbf{Y}_1 \mathbf{s})^T (\mathbf{b} - \mathbf{Y}_1 \mathbf{s}) = \mathbf{b}^T \mathbf{b} - 2\mathbf{s}^T \mathbf{Y}_1^T \mathbf{b} + \mathbf{s}^T \mathbf{Y}_1^T \mathbf{Y}_1 \mathbf{s}, \quad (31)$$

zu minimieren [GCC⁺11], wobei $\mathbf{Y}_1 = [\mathbf{y}_1(\theta_1), \mathbf{y}_1(\theta_2), \mathbf{y}_1(\tilde{\theta}_3), \mathbf{y}_1(\tilde{\theta}_4)]$ die Matrix der Kugelflächenfunktionen bis zur ersten Ordnung, an $\theta_{1,2}$ und $\tilde{\theta}_{3,4}$ ausgewertet, ist. Die Lösung für den minimalen Fehler kann durch Ableitung nach der Variable \mathbf{s} mit anschließender Nullsetzung der Gleichung berechnet werden,

$$\begin{aligned} \frac{d\|e\|^2}{d\mathbf{s}} &= -2\mathbf{b}^T \mathbf{Y}_1 + 2\mathbf{Y}_1^T \mathbf{Y}_1 \mathbf{s} \stackrel{!}{=} \mathbf{0} \\ \implies \mathbf{s} &= (\mathbf{Y}_1^T \mathbf{Y}_1)^{-1} \mathbf{Y}_1^T \mathbf{b}. \end{aligned} \quad (32)$$

Es ist sehr anschaulich, dass der Ausdruck $\mathbf{Y}_1^T \mathbf{b}$ einen Beamformer mit Hypernierencharakteristik darstellt, der mit einer Übersprechdämpfungsmatrix $(\mathbf{Y}_1^T \mathbf{Y}_1)^{-1}$ ergänzt wird. Damit wird sichergestellt, dass Signale an anderen Richtungen das Ergebnis wenig beeinflussen.

Falls der Öffnungswinkel φ sehr klein ist, kommt es bei der Inversion zu Problemen, weil die Matrix $\mathbf{Y}_1^T \mathbf{Y}_1$ (annähernd) singular ist. Um dies zu umgehen, ist eine Regularisierung notwendig. Eine Singulärwertzerlegung $\mathbf{Y}_1^T \mathbf{Y}_1 = \mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{V} \mathbf{S} \mathbf{V}^T$ erlaubt die Singulärwerte der Matrix $\mathbf{Y}_1^T \mathbf{Y}_1$ direkt zu beeinflussen,

$$[\mathbf{Y}_1^T \mathbf{Y}_1]^{-1} = [\mathbf{V}^T]^{-1} \mathbf{S}^{-1} \mathbf{V}^{-1} = \mathbf{V} \mathbf{S}^{-1} \mathbf{V}^T \quad (33)$$

Somit ergibt sich als Lösung für den Beamformer

$$\mathbf{s} = \mathbf{V} \mathbf{S}^{-1} \mathbf{V}^T \mathbf{Y}_1^T \mathbf{b}. \quad (34)$$

Um die Inversion einer singulären Matrix \mathbf{S} zu vermeiden, wird ein linearer Überblendparameter c eingeführt, sodass sich für die Singulärwerte

$$\tilde{\sigma}_i = (1 - c) \sigma_i + \frac{c}{\sqrt{\pi}} \quad \sigma = \{1, 2, 3, 4\} \quad (35)$$

und schließlich für die Matrix $\tilde{\mathbf{S}} = \text{diag}\{\tilde{\sigma}\}$ ergibt. Der Wertebereich dieses Regularisierungsparameters reicht von 10^{-12} bis 10^{-4} , damit der Einfluss des Beamformers ausreichend niedrig gehalten werden kann, wie es im Kapitel 3.5 diskutiert wird.

Es kann nun eine Impulsantwort beliebiger ambisonischer Ordnung N mit der Berechnungsvorschrift

$$\mathbf{b}_N = \mathbf{Y}_N \mathbf{V} \tilde{\mathbf{S}}^{-1} \mathbf{V}^T \mathbf{Y}_1^T \mathbf{b} \quad (36)$$

bestimmt werden, wobei $\mathbf{Y}_N = [\mathbf{y}_N(\boldsymbol{\theta}_1), \mathbf{y}_N(\boldsymbol{\theta}_2), \mathbf{y}_N(\tilde{\boldsymbol{\theta}}_3), \mathbf{y}_N(\tilde{\boldsymbol{\theta}}_4)]$ die Matrix der Kugelflächenfunktionen bis zur N -ten Ordnung an $\boldsymbol{\theta}_{1,2}$ und $\tilde{\boldsymbol{\theta}}_{3,4}$ ausgewertet ist.

Anhang C zeigt für die effiziente Implementierung die Möglichkeit auf, die Singulärwertzerlegung durch statische Transformationen analytisch durchzuführen, die zwei Dimensionen sehr leicht entkoppelt. Darüber hinaus bleibt eine von den Richtungen abhängige 2×2 Singulärwertzerlegung zu rechnen.

2.7 Spektrale Korrektur

Die in Gleichung 36 eingeführte Enkodierung führt zu einer typischer Weise unnatürlich klingenden Anhebung höherer Frequenzen, welche hauptsächlich den diffusen Nachhall beeinflusst.

Um diesem Effekt gegenzusteuern, wird in [ZF19] und [ZFZ20] die Pseudoallpasseigenschaft der Kugelflächenfunktionen zur Problemlösung bei der ASDM erwähnt. Um diese Eigenschaft ausnutzen zu können, werden die Signale im Vektor $\mathbf{b}(t) = [b_n^m(t)]_{nm}$ in Subbänder aufgeteilt, wodurch sich die Subbandimpulsantworten $b_n^m(t, b)$ ergeben. Für jedes Subband b und Ordnung n kann nun der spektrale Fehler ausgeglichen werden, indem die bereits erwähnte Pseudoallpasseigenschaft

$$\sum_{m=-n}^n \mathcal{E}\{|b_n^m(t, b)|^2\} = (2n + 1) \mathcal{E}\{|b_0^0(t, b)|^2\} \quad (37)$$

ausgenutzt wird, wobei $\mathcal{E}\{(\cdot)^2\}$ für den Erwartungswert der Einhüllenden steht. Dadurch ergibt sich als Berechnungsvorschrift für die spektrale Korrektur

$$\tilde{b}_n^m(t, b) = b_n^m(t, b) \sqrt{\frac{(2n+1) \mathcal{E}\{|b_n^0(t, b)|^2\}}{\sum_{m=-n}^n \mathcal{E}\{|b_n^m(t, b)|^2\}}}. \quad (38)$$

Aufgrund der Frequenzauflösung des menschlichen Gehörs, ist eine Aufteilung in Terzbändern günstig und für audiotecnische Anwendungen typisch.

2.8 Quellaufweitung für diffusen Nachhall

Für bestimmte Anwendungen kann es sinnvoll sein, den Anteil der Impulsantwort für den diffusen Nachhall aufzuweiten. Dazu wird die Rotationsmatrix \mathbf{R} , die eine frequenzabhängige Rotation um die z-Achse beschreibt, aus [ZF19] verwendet,

$$\mathbf{R}(m\hat{\phi} \cos(\omega t)) = \begin{bmatrix} \cos(m\hat{\phi} \cos(\omega t)) & \sin(m\hat{\phi} \cos(\omega t)) \\ -\sin(m\hat{\phi} \cos(\omega t)) & \cos(m\hat{\phi} \cos(\omega t)) \end{bmatrix}, \quad (39)$$

wobei $\hat{\phi}$ und ω die Stärke bzw. die Änderungsrate der Rotation über der Frequenz beschreiben. In [ZF13] wird ein Filter, welches zur Quellaufweitung verwendet werden kann, in zwei Teile

$$H_1(\omega) = \cos(\alpha \cos \omega \tau), \quad H_2(\omega) = \sin(\alpha \cos \omega \tau), \quad (40)$$

aufgeteilt, wobei $\alpha = m\hat{\phi}$, wodurch die Allpasseigenschaft

$$|H_1(\omega)|^2 + |H_2(\omega)|^2 = 1 \quad (41)$$

sichergestellt wird. Für die zwei Teile im Zeitbereich

$$\begin{aligned} h_1(t) &= \sum_{q=-\infty}^{\infty} J_{|q|}(\alpha) \cos\left(\frac{\pi}{2}|q|\right) \delta(t - q\tau) \\ h_2(t) &= \sum_{q=-\infty}^{\infty} J_{|q|}(\alpha) \sin\left(\frac{\pi}{2}|q|\right) \delta(t - q\tau), \end{aligned} \quad (42)$$

mit den Besselfunktionen J ist es ausreichend, Werte für $-5 \leq q \leq 5$ zu berechnen. Der Algorithmus platziert aufeinanderfolgende Frequenzen in leicht versetzten Richtungen. Dadurch nimmt die Breite der Phantomschallquelle zu. Der frequenzabhängige Teil bewirkt außerdem ein Verschmieren der zeitlichen Feinstruktur im Signal. In [ZF17] wird angemerkt, dass eine kausale Implementierung $0 \leq q \leq 5$ natürlicher klingt.

Für die Anwendung bei Raumimpulsantworten erweist sich die Wahl der Parameter $\hat{\phi} = 50^\circ$, $\tau = 5$ und $m = 5$ als nützlich, wobei 100 ms nach dem Maximum der Impulsantwort mit einer Dauer von 50 ms auf das aufgeweitete Signal überblendet wird. Davor wird die unaufgeweitete Impulsantwort verwendet.

3 Technische Analyse

Um den in Abschnitt 2 präsentierten Algorithmus zu testen, wurde eine einfache Testumgebung erstellt, bei der die Richtung von Hann-Impulsen detektiert werden soll. Mit Hilfe von sogenannten Weltkartendarstellungen kann die Detektion des Doppelrichtungsschätzers (engl. *Dual-Direction Estimator*, DDE) graphisch dargestellt und mit dem Ergebnis des Pseudointensitätsvektors (PIV), welcher für die ASDM verwendet wird, gegenübergestellt werden.

In allen folgenden Testszenarien werden den Hann-Impulsen mit einer Länge von 20 Samples ein Rauschsignal mit Amplitude 10^{-5} beigemischt. Der Hann-Impuls, der zum Zeitpunkt $t = 0$ startet, wird mit Amplitude 1 erstellt, wodurch das Signal um 100 dB lauter als das Rauschen abgebildet wird. Für die Mittelungszeit aus Abschnitt 2.2 wurde 4 Samples gewählt. Der Regularisierungsparameter c aus Gleichung 35 wurde dabei auf $c = 10^{-12}$ festgesetzt.

3.1 Richtungsdetektionsquote bei einer aktiven Wiedergabe- richtung

Zunächst wurde überprüft, ob die Richtungsdetektion korrekt funktioniert. Dazu wurde ein Hann-Impuls erstellt, dessen Richtungen nacheinander schrittweise auf jene eines virtuellen 10-Designs mit 60 Punkten gelegt wurde, Abbildung 7, [Gra]. Weil der Median des horizontalen Lokalisationsfehler unter 3° bei Wiedergabe mit Ambisonics 5. Ordnung liegt [FZ08], können Richtungen, deren Richtungsfehler unter 3° zur gewünschten Position aufweisen, als *korrekt detektiert* eingestuft werden. Das Ergebnis, über 1000 Iterationen gemittelt, ist in Abbildung 8 dargestellt.

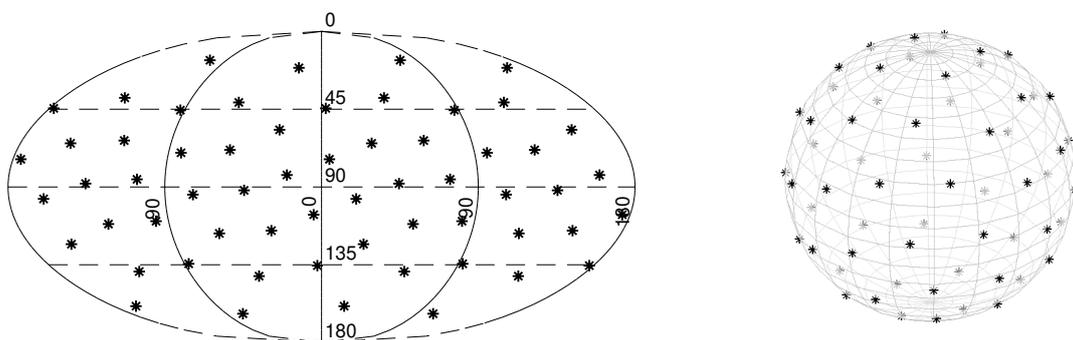


Abbildung 7 – 10-Design mit 60 Punkten zur Auswertung der detektierten Richtungen.

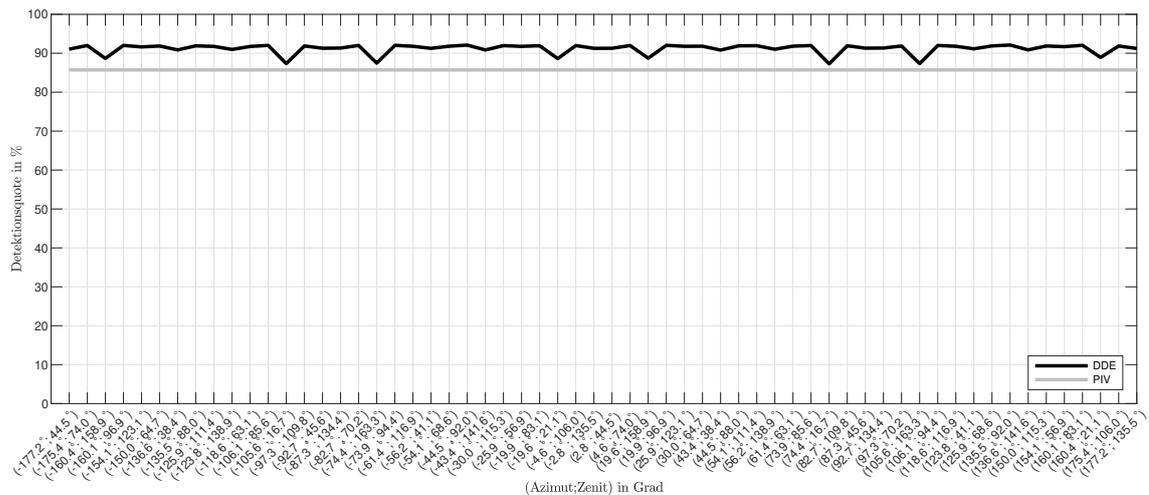


Abbildung 8 – Detektionsquote in % für PIV und DDE, wobei eine maximale Abweichung von 3° zur gewünschten Richtung erlaubt ist.

Für die Richtungsdetektion mit dem Pseudointensitätsvektor ergibt sich eine Detektionsquote von 85,7%. Für den Fall der Schätzung mit dem DDE schwankt die Quote richtungsabhängig um bis zu 4,7%. Sie ist aber dennoch um 1,6% bis zu 6,3% höher als jene des Pseudointensitätsvektor.

3.2 Ergebnis für eine aktive Wiedergaberichtung

Weiters ist es interessant, die Ergebnisse des Rekonstruktions- und Energiefehlers für verschiedene Ambisonics Ordnungen darzustellen. In diesem Beispiel wurde das Signal bei $(-90^\circ; 90^\circ)$ abgebildet. Diese Position ist in Abbildung 10 als Kreis dargestellt, die detektierten Richtungen als gefüllte Kreise, wobei die Amplitude des Signals die Farbe der Füllung, sowie die Größe bestimmt. Je größer und dunkler die Markierung ist, desto lauter wird das Signal abgebildet.

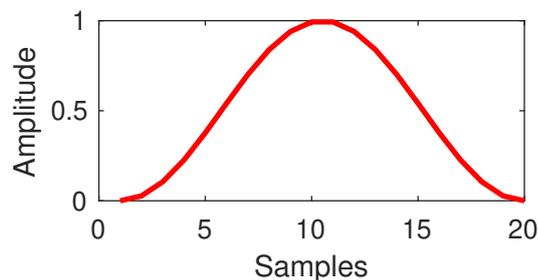


Abbildung 9 – Hann-Impuls zur Ergebnisdarstellung für eine aktive Wiedergaberichtung

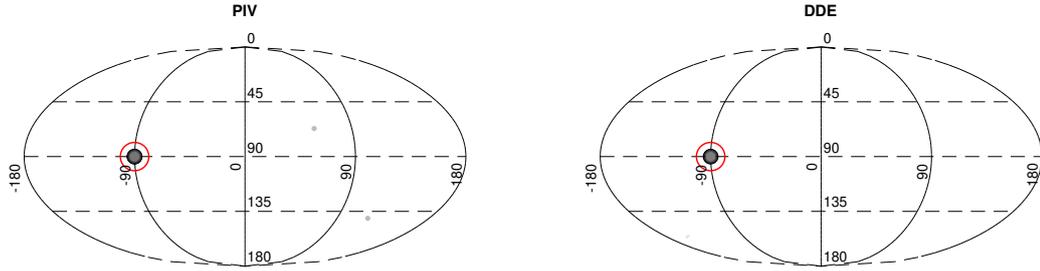


Abbildung 10 – Mit PIV (links) und DDE (rechts) detektierte Richtungen, Amplitude bestimmt Größe und Farbe.

Es kann festgestellt werden, dass die gewünschte Richtung stabil detektiert wird.

Abbildung 11 zeigt den Rekonstruktions- (E_R) und Energiefehler (E_E) zwischen rekonstruierter Impulsantwort \mathbf{b}_{rek} und originaler Impulsantwort \mathbf{b} für Ambisonics Ordnungen 1 und 5

$$E_R = 20 \cdot \log_{10} (\|\mathbf{b}_{rek} - \mathbf{b}\|) \quad E_E = 10 \cdot \log_{10} \left(\frac{\|\mathbf{b}_{rek}\|_F^2}{\|\mathbf{b}\|_F^2} \right), \quad (43)$$

wobei $\|\cdot\|_F$ für die Frobeniusnorm steht. Bei Berechnung der Fehler für 5. Ordnung wurde ein Hann-Impuls 5. Ordnung erstellt, wobei für die Richtungsdetektion die Signalanteile bis zur 1. Ordnung entnommen wurden.

Interessant ist, dass sich der Rekonstruktionsfehler beim Hochmischverfahren mit der 2DSE2 für 1. Ordnung Ambisonics deutlich von dem des Hochmischverfahren mit dem Pseudointensitätsvektor absetzt. Größtenteils ist der Rekonstruktionsfehler bei Verwendung Ambisonics 5. Ordnung für die 2DSE2 besser. Hier ist aber anzumerken, dass der Fehler für beide Algorithmen auf Grund seiner Größenordnung vernachlässigt werden kann.

In der Darstellung für den Energiefehler sind die Plots für Ambisonics 1. Ordnung nicht angeführt, weil diese sehr klein sind und in der Darstellung nicht sichtbar wären. Für beide Algorithmen kann ebenso festgehalten werden, dass die Größenordnung des Energiefehlers nicht relevant ist.

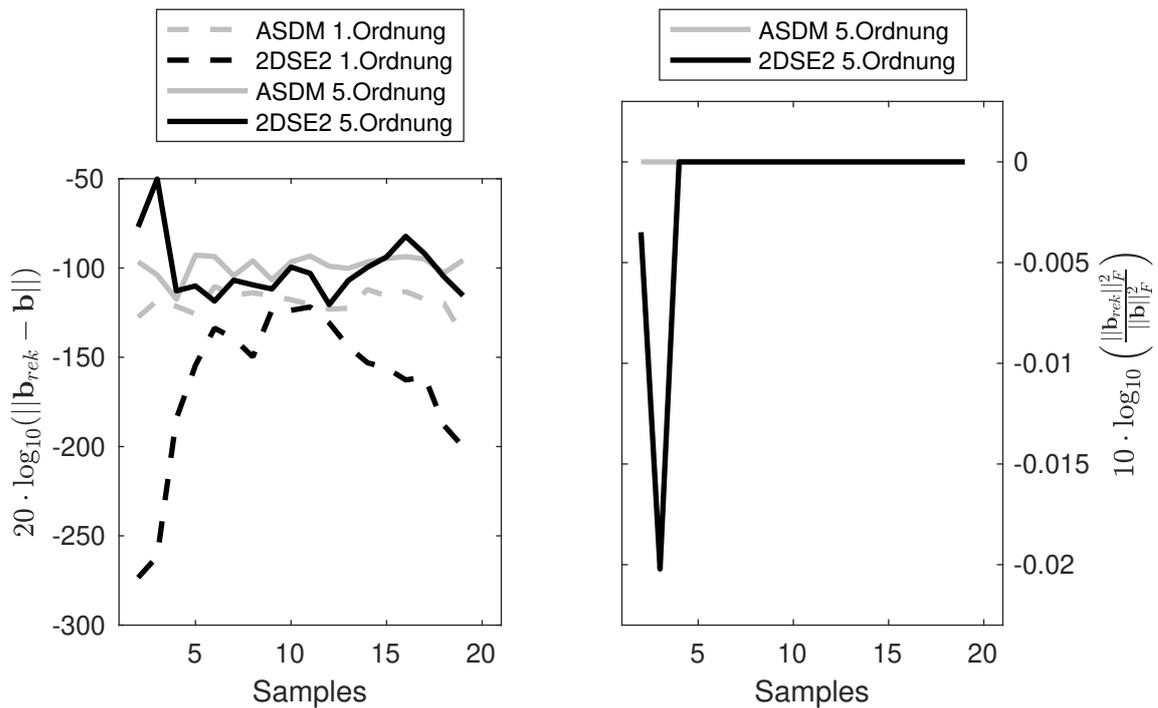


Abbildung 11 – Rekonstruktions- (links) und Energiefehler (rechts) der ASDM und 2DSE2 für Ambisonics Ordnungen 1 und 5.

3.3 Ergebnis für zwei aktive Wiedergaberichtungen

Sofern nur eine Richtung im Signal präsent ist, sind die Ergebnisse für beide Richtungsdetektionsalgorithmen bei Ambisonics 5. Ordnung ähnlich. Im nächsten Schritt wird die Richtungsdetektion für ein Signal, das aus zwei leicht zeitlich versetzten Hann-Impulsen (1 Sample) mit Richtungen $\theta_1 = (-90^\circ; 90^\circ)$ und $\theta_2 = (-93^\circ; 93^\circ)$ gleicher Amplitude besteht, analysiert.

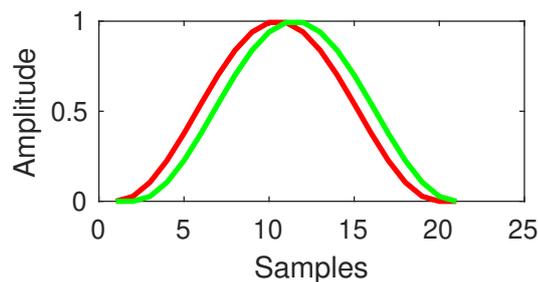


Abbildung 12 – Hann-Impulse zur Ergebnisdarstellung für zwei leicht versetzte Wiedergaberichtungen

Abbildung 13 zeigt die Detektionsergebnisse für beide Algorithmen. Beim genauen Betrachten fällt auf, dass der DDE zwei leicht versetzte Richtungen erkennt, während der Pseudointensitätsvektor stabil in eine Richtung zeigt.

In Abbildung 14 sind die zugehörigen Rekonstruktions- und Energiefehler dargestellt.

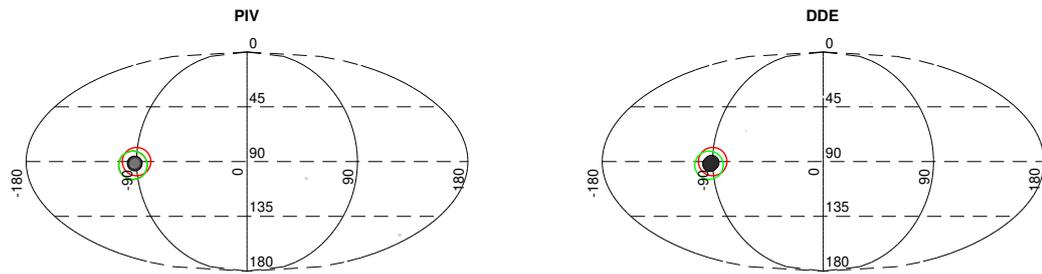


Abbildung 13 – Mit PIV (links) und DDE (rechts) detektierte Richtungen, Amplitude bestimmt Größe und Farbe.

Wiederum kann festgehalten werden, dass der Rekonstruktionsfehler der 2DSE2 bei Reenkodierung mit Ambisonics 1. Ordnung deutlich geringer ausfällt als jener der ASDM. Für 5. Ordnung Ambisonics fallen die Ergebnisse ähnlich aus, jedoch ist der Fehler im interessanten Zeitbereich (große Signalamplitude) für die 2DSE2 um circa 1.5 dB niedriger. Für beide Algorithmen ist dennoch der Energiefehler sehr klein und kann daher als irrelevant eingestuft werden.

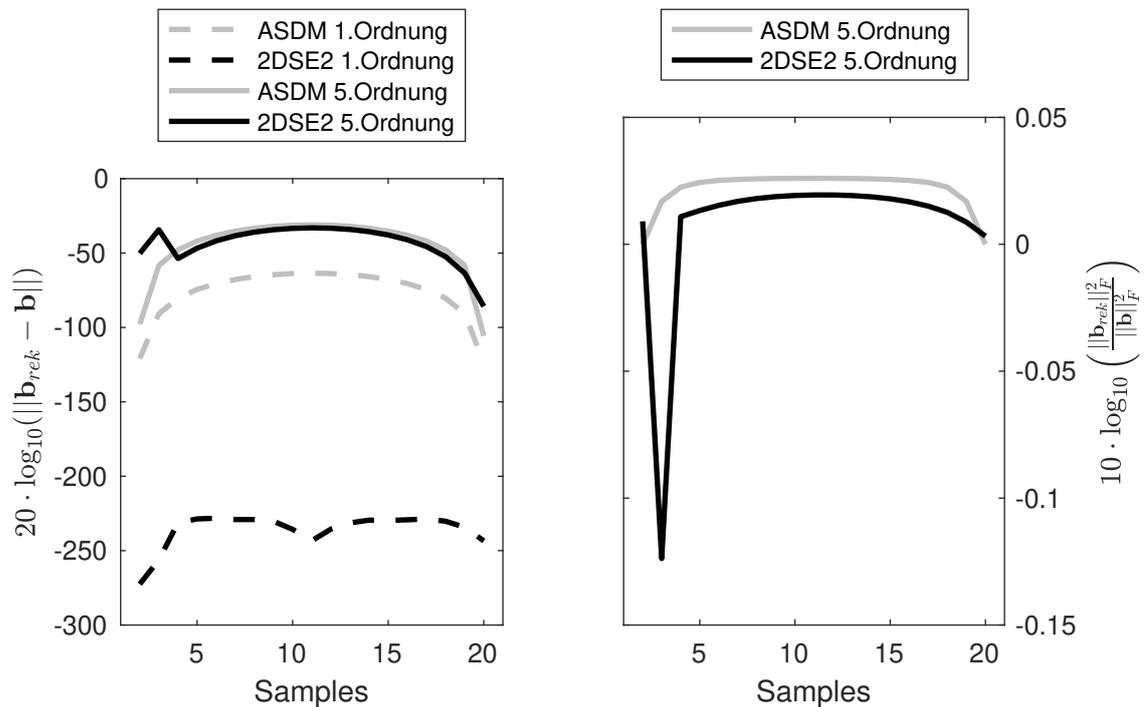


Abbildung 14 – Rekonstruktions- (links) und Energiefehler (rechts) der ASDM und 2DSE2 für Ambisonics Ordnungen 1 und 5.

Um die Grenzen der Schätzung mit dem PIV aufzuzeigen, wird nun ein um 4 Sample verzögerter Hann-Impuls mit Amplitude 0.75 beigemischt, wobei die Richtungen $\theta_1 = (-90^\circ; 90^\circ)$ und $\theta_2 = (45^\circ; 45^\circ)$ verwendet werden. Es ergeben sich die Abbildungen 16 und 17. Es ist zu erkennen, dass sich die Richtungsdetektion mit dem Pseudointensitätsvektor durch die zweite Wiedergaberichtung beeinflussen lässt, während der DDE zwei stabile richtig erkannte Richtungen liefert.

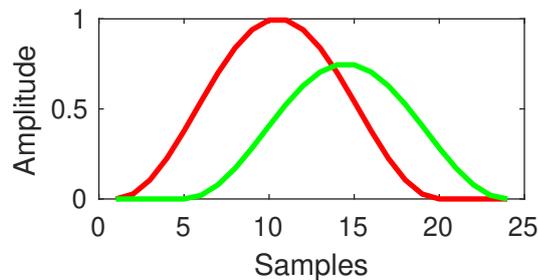


Abbildung 15 – Hann-Impulse zur Ergebnisdarstellung für zwei aktive Wiedergaberichtungen

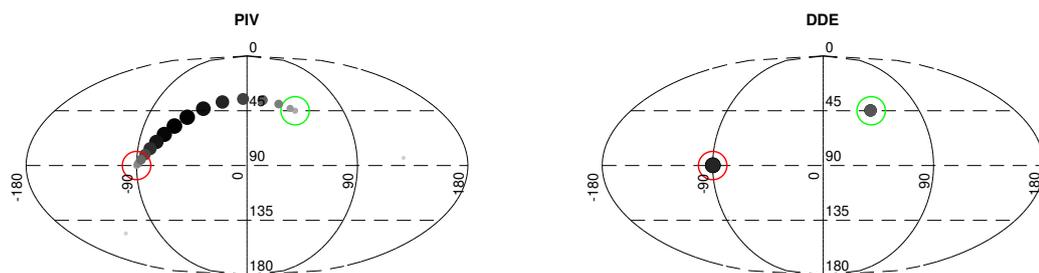


Abbildung 16 – Mit PIV (links) und DDE (rechts) detektierte Richtungen, Amplitude bestimmt Größe und Farbe.

Beim Betrachten des Rekonstruktionsfehlers fällt auf, dass die ASDM für den frühen Anteil der Impulsantwort etwas besser abschneidet. Jedoch für jene Zeitpunkte mit starker Signalamplitude bringt die 2DSE2 deutlich bessere Ergebnisse. Ebenso sieht man in der Darstellung des Energiefehlers, dass die 2DSE2 hier den kleineren energetischen Fehler verursacht, welcher im Gegensatz zur ASDM vernachlässigbar klein ist.

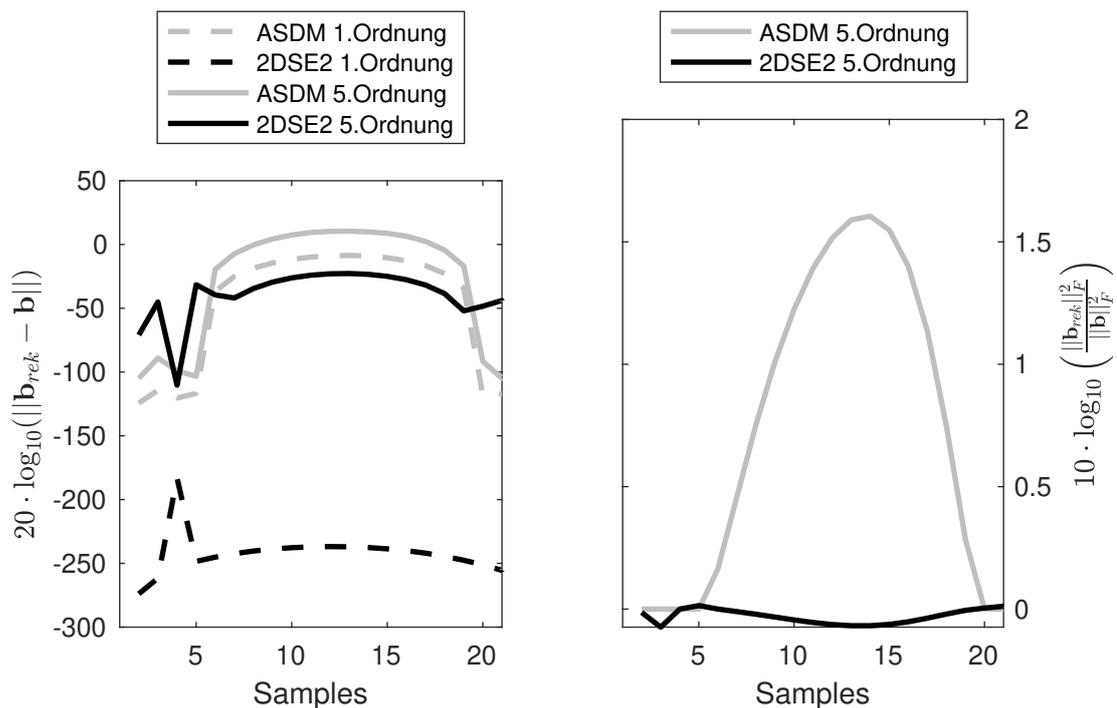


Abbildung 17 – Rekonstruktions- (links) und Energiefehler (rechts) der ASDM und 2DSE2 für Ambisonics Ordnungen 1 und 5.

3.4 Ergebnis für drei aktive Wiedergaberichtungen

Wenn dem Beispiel von zuvor eine weitere Richtung bei $(100^\circ; 135^\circ)$ mit Amplitude 0.25 und einem Zeitversatz von 4 Samples zum zweiten Impuls beigemischt wird, so ändern sich auch die Verhältnisse für den DDE.

Im Gegensatz zur Richtungsdetektion mit dem Pseudointensitätsvektor detektiert der DDE das Signal, welches am stärksten auftritt, einigermaßen stabil. Der PIV hingegen lässt sich, wie in den vorigen Ergebnissen, von den zwei übrigen Richtungen stark beeinflussen. Bei dem DDE wird augenscheinlich ein Kompromiss zwischen zweiten und dritten Hann-Impuls gesucht.

Der Rekonstruktionsfehler ist im Fall einer Reenkodierung mit Ambisonics 1. Ordnung für die 2DSE2 deutlich niedriger, bei 5. Ordnung liefert auch vorwiegend die 2DSE2 bessere Ergebnisse. Es kann festgestellt werden, dass bei der 2DSE2 der Rekonstruktionsfehler einigermaßen niedrig gehalten werden kann, sofern 2 Richtungen im Signal präsent sind. Wenn allerdings eine dritte dazukommt, stößt auch die 2DSE2 an ihre Grenzen.

Beim Betrachten des energetischen Fehlers lässt sich ebenso feststellen, dass die 2DSE2 den kleineren Fehler verursacht. Jedoch lässt sich auch hier bemerken, dass dieser im Vergleich zu den Test mit einer bzw. zwei Richtungen zuvor deutlich höher ausfällt.

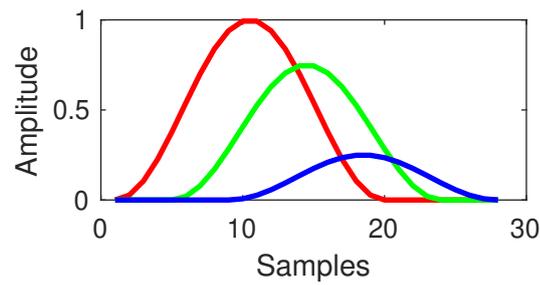


Abbildung 18 – Hann-Impulse zur Ergebnisdarstellung für drei aktive Wiedergaberichtungen

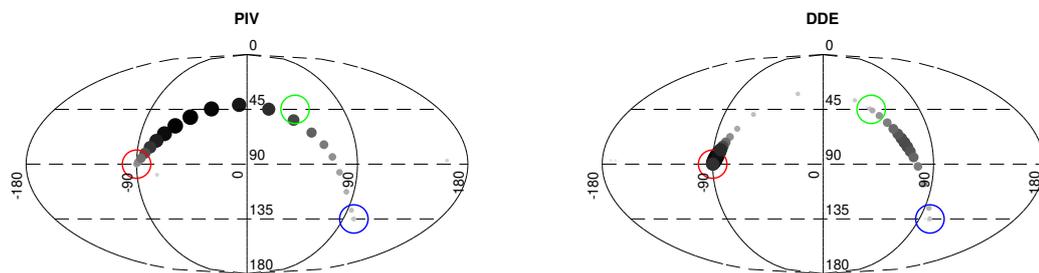


Abbildung 19 – Mit PIV (links) und DDE (rechts) detektierte Richtungen, Amplitude bestimmt Größe und Farbe.

3.5 Winkelabhängiger Fehler

Ein weiterer Punkt, der nicht vergessen werden sollte, ist der richtungsabhängige Rekonstruktions- und Energiefehler für die 2DSE2. Dieser ist in Abbildung 21 für drei verschiedene Ambisonics Ordnungen dargestellt. Für alle drei Fälle wurde wiederum dieselbe Regularisierungskonstante $c = 1e^{-12}$ verwendet. Man sieht deutlich, dass der Rekonstruktionsfehler mit steigender Ambisonics-Ordnung auch ansteigt, dennoch sehr klein ist. Das Optimum liegt bei $\psi = 110^\circ$. Es würde sich mit den weiteren zwei Richtungen $\theta_{3,4}$ ein regelmäßiges Tetraeder bilden und somit den optimalen Fall darstellen. Der Energiefehler ist in allen Situation vernachlässigbar klein.

Für eine Reenkodierung mit Ambisonics 5. Ordnung ist der winkelabhängige Fehler mit Betracht der Variation des Regularisierungsparameters c in Abbildung 22 angeführt. In der Grafik wird deutlich, dass zu hohe Werte für c negativen Einfluss mit sich ziehen. Der in den vorigen Abbildung verwendete Regularisierungsparameter $c = 1e^{-12}$ verursacht sehr kleine Fehler. Um den Energiefehler, als auch den Rekonstruktionsfehler einigermaßen niedrig halten zu können, darf dieser Parameter bis zur Obergrenze von $c = 1e^{-4}$ gewählt werden.

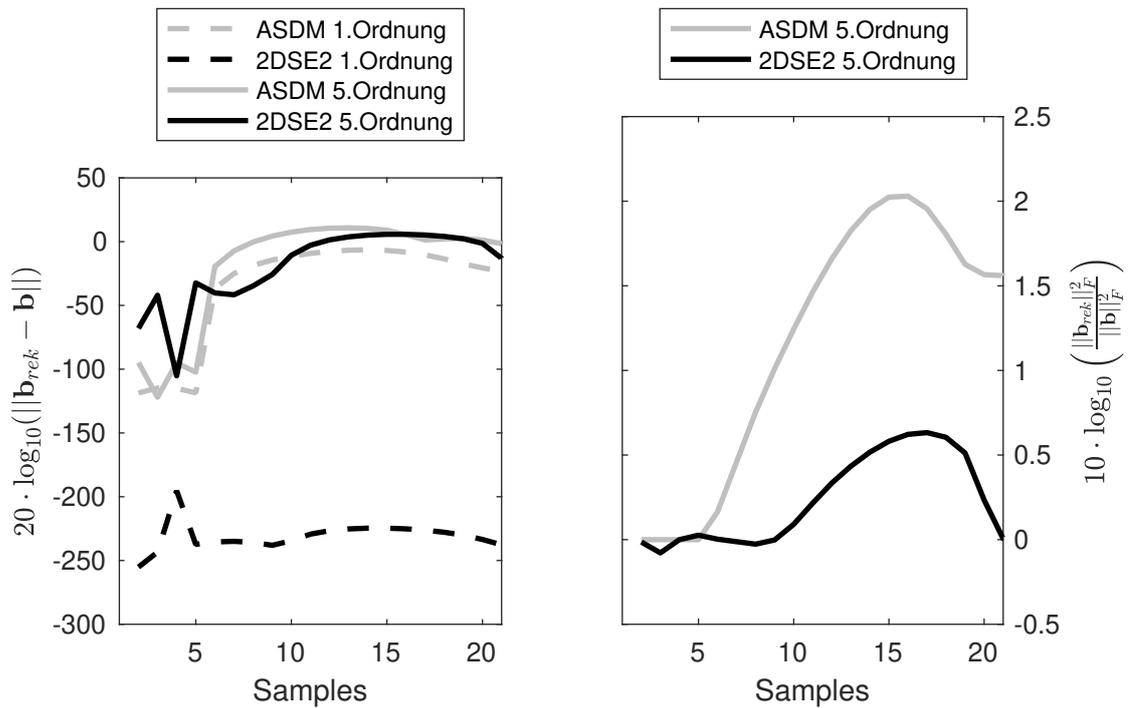


Abbildung 20 – Rekonstruktions- (links) und Energiefehler (rechts) der ASDM und 2DSE2 für Ambisonics Ordnungen 1 und 5.

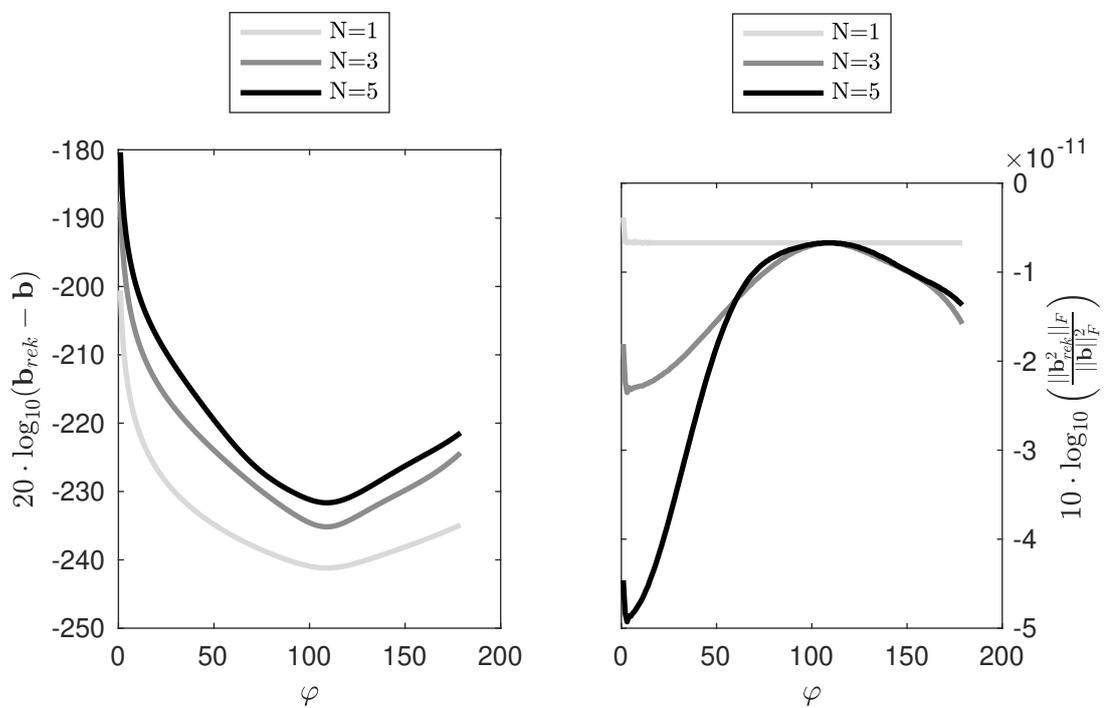


Abbildung 21 – Rekonstruktions- (links) und Energiefehler (rechts) für Hochmischverfahren mit der 2DSE2 für Ambisonics Ordnungen 1, 3 und 5.

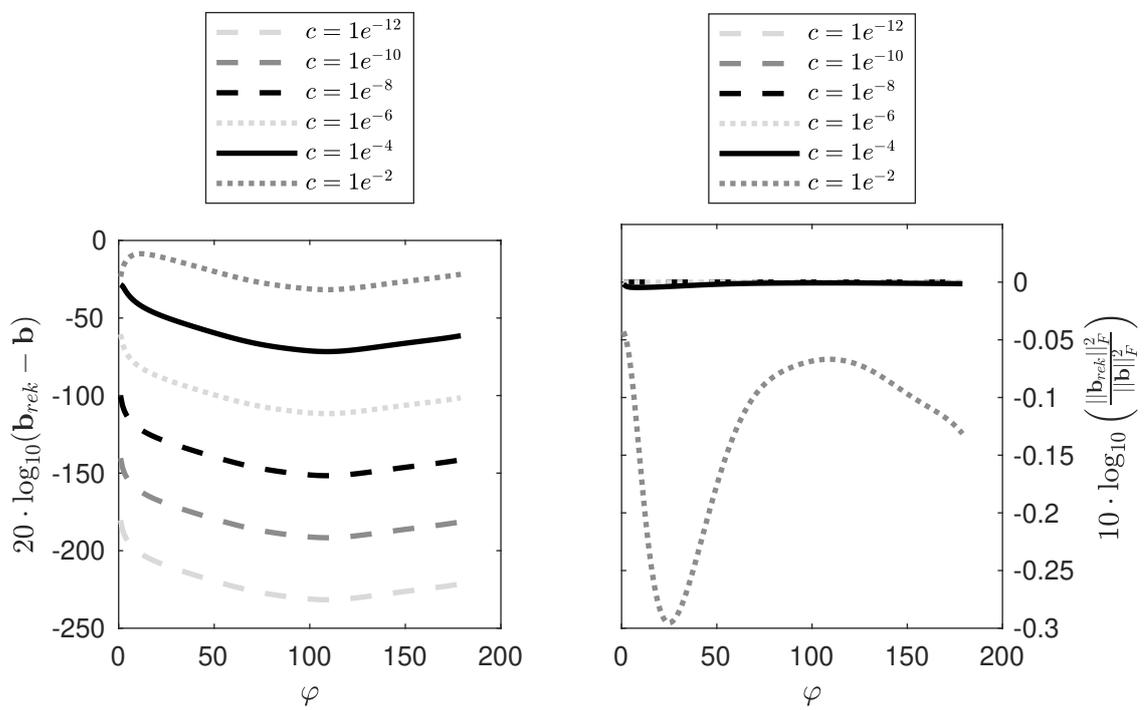


Abbildung 22 – Rekonstruktions- (links) und Energiefehler (rechts) für Hochmischverfahren mit der 2DSE2 für verschiedene Regularisierungsparameter c .

4 Hörversuch

Um die theoretischen Ergebnisse zu unterstreichen, wurde ein Hörversuch durchgeführt. Eine einfache Methode, um diesen durchzuführen, bietet der *Multi-Stimulus Test with Hidden Reference and Anchor* (MUSHRA) [Int15].

4.1 Vorbereitung

Um vergleichbare Ergebnisse, wie zum Beispiel aus [TPKL13], zu erhalten, wurden zwei verschiedene Signale verwendet, Musik und Sprache. Als Musiksinal wurde ein Ausschnitt aus *What's Trumps* (0:38-0:42, Komponist: Lukas Lohner, gespielt von der RhythmusSportGruppe) [Deu] gewählt, als Sprachsignal diente ein Ausschnitt aus *EBUs female English speech reference recording* aus [EBU08]. Damit auch Räume unterschiedlicher Raumakustik in den Hörversuch einfließen, wurden empfänger- und queldirektionale Impulsantwortmessungen des Györgi-Ligeti-Saales der Kunstuniversität Graz, sowie des Grazer Dom im Bergs verwendet. Diese Raumimpulsantworten wurden aus aufgezeichneten Sweepantworten mit dem Mikrofon *Soundfield ST450* berechnet.

Beide Räume verfügen über ein ambisonisches Beschallungssystem. Jedoch sind für den Hörversuch nur Schallereignisse von vorne interessant, weil sie häufig die deutlichsten Unterschiede zur Kuntskopfaufnahme aufweisen. Mit Hilfe des Plugins *StereoEncoder* aus der IEM Plugin-Suite [Rud20] wurden die Signale entsprechend enkodiert, mit dem Plugin *AllRADDecoder*, ebenso aus [Rud20], auf die entsprechende Lautsprecheranordnung dekodiert. Damit Lautsprecher, die nicht frontal auf Kopfhöhe aufgestellt waren, keinen Einfluss nehmen, wurden für den György-Ligeti-Saal nur drei Lautsprecher, für den Dom im Berg fünf Lautsprecher virtuell mit einem Signal versorgt. Geometrie und Lautsprecheranordnung sind in Abbildung 23 dargestellt, wobei nur die virtuell bespielten Lautsprecher eingezeichnet sind.

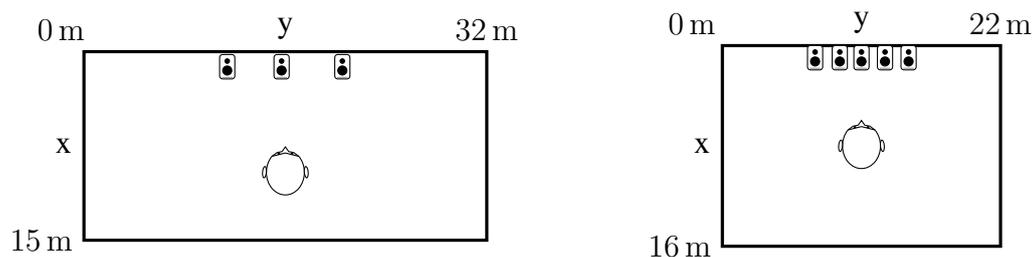


Abbildung 23 – Geometrie, Hör- und Lautsprecherposition(en) im virtuellen György-Ligeti-Saal (links) und im virtuellen Dom im Berg (rechts).

Die ambisonischen Raumimpulsantworten erster Ordnung (First Order Ambisonic Room Impulse Response, ARIR1) wurden mit den unterschiedlichen Hochmischverfahren bearbeitet, sodass ambisonische Raumimpulsantworten fünfter Ordnung (ARIR5) für den Hörversuch verwendet werden konnten. Die zu untersuchenden Algorithmen waren ASDM und 2DSE2, wobei beide jeweils ohne und mit zusätzlicher Quellauflösung im späten

Diffusteil zu vergleichen waren. Als Referenz dienten Impulsantwortmessungen mit dem Kunstkopf *Neumann KU100*, die jedoch nur für den Hörversuch im virtuellen György-Ligeti-Saal zu Verfügung standen.

Zur Faltung der ambisonischen Raumimpulsantworten mit den Signalen wurde Matthias Kronlachers Plugin `mcfx_convolver36` aus seiner Multichannel Audio Plugin Suite (`mcfx`) [Kro20] verwendet. Für die anschließende Binauraldekodierung diente der `BinauralDecoder` aus der IEM Plugin Suite [Rud20]. Für die Faltung der Signale mit den binauralen Raumimpulsantworten des Kunstkopfes (Binaural Room Impulse Response, BRIR) wurde ebenso `mcfx_convolver36` verwendet. Den Versuchspersonen konnten anschließend fertig gerenderte Binauralaufnahmen präsentiert werden, wodurch Kopfbewegungen während des Versuchs nicht berücksichtigt werden konnten.

4.2 Durchführung

4.2.1 György-Ligeti-Saal

Die Versuchspersonen wurden gebeten, zunächst die Natürlichkeit zwischen den unterschiedlichen Binauraldateien des Musikbeispiels auf einer Skala von 0 (unnatürlich) bis 10 (natürlich) zu bewerten. Anschließend galt es, die Hochmischalgorithmen mit der Kunstkopfaufnahme auf einer Skala von 0 (unterschiedlich) bis 10 (ident) zu vergleichen. Hier wurde auch die Kunstkopfaufnahme in den zu bewerteten Beispielen versteckt, um nachträglich die Vertrauenswürdigkeit der Versuchsperson einschätzen zu können.

Dieselbe Fragestellung wurde für das Sprachklangbeispiel verwendet. Zusätzlich wurde für den György-Ligeti-Saal als Testsignal ein Schnipsen gewählt, um den Höreindruck für transiente Signale überprüfen zu können. Weiters kann man mit Hilfe dieses Signals den diffusen Nachhall zwischen den unterschiedlichen Hochmischalgorithmen und der Kunstkopfaufnahme vergleichen.

4.2.2 Dom im Berg

Da in der Messreihe für den Dom im Berg keine Kunstkopfaufnahmen zur Verfügung standen, beschränkte sich der Hörversuch auf die Beurteilung der Natürlichkeit der Binauralbeispiele der ASDM und der 2DSE2. Wiederum wurden die Versuchspersonen gebeten, Musik und Sprache auf einer Skala von 0 (unnatürlich) und 10 (natürlich) zu bewerten. Da bei trockenerer Raumakustik Unterschiede für transiente Signale zwischen den einzelnen Algorithmen klein ausfallen, entfiel auch dieser Teil, sodass für diesen Abschnitt des Hörversuchs nur die Natürlichkeit für zwei unterschiedliche Signale bewertet werden mussten.

4.3 Ergebnisse

Am Hörversuch nahmen 12 Personen (alle männlich) teil, wobei das Durchschnittsalter 29 Jahre und die durchschnittlich benötigte Zeit 12 min 29 s betrug.

Abbildung 24 und 25 zeigen die Ergebnisse des Hörversuchs für den György-Ligeti-Saal. Auffallend ist, dass der Kunstkopf (KK) für alle drei Signale (Musik, Sprache und Schnipsen) am natürlichsten empfunden wurde. Wie zu erwarten, liefert die 2DSE2 in allen drei Fällen bessere Ergebnisse als die ASDM. Für Musik- und Sprachanwendungen macht es augenscheinlich keinen Unterschied, ob der spätere Teil der ambisonischen Impulsantwort aufgeweitet wird oder nicht. Das lässt sich nicht nur beim Betrachten des Medians feststellen, sondern spiegelt sich auch in den Ergebnissen eines Signifikanztests mit Bonferroni Korrektur wieder. Während sich für die ASDM und die ASDM mit Quellaufweitung bzw. 2DSE2 und 2DSE2 mit Quellaufweitung für Musik und Sprache keine signifikanten Unterschiede ($p > 2.5$) ergeben, können ansonsten größtenteils signifikante Unterschiede ($p < 0.02$) festgestellt werden. Jedoch für transiente Signale, wie das Schnipsen, verhilft ein Aufweiten des späten Anteils der Impulsantwort zu besseren Resultaten. Hier sind auch die Resultate durchwegs signifikant unterschiedlich ($p < 0.01$).

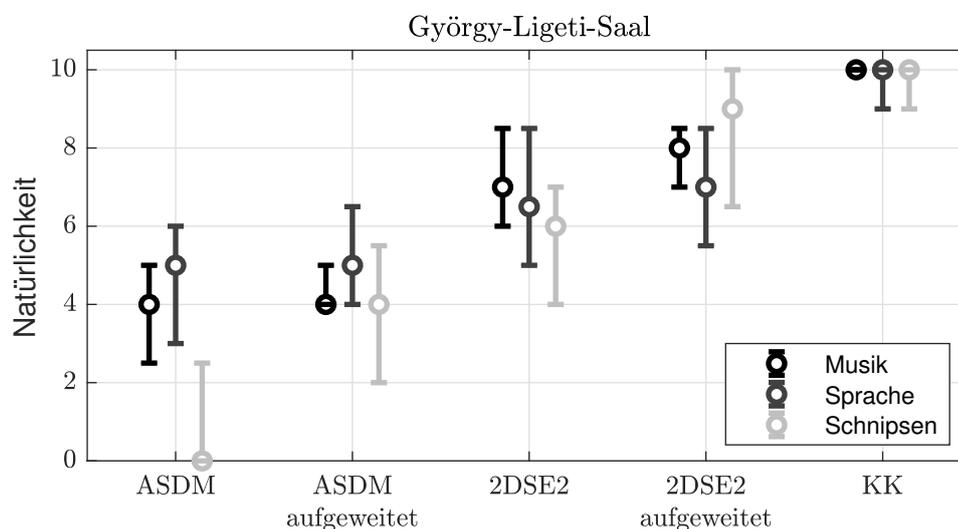


Abbildung 24 – Ergebnis: Bewertung der Natürlichkeit für den György-Ligeti-Saal.

Abbildung 25 zeigt die Ergebnisse für die Bewertung der Ähnlichkeit mit dem Kunstkopf. Der Unterschied zwischen der 2DSE2 und der ASDM ist für diesen Teil des Hörversuchs nicht so stark ausgeprägt. Dennoch wird die 2DSE2 als ähnlicher zum Kunstkopf eingestuft als die ASDM. Ein Grund für dieses Resultat ist, dass die Klangfarbe zwischen SDM/2DSE2 und Kunstkopf nicht völlig identisch ist. Es wurde weiters bemerkt, dass das klangliche Abbild des Kunstkopfes stets die wenigsten hörbaren stark seitlichen Anteile beinhaltet hat. Einige Versuchspersonen haben im Gespräch angemerkt, dass sie bei der schlechter bewerteten Variante stärkere seitliche Anteile gehört haben.

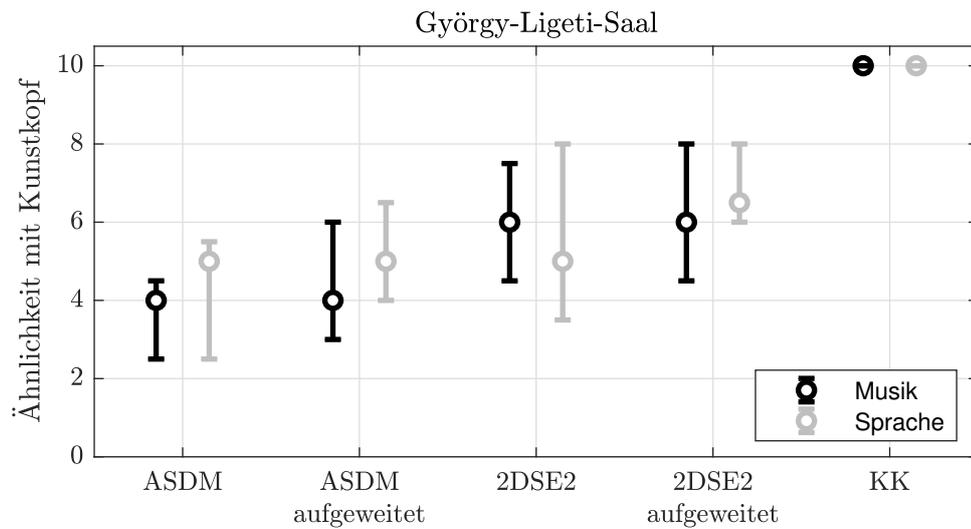


Abbildung 25 – Ergebnis: Bewertung der Ähnlichkeit mit dem Kunstkopf für den György-Ligeti-Saal.

Für den Dom im Berg hingegen lassen sich sehr deutliche Ergebnisse darstellen, Abbildung 26. Während die 2DSE2 als sehr natürlich bewertet wurde, wurde die ASDM sehr oft als unnatürlich eingestuft. Wiederum zeigt sich in den Ergebnissen, dass für Musik- und Sprachanwendungen eine Aufweitung im diffusen Nachhall nicht notwendig ist. Weiters sieht man in den Resultaten des Signifikanztest aus ASDM und ASDM mit Quellaufweitung bzw. 2DSE2 und 2DSE2 mit Quellaufweitung keine signifikanten Unterschiede ($p > 2.3$), während die anderen Ergebnisse zueinander deutlich signifikant unterschiedlich sind ($p < 0.01$).

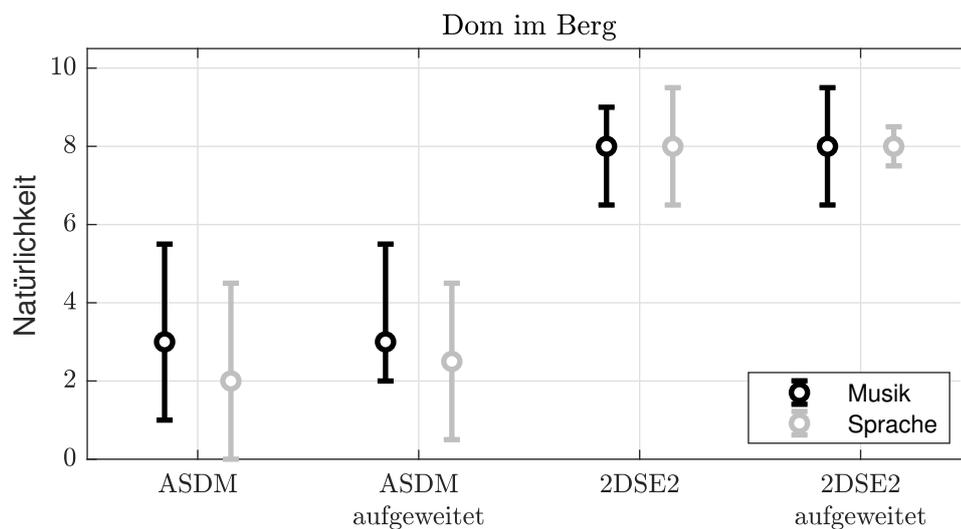


Abbildung 26 – Ergebnis: Bewertung der Natürlichkeit für den Dom im Berg.

5 Schlussfolgerung und Ausblick

In dieser Projektarbeit konnte ein neues Verfahren zur Richtungsauflösungsverbesserung von ambisonischen Impulsantworten erster Ordnung gezeigt und getestet werden. Ein Vergleich mit der *Spatial Decomposition Method* im Abschnitt Technische Analyse zeigte die Vorteile der Richtungsschätzung mit dem DDE, speziell wenn es um Richtungsdetektion von zwei Richtungen im Signal geht. Mit den eingeführten Fehlermaßen, Rekonstruktions- und Energiefehler, konnte für beide Algorithmen die Funktionsweise überprüft und miteinander verglichen werden.

Ein Hörversuch unterstreicht die theoretisch angeführten Ergebnisse. Hier konnte festgestellt werden, dass die 2DSE2 für Impulsantworten aus dem György-Ligeti-Saal sowie aus dem Grazer Dom im Berg bessere Ergebnisse bezüglich Natürlichkeit als die ASDM liefert. Weiters konnte erkannt werden, dass für transiente Signale eine Aufweitung im späten Diffusanteil zu natürlicheren Resultaten führt, die jedoch für Musik- und Sprachanwendungen nicht zwingend notwendig ist.

Da der Hörversuch mit Kopfhörern durchgeführt wurde und derselbe Kopfhörer für alle Versuchspersonen nicht gefordert werden konnte, wäre ein Hörversuch, wobei die Bewertung über eine Lautsprecherkonstruktion erfolgen könnte, anzustreben. Somit kann ein einheitliche Hörumgebung für alle Teilnehmenden geschaffen werden. Um auch hier einen Vergleich mit einem Referenzsignal durchzuführen, könnten ambisonische Raumimpulsantworten mit dem *em32 Eigenmike* der Firma *mh acoustics*, welches Ambisonics 4. Ordnung unterstützt, gemessen werden. Für die Hochmischalgorithmen wird das Signal bis zur ersten Ordnung entnommen, hochgemischt und anschließend mit dem Original verglichen.

Nichtsdestotrotz konnte mit Hilfe der technischen Analyse sowie eines anschließenden Hörversuchs gezeigt werden, dass die 2DSE2 als Methode zur Lautsprechervirtualisierung in einem Raum gegenüber der bekannten ASDM bevorzugt werden kann.

A Soundfield ST450: Signalamplituden nullter und erster Ordnung

Abbildung 27 zeigt die Signalamplituden eines mit dem *Soundfield ST450* aufgenommenen Signal bei diversen Frequenzen, wodurch ersichtlich wird, dass im Frequenzbereich, der für die Richtungsdetektion relevant ist, eine Multiplikation des omnidirektionalen w-Kanals mit $\sqrt{3}$ notwendig ist.

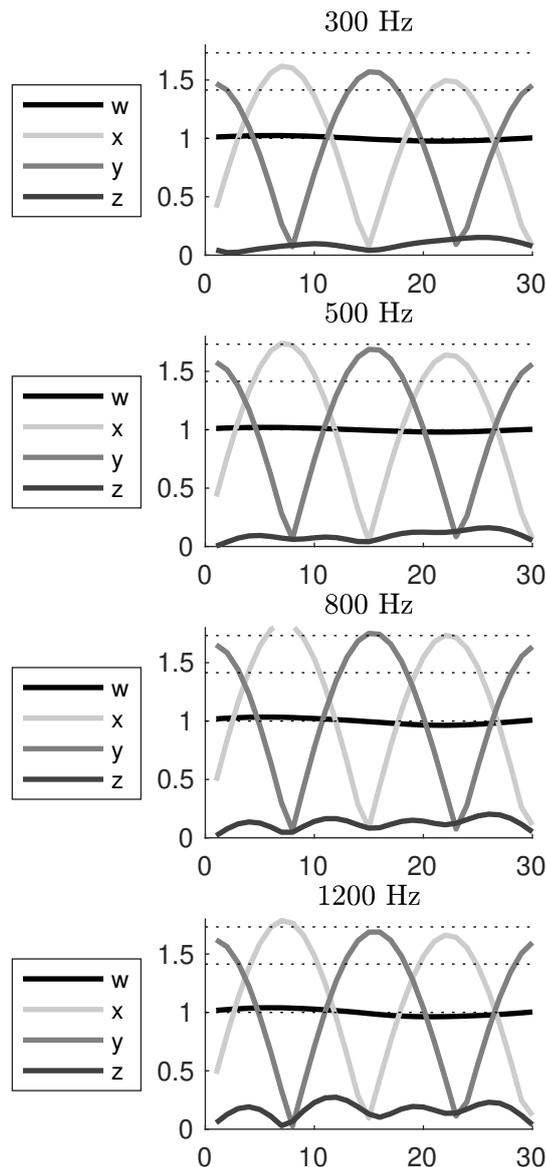


Abbildung 27 – Amplituden eines mit dem *Soundfield ST450* aufgenommenen Signal bei diversen Frequenzen.

B Konditionszahl der Tetraederenkodierungsmatrix und optimaler Öffnungswinkel ψ

Die Konditionszahl der Tetraederenkodierungsmatrix $\tilde{\mathbf{Y}}_1$ kann über dessen Singulärwerte $\sigma_{\tilde{\mathbf{Y}}_1}$ berechnet werden,

$$\text{cond}(\tilde{\mathbf{Y}}_1) = \frac{\max(\sigma_{\tilde{\mathbf{Y}}_1})}{\min(\sigma_{\tilde{\mathbf{Y}}_1})}. \quad (44)$$

Da die Eigenwerte der quadrierten transponierten Enkodierungsmatrix $\tilde{\mathbf{Y}}_1 \tilde{\mathbf{Y}}_1^T$ die quadrierten Eigenwerte der Enkodierungsmatrix $\tilde{\mathbf{Y}}_1$ liefern, werden zur Berechnung des optimalen Öffnungswinkels ψ jene der quadrierten transponierten Enkodierungsmatrix berechnet,

$$\begin{aligned} \tilde{\mathbf{Y}}_1 \tilde{\mathbf{Y}}_1^T &= \begin{bmatrix} a & a & a & a \\ c_1 & c_1 & -c_2 & -c_2 \\ s_1 & -s_1 & 0 & 0 \\ 0 & 0 & s_2 & -s_2 \end{bmatrix} \begin{bmatrix} a & c_1 & s_1 & 0 \\ a & c_1 & -s_1 & 0 \\ a & -c_2 & 0 & s_2 \\ a & -c_2 & 0 & -s_2 \end{bmatrix} \\ &= \begin{bmatrix} 4a^2 & 2ac_1 - 2ac_2 & 0 & 0 \\ 2ac_1 - 2ac_2 & 2c_1^2 + 2c_2^2 & 0 & 0 \\ 0 & 0 & 2s_1^2 & 0 \\ 0 & 0 & 0 & 2s_2^2 \end{bmatrix}, \end{aligned} \quad (45)$$

wobei $a = \frac{1}{\sqrt{3}}$, $c_1 = \cos \frac{\varphi}{2}$, $s_1 = \sin \frac{\varphi}{2}$, $c_2 = \cos \frac{\psi}{2}$ und $s_2 = \sin \frac{\psi}{2}$. Zwei Eigenwerte können direkt abgelesen werden, $\sigma_{3, \tilde{\mathbf{Y}}_1 \tilde{\mathbf{Y}}_1^T} = 2s_1^2$ und $\sigma_{4, \tilde{\mathbf{Y}}_1 \tilde{\mathbf{Y}}_1^T} = 2s_2^2$. Die beiden anderen Singulärwerte erhält man durch Lösen des Eigenwertproblems der links oberen 2×2 Matrix,

$$\begin{aligned} (4a^2 - x)(2c_1^2 + 2c_2^2 - x) - (2ac_1 - 2ac_2)^2 &= 0 \\ x^2 - x(4a^2 + 2c_1^2 - 2c_2^2) - 4a^2c_1^2 + 4a^2c_2^2 + 8a^2c_1c_2 &= 0 \\ 2a^2 + c_1^2 + c_2^2 \pm \sqrt{(2a^2 + c_1^2 + c_2^2)^2 - 4a^2c_1^2 - 4a^2c_2^2 - 8a^2c_1c_2} &= x_{1,2} \\ 2a^2 + c_1^2 + c_2^2 \pm \sqrt{4a^4 - 8a^2c_1c_2 + c_1^4 + 2c_1^2c_2^2 + c_2^4} &= x_{1,2}. \end{aligned} \quad (46)$$

Somit sind die Singulärwerte der quadrierten transponierten Tetraeder-Enkodierungsmatrix gegeben,

$$\boldsymbol{\sigma} = \begin{bmatrix} 2a^2 + c_1^2 + c_2^2 - \sqrt{4a^4 - 8a^2c_1c_2 + c_1^4 + 2c_1^2c_2^2 + c_2^4} \\ 2a^2 + c_1^2 + c_2^2 + \sqrt{4a^4 - 8a^2c_1c_2 + c_1^4 + 2c_1^2c_2^2 + c_2^4} \\ 2s_1^2 \\ 2s_2^2 \end{bmatrix}. \quad (47)$$

Zur Berechnung des optimalen Vertikalwinkels in Abhängigkeit vom Horizontalwinkel, muss die Konditionszahl minimiert werden. Für Horizontalwinkel bis $\varphi = 110^\circ$ berechnet sich die Konditionszahl aus $\frac{\sigma_2}{\sigma_3}$,

$$\text{cond}(\tilde{\mathbf{Y}}_1 \tilde{\mathbf{Y}}_1^T) = \frac{2a^2 + c_1^2 + c_2^2 + \sqrt{4a^4 - 8a^2c_1c_2 + c_1^4 + 2c_1^2c_2^2 + c_2^4}}{2s_1^2}. \quad (48)$$

Um die minimale Konditionszahl zu erhalten, muss der Ausdruck im Zähler minimiert werden,

$$x = 2a^2 + c_1^2 + c_2^2 + \sqrt{4a^4 - 8a^2c_1c_2 + 2c_1^2c_2^2 + c_1^4 + c_2^4}$$

$$\frac{dx}{d\psi} = \frac{d \cos \psi}{d\psi} \frac{dx}{dc_2} = -\sin \psi \left[2c_2 + \frac{-8a^2c_1 + 4c_1^2c_2 + 4c_2^3}{2\sqrt{4a^4 - 8a^2c_1c_2 + 2c_1^2c_2^2 + c_1^4 + c_2^4}} \right] = 0. \quad (49)$$

Die nicht triviale Lösung ist dabei interessant,

$$2c_2 + \frac{-8a^2c_1 + 4c_1^2c_2 + 4c_2^3}{2\sqrt{4a^4 - 8a^2c_1c_2 + 2c_1^2c_2^2 + c_1^4 + c_2^4}} = 0$$

$$c_2\sqrt{4a^4 - 8a^2c_1c_2 + 2c_1^2c_2^2 + c_1^4 + c_2^4} = 2a^2c_1 - c_1^2c_2 - c_2^3$$

$$4a^4c_2^2 - 8a^2c_1c_2^3 + \cancel{2c_1^2c_2^4} + \cancel{c_1^4c_2^2} + \cancel{c_2^6} = 4a^4c_1^2 + \cancel{c_1^4c_2^2} + \cancel{c_2^6} - 4a^2c_1^3c_2 - 4a^2c_1c_2^3 + \cancel{2c_1^2c_2^4}$$

$$a^2c_2^2 - c_1c_2^3 + c_1^3c_2 - a^2c_1^2 = 0. \quad (50)$$

Es ist offensichtlich, dass eine Lösung der Gleichung $c_2 = c_1$ ist, d.h. $\varphi = \psi$. Dies erlaubt eine Polynomdivision,

$$\begin{array}{r} (-c_1c_2^3 + a^2c_2^2 + c_1^3c_2 + a^2c_1^2) : (c_2 - c_1) = -c_1c_2^2 + (a^2 - c_1^2)c_2 + a^2c_1 \\ \underline{-c_1c_2^3 + c_1^2c_2^2} \\ (a^2 - c_1^2)c_2^2 + c_1^3c_2 + a^2c_1^2 \\ \underline{(a^2 - c_1^2)c_2^2 - (a^2 - c_1^2)c_1c_2} \\ a^2c_1c_2 - a^2c_1^2 \end{array} \quad (51)$$

Diese quadratische Gleichung $c_2^2 + \frac{c_1^2 - a^2}{c_1}c_2 - a^2 = 0$ löst sich mit der bekannten Lösungsformel,

$$c_2 = \frac{a^2 - c_1^2}{2c_1} \pm \sqrt{\frac{a^4 - 2a^2c_1^2 + c_1^4 + 4a^2c_1^2}{4c_1^2}}$$

$$= \frac{a^2 - c_1^2}{2c_1} \pm \sqrt{\frac{(a^2 + c_1)^2}{(2c_1)^2}} = \left\{ \frac{a^2}{c_1}, -c_1 \right\}. \quad (52)$$

Dabei ist die positive Lösung $c_2 = \frac{a^2}{c_1}$ interessant,

$$\psi = 2 \arccos \left(\frac{a^2}{\cos \frac{\varphi}{2}} \right). \quad (53)$$

Somit ergibt sich ein optimaler Öffnungswinkel $\psi = 2 \arccos \left(\frac{a^2}{\cos \frac{\varphi}{2}} \right)$ unter der Voraussetzung, dass $\varphi < 110^\circ$.

C Singulärwertzerlegung der quadrierten Enkodierungsmatrix

Die quadrierte Enkodierungsmatrix, für die folgend die Singulärwertzerlegung durchgeführt wird, ist als

$$\tilde{\mathbf{Y}}_1^T \tilde{\mathbf{Y}}_1 = \begin{bmatrix} a^2 + c_1^2 + s_1^2 & a^2 + c_1^2 - s_1^2 & a^2 - c_1 c_2 & a^2 - c_1 c_2 \\ a^2 + c_1^2 - s_1^2 & a^2 + c_1^2 + s_1^2 & a^2 - c_1 c_2 & a^2 - c_1 c_2 \\ a^2 - c_1 c_2 & a^2 - c_1 c_2 & a^2 + c_2^2 + s_2^2 & a^2 + c_2^2 - s_2^2 \\ a^2 - c_1 c_2 & a^2 - c_1 c_2 & s^2 + c_2^2 - s_2^2 & a^2 + c_2^2 + s_2^2 \end{bmatrix}, \quad (54)$$

festgelegt. Durch Einführung zweier orthogonaler Matrizen

$$\mathbf{Q}_1 = \frac{1}{\sqrt{2}} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & -1 \end{bmatrix} \quad \mathbf{Q}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (55)$$

die jeweils von links und von rechts multipliziert werden, ergibt sich folgende Struktur,

$$\mathbf{Q}_2^T \mathbf{Q}_1^T \tilde{\mathbf{Y}}_1^T \tilde{\mathbf{Y}}_1 \mathbf{Q}_1 \mathbf{Q}_2 = \begin{bmatrix} 4a^2 + (c_1 - c_2)^2 & c_1^2 - c_2^2 & 0 & 0 \\ c_1^2 - c_2^2 & (c_1 + c_2)^2 & 0 & 0 \\ 0 & 0 & 2s_1^2 & 0 \\ 0 & 0 & 0 & 2s_2^2 \end{bmatrix} \quad (56)$$

Als nächsten Schritt ist es notwendig, die links obere 2×2 Matrix mit Hilfe von orthogonalen Matrizen zu diagonalisieren,

$$\begin{aligned} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} 4a^2 + (c_1 - c_2)^2 & c_1^2 - c_2^2 \\ c_1^2 - c_2^2 & (c_1 + c_2)^2 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} &= \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} e & f \\ f & h \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \\ &= \begin{bmatrix} ec^2 + efs c + hs^2 & f(c^2 - s^2) + (h - e)sc \\ f(c^2 - s^2) + (h - e)sc & hc^2 - 2fsc + es^2 \end{bmatrix}. \end{aligned} \quad (57)$$

Mit $\begin{bmatrix} \sqrt{1-x^2} & -x \\ x & \sqrt{1-x^2} \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$ als orthogonale Matrix und der Bedingung, dass die Nebendiagonale verschwinden muss, folgen Berechnungsvorschriften für x ,

$$\begin{aligned} (1 - 2x^2)f + x\sqrt{1-x^2}(h - e) &= 0 \\ x\sqrt{1-x^2}(h - e)^2 &= (2x^2 - 1)f \\ -(x^4 - x^2)(h - e)^2 &= \underbrace{4x^4 f^2 - 4x^2 f^2}_{(x^4 - x^2)4f^2} + f^2 \\ [(h - e)^2 + 4f^2]x^4 - [(h - e)^2 + 4f^2]x^2 + f^2 &= 0 \\ \pm \sqrt{\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{f^2}{(h - e)^2 + 4f^2}}} &= x \end{aligned} \quad (58)$$

Es wird zur Berechnung jenes Vorzeichen für x verwendet, für das

$$(1 - 2x^2)f + x\sqrt{1 - x^2}(h - e) = 0 \quad (59)$$

gilt. Für das innere Vorzeichen kann stets das negative Vorzeichen verwendet werden (die geringstmögliche Rotation). Mit $e = 4a^2 + (c_1 - c_2)^2$, $f = c_1^2 - c_2^2$, $h = (c_1 + c_2)^2$ kann ein einfacher Ausdruck gefunden werden,

$$\begin{aligned} x &= \pm \sqrt{\frac{1}{2} \pm \frac{1}{2} \sqrt{1 - \frac{(c_1^2 - c_2^2)^2}{(c_1 c_2 - a^2)^2 + (c_1^2 - c_2^2)^2}}} \\ &= \pm \sqrt{\frac{1}{2} \pm \frac{1}{2} \sqrt{\frac{(c_1 c_2 - a^2)^2}{(c_1 c_2 - a^2)^2 + (c_1^2 - c_2^2)^2}}} \\ &= \pm \frac{1}{\sqrt{2}} \sqrt{1 \pm \sqrt{1 + \frac{1}{\frac{(c_1^2 - c_2^2)^2}{(c_1 c_2 - a^2)^2}}}}. \end{aligned} \quad (60)$$

Damit kann die Singulärwertzerlegung der quadrierten Enkodiermatrix $\tilde{\mathbf{Y}}_1^T \tilde{\mathbf{Y}}_1$ als

$$\begin{aligned} \underbrace{\mathbf{Q}_3^T \mathbf{Q}_2^T \mathbf{Q}_1^T}_{\mathbf{U}^T = \mathbf{V}^T} \tilde{\mathbf{Y}}_1^T \tilde{\mathbf{Y}}_1 \underbrace{\mathbf{Q}_1 \mathbf{Q}_2 \mathbf{Q}_3}_{\mathbf{V}} &= \mathbf{S} \\ \tilde{\mathbf{Y}}_1^T \tilde{\mathbf{Y}}_1 &= \mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{V} \mathbf{S} \mathbf{V}^T \end{aligned} \quad (61)$$

mit

$$\begin{aligned} \mathbf{V} &= \mathbf{Q}_1 \mathbf{Q}_2 \mathbf{Q}_3 \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{1 - x^2} & -x & 0 & 0 \\ x & \sqrt{1 - x^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{x + \sqrt{1 - x^2}}{2} & \frac{\sqrt{1 - x^2} - x}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{x + \sqrt{1 - x^2}}{2} & \frac{\sqrt{1 - x^2} - x}{2} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{\sqrt{1 - x^2} - x}{2} & \frac{-x - \sqrt{1 - x^2}}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{\sqrt{1 - x^2} - x}{2} & \frac{-x - \sqrt{1 - x^2}}{2} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \\ \mathbf{S} &= \text{diag} \left\{ \begin{bmatrix} 2a^2 + c_1^4 + c_2^4 - \sqrt{4a^4 - 8a^2 c_1 c_2 + c_1^4 + 2c_1^2 c_2^2 + c_2^4} \\ 2a^2 + c_1^4 + c_2^4 + \sqrt{4a^4 - 8a^2 c_1 c_2 + c_1^4 + 2c_1^2 c_2^2 + c_2^4} \\ 2s_1^2 \\ 2s_2^2 \end{bmatrix} \right\} \end{aligned} \quad (62)$$

angeschrieben werden.

Literatur

- [Ahr19] J. Ahrens, “Auralization of Omnidirectional Room Impulse Responses Based on the Spatial Decomposition Method and Synthetic Spatial Data,” in ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 146–150.
- [BB10a] N. Barrett and S. Berge, “A New Method for B-Format to Binaural Transcoding,” in Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space, Oct 2010. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=15527>
- [BB10b] S. Berge and N. Barrett, “High Angular Plane Wave Expansion,” in Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics, 05 2010.
- [Deu] Deutsche Gesellschaft für Akustik, “Stimulus-Datenbank: Jazz/Funk „Rhythmusgruppe“: What’s Trumps,” accessed 2020-04-16. [Online]. Available: <https://www.dega-akustik.de/fachausschuesse/va/stimulus-datenbank/>
- [Deu06] —, “DEGA-Empfehlung 101,” 03 2006.
- [EBU08] EBU, “EBU SQAM CD: Sound Quality Assessment Material recordings for subjective tests,” 2008.
- [FZ08] M. Frank and F. Zotter, “Localization Experiments Using Different 2D Ambisonics Decoders,” 01 2008.
- [GCC⁺11] P.-A. Gauthier, E. Chambatte, C. Camier, Y. Pasco, and A. Berry, “Derivation of Ambisonics Signals and Plane Wave Description of Measured Sound Field Using Irregular Microphone Arrays and Inverse Problem Theory,” 06 2011.
- [Gra] M. Graef, “Quadrature Rules on Manifolds,” accessed 2020-04-16. [Online]. Available: <https://homepage.univie.ac.at/manuel.graef/quadrature.php>
- [Int15] International Telecommunication Union, “Recommendation ITU-R BS.1534-3, Method for the subjective assessment of intermediate quality level of audio systems,” 10 2015.
- [Kro20] M. Kronlachner, “mcfx Plugin Suite,” 2020, accessed 2020-04-16. [Online]. Available: <http://www.matthiaskronlachner.com/?p=1910>
- [MP05] J. Merimaa and V. Pulkki, “Spatial Impulse Response Rendering I: Analysis and Synthesis,” J. Audio Eng. Soc, vol. 53, no. 12, pp. 1115–1127, 2005. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=13401>
- [Oct] Octava, “Octava Shop: MK-4012 4-D Ambisonic A-format microphone,” accessed 2020-04-17. [Online]. Available: <http://www.oktava-shop.com/Small-and-medium-diaphragm-condenser-mics/Oktava-MK-012-01-12.html?language=en>
- [PM06] V. Pulkki and J. Merimaa, “Spatial Impulse Response Rendering II: Reproduction of Diffuse Sound and Listening Tests,” J. Audio Eng. Soc, vol. 54, no. 1/2, pp. 3–20, 2006. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=13664>

- [Rai] M. Raindl, “Gram-Schmidt’sches Orthogonalisierungsverfahren,” accessed 2020-04-19. [Online]. Available: <https://www.opt.math.tugraz.at/~raindl/archiv/ortho.pdf>
- [Rod] Rode, “Røde NT-SF1 Produktseite,” accessed 2020-04-17. [Online]. Available: <https://www.rote.com/ntsf1>
- [Rud20] D. Rudrich, “IEM Plugin Suite,” 2020, accessed 2020-04-16. [Online]. Available: <https://plugins.iem.at>
- [Sen] Sennheiser, “Sennheiser Ambeo VR Mic Produktseite,” accessed 2020-04-17. [Online]. Available: <https://de-at.sennheiser.com/mikrofon-3d-audio-ambeo-vr-mic>
- [Soua] Soundfield, “Soundfield SPS200 Bedienungsanleitung ,” accessed 2020-04-17. [Online]. Available: <http://cdn.soundfield.com/assets/downloads/manual/SPS200.pdf>
- [Soub] —, “Soundfield ST450 Bedienungsanleitung,” accessed 2020-04-17. [Online]. Available: <https://www.teltec.de/out/media/ST450-MKII-manual.pdf>
- [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>
- [ZF13] F. Zotter and M. Frank, “Efficient Phantom Source Widening,” *Archives of Acoustics*, vol. 38, pp. 27–37, 03 2013.
- [ZF17] —, “Phantom Source Widening by Filtered Sound Objects,” in *Audio Engineering Society Convention 142*, May 2017. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=18605>
- [ZF19] —, *Ambisonics - A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, first edition ed. Heidelberg (Deutschland): Springer, 05 2019.
- [ZFZ20] M. Zaunschirm, M. Frank, and F. Zotter, “Binaural Rendering with Measured Room Responses: First-Order Ambisonic Microphone vs. Dummy Head,” *Applied Sciences*, vol. 10, no. 5, p. 1631, Feb 2020. [Online]. Available: <http://dx.doi.org/10.3390/app10051631>
- [Zoo] Zoom, “Zoom H3-VR Produktseite,” accessed 2020-04-17. [Online]. Available: <https://www.zoom-na.com/products/field-video-recording/field-recording/zoom-h3-vr-handly-recorder#media>