Vorhersage des Hörbereichs für Surround-Wiedergabesysteme auf Basis des Energievektormodells

Toningenieur-Projekt

Eric Kurz

Betreuung: Dr. Matthias Frank

Graz, 22. November 2017

institut für elektronische musik und akustik

Zusammenfassung

Bei der Reproduktion dreidimensionaler Schallfelder stellt Higher-Order Ambinsonics schon seit längerer Zeit einen Forschungsgegenstand dar. Ziel der Forschung ist u.a. die Schallfeldreproduktion auf einen möglichst großen Zuhörerbereich auszuweiten. Zur Prädiktion der Schallfeldreproduktion erfreut sich der sogenannte Energievektor ($\mathbf{r_E}$-Vektor) großer Beliebtheit. Dieser hat sich bereits zur Vorhersage der wahrgenommenen Hörereignisrichtung und -breite an der zentralen Hörposition bewährt. Die Vorhersage an dezentralen Hörpostionen mit Hilfe des Energievektors gestaltet sich jedoch schwieriger. In dieser Projektarbeit werden verschiedene Erweiterungen des Energievektors miteinander verglichen. Die Erweiterungen versuchen Laufzeit- und Pegelunterschiede, sowie den Präzedenz-Effekt in Abhängigkeit der Hörposition zu berücksichtigen. Anhand existierender Hörversuchsdaten werden die Erweiterungen evaluiert und Parameter dieser optimiert. Simulationen untersuchen das Lautsprecher-Array des IEM CUBE mit ambisonischer Wiedergabe 1. bis 5. Ordnung. Es werden die erweiterten Energievektoren an diskreten Punkten im gesamten Zuhörerbereich berechnet. Durch diese Berechnungen kann die Lokalisationsgenauigkeit der erweiterten Energievektoren an dezentralen Hörpositionen verglichen werden und der Hörbereich mit ausreichender Lokalisationsgenauigkeit abgeschätzt werden.

Abstract

For some time, Higher-Order Ambisonics is a research topic in reproducing three-dimensional sound fields. Among others objectives, research aims at sound field reproduction over a large listening area. For prediction of a sound field the so-called energy vector ($\mathbf{r_E}$-vector) is a commonly used tool. At the central listening position, the energy vector provides good results for predicting position and width of a phantom source. However, at off-center listening positions, the performance of the energy vector becomes worse. This project work compares different extensions of the energy vector model. The extensions try to include time and level differences, as well as the precedence effect in dependence of the listening position. Existing data from listening experiments is used to evaluate the extensions and to optimize some parameters. Simulations investigate the loudspeaker array of the IEM CUBE for Ambisonic playback of 1st to 5th order. Extended energy vectors for discrete points within the whole listening area are computed. The resulting localization accuracy of the extended energy vectors at off-center listening positions can be compared to further predict areas with sufficient localization accuracy.

Inhaltsverzeichnis

1 Einleitung

Versucht man ein immersives Schallereignis für ein möglichst großes Auditorium zu reproduzieren kann man heutzutage auf verschiedenste Verfahren zurückgreifen. Es gibt kanalbasierte Verfahren, wie z.B. Dolby Atmos, Auro 3D oder Vector-Base Amplitude Panning (VBAP), objektbasierte Verfahren (Wellenfeldsynthese) und szenenbasierte Verfahren, wie z.B. Higher Order Ambisonics. Die verschiedenen Verfahren besitzen ihre Vor- und Nachteile. Die kanalbasierten Verfahren benötigen für die Reproduktion des Schallfelds eine standardisierte Lautsprecheranordnung. In der ITU-R BS.2051 [23] sind für diese Verfahren Anordnungen mit bis zu 22 Lautsprechern definiert. Kanalbasierte Verfahren erfreuen sich vor allem in der Beschallung von Kinosälen großer Beliebtheit. Allerdings beschränkt sich der Abhörbereich dieser Verfahren, der sogenannte Sweetspot, auf einen relativ kleinen Bereich um den Mittelpunkt der Lautsprecheranordnung. Bei der Wellenfeldsynthese werden durch die Überlagerung sogenannter Elementarwellen beliebige Wellenfronten unterhalb einer Grenzfrequenz physikalisch korrekt synthetisiert [1,2]. Dabei ist man bei der Platzierung der virtuellen Schallquelle nicht auf die Lautsprecherebene beschränkt. Ein großer Vorteil dieses Verfahren ist die Ausdehnung des Abhörbereichs. Dieser erstreckt sich über den gesamten Bereich vor dem Lautsprecher-Array. Allerdings benötigt man für eine genaue Reproduktion des Schallfeldes im gesamten Hörfrequenzbereich eine sehr große Anzahl an Lautsprechern. So ist z.B. das Wellenfeldsynthesesystem im Hörsaal 104 der TU Berlin mit 2700 Lautsprecher an 832 unabhängigen Kanälen ausgestattet [32]. Higher Order Ambisonics (HOA) kommt hingegen mit wesentlich weniger Lautsprechern aus. Im sogenannten CUBE des Instituts für Elektronische Musik und Akustik an der Universität für Musik und darstellende Kunst Graz (IEM CUBE) sind z.B. nur 24 Lautsprecher verbaut [43]. Die Lautsprecher eines solchen Arrays können relativ beliebig angeordnet sein solange die Hüllfläche der Lautsprecherkuppel gleichmäßig besetzt ist. Der physikalische Sweetspot von HOA-Systemen beschränkt sich, ähnlich wie bei den kanalbasierten Verfahren, auf einen relativ kleinen Bereich um die Mittenposition. Mit ansteigender Ordnung der Wiedergabe vergrößert sich dieser Bereich. Wenn eine ambisonische Wiedergabe erster Ordnung gewählt wird ergibt sich bei einer Frequenz von $f = 700$ Hz ein physikalischer Sweetspot, welcher der Kopfgröße eines einzelnen Hörers im Zentrum der Anordnung entspricht [42].

Frank und Zotter konnten jedoch zeigen, dass der Sweetspot der Hörwahrnehmung, im speziellen für plausible Lokalisation, wesentlich größer ist [16]. Die Diskrepanz zwischen Theorie der Schallfeldreproduktion und der menschlichen Hörwahrnehmung sorgt dafür, dass der Hörbereich eines Lautsprecher-Arrays für ambisonische Wiedergabe durch Hörversuche (ähnlich zu [16]) ausgemessen werden muss. Diese Hörversuche sind jedoch sehr zeitaufwändig. Vor allem wenn der Einfluss verschiedener Parameter wie z.B. der Ordnung der ambisonischen Wiedergabe oder der Dekodergewichtung untersucht werden soll. Motivation dieser Projektarbeit ist es ein technisches Hilfsmittel zu entwickeln, welches zur Vorhersage des Hörbereichs von Surround-Wiedergabesystemen eingesetzt werden kann.

Frank konnte zeigen, dass sich der $\mathbf{r_E}$-Vektor zur Prädiktion der Lokalisationsrichtung und Quellenbreite einer virtuellen Schallquelle an der Mittenposition bei ambisonischer

Wiedergabe eignet [12–14]. Weiterhin konnte er zeigen, dass die zeitliche Längenänderung des $\mathbf{r_E}$-Vektors die Klangfarbenänderung beim Panning der virtuellen Schallquelle beschreibt. Um die Hörwahrnehmung und im speziellen die Lokalisation an dezentralen Hörpositionen beschreiben zu können entwickelte Stitt ein erweitertes $\mathbf{r_E}$-Vektormodell, welches den Präzendenzeffekt und den „Cone-of-Confusion“ in die Modellierung einbezieht [38].

Diese Projektarbeit soll aufbauend auf den Arbeiten von Frank und Stitt Erweiterungen des $\mathbf{r_E}$-Vektormodells untersuchen, welche für die Prädiktion des Hörbereichs von immersiven Wiedergabesystemen verwendet werden sollen. Hierfür wird zunächst der Algorithmus des einfachen $\mathbf{r_E}$-Vektors implementiert. Weiterhin werden drei aufeinander aufbauende Erweiterungen des Vektormodells implementiert. Es wird zunächst die Schallausbreitungsdämpfung integriert. Dann folgen zeitabhängige Gewichtungsterme sowie die Einbeziehung des Präzendenzeffekts. Die Komplexität der Algorithmen nimmt mit jeder Erweiterung zu. Zum Vergleich wird der von Stitt in [38] verwendete Algorithmus in die Simulationen integriert. Als Simulationsumgebung wird zunächst der untere Ring der sphärischen Lautsprecheranordnung des IEM CUBE verwendet (siehe Abb. 1).

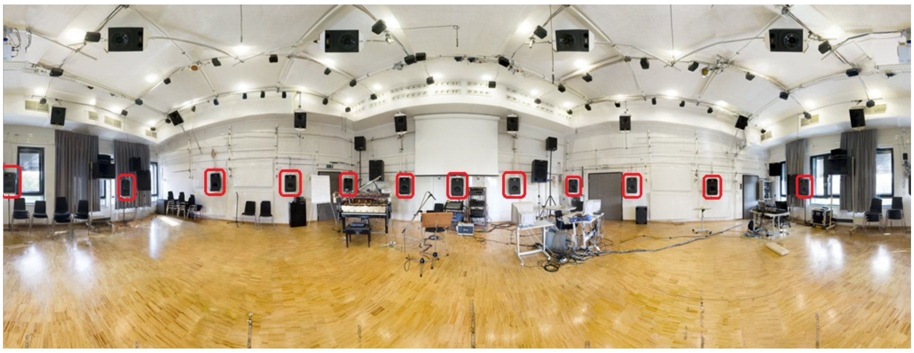


Abbildung 1 – $360°$-Panorama des IEM CUBE mit den für die Simulation verwendeten Lautsprechern (rot markiert). IEM CUBE Panorama ©2004 Alexandre Castonguay.

Stitt führt in [38] den Parameter α ein. Dieser wird für das Überblenden zwischen gewöhnlicher und erweiterter Berechnung des $\mathbf{r_E}$-Vektors verwendet. Auch in dieser Projektarbeit wird α für die Simulation verwendet. Um ein geeignetes α für die weiteren Simulationen zu finden wird eine Optimierung des Parameters anhand der von Frank und Zotter experimentell ermittelten Hörbereiche für den IEM CUBE durchgeführt [16]. Anschließend wird mit dem gefundenen α ein Performance-Vergleich der verschiedenen $\mathbf{r_E}$-Algorithmen mit den Daten des Hörversuchs von [16] und [10] durchgeführt.

2 Higher Order Ambisonics (HOA)

Ambisonics ist ein Mehrkanalwiedergabeverfahren, das versucht ein Schallfeld in einem möglichst großen Hörbereich innerhalb zirkulärer oder sphärischer Lautsprecher-Arrays zu reproduzieren. Die ersten Versuche zur ambisonischen Signalverarbeitung und Aufnahme von zweidimensionalen Schallfeldern wurden von Cooper und Shiga durchgeführt [3]. Cooper arbeitete an der Erweiterung der Mid/Side-Stereofonie und experimentierte mit der Kombination von Kugel- und rotierten Achterrichtcharakteristiken. Ausgehend von dieser Arbeit entwickelten maßgeblich Felgett [9], Gerzon [19] und Craven [4] Ambisonics. Ein entscheidender Vorteil von Ambisonics, gegenüber anderer Wiedergabeverfahren wie Wellenfeldsynthese (WFS) oder Vector-Base Amplitude Panning (VBAP), ist die Trennung der Signalverarbeitung in En- und Dekodierstufe. Somit wird ein sehr flexibles System der Aufnahme als auch Wiedergabe dreidimensionaler Schallfelder erschaffen. Der Ansatz von Ambisonics ist es eine einfallende Schallwelle in eine Anzahl von Kugelflächenfunktionen zu zerlegen. Dieser Vorgang ist in der ambisonischen Signalverarbeitung im Enkoder implementiert. Theoretisch werden für eine perfekte Rekonstruktion der Schallwelle unendlich viele Kugelflächenfunktionen benötigt. In der Praxis ist die Anzahl der Kugelflächenfunktionen N limitiert. Sie ergibt sich aus der verwendeten Ordnung M von Ambisonics (Gl. 1 und 2) [6]. Ambisonische Signalverarbeitung mit einer Ordnung $M > 1$ wird unter dem Begriff „Higher Order Ambisonics“ (HOA) zusammengefasst. Die Anzahl der Kugelflächenfunktionen entspricht der benötigten Kanalanzahl des Enkoders. Die enkodierten Ambisonics-Kanäle können gespeichert und manipuliert werden [25].

$$\text{3D-Ambisonics:} \quad N = (M+1)^2 \tag{1}$$

$$\text{2D-Ambisonics:} \quad N = 2M+1 \tag{2}$$

Sollen Ambisonics-Aufnahmen bzw. -Produktionen wiedergegeben werden wird ein Dekoder benötigt. Dieser trägt die ambisonischen Signale mit einer entsprechenden Gewichtung auf die Lautsprecher eines sphärischen Arrays auf. Durch die Wiedergabe der dekodierten Signale über die Lautsprecher wird das Schallfeld reproduziert. Das Lautsprecher-Array muss keine standardisierte Form besitzen, da die Gewichte des Dekoders entsprechend der Lautsprecheraufstellung angepasst werden können. Jedoch müssen die Lautsprecher gleichmäßig auf der Sphäre angeordnet sein. Einige Methoden zur Abtastung einer sphärischen Oberfläche wurden von Zotter in [44] und [45] beschrieben. Analog zur Enkodierung ist eine Mindestanzahl von Lautsprecher nötig um Ambisonics mit einer bestimmten Ordnung wiederzugeben (siehe Gl. 1 und 2). Weiterhin kann bei der Dekodierung eine Gewichtung vorgenommen werden. Am gebräuchlichsten sind die max-$\mathbf{r_E}$- und die in-phase-Gewichtung [20, 31].

	standard	**max $\mathbf{r_E}$**	**in-phase**
$\mathbf{g_m}$	1	$cos\left(\frac{m\pi}{2M+2}\right)$	$\frac{M!^2}{(M-m)!(M+m)!}$

Tabelle 1 – 2D-Ambisonics Gewichte

Außerdem ist es möglich bei der Wiedergabe von Ambisonics über ein sphärisches Lautsprecher-Array die Laufzeitunterschiede τ der einzelnen Lautsprecher bezüglich des Mittelpunktes der sphärischen Anordnung zu kompensieren. Frank konnte zeigen, dass sich die Laufzeit-Kompensation negativ auf die wahrgenommene Wiedergabequalität und Lokalisation an detzentralen Hörpositionen auswirken kann [10].

3 Vektormodelle der Lokalisation

Vektorbasierte Modelle haben eine große Bedeutung in der Beschreibung der menschlichen Hörwahrnehmung, im Speziellen der Lokalisation. An der Hörposition innerhalb eines sphärischen/zirkulären Arrays wird ein Aufpunkt definiert. An diesem Punkt werden die Vektoren in Richtung der Lautsprecher des Arrays entsprechend gewichtet aufgetragen. Anschließend werden die Vektoren addiert, eine Normalisierung wird durchgeführt und es ergibt sich ein resultierender Vektor.

Gerzon teilt Vektormodelle nach drei verschiedenen Kategorien ein [21]. Der Grad eines Modells beschreibt seine Nichtlinearität. Somit haben lineare Modelle ersten Grad, quadratische Modelle zweiten Grad usw.. Die Ordnung eines Modells beschreibt die Ordnung M der verwendeten Kugelflächenfunktionen. Ein Vektormodell nullter Ordnung ist omnidirektional wohin gegen ein Modell erster Ordnung direktional hinsichtlich der drei Raumrichtungen ist. Der dritte Parameter beschreibt den Einfluss von Kopfbewegungen auf das verwendete Lokalisationsmodell.

Die beiden bekanntesten Vektormodelle der Lokalisation sind der Schnellevektor ($\mathbf{r_V}$-Vektor) und der Energievektor ($\mathbf{r_E}$-Vektor). Diese werden in den folgenden Abschnitten (3.1 und 3.2) vorgestellt.

3.1 Schnellevektor ($\mathbf{r_V}$-Vektor)

Der Schallschnellevektor, kurz $\mathbf{r_V}$-Vektor, steht in Zusammenhang mit der Schnelle der Teilchen des schwingenden Mediums (Luft) an der Hörposition und lässt sich über folgende Formel ermitteln [21]:

$$\mathbf{r_V} = \frac{\sum\limits_{l=1}^{L} g_l \boldsymbol{\theta}_l}{\sum\limits_{l=1}^{L} g_l} \tag{3}$$

Hierbei sind die Variablen wie folgt definiert:

g_l	...	komplexes Gewicht (Betrag und Phase) des l-ten Lautsprechers
$\boldsymbol{\theta}_l$	...	Einheitsvektor in Normalrichtung des l-ten Lautsprechers

Somit ist der Schnellevektor ein Modell erster Ordnung und 1. Grades, welches den Einfluss von Kopfbewegungen nicht beachtet. Er eignet sich für die Vorhersage der Lokalisation von virtuellen Schallquellen bei Frequenzen von $f < 700$ Hz. Der Schnellevektor ist eng mit dem Vector-Base Amplitude Panning (VBAP) [33] und dem Tangensgesetz [26] verknüpft.

Die Gleichung 3 nimmt die Ausbreitung ebener Wellen, welche aus dem Unendlichen einlaufen, an. In der Realität ergibt sich jedoch durch die endliche Distanz der Lautsprecher zur Hörposition und die Ausbreitung von eher kugelförmigen Wellen eine Anhebung der tiefen Frequenzen.

3.2 Energievektor ($\mathbf{r_E}$-Vektor)

Ein weiteres weit verbreitetes Vektormodell zur Prädiktion der wahrgenommenen Phantomschallquellenrichtung ist der Energievektor. Er wird auch kurz $\mathbf{r_E}$-Vektor genannt. Er ist definiert als die Superposition der einzelnen Energievektoren aller Lautsprecher an einem Aufpunkt.

$$\mathbf{r_E} = \frac{\sum_{l=1}^{L} |g_l|^2 \boldsymbol{\theta}_l}{\sum_{l=1}^{L} |g_l|^2} \tag{4}$$

Laut Gerzon handelt es sich bei dem Energievektor um ein Modell erster Ordnung und zweiten Grades, das den Einfluss von Kopfbewegungen nicht beachtet. Er eignet sich für die Vorhersage der Lokalisation von virtuellen Schallquellen bei Frequenzen von 500 Hz $< f <$ 5 kHz.

Jede Phantomschallquelle, welche von mehr als einem Lautsprecher erzeugt wird, besitzt per Definition eine Amplitude von $|\mathbf{r_E}| < 1$. Für die Wiedergabe mit Ambisonics ist es allerdings wünschenswert, dass der $\mathbf{r_E}$-Vektor über alle Panning-Richtungen einen konstanten und möglichst großen Betrag besitzt. Dieses Kriterium wird bei der Optimierung von Ambisonics-Dekodern eingesetzt (siehe Abschnitt 2).

Weiterhin eignet sich der $\mathbf{r_E}$-Vektor sehr gut als Lokalisationsprädiktor für breitbandige Phantomschallquellen, wie bei Versuchen in [12, 14, 18] für frontale und horizontal umhüllende Wiedergabesituationen, u.a. erzeugt mit VBAP und Ambisonics, gezeigt wurde. Die wesentlich komplexeren binauralen Modelle, wie z.B. das Modell nach Jeffress [24], Dietz [7] oder Lindemann [28, 29], erzielen bei der Prädiktion der Lokalisation wesentlich schlechtere Ergebnisse (siehe [14, 18]). Sowohl der Intensitätsvektor als auch der $\mathbf{r_V}$-Vektor können die Vorhersagegenauigkeit des $\mathbf{r_E}$-Vektors nicht erreichen. Auch die wahrgenommene Quellenbreite W kann in direkten Zusammenhang mit dem $\mathbf{r_E}$-Vektor gesetzt werden [13]. Für frontale Phantomschallquellen ist sie indirekt proportional zum Betrag des Energievektors (vgl. Gl. 5).

$$W = 186.4^\circ \cdot (1 - |\mathbf{r_E}|) + 10.7^\circ \tag{5}$$

Der Korrekturterm von 10.7° ist von Parametern (z.B. Nachhallzeit T_{60}) des Raumes abhängig und wurde für den IEM CUBE [43] ermittelt. Weiterhin kann die Längenänderung des Energievektors bei einer horizontal rotierenden virtuellen Schallquelle mit der Klangfarbenänderung dieser verbunden werden [12].

Auch für die Lokalisationsprädiktion an dezentralen Hörpositionen wurde der Energievektor schon vorgeschlagen [6]. Durch variierende Distanzen zwischen den Lautsprechern des Ambisonics-Arrays und der Hörposition unterliegen die Phasen der einzelnen Lautsprechersignale zueinander einer zufälligen Verteilung. Diese werden bei der Berechnung des Energievektors jedoch nicht berücksichtigt. Um die Qualität der Vorhersage des Energievektors für diesen Anwendungsfall zu untersuchen, müssen allerdings noch Hörversuche durchgeführt werden.

4 Adaption des $\mathbf{r_E}$-Vektors

4.1 Gesetz der ersten Wellenfront

Das Gesetz der ersten Wellenfront ist ein wichtiger Aspekt der menschlichen Hörwahrnehmung und hat große Bedeutung bei der Lokalisation. Es besagt, dass ein identisches Schallereignis, welches einem Hörer zeitverzögert aus verschiedenen Richtungen präsentiert wird, aus der Richtung des ersten eintreffenden Schallwelle lokalisiert wird [5]. Das Gesetz ist auch unter dem Begriff „Präzedenz-Effekt" bekannt [30, 41]. Ab einer zeitlich Verzögerung von der ersten zur folgenden Wellenfront von 2 ms bis ca. 50 ms (Sprache) bzw. 100 ms (Musik) ist der Präzedenz-Effekt von Bedeutung. Unterhalb der Zeitverzögerung von 2 ms tritt Summenlokalisation auf. Oberhalb der Grenzen für Sprache bzw. Musik ist der später eintreffende Schall als Echo wahrnehmbar. Deshalb wird die obere Zeitverzögerungsgrenze auch Echoschwelle genannt. Die Echoschwelle ist sehr stark abhängig von der Länge des präsentierten Stimulus und seinen spektralen Komponenten. Selbst wenn der Pegel des verspätet eintreffenden Schalls um 10 dB höher ist als der der ersten Wellenfront, besitzt das Gesetz der ersten Wellenfront noch Gültigkeit [22]. Dieses Verhalten ist als Haas-Effekt bekannt und gilt allerdings nur für Zeitverzögerungen von 10 ms bis ca. 30 ms.

Das Gesetz der ersten Wellenfront wird häufig mit einem sogenannten Lead-Lag-Experiment untersucht. Hierbei wird über ein frontal symmetrisch angeordnetes Lautsprecherpaar einer Versuchsperson der selbe Stimulus zugespielt. Jedoch wird das Signal an einem Lautsprecher zeitlich verzögert wiedergegeben. Für gewöhnlich wird als Stimulus ein kurzes Klicken verwendet, sodass für die meisten Zeitverzögerungen ein sehr kleines bzw. gar kein zeitliches Überlappen der Signale der Lautsprecher statt findet. Ein Lead-Lag-Experiment kann auch mit Kopfhörern durchgeführt werden. Hierbei wird jeder der beiden Klicks mit einer entsprechenden synthetischen interauralen Zeitdifferenz (ITD) wiedergegeben. Somit kann die Richtung der Wiedergabe des Lead- bzw. Lag-Klicks simuliert werden.

Aus früheren Untersuchungen geht hervor, dass die Lokalisationsschärfe bei gleicher ITD von Lead- und Lag-Stimulus abnimmt [37,41]. Weiterhin wurde herausgefunden, dass der Präzedenz-Effekt nicht mehr vorliegt wenn bei seitlicher Wiedergabe die zwei Lautsprecher auf dem selben Cone-of-Confusion platziert werden [34]. Diese Erkenntnis erlangt besondere Bedeutung bei der Lokalisation an dezentralen Hörpositionen bei Wiedergabe in zirkulären bzw. sphärischen Lautsprecheranordnungen. Die Erweiterung des $\mathbf{r_E}$-Vektors nach Stitt versucht diesen Zusammenhang in das $\mathbf{r_E}$-Vektormodell zu integrieren (siehe Abschnitt 4.5).

4.2 Ausbreitungsdämpfung - $\mathbf{r_E}$-Algorithmus 2

Um die Dissipation des Schalls bei der Ausbreitung in Luft zu beschreiben, wird eine Ausbreitungsdämpfung von 6dB je Abstandverdopplung implementiert. Dazu muss zunächst der Abstand des nähesten Lautsprechers des Wiedergabesystems (Lautsprecher 1 bis l)

zur Hörposition d_{min} ermittelt werden.

$$d_{min} = \min(|\boldsymbol{D}_{1...l}|) \tag{6}$$

$\boldsymbol{D}_{1...l}$ stellen hierbei die Vektoren zwischen der Hörposition und den Lautsprechern 1 bis l dar. Der Gewichtungsfaktor $w_{d,l}$ für den l-ten Lautsprecher ergibt sich durch folgenden Quotient.

$$w_{d,l} = \frac{d_{min}}{|\boldsymbol{D}_l|} \tag{7}$$

$w_{d,l}$ wird nun zum Gewicht g_l des l-ten Lautsprechers multipliziert.

$$\mathbf{r_E} = \frac{\sum\limits_{l=1}^{L} (w_{d,l} g_l)^2 \boldsymbol{\theta}_l}{\sum\limits_{l=1}^{L} (w_{d,l} g_l)^2} \tag{8}$$

Eine kleine Simulation soll die Funktionsweise der Erweiterung verdeutlichen. Die Simulation wird mit den 12 Lautsprecherpositionen des unteren Lautsprecherrings des IEM CUBE [43] an der angegebenen Hörposition durchgeführt (vgl. Abb. 2). Die Gewichte der Lautsprecher ergeben sich für eine ambisonische Wiedergabe der virtuellen Schallquelle an der in Abb. 2 angegebenen Position. Es wurde mit 1. Ordnung Ambisonics, max-$\mathbf{r_e}$-Gewichtung und ohne Delay-Kompensation wiedergegeben.

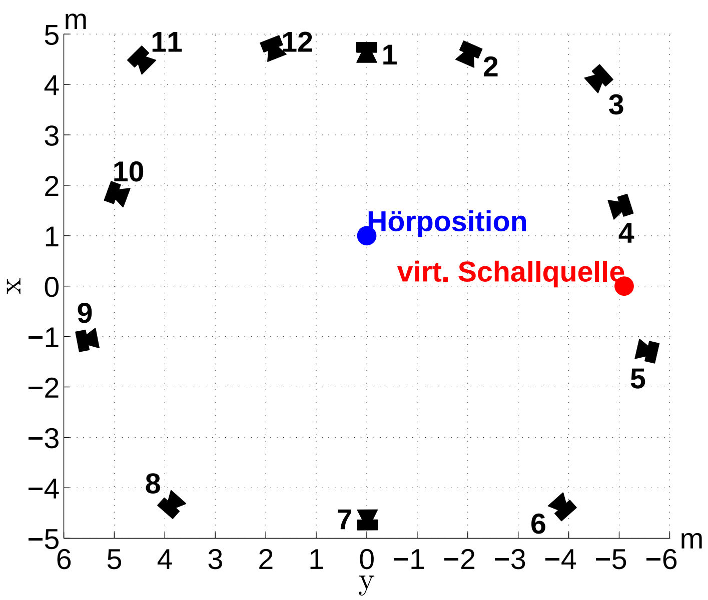
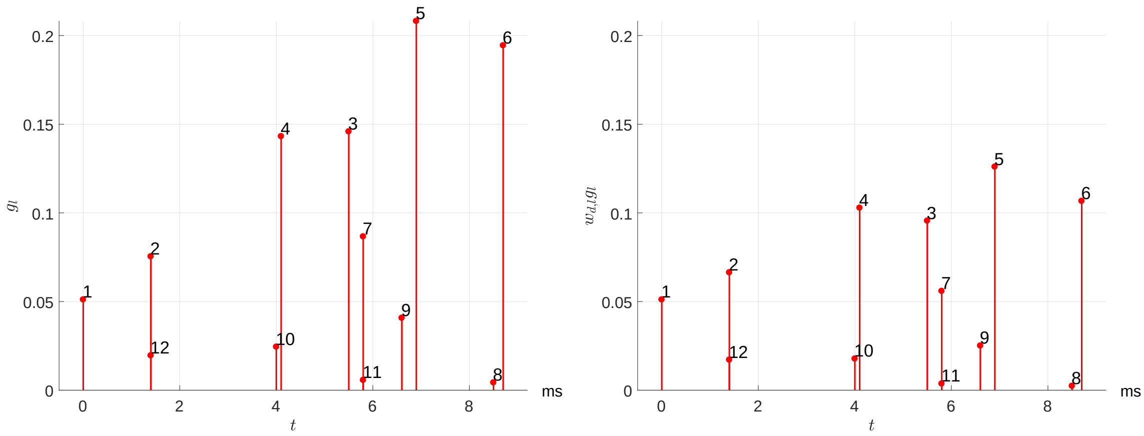


Abbildung 2 – Simulationsumgebung IEM CUBE mit Hörposition und Position der virtuellen Schallquelle.

Abbildung 3 (a) zeigt die Gewichte g_l der Lautsprecher als Impulse entsprechender Höhe, welche mit einer entsprechenden zeitlichen Verzögerung an der Hörposition eintreffen. Werden die Gewichte g_l mit den entsprechenden Dämpfungsfaktoren $w_{d,l}$ nach Gl. 8 beaufschlagt ergibt sich der in Abb. 3(b) ersichtliche Impulszug.



(a) (b)

Abbildung 3 – (a): Impulszug anhand der Gewichte g_l für die Simulationsumgebung des IEM CUBE (vgl. Abb. 2). (b): Berücksichtigung der Ausbreitungsdämpfung anhand Gl. 8.

4.3 Zeitliche Gewichtung - $\mathbf{r_E}$-Algorithmus 3

Ein relativ einfacher Ansatz (vgl. 4.5) das Gesetz der ersten Wellenfront in die Berechnung des $\mathbf{r_E}$-Vektors zu integrieren ist es, verspätet an der Hörposition eintreffende Schallwellen entsprechend ihrer Verzögerung weniger stark zu gewichten.

Es wird ein weiterer Faktor $w_{\tau,l}$ zu dem Gewicht g_l des l-ten Lautsprechers multipliziert.

$$\mathbf{r_E} = \frac{\sum_{l=1}^{L} (w_{\tau,l} w_{d,l} g_l)^2 \boldsymbol{\theta}_l}{\sum_{l=1}^{L} (w_{\tau,l} w_{d,l} g_l)^2} \tag{9}$$

Nun werden die eintreffenden Signale nach ihrer Ankunftszeit T_1 bis T_l aufsteigend geordnet. Die Verzögerung des l-ten Signals ergibt sich durch Subtraktion der Ankunftszeit des ersten Signals T_1 von der des l-ten Signals T_l.

$$\tau_l = T_l - T_1 \quad \text{mit} \quad T_1, T_l \text{ in ms} \tag{10}$$

Die Faktoren $w_{\tau,l}$ können jetzt mit dem in [46] verwendeten Faktor $w_\tau = -0.25$ dB/ms berechnet werden.

$$w_{\tau,l} = w_\tau \cdot \tau_l \tag{11}$$

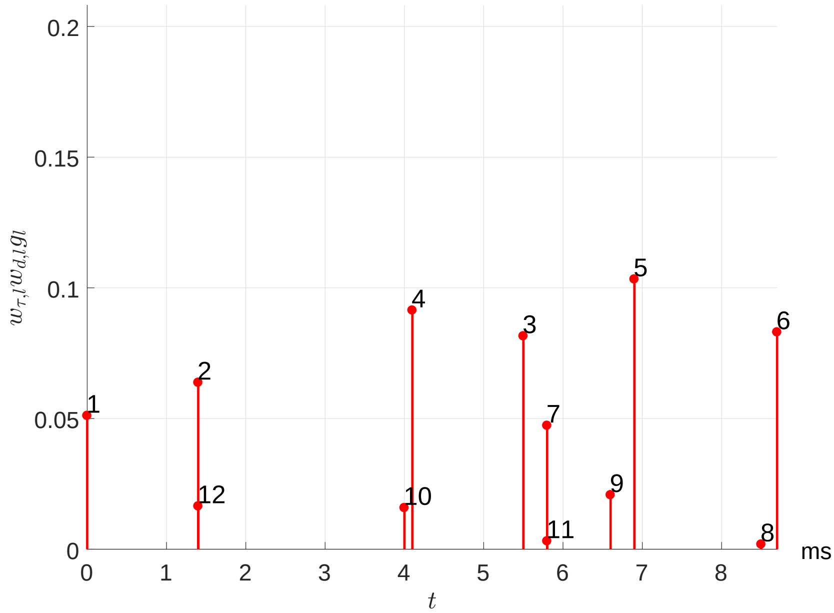


Abbildung 4 – Impulszug anhand der Gewichte g_l für die Simulationsumgebung des IEM CUBE (vgl. Abb. 2) mit Berücksichtigung der Ausbreitungsdämpfung und zeitlicher Gewichtung anhand Gl. 9.

4.4 Schwelle für den Präzedenzeffekt - $\mathbf{r_E}$-Algorithmus 4

Ein wesentlich aufwändigerer Ansatz das Gesetz der ersten Wellenfront in die Berechnung des $\mathbf{r_E}$-Vektors zu integrieren ist es jedem an der Hörposition eintreffenden Signal eine zeitabhängige Maskierungsfunktion aufzuprägen (siehe Abb. 12). Diese Maskierungsfunktion gliedert sich in zwei Bereiche. Im Intervall von $0 \leq t \leq 1$ms steigt die Funktion mit 60dB/ms von -60dB auf 0dB an. Durch diesen Anstieg kann die Summenlokalisation modelliert werden. Im Intervall $t > 1$ms fällt die Maskierungsfunktion mit dem Faktor $w_\tau = -0.25$ dB/ms ab. Somit wird die so genannte „Präzedenzschwelle“modeliert [35,36]. Abbildung 5 zeigt die ersten 50ms dieser Maskierungsfunktion.

$$f_{mask}(t) = \begin{cases} 60\text{dB/ms} & , 0 \leq t \leq 1\text{ms} \\ -0.25\text{dB/ms} & , t > 1\text{ms} \end{cases} \tag{12}$$

Die Verstärkungsfaktoren der Lautsprecher können wieder als Impulse entsprechender Höhe modelliert werden. Diese treffen mit einer zeitlichen Verzögerung an der Hörposition ein (vgl. Abb. 6 (a)). Die Impulse werden nun, wie in Abschnitt 4.2 beschrieben, entsprechend der Hörposition mit der jeweiligen Ausbreitungsdämpfung beaufschlagt (vgl. Abb. 6 (b)). Die Maskierungsfunktion kann nun mit den bedämpften Impulsen gefaltet werden. Es ergibt sich eine entsprechend skalierte und zeitlich verschobene Maskierungsfunktion je Impuls. Diese zeitlich verschobenen Maskierungsfunktionen können nun übereinander gelegt werden und es ergibt sich eine resultierende Maskierungsfunktion $F_{mask}(t)$ (vgl. Abb. 6 (c) und Gl. 13).

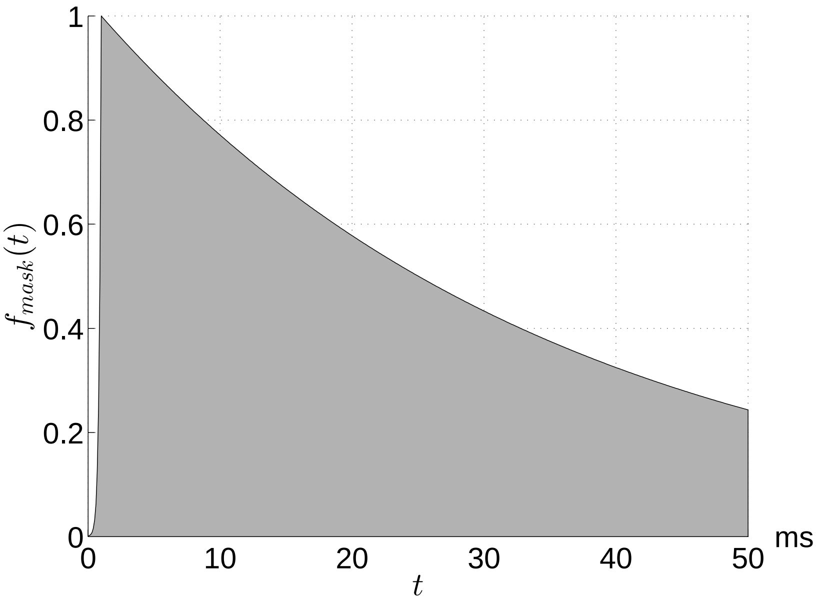


Abbildung 5 – Maskierungsfunktion $f_{mask}(t)$ im Zeitfenster von 0 bis 50ms.

$$F_{mask}(T_l) = \max(g(T_l) * f_{mask}(t)), \quad \text{mit} \quad g(T_l) = g_l \tag{13}$$

Impulse bzw. Gewichte unterhalb dieser resultierenden Präzedenzschwelle werden mit den Gewichten $w_{thresh,l}$ multipliziert. Die Gewichte $w_{thresh,l}$ werden entsprechend der Gl. 14 für den Fall $g(T_l) < F_{mask}(T_l)$ ermittelt(vgl. Abb. 6 (d)).

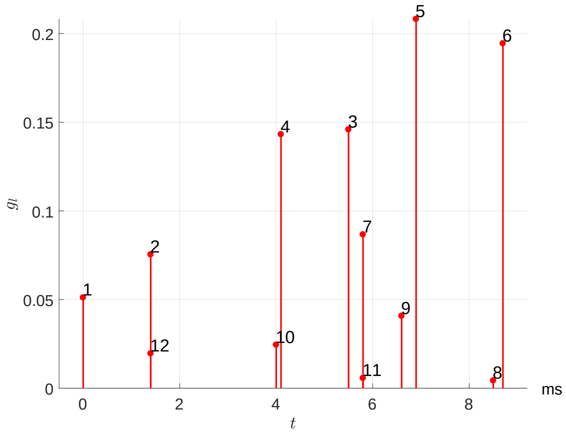
$$w_{thresh,l} = \begin{cases} 1 & , g(T_l) > F_{mask}(T_l) \\ \left(\frac{w_{d,l} g_l}{F_{mask}(T_l)}\right)^2 & , g(T_l) < F_{mask}(T_l) \end{cases} \tag{14}$$

Weiterhin werden die Gewichte g_l mit der zeitlichen Gewichtung von Abschnitt 4.3 versehen. Abbildung 6 (e) soll dies verdeutlichen. Letztendlich können die einzelnen Gewichte entsprechend der Gleichung 15 zusammengeführt werden und es wird ein Parameter α eingeführt. Dieser ermöglicht es zwischen dem mit der Ausbreitungsdämpfung beaufschlagten $\mathbf{r_E}$-Vektor ($\alpha = 1$) und der Erweiterung ($\alpha = 0$) für transiente Signale zu überblenden. α ist abhängig von spektraler Verteilung und Einhüllender des wiedergegebenen Signals (siehe Abschnitt 4.7). Abbildung 6 (f) verdeutlicht den Einfluss auf die Gewichte g_l für $\alpha = 0.5$.

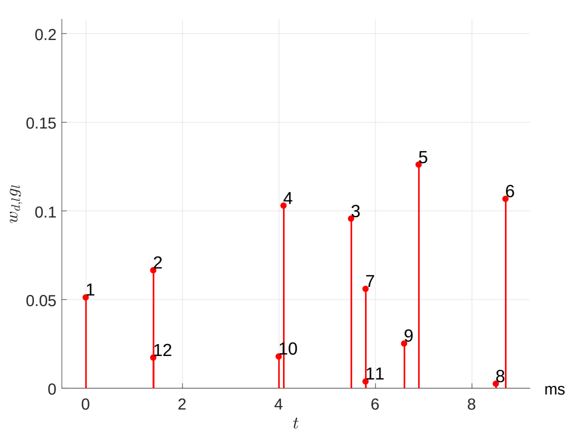
$$\bar{w}_l = ((1 - \alpha) w_{\tau,l} w_{thresh,l} + \alpha) w_{d,l} \tag{15}$$

Äquivalent zu den Abschnitten 4.2 und 4.3 kann der resultierende Faktor $\bar{w}_l$ zu den Gewichten g_l der Lautsprecher multipliziert werden (siehe Gl. 16).

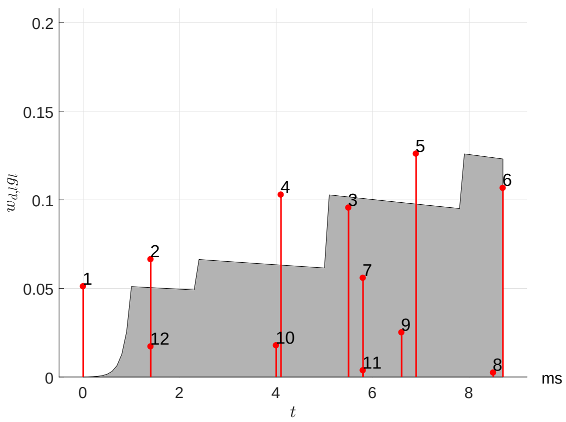
$$\mathbf{r_E} = \frac{\sum\limits_{l=1}^{L} (\bar{w}_l g_l)^2 \boldsymbol{\theta}_l}{\sum\limits_{l=1}^{L} (\bar{w}_l g_l)^2} \tag{16}$$



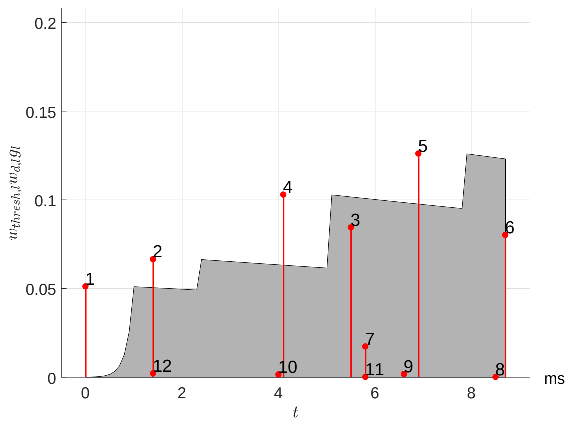
(a)



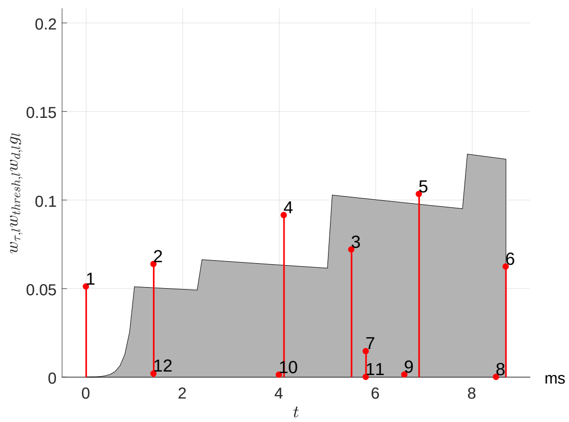
(b)



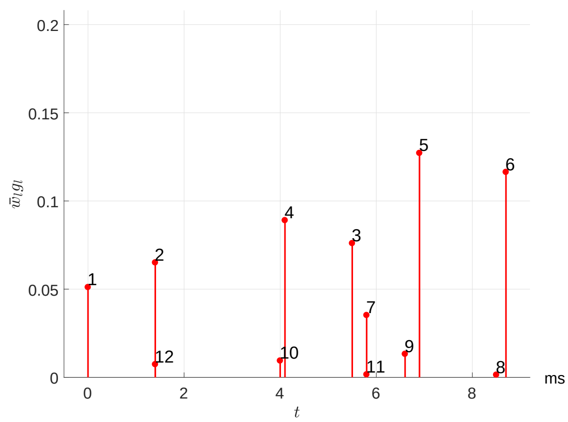
(c)



(d)



(e)



(f) $\alpha = 0.5$

Abbildung 6 – Verarbeitung der Impulsabfolge je Hörposition.

Eine Integration des sog. „Cone-of-Confusion“ wurde nicht durchgeführt, da der Hörer in der Aufführungspraxis seinen Kopf immer leicht bewegt, auch wenn seine Hauptblickrichtung nach vorne gerichtet ist. Des weiteren war auch in den Hörversuchen von [16, 38] die Bewegung des Kopfs möglich.

4.5 Erweiterter $\mathbf{r_E}$-Vektor - $\mathbf{r_E}$-Algorithmus 5

Stitt hat das Gesetz der ersten Wellenfront in seiner Doktorarbeit wie folgt in das $\mathbf{r_E}$-Vektormodell integriert (vgl. [38]).

Zunächst wird ein Faktor $\bar{w}_l$ zu dem Gewicht g_l des l-ten Lautsprechers multipliziert (vgl. Abschnitt 4.4).

$$\mathbf{r_E} = \frac{\sum\limits_{l=1}^{L} (\bar{w}_l g_l)^2 \boldsymbol{\theta}_l}{\sum\limits_{l=1}^{L} (\bar{w}_l g_l)^2} \tag{17}$$

Jeder Faktor $\bar{w}_l$ ist wie folgt definiert:

$$\bar{w}_l = (1-\alpha) w_{\tau R \Theta}^{(l)} + \alpha \tag{18}$$

Der Parameter α ermöglicht es, wie auch in Abschnitt 4.4, stufenlos zwischen dem erweiterten Vektormodell ($\alpha = 0$) und dem $\mathbf{r_E}$-Vektor unter Berücksichtigung der Ausbreitungsdämpfung ($\alpha = 1$) zu überblenden. Stitt schlägt ein $\alpha = 0.5$ für gewöhnliche Signaltypen (z.B. Sprache). Weiterhin ist $w_{\tau R \Theta}^{(l)}$ der Faktor des l-ten Lautsprechers als eine Funktion der Zeitverzögerung τ der eintreffenden Signale, dem relativen Pegel R zwischen diesen und der relativen Richtung Θ der Signale zueinander. Jedes $w_{\tau R \Theta}^{(l)}$ beinhaltet auch die Interaktion aller Lautsprecher miteinander für die jeweilige Hörposition.

Der folgende Abschnitt soll die Herleitung des Faktors $w_{\tau R \Theta}^{(l)}$ näher beleuchten. Es wird zunächst ein 2-kanaliges Lead-Lag-Experiment angenommen mit dessen Hilfe ein zeitabhängiger, ein pegelabhängiger sowie ein richtungsabhängiger Faktor hergeleitet werden. Die Kombination der Faktoren sowie die Verallgemeinerung auf mehrkanalige Systeme führt uns zum Faktor $w_{\tau R \Theta}^{(l)}$.

4.5.1 Zeitabhängigkeit w_τ

Der zeitabhängige Faktor gibt später eintreffenden Signalen ein geringeres Gewicht bei der Berechnung des $\mathbf{r_E}$-Vektors. Stitt schlägt die komplementäre gaußsche Fehlerfunktion mit einem Faktor von 1.09 als Gewichtungsfunktion vor:

$$w_\tau = 1 - \frac{2}{\sqrt{\pi}} \int_0^\tau e^{-(1.09t)^2} dt \tag{19}$$

Hierbei ist τ die Zeitverzögerung zwischen Lead- und Lag-Stimulus in Millisekunden. Die Funktion ist genau so gewählt, dass eine Zeitverzögerung von 1 ms eine Abschwächung des Lag-Stimulus von -18 dB verursacht. Diese kritische Pegeldifferenz sorgt bei Amplituden-Panning für das Kippen der Phantomschallquelle in einen Lautsprecher der Stereoanordnung [27].

4.5.2 Pegelabhängigkeit w_R

Wird der Pegel des nachlaufenden Signals hingegen über einen bestimmten relativen Pegel angehoben bewegt sich die wahrgenommene Phantomschallquelle in Richtung der

Quelle des nachlaufenden Signals. Somit ist die Gewichtung der zeitlich verzögerten Signale auch eine Funktion der Lautsprecherpegel zueinander.

$$w_R = \frac{1}{2} + \mathsf{sgn}(10^{R/20} - 4.25)\frac{\gamma\left(\frac{1}{10}, \left(\frac{|10^{R/20}-4.25|}{3.75}\right)^{10}\right)}{\Gamma(1/10)} \tag{20}$$

Die von Stitt vorgeschlagene Funktion (Gl. 20) ist so gewählt, dass bei einem Pegelunterschied des nachlaufenden Signals von $R = +18$ dB oder mehr der Faktor $w_R = 1$ ist. Γ ist die Eulersche Gammafunktion und γ die unvollständige Gammafunktion der unteren Grenze. Auf den Haas-Effekt wird in diesem Zusammenhang allerdings nicht eingegangen (vgl. Abschnitt 4.1).

4.5.3 Kombinierte Zeit-Pegel-Abhängigkeit

Um die Zeit- als auch die Pegelabhängigkeit des nachlaufenden Signals in einem Faktor zu vereinen schlägt Stitt folgende Formel vor:

$$w_{\tau R} = (1 - w_\tau)w_R + w_\tau \tag{21}$$

Somit ist $0 < w_{\tau R} < 1$ für $0 < w_\tau, w_R < 1$.

4.5.4 Zeit-Pegel-Richtungs-Abhängigkeit $w_{\tau R\Theta}$

Wie schon in Abschnitt 4.1 beschrieben wurde ist der Präzedenz-Effekt für bestimmte laterale Lautsprecheranordnungen nicht mehr vorhanden. Stitt versucht diesen Umstand in die Berechnung der Faktoren einfließen zu lassen (Gl. 22, 23 und 24).

$$w_{\tau R\Theta} = (1 - w_\Theta)w_{\tau R} + w_\Theta \tag{22}$$

Der Richtungsfaktor w_Θ ist abhängig von θ_{lead} und θ_{lag}, den Azimut-Winkel der Lautsprecher des Lead- bzw. Lag-Signals (Gl. 23).

$$w_\Theta(\theta_{lead}, \theta_{lag}) = e^{-0.5(\Delta\Theta/\sigma)^2} \tag{23}$$

Wobei sich die Winkeldifferenz wie folgt berechnet:

$$\Delta\Theta = \sin(\theta_{lead}) - \sin(\theta_{lag}) \tag{24}$$

Gl. 24 stellt sicher, dass der Lautsprecher des Lag-Signals gleiches Gewicht bekommt wie der des Lead-Signals, wenn dieser auf dem gleichen Cone-of-Confusion liegt. Somit hat der Präzedenz-Effekt für diesen Fall der Wiedergabe keinen Einfluss auf die Berechnung des $\mathbf{r_E}$-Vektors.

4.5.5 Berechnung für Mehrkanalsysteme

Sollen die vorangegangenen Überlegungen auf Mehrkanalsysteme übertragen werden ist es zunächst wichtig die an der Hörposition eintreffenden Signale nach ihrer Ankunftszeit zu ordnen. Ausgehend vom frühesten Signal kann die zeitliche Verzögerung des i-ten Signals angegeben werden.

$$\tau_l = T_l - T_1 \tag{25}$$

Das Pegelverhältnis R_l des l-ten Signals zu der Summe der vorangegangen Signale lässt sich wie folgt berechnen:

$$R_i = 10\log_{10}\left(\frac{g_l^2}{\sum_{m\in\Omega_i} g_m^2}\right) \tag{26}$$

Wobei $\sum_{m\in\Omega_l} g_m^2$ die Summe der Gewichte der Lautsprecher, deren Signale kleinere Verzögerungen als τ_l aufweisen, ist. Die Gleichungen 19, 20 und 23 können somit verallgemeinert werden.

$$\begin{aligned}
w_\tau^{(l,m)} &= 1 - \frac{2}{\sqrt{\pi}}\int_0^{\tau_i-\tau_j} e^{-(1.09t)^2}dt && (27)\\
w_R^{(l)} &= \frac{1}{2} + \mathsf{sgn}(10^{R_l/20} - 4.25)\frac{\gamma\left(\frac{1}{10}, \left(\frac{|10^{R_l/20}-4.25|}{3.75}\right)^{10}\right)}{\Gamma(1/10)} && (28)\\
w_\Theta^{(l,m)}(\theta_m,\theta_l) &= e^{-0.5((\sin(\theta_m)-\sin(\theta_l))/\sigma)^2} && (29)
\end{aligned}$$

Unter Verwendung von Gl. 21 und 22 ergibt sich

$$w_{\tau R\Theta}^{(l,m)} = w_\tau^{(l,m)} + w_R^{(l)} + w_\Theta^{(l,m)} - (w_\tau^{(l,m)}w_R^{(l)} + w_\tau^{(l,m)}w_\Theta^{(l,m)} + w_R^{(l)}w_\Theta^{(l,m)}) + w_\tau^{(l,m)}w_R^{(l)}w_\Theta^{(l,m)}. \tag{30}$$

Letztendlich ergibt sich der Faktor $w_{\tau R\Theta}^{(l)}$ durch eine gewichtete Mittelung der kombinierten Faktoren $w_{\tau R\Theta}^{(l,m)}$.

$$w_{\tau R\Theta}^{(l)} = \frac{\sum_{m\in\Omega_l}\zeta_m w_{\tau R\Theta}^{(l,m)}}{\sum_{m\in\Omega_l}\zeta_m} \quad \text{mit} \quad \zeta_m = 10^{\bar{R}_m/20} \tag{31}$$

$\bar{R}_m$ beschreibt dabei das gewichtete Pegelverhältnis des Signals mit Verzögerung τ_m zu den vorangegangenen Signalen (Gl. 32).

$$\bar{R}_m = 10\log_{10}\left(\frac{(w_{\tau R\Theta}^{(m)}g_m)^2}{\sum_{m\in\Omega_l}(w_{\tau R\Theta}^{(k)}g_k)^2}\right) \tag{32}$$

4.6 Überblick und Laufzeitanalyse

Um eine Einschätzung für den Rechenaufwand der einzelnen Algorithmen zu bekommen wurde eine kurze Laufzeitanalyse durchgeführt. Tabelle 2 zeigt die fünf verschiedenen Algorithmen zur Berechnung des $\mathbf{r_E}$-Vektors. $\mathbf{r_E}$-Alg. 1 basiert auf der Berechnung des gewöhnlichen $\mathbf{r_E}$-Vektors nach Gl. 4. Es wird die Ausbreitungsdämpfung nach Gl. 8 ($\mathbf{r_E}$-Alg. 2) und ein zeitabhängiges Gewicht nach Gl. 9 ($\mathbf{r_E}$-Alg. 3) integriert. $\mathbf{r_E}$-Alg. 4 bezieht den Präzedenzeffekt in die Berechnungen ein (Abschnitt 4.4). Der von P. Stitt entworfene Algorithmus (Abschnitt 4.5) wird mit $\mathbf{r_E}$-Alg. 5 abgekürzt. Vergleicht man die relativen Laufzeiten der $\mathbf{r_E}$-Alg. 2 und 3 mit der des $\mathbf{r_E}$-Alg. 1 ist zu erkennen, dass diese ca. 10% mehr Berechnungszeit benötigen. Die komplexeren $\mathbf{r_E}$-Alg. 4 und 5 benötigen hingegen das 4-fache der Laufzeit.

$\mathbf{r_E}$-Alg.	1	2	3	4	5
Ausbreitungsdämpfung	-	$\surd$	$\surd$	$\surd$	$\surd$
zeitliche Gewichtung	-	-	$\surd$	$\surd$	$\surd$
Präzedenzeffekt	-	-	-	$\surd$	$\surd$
Cone-of-Confusion	-	-	-	-	$\surd$
Laufzeit in in ms	0.20	0.23	0.22	0.79	0.81
rel. Laufzeit	1	1.11	1.08	3.85	3.97

Tabelle 2 – $\mathbf{r_E}$-Algorithmen 1 bis 5 sowie absolute und relative Laufzeiten zur Berechnung eines Auswertungspunktes bei einer Simulation mit den 12 Lautsprechern der Horizontalebene des IEM CUBE.

In Abschnitt 5 soll anhand zweier Hörversuche ermittelt werden, ob der rechnerische Mehraufwand der $\mathbf{r_E}$-Alg. 4 und 5 durch eine genauere Vorhersage der Versuchsdaten zu rechtfertigen ist.

4.7 Parameter α

Wie in den Abschnitten 4.4 und 4.5 beschrieben, wurde für das Überblenden zwischen dem normalen $\mathbf{r_E}$-Vektor und dem jeweiligen erweiterten $\mathbf{r_E}$-Vektor ein Parameter α eingeführt. Untersuchungen in [39] ergaben, dass α stark abhängig von der spektralen Verteilung und der Einhüllenden des ambisonisch wiedergegebenen Signals ist.

Abbildung 7 (a) bis (c) zeigen beispielhaft $\mathbf{r_E}$-Vektoren und den Betrag der horizontalen Winkelabweichung $|\Delta\varphi|$ im Auditorium des IEM CUBE für ein ansteigendes α. Es wurde eine virtuelle Schallquelle an der Lautsprecherposition 1 ($\varphi = \theta = 0°$) simuliert. Die Simulation erfolgte mit dem $\mathbf{r_E}$-Alg. 5, 3. Ordnung Ambisonics mit max-$\mathbf{r_E}$-Gewichtung und ohne Delay-Kompensation. $|\Delta\varphi|$ berechnet sich wie folgt:

$$|\Delta\varphi| = |\varphi_{virt.} - \varphi_{\mathbf{r_E}}|. \tag{33}$$

Wobei $\varphi_{virt.}$ der Panning-Winkel der virtuellen Schallquelle und $\varphi_{\mathbf{r_E}}$ den Winkel des $\mathbf{r_E}$-Vektors im jeweiligen Punkt angibt. Rot gefärbte Flächenteile weisen ein $|\Delta\varphi| \geq 23°$

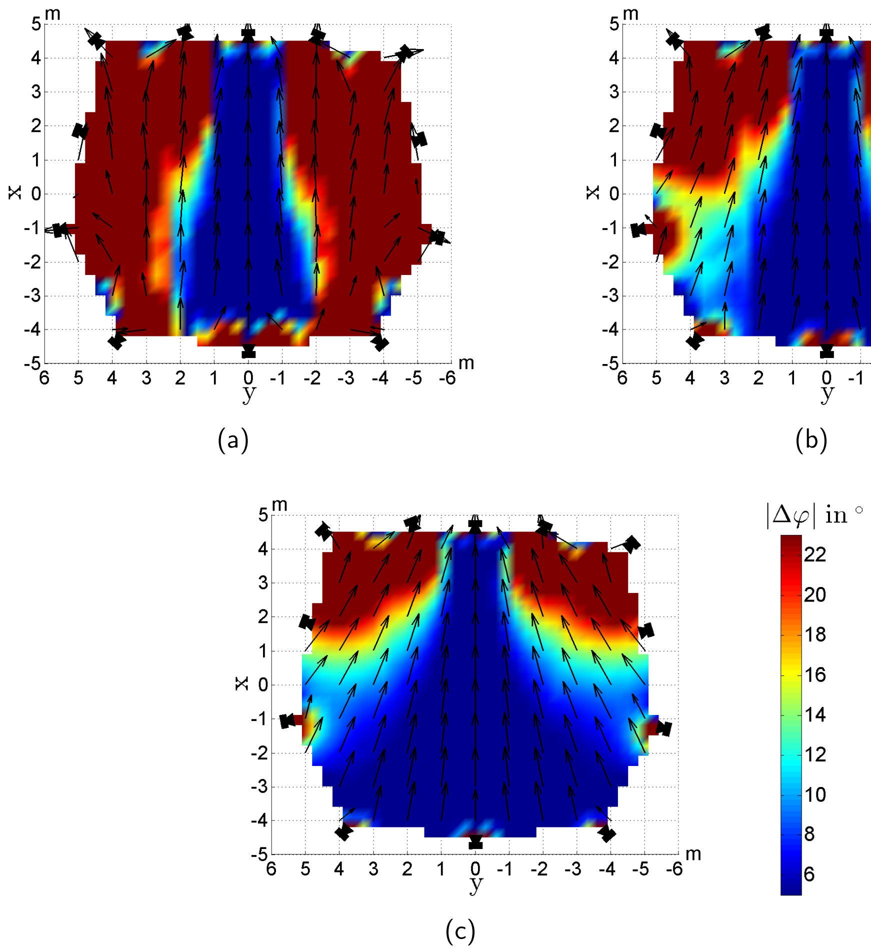


Abbildung 7 – Einfluss des Parameters α auf die Prädiktion der $\mathbf{r_E}$-Vektoren und des absoluten Winkelfehlers $|\Delta\varphi|$. Simulation mit $\mathbf{r_E}$-Alg. 5, max-$\mathbf{r_E}$-Gewichtung und ohne Delay-Kompensation. (a) $\alpha = 0.1$, (b) $\alpha = 0.5$ (c) $\alpha = 0.9$.

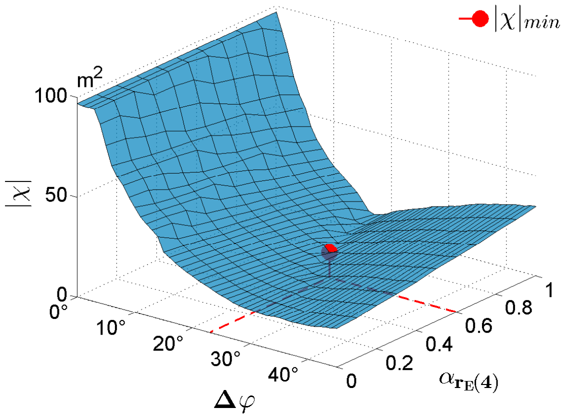
auf. In den blau gefärbten Bereichen ist $|\Delta\varphi| \leq 5°$. Für plausible Lokalisation sollte $|\Delta\varphi| < 23°$ sein, da der Winkel zwischen den Lautsprechern 1 und 2 bzw. 1 und 12 des IEM CUBE ca. $23°$ beträgt (vgl. [16]). Die Lokalisationsgenauigkeit steigt mit kleiner werdendem $|\Delta\varphi|$. Die Richtung der $\mathbf{r_E}$-Vektoren im jeweiligen Punkt zeigt die Richtung der „wahrgenommen" Phantomschallquelle an. $|\mathbf{r_E}|$ ist ein indirektes Maß für die Breite der Phantomschallquelle. Abb. 7 (a) zeigt einen Hörbereich der, fast nur auf die Hauptabstrahlrichtung des Lautsprechers 1 beschränkt ist. Es wurde ein $\alpha = 0.1$ für stark transiente Signale simuliert. In den lateralen Bereichen der bespielbaren Fläche ist die Lokalisation laut Simulation nicht mehr plausibel. Es ergeben sich breitere Phantomschallquellen an falschen Positionen. Abb. 7 (b) zeigt das Simulationsergebnis für das von Stitt empfohlene $\alpha = 0.5$. Es ist ersichtlich, dass sich für die hinteren seitlichen Bereiche die Lokalisation stark verbessert. Die schlechte Lokalisation in den vorderen seitlichen Bereichen bleibt jedoch erhalten. Dieses Verhalten ist jedoch nachvollziehbar, da bei der ambisonischen Wiedergabe mit 3. Ordnung trotz max-$\mathbf{r_E}$-Gewichtung Nebenkeulen noch nicht so stark bedämpft werden. Das Ergebnis der Simulation für ein eingeschwungenes bzw. stationäres Schallfeld ($\alpha = 0.9$) ist in Abb. 7 (c) zu sehen. Die virtuelle Schallquelle ist jetzt nahezu im gesamten hinteren Bereich sehr gut lokalisierbar. Die Bereiche unplausibler Lokalisation links und rechts der virtuellen Schallquelle bleiben aufgrund der Wiedergabeordnung erhalten, werden jedoch wesentlich kleiner. $|\Delta\varphi|$ der $\mathbf{r_E}$-Vektoren in diesen Bereichen fällt entsprechend geringer aus.

Um einen optimalen Wert für α zu finden wurde eine Optimierung anhand der Daten des von Frank und Zotter durchgeführten Hörversuchs zur Untersuchung der Größe des Sweet-Spots vorgenommen [16]. Hierbei wurde die Norm des Flächenfehlers $|\chi(\theta, \alpha)|$ entsprechend der Gl. 34 minimiert.

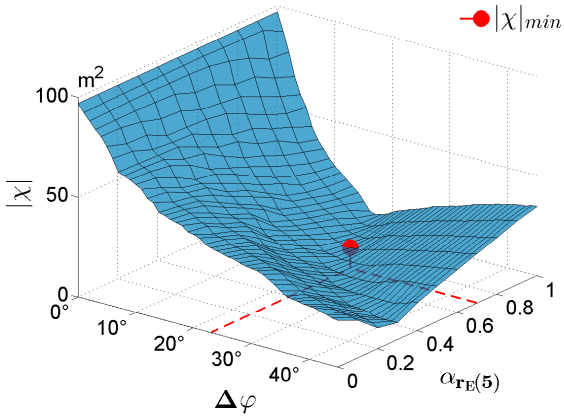
$$\min(|\chi(\theta, \alpha)|) = \min(|A_{sweet} - A_{sweet,sim}(\theta, \alpha)|) \quad \Rightarrow \quad \alpha_{opt} \tag{34}$$

Die Hörfläche A_{sweet} wurde in dem Hörversuch von [16] für die Wiedergabe einer frontalen Schallquelle über den unteren Lautsprecherring des IEM CUBE ermittelt (vgl. Lautsprecherposition 1 in Abb. 2). Die ambisonische Wiedergabe erfolgte mit Hilfe des AllRAD-Dekoders [11]. Es wurde mit 1., 2., 3. und 5. ambisonischer Ordnung wiedergegeben. Die Wiedergabe erfolgte mit max-$\mathbf{r_E}$-Gewichtung des Dekoders, da diese gute Ergebnisse für Lokalisation und Klangfärbung an dezentralen Hörpositionen erzielt [10, 12]. Bei der Wiedergabe wurde keine Delay-Kompensation vorgenommen. Die Versuchspersonen bewegten sich entlang der Radien zwischen Mittelpunkt und jeweiligem Lautsprecher (1 bis 12) der Anordnung. Sie sollten evaluieren, ab welchem Radius ein wahrgenommenes Schallereignis aus der Richtung von Lautsprecher 1 nicht mehr plausibel war, d.h. ab welcher Entfernung zum Mittelpunkt die Schallquelle nicht mehr zwischen dem frontalen Lautsprechertriplet 12-1-2 zu lokalisieren war ($|\Delta\varphi| > 23°$). Als Stimulus wurde ein Ausschnitt der englischen EBU Sprachreferenzaufnahmen verwendet [8]. A_{sweet} ist die Fläche der konvexen Hüllkurve, welche durch die 12 Radien erzeugt wird. Die Simulationen mit $\mathbf{r_E}$-Alg. 4 und 5 für die Bestimmung der Hörfläche $A_{sweet,sim}(\Delta\varphi, \alpha)$ wurden in Analogie zum Hörversuch [16] durchgeführt. Die Lautsprecher des IEM CUBE wurden mit frequenzunabhängiger omnidirektionaler Richtcharakteristik simuliert. Die Schallausbreitung wurde mit Freifeldbedingungen simuliert. Aus den berechneten Radien für einen maximal zulässigen Winkelfehler der Lokalisation $\Delta\varphi$ und dem Parameter α wurde die Hörfläche $A_{sweet,sim}(\Delta\varphi, \alpha)$ analog zum Hörversuch berechnet.

Abb. 8 zeigt die Fehlerflächen von $\mathbf{r_E}$-Alg. 4 und $\mathbf{r_E}$-Alg. 5. Sie ähneln einander stark und weisen beide einen ausgezeichneten Bereich für ein $|\chi|_{min}$ auf. Das Minimum der Fehlerfläche für den $\mathbf{r_E}$-Alg. 5 scheint jedoch ausgeprägter zu sein. Für $\alpha = 1$ ist der Verlauf der Fehlerflächen von $\mathbf{r_E}$-Alg. 4 und 5 identisch da beide Algorithmen nur mehr den einfachen $\mathbf{r_E}$-Vektor (nach $\mathbf{r_E}$-Alg. 2) berechnen.



(a)



(b)

Abbildung 8 – Norm des Flächenfehlers $|\chi|$ in Abhängigkeit des zulässigen Winkelfehlers θ und des Parameters α. $|\chi|_{min}$ (rot) für ein maximales $\Delta\varphi = 23°$ und $\alpha_{\mathbf{r_E}(4)} = 0.6$ (a) bzw. $\alpha_{\mathbf{r_E}(4)} = 0.7$ (b).

Der Winkelfehler $\Delta\varphi$ wurde wieder entsprechend der Plausibilitätsdefinition aus [16] gewählt. Abb. 9 zeigt den Verlauf des Flächenfehlers $|\chi|$ in Abhängigkeit von α bei eben diesem Winkelfehler. Beim Vergleich der beiden Verläufe von $|\chi(\alpha)|$ fällt auf, dass $\mathbf{r_E}$-Alg. 5 für stark transiente Stimuli ($\alpha = 0$) einen Flächenfehler erzeugt, der ca. das doppelte des von $\mathbf{r_E}$-Alg. 4 erzeugten Flächenfehlers beträgt. Das $|\chi|_{min}$ beider Algorithmen ist hingegen in einem ähnlichen Wertebereich angesiedelt (siehe auch Tab. 3). Da beim Plot von $|\chi|$ eine geringere Auflösung von α gewählt wurde als bei der Optimierung, sind in Tab. 3 die Ergebnisse beider Verfahren dargestellt. Für die weiteren Betrachtungen werden $\alpha_{\mathbf{r_E}(4)}$ bzw. $\alpha_{\mathbf{r_E}(5)}$ verwendet.

	α	$\|\chi\|_{\mathbf{min}}$	$\alpha_{\mathbf{opt}}$	$\|\chi\|_{\mathbf{min,opt}}$
$\mathbf{r_E}$**-Alg. 4**	0.6	$12.6m^2$	0.62	$12.47m^2$
$\mathbf{r_E}$**-Alg. 5**	0.7	$10.5m^2$	0.76	$10.13m^2$

Tabelle 3 – Parameter α und residualer Flächenfehler $|\chi|_{min}$ der Fehlerflächenberechnung sowie optimaler Parameter α_{opt} und zugehöriger residualer Flächenfehler $|\chi|_{min,opt}$ für beide $\mathbf{r_E}$-Algorithmen.

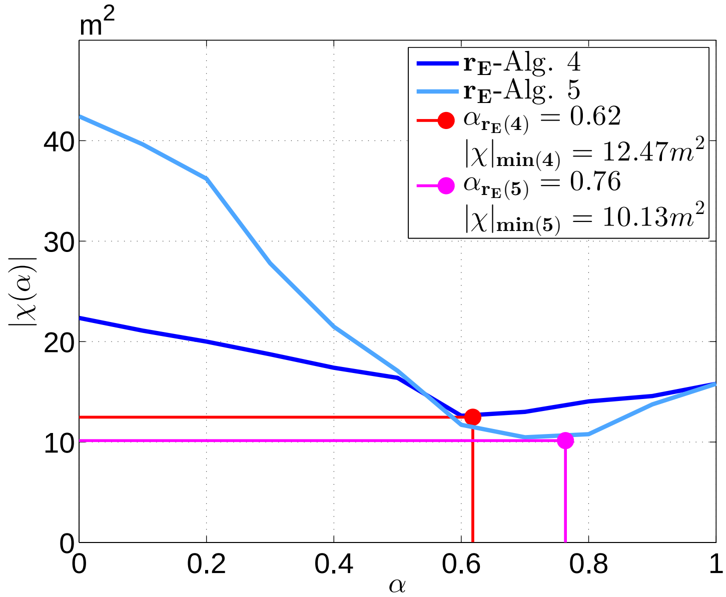


Abbildung 9 – Norm des Flächenfehlers $|\chi|$ in Abhängigkeit des Parameters α bei einem Winkelfehler von $\Delta\varphi = 23°$.

5 Modellierung existierender Hörversuche

5.1 Hörversuch 1

Die Modellierung der Ergebnisse des Hörversuchs von [16] wurde mit den selben Einstellungen, welche zur Bestimmung des Parameters α verwendet wurden, durchgeführt.

Abbildung 10 (a) zeigt die Größe A_{sweet} der mit den $\mathbf{r_E}$-Alg. 2 bis 5 berechneten Hörflächen im Vergleich mit den erhobenen Daten von [16] für die Wiedergabe mit Ambisonics 1., 2., 3. und 5. Ordnung. Diese Daten zeigen die im Hörversuch erhobenen Medianwerte für A_{sweet} sowie deren zugehörige 95%-Konfidenzintervalle. Es ist zu erkennen, dass der Anstieg der Kurven für $\mathbf{r_E}$-Alg. 3 bis 5 größer ist als der Anstieg der Hörversuchsdaten. So wird A_{sweet} bei diesen drei Algorithmen bei ambisonischer Wiedergabe mit 1. Ordnung signifikant unter- und bei Wiedergabe mit 5. Ordnung signifikant überschätzt. Lediglich der $\mathbf{r_E}$-Alg. 2 prädiziert für die Wiedergabe mit 1. Ordnung Ambisonics A_{sweet} exakt. Bei Wiedergabe mit 2. bis 5. Ordnung von Ambisonics überschätzt dieser Algorithmus A_{sweet} allerdings signifikant. Eine weitere Möglichkeit um die verschiedenen $\mathbf{r_E}$-Algorithmen mit dem Hörversuch von [16] zu vergleichen ist es den mittleren relativen Sweetspot-Radius $\bar{r}_{sweet}$ zu betrachten (siehe Abb. 10 (b)). Es ist ersichtlich, dass der Anstieg der Kurven aller $\mathbf{r_E}$-Algorithmen mit steigender ambisonischer Ordnung abnimmt. Die Kurve aus den Daten des Hörversuchs zeigt ein gegensätzliches Verhalten. So wird, ähnlich zu der Abbildung 10 (a), $\bar{r}_{sweet}$ für die 1. Ordnung signifikant unterschätzt. Bei Wiedergabe mit 2. Ordnung prädiziert nur $\mathbf{r_E}$-Alg. 3 $\bar{r}_{sweet}$ richtig. Überschätzten bei ambisonischer Wiedergabe mit 3. Ordnung alle Algorithmen $\bar{r}_{sweet}$ signifikant, können sie hingegen $\bar{r}_{sweet}$ bei 5. Ordnung ziemlich genau vorhersagen. Für eine weitere Analyse scheint es sinnvoll sich die Gestalt der berechneten Hörflächen genauer anzuschauen (siehe Abb. 11).

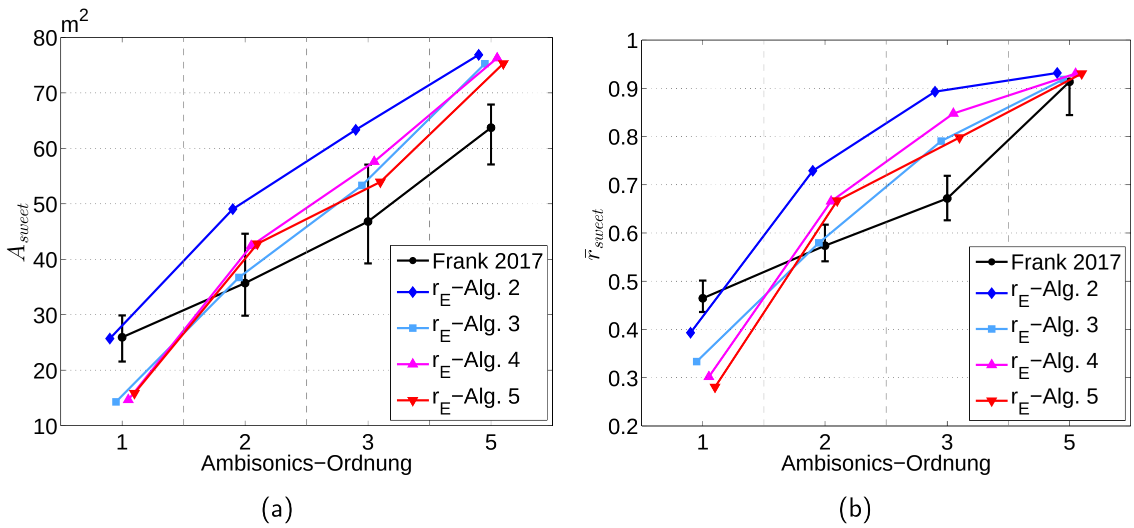


Abbildung 10 – Größe der Hörfläche A_{sweet} (a) und mittlerer relativer Hörradius der Hörfläche $\bar{r}_{sweet}$ (b) im Vergleich.

Abbildung 11 (a) bis (d) zeigt die Hörflächen für die ambisonische Wiedergabe mit 1., 2., 3., und 5. Ordnung, welche sich durch die Simulation mit den $\mathbf{r_E}$-Alg. 2 bis 5 ergeben. Zum Vergleich sind in Abbildung 11 (e) die Hörflächen aus dem Hörversuch von [16] abgebildet. Der Plot von $\mathbf{r_E}$-Alg. 1 ist nicht dargestellt, da sich die Konturen der Hörflächen für die Simulation mit verschiedenen Ambisonics-Ordnungen nicht unterscheiden. Vergleicht man die Ergebnisse der vier Algorithmen miteinander stellt man mitunter starke Ähnlichkeiten fest. Vor allem bei der Simulation mit 5. Ordnung Ambisonics scheinen die Algorithmen quasi identisch abzuschneiden. Es ist zu erkennen, dass lediglich die Berücksichtigung der Ausbreitungsdämpfung von 6 dB je Abstandverdopplung ($\mathbf{r_E}$-Alg. 2) schon eine starke Auswirkung auf die Ausformung der Hörflächen besitzt. Da jedoch der Präzendenz-Effekt in diesem Berechnungsmodell keinen Einfluss findet, wird der hintere Bereich der Hörflächen stark vergrößert prädiziert (vgl. Abb. 11 (e)). Dies trifft im speziellen für die Wiedergabe mit 2. und 3. Ordnung Ambisonics zu. $\mathbf{r_E}$-Alg. 2 überschätzt auch die lateralen Bereiche von A_{sweet} für 3. Ordnung Ambisonics. Im Gegensatz dazu werden für die Wiedergabe mit 1. Ordnung die lateralen Bereiche unterschätzt.

Die relativ einfache Erweiterung durch das Hinzufügen einer zeitlichen Gewichtung, vgl. Abschitt 4.3, hat einen großen Einfluss auf die Qualität der prädizierten Hörflächen (Abb. 11 (b)). Das fällt vor allem bei den Hörflächen für Wiedergabe mit 2. und 3. Ordnung Ambisonics auf. Die Konturen nähern sich stark denen von Abb. 11 (e) an. A_{sweet} bei Wiedergabe mit 1. Ordnung Ambisonics wird hingegen stark unterschätzt, vor allem in den lateralen Bereichen.

Der $\mathbf{r_E}$-Alg. 4 betreibt einen erheblichen Mehraufwand um den Präzedenz-Effekt in die Berechnung zu integrieren. Es ist jedoch festzustellen, dass sich für diese Simulation keine qualitative Verbesserung der prädizierten Hörflächen einstellt. Die Hörfläche für die Wiedergabe mit 1. Ordnung Ambisonics wird im Vergleich zum Hörversuch eher seitlich unter- und hinten überschätzt. Die Hörflächen für die Wiedergabe mit 2. bzw. 3. Ordnung sind im hinteren Auditorium stark bzw. mäßig vergrößert.

Die Berechnung des $\mathbf{r_E}$-Vektors nach Stitt ($\mathbf{r_E}$-Alg. 5) liefert im Vergleich zu $\mathbf{r_E}$-Alg. 4 sehr ähnliche Ergebnisse. Die Hörfläche für die ambisonische Wiedergabe 1. Ordnung wird vor allem im lateralen Bereich unterschätzt. Bei Wiedergabe mit 2. und 3. Ordnung nähern sich die Konturen der Hörflächen denen der Abb. 11 (d) leicht an. Der hintere Bereich wird allerdings wieder überschätzt.

Generell ist zu beobachten, dass alle Algorithmen die Hörfläche im hinteren Bereich eher über- und im lateralen Bereich eher unterschätzen. Die Simulation wurde unter der Annahme von Freifeldbedingungen durchgeführt, der Hörversuch fand hingegen unter Einfluss von Reflexionen des Raumes statt. Somit wäre es möglich, dass die beim Hörversuch auftretenden Reflexionen die seitlichen Lautsprechersignale maskiert haben und somit die Lokalisation in den lateralen Bereichen stabiler empfunden wurde.

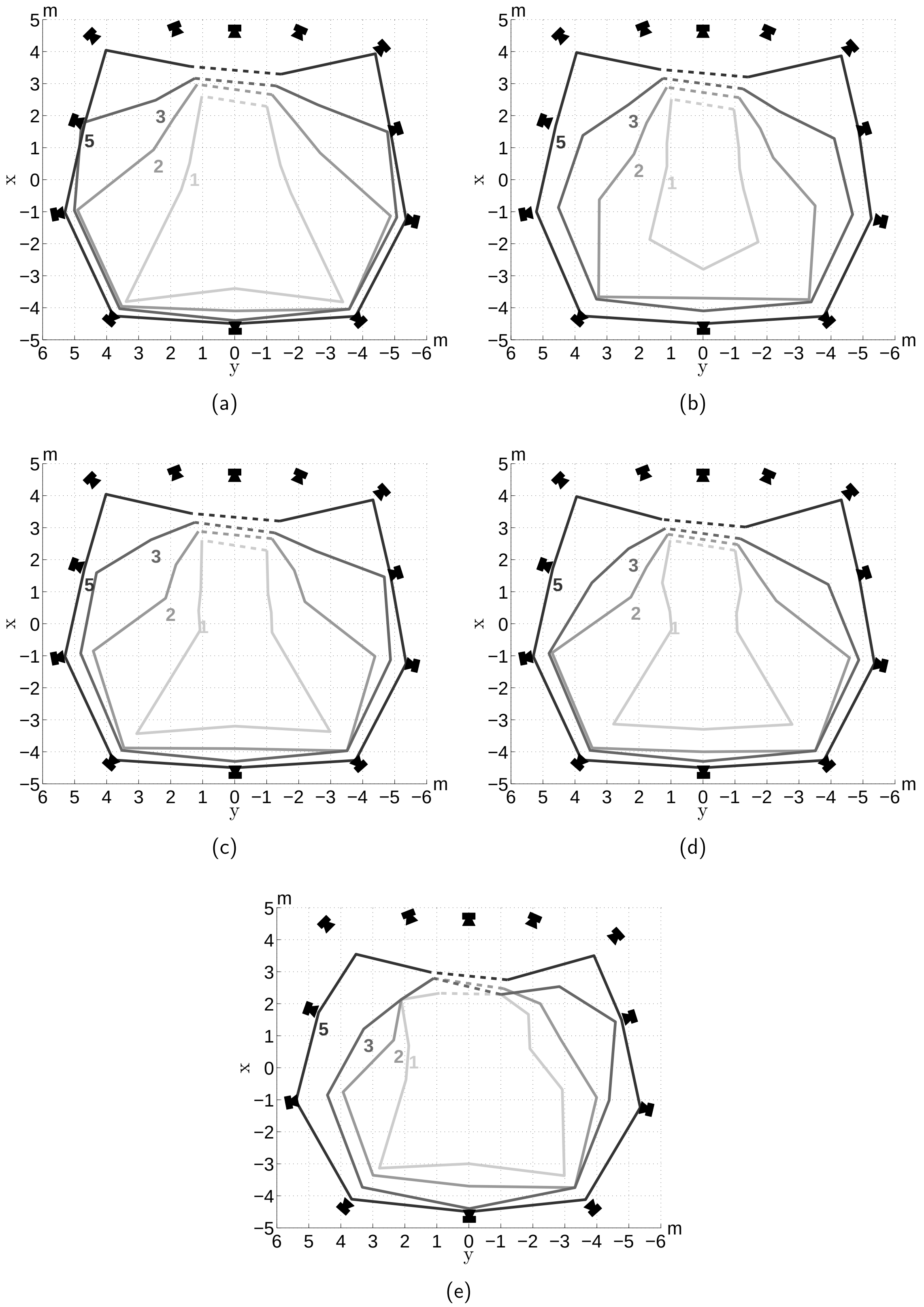


Abbildung 11 – Hörflächen für die ambisonische Wiedergabe mit 1., 2., 3., und 5. Ordnung. (a) $\mathbf{r_E}$-Alg. 2, (b) $\mathbf{r_E}$-Alg. 3, (c) $\mathbf{r_E}$-Alg. 4, (d) $\mathbf{r_E}$-Alg. 5, (e) Hörversuch von [16].

Abbildung 12 zeigt beispielhaft $\mathbf{r_E}$-Vektoren und den Betrag der horizontalen Winkelabweichung $|\Delta\varphi|$ im gesamten bespielbaren Bereich des IEM CUBE. Es wird die ambisonische Wiedergabe 3. Ordnung einer Schallquelle an Lautsprecherposition 1 ($\varphi = \theta = 0°$, vgl. Abb. 2) simuliert. Die resultierenden Hörbereiche der $\mathbf{r_E}$-Alg. 2 bis 5 unterscheiden sich nur marginal. Alle Simulationen prädizieren eine sehr gute Lokalisation im hinteren Hörbereich und auf der Lautsprechernormalen des Lautsprechers 1. Der Einfluss auf den hinteren Hörbereich der Lautsprecher 5 bis 9 (vgl. Abb. 2) ist am stärksten in der Simulation mit $\mathbf{r_E}$-Alg. 3 (Abb. 12 (b)) zu erkennen. Der geringste Einfluss wird durch $\mathbf{r_E}$-Alg. 2 prädiziert. Alle $\mathbf{r_E}$-Alg. prädizieren eine unplausible Lokalisation in den vorderen seitlichen Bereichen des Auditoriums. Die Nebenkeulen bei Wiedergabe mit 3. Ordnung Ambisonics sind verantwortlich für dieses Verhalten. Bei $\mathbf{r_E}$-Alg. 3 bis 5 (Abb. 12 (b), (c), (d)) sind diese Bereiche ähnlich stark ausgeprägt. $\mathbf{r_E}$-Alg. 2 prädiziert diese weniger ausgeprägt.

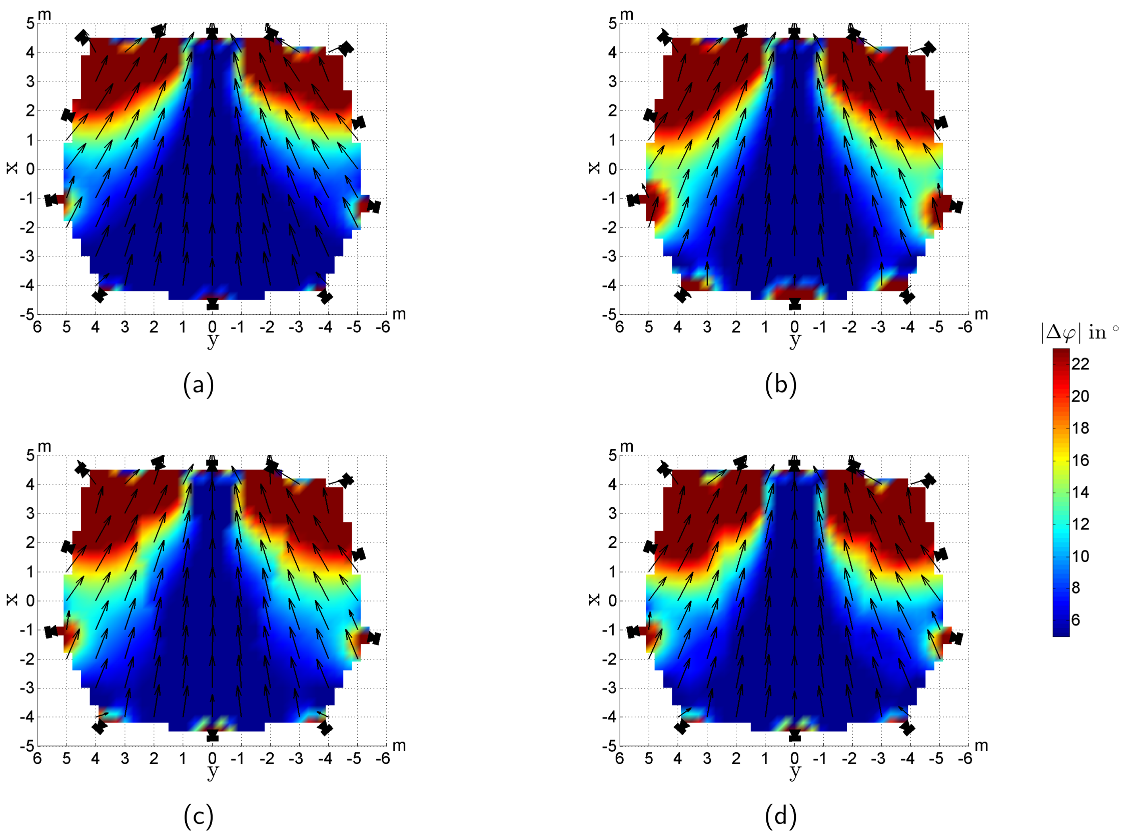


Abbildung 12 – $\mathbf{r_E}$-Vektor und $|\Delta\varphi|$ für die ambisonische Wiedergabe mit 3. Ordnung. (a) $\mathbf{r_E}$-Alg. 2, (b) $\mathbf{r_E}$-Alg. 3, (c) $\mathbf{r_E}$-Alg. 4, $\mathbf{r_E}$-Alg. 5. Die Simulation wurde mit max-$\mathbf{r_E}$-Gewichtung und ohne Delay-Kompensation durchgeführt.

5.2 Hörversuch 2

In dem 2008 von Frank durchgeführten Hörversuch [10] wurde die Lokalisation bei frontaler Wiedergabe mit verschiedenen Ambisonics-Dekodern untersucht. 15 Versuchspersonen wurden Bursts mit breitbandigem rosa Rauschen in der horizontalen Lautsprecherebene des IEM CUBE präsentiert. Die Rausch-Bursts wurden im horizontalen Winkelbereich von $\varphi_{virt.} = \pm 40°$ um die $\varphi_0 = 0°$-Richtung in $5°$-Schritten präsentiert (vgl. Abb. 13). Jeder Vesuchsperson wurden insgesamt 78 Stimuli zugespielt. Die Wiedergabe wurde mit basic-, max-$\mathbf{r_E}$- und in-phase-Dekodern 1., 3. und 5. Ordnung mit und ohne Delay-Kompensation durchgeführt. Die Auswahl der Wiedergaberichtungen erfolgte zufällig und für jede Versuchsperson personalisiert. Die Versuchspersonen sollten für zwei Hörpositionen, innerhalb ($[x, y] = [0, 0]$ m) und außerhalb ($[x, y] = [-0.23, 2.1]$ m) des Zentrums, mittels eines Zeigegerätes die Richtung der Wiedergabe der Stimuli angeben (vgl. Abb. 13). Mittels der so gewonnen Daten wurde der Winkelfehler $|\Delta\varphi|$ bestimmt.

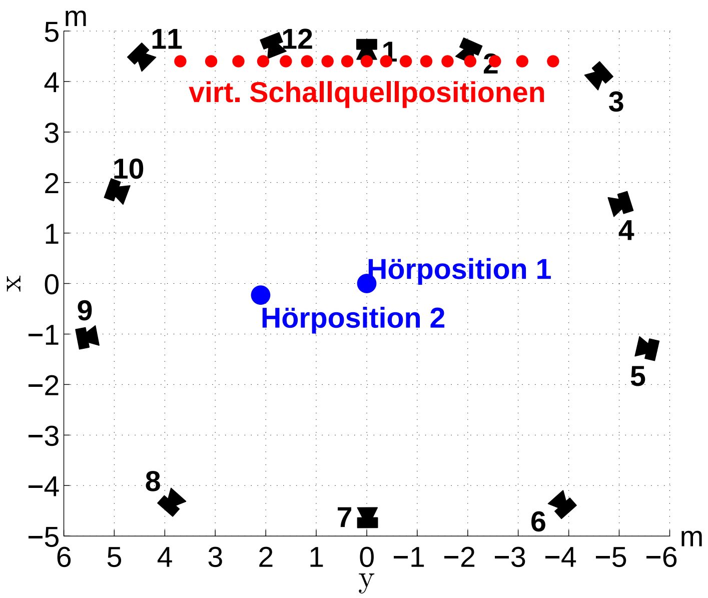


Abbildung 13 – Hörpositionen und Positionen der virtuellen Schallquellen in dem 2008 von Frank durchgeführten Hörversuch [10].

Um die Performance der verschiedenen Algorithmen zur Berechnung des $\mathbf{r_E}$-Vektors möglichst genau beurteilen zu können, wurde die Simulation dem Hörversuch nachempfunden. Wieder wurden die Lautsprecher als frequenzunabhängige Punktschallquellen modelliert. Bei der Simulation wurde nur der Direktschall der Lautsprecher berücksichtigt. Um einen vergleichbaren Datensatz zu generieren wurde je Hörposition ein Raster von $3 \times 5 = 15$ Punkten ausgewertet. Diese wurden entsprechend der Abb. 14 um die Hörposition angeordnet.

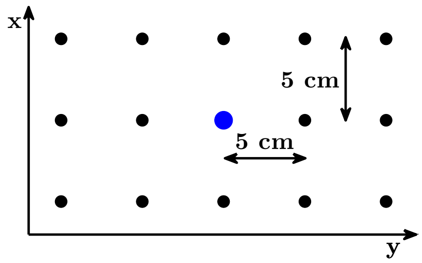
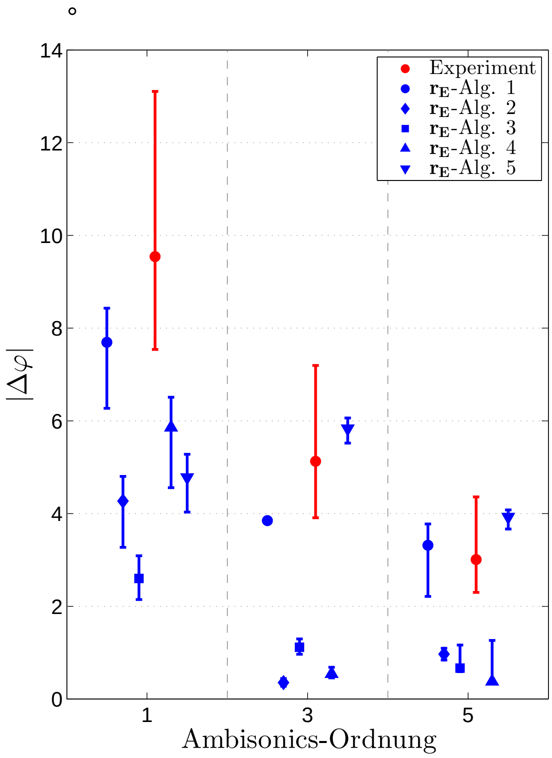
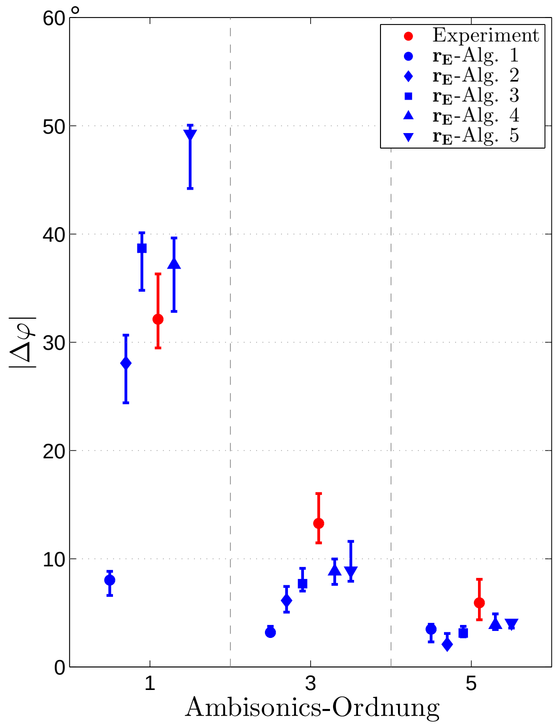


Abbildung 14 – Schematische Darstellung des Rasters der Auswertungspunkte (schwarz) an der Hörposition (blau).

Die Abbildungen 15, 18 und 19 zeigen die im Hörversuch ermittelten Medianwerte sowie die 95%-Konfidenzintervalle des Betrags des Winkelfehlers $|\Delta\varphi|$ in rot. Die Medianwerte für $|\Delta\varphi|$ und die zugehörigen 95%-Konfidenzintervalle durch die verschiedenen $\mathbf{r_E}$-Algorithmen (blau) ergeben sich durch die Auswertung der Winkelfehler für die verschiedenen virtuellen Schallquellpositionen und alle Auswertungspunkte je Hörposition. Es wurden nur die Winkelfehler dieser Schallquellpositionen ausgewertet für die Hörversuchsdaten vorlagen. Ganz allgemein gilt, je kleiner ein Konfidenzintervall ist, desto zuverlässiger ist der darin liegende Medianwert des Winkelfehlers $|\Delta\varphi|$. Überlappen sich die Konfidenzintervalle der Winkelfehler des Hörversuchs und der jeweiligen Simulation nicht, unterscheiden sich die Ergebnisse signifikant. Dieser Zusammenhang gilt natürlich auch für die Simulationsergebnisse untereinander.



(a)



(b)

Abbildung 15 – Einfluss der Dekoderordnung auf den absoluten Winkelfehler $|\Delta\varphi|$ für zentrale (a) und dezentrale Hörposition (b). Simulation mit max-$\mathbf{r_E}$-Gewichtung und ohne Delay-Kompensation.

Abbildung 15 zeigt den Einfluss der Dekoderordnung auf den Winkelfehler $|\Delta\varphi|$ für die Hörposition inner- und außerhalb des Sweetspots. Zur Auswertung wurden die Versuchsdaten der Wiedergabe mit max-$\mathbf{r_E}$-gewichteten Dekoder ohne Delay-Kompensation verwendet. Für die zentrale Hörposition 1 ist zunächst erkenntlich, dass die verschiedenen $\mathbf{r_E}$-Alg. ähnlich geringe $|\Delta\varphi|$ wie die Versuchsdaten produzieren. Mit zunehmender Ordnung der Wiedergabe nimmt $|\Delta\varphi|$ der Versuchdaten ab und die zugehörigen Konfidenzintervalle verkleinern sich. Das heißt, dass die Genauigkeit der Lokalisation der Versuchspersonen zunimmt, da mit steigender Ordnung der Wiedergabe die wahrgenommene Quellenbreite sinkt. Tendenziell ist auch bei den $\mathbf{r_E}$-Alg. ein Abnehmen von $|\Delta\varphi|$ und ein Verkleinerung der Konfidenzintervalle für höhere Ordnungen der Wiedergabe zu erkennen. Im Allgemeinen unterschätzen die Algorithmen $|\Delta\varphi|$ eher (ausgenommen $\mathbf{r_E}$-Alg. 1 und 5). Bei Wiedergabe mit 1. Ordnung Ambisonics unterschätzen die Alg. 2 bis 5 $|\Delta\varphi|$ signifikant. Der geringste Winkelfehler wird durch $\mathbf{r_E}$-Alg. 3 prädiziert. Nur $\mathbf{r_E}$-Alg. 1 liegt innerhalb des Konfidenzintervalls der Versuchsdaten. Wird mit 3. Ordnung Ambisonics simuliert, vermindert sich $|\Delta\varphi|$ für die Alg. 1 bis 4 stark. Die resultierenden Konfidenzintervalle aller Simulationen verkleinern sich erheblich. $|\Delta\varphi|$ für $\mathbf{r_E}$-Alg. 5 steigt an und prädiziert somit die Winkelabweichung der Messdaten relativ akkurat. $|\Delta\varphi|$ für $\mathbf{r_E}$-Alg. 1 liegt am unteren Ende des Konfidenzintervalls der Messdaten. Die $\mathbf{r_E}$-Alg. 2 bis 4 unterschätzen $|\Delta\varphi|$ signifikant. Für die Wiedergabe mit 5. Ordnung zeigt sich ein ähnliches Verhalten. $\mathbf{r_E}$-Alg. 1 und 5 überschätzen die Versuchsergebnisse leicht. $\mathbf{r_E}$-Alg. 2 bis 4 unterschätzen $|\Delta\varphi|$ signifikant. Es ist festzustellen, dass der physikalisch unkorrekte $\mathbf{r_E}$-Alg. 1, der die Ausbreitungsdämpfung nicht berücksichtigt, im Sweetspot als Prädiktor verwendet werden kann. Dieses Ergebnis bestätigt die von Frank in [13] getroffene Aussage, dass sich der $\mathbf{r_E}$-Vektor sehr gut zur Modellierung der menschlichen Hörwahrnehmung im Sweetspot einer Lautsprecheranordnung eignet. Um die Performance der anderen $\mathbf{r_E}$-Alg. an der zentralen Hörposition zu verbessern könnte man versuchen die größere Unsicherheit für geringere Ordnung der Wiedergabe über den Betrag des $\mathbf{r_E}$-Vektors zu modellieren. Da $|\mathbf{r_E}|$ verwendet werden kann um die Quellenbreite zu berechnen, könnte diese genutzt werden um eine Lokalisationsunsicherheit und somit ein größeres $|\Delta\varphi|$ zu prädizieren.

Betrachtet man die Ergebnisse für die Hörposition 2 (Abb. 15 (b)) sind generell stärkere Unterschiede in der Performance der einzelnen $\mathbf{r_E}$-Algorithmen zu erkennen. Für die ambisonische Wiedergabe mit erster Ordnung ist festzustellen, dass der $\mathbf{r_E}$-Alg. 1 den Winkelfehler $|\Delta\varphi|$ signifikant unterschätzt. Lediglich mit der Berücksichtigung der Ausbreitungsdämpfung des Schalls ($\mathbf{r_E}$-Alg. 2) kann man jedoch θ schon relativ gut prädizieren. Durch Einbindung der zeitliche Gewichtung von -0.25 dB/ms ($\mathbf{r_E}$-Alg. 3) lässt sich die Prädiktion weiter verbessern. Die $\mathbf{r_E}$-Algorithmen 4 und 5, welche den Präzendenzeffekt relativ aufwändig in die Berechnung des $\mathbf{r_E}$-Vektors einbeziehen, können die Vorhersage jedoch kaum verbessern. Für die Wiedergabe mit 3. Ordnung Ambisonics ist zu beobachten, dass alle $\mathbf{r_E}$-Algorithmen die Versuchsergebnisse unterschätzen. Bei $\mathbf{r_E}$-Alg. 1 geschieht dies signifikant. Alle anderen Algorithmen produzieren einen Prädiktionsfehler $< 10°$. Die Berücksichtigung der Ausbreitungsdämpfung und die zeitliche Gewichtung erzielen wieder einen Performance-Gewinn. Der Mehraufwand der $\mathbf{r_E}$-Algorithmen 4 und 5 bringt nur eine minimale Verbesserung der Simulationsergebnisse. Wird mit Ambisonics 5. Ordnung wiedergegeben, werden die Prädiktionsfehler nocheinmal kleiner. Alle Simulationsergebnisse liegen knapp unter dem 95%-Konfidenzintervall des Versuchsergebnisses. Die Simulationsergebnisse untereinander weisen gleiche Tendenzen auf wie bei der Wiedergabe mit 3. Ordnung Ambisonics, wenn auch weniger stark ausgeprägt. Einzig $\mathbf{r_E}$-Alg. 1 prädiziert nicht mehr das kleinste $|\Delta\varphi|$.

Exemplarisch sind die Simulationen mit den $\mathbf{r_E}$-Alg. 2 und 5 für das gesamte Auditorium des IEM CUBE in der Abb. 16 ersichtlich. Es ist der Medianwert des absoluten Winkelfehlers $|\widetilde{\Delta\varphi}|$ farblich codiert. Der Medianwert wurde über die 17 Panning-Positionen mit $-40° \leq \varphi_{virt.} \leq +40°$ in $5°$-Schritten ermittelt. Es ist zu erkennen, dass sich eine höhere Ordnung der ambisonischen Wiedergabe erheblich auf die Gestalt und Größe des Hörbereichs auswirkt. Bei 1. Ordnung kann gute Lokalisation nur unmittelbar an der zentralen Hörpostion und dahinter erreicht werden ($|\widetilde{\Delta\varphi}| \lesssim 10°$). An der dezentralen Hörposition ist keine plausible Lokalisation möglich. Bei 3. Ordnung ist die Lokalisation an der dezentralen Hörposition bereits annehmbar ($|\widetilde{\Delta\varphi}| \approx 10°$ für $\mathbf{r_E}$-Alg. 2 und $|\widetilde{\Delta\varphi}| \approx 11°$ für $\mathbf{r_E}$-Alg. 2, vgl. Abb. 15 (b)). Der Hörbereich bei Wiedergabe mit 5. Ordnung erstreckt sich, mit Ausnahme des Bereichs unmittelbar vor den Lautsprechern 1, 2 und 12, über das gesamte Auditorium. Es ist beachtlich, dass nur die Berücksichtigung der Ausbreitungsdämpfung ($\mathbf{r_E}$-Alg. 2) ab 3. Ordnung Ambisonics ähnliche Ergebnisse wie $\mathbf{r_E}$-Alg. 5 liefert.

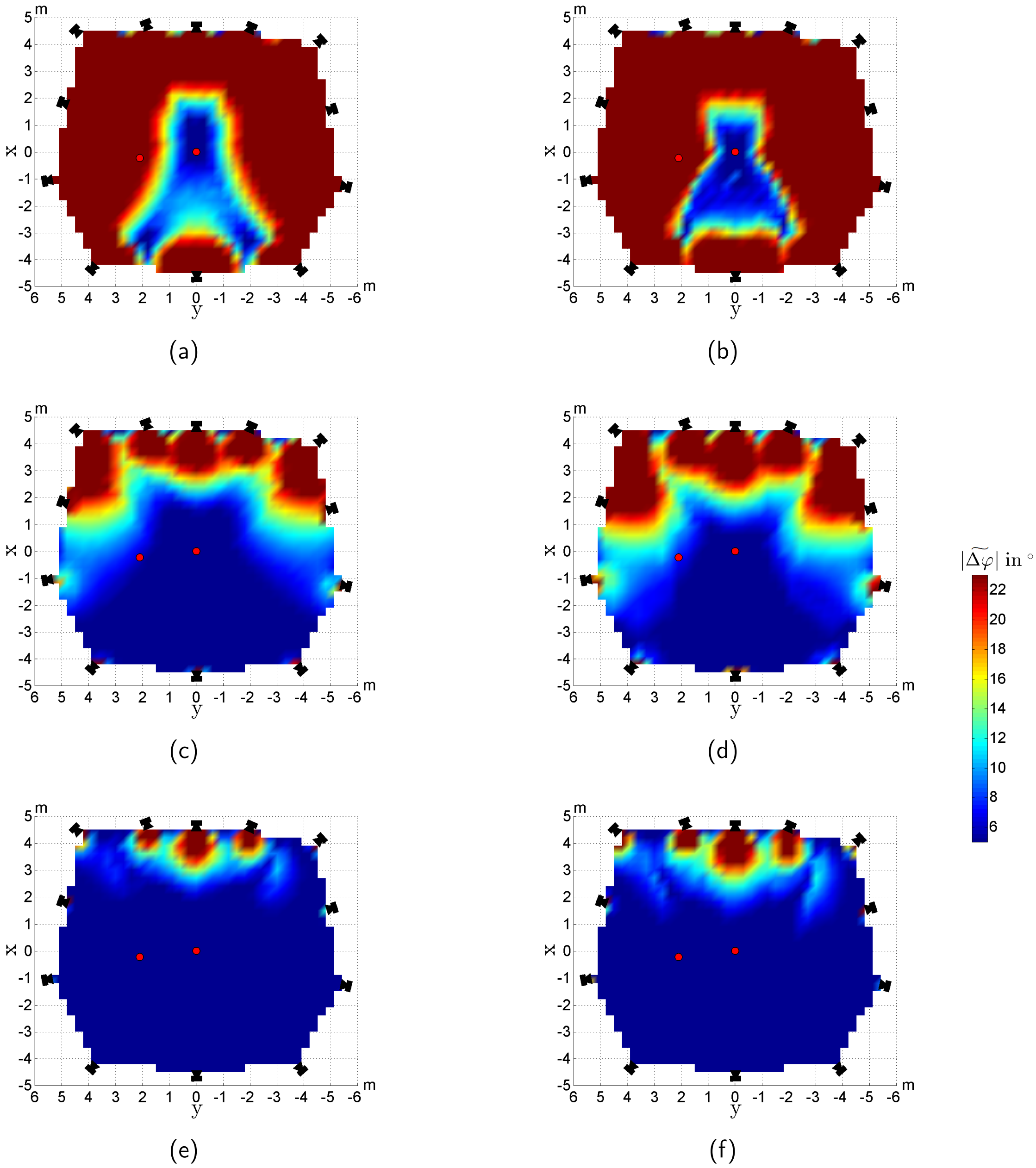
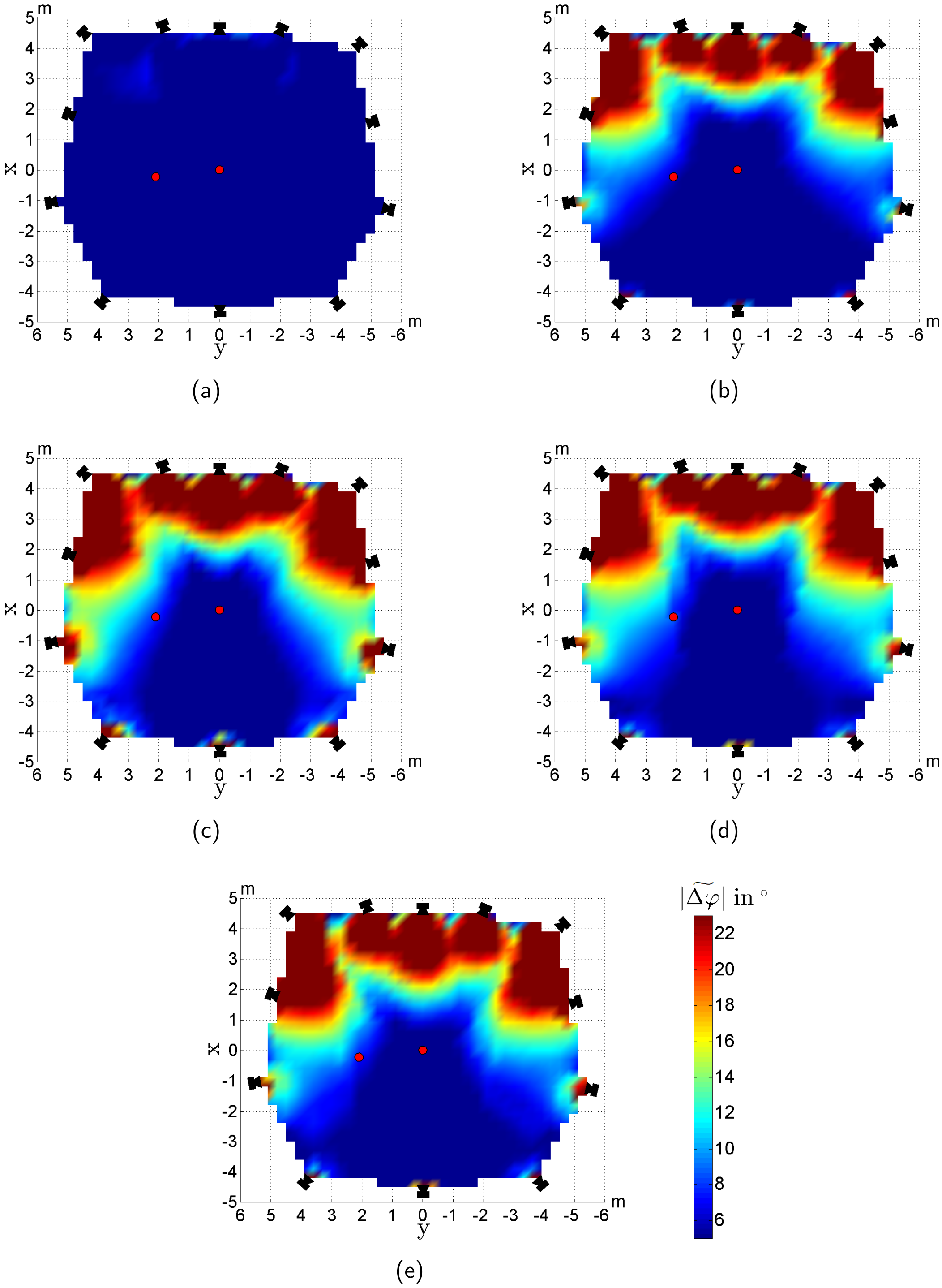


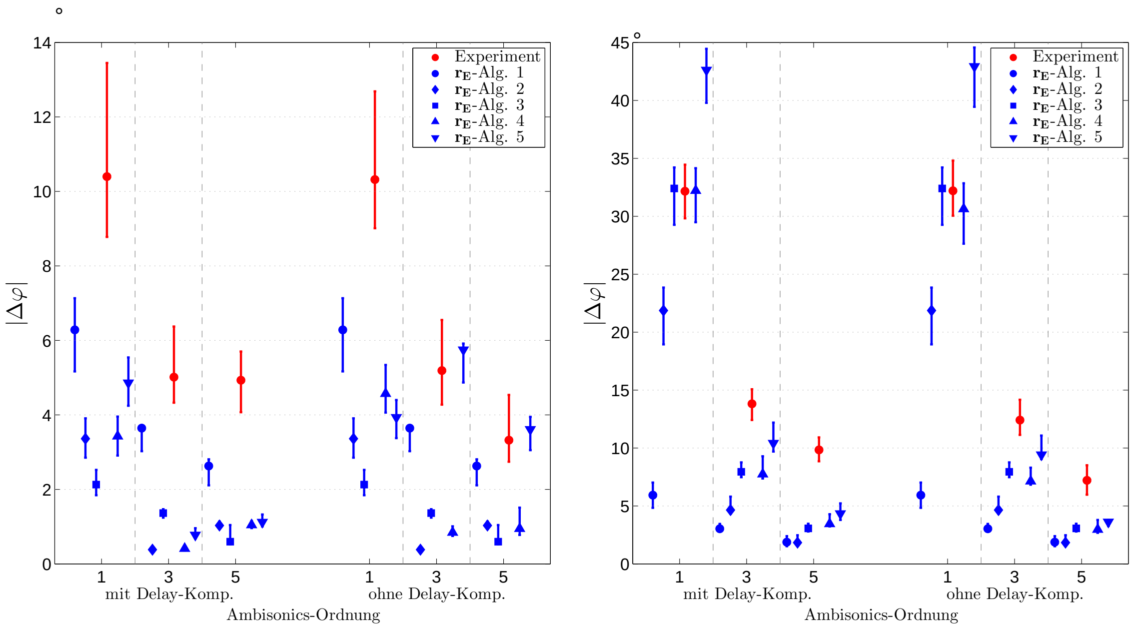
Abbildung 16 – Auditorium des IEM CUBE für die ambisonische Wiedergabe mit 1. (a,b), 3. (c,d) und 5. (e,f) Ordnung mit $\mathbf{r_E}$-Alg. 2 (a,c,e) und 5 (b,d,f). Median des absoluten Winkelfehlers $|\widetilde{\Delta\varphi}|$ über die 17 Panning-Positionen ($-40° \leq \varphi_{virt.} \leq +40°$ in $5°$-Schritten). Die Simulation wurde mit max-$\mathbf{r_E}$-Gewichtung und ohne Delay-Kompensation durchgeführt. Zentrale und dezentrale Hörposition sind durch die roten Punkte gekennzeichnet.



(a) (b) (c) (d) (e)

Abbildung 17 – Auditorium des IEM CUBE für die ambisonische Wiedergabe mit 3. Ordnung. (a) $\mathbf{r_E}$-Alg. 1, (b) $\mathbf{r_E}$-Alg. 2, (c) $\mathbf{r_E}$-Alg. 3, (d) $\mathbf{r_E}$-Alg. 4, (e) $\mathbf{r_E}$-Alg. 5. Die Simulation wurde mit max-$\mathbf{r_E}$-Gewichtung und ohne Delay-Kompensation durchgeführt. Zentrale und dezentrale Hörposition sind durch die roten Punkte gekennzeichnet.

Abbildung 17 zeigt wieder den Medianwert des absoluten Winkelfehlers $|\Delta\varphi|$ im IEM CUBE ermittelt über die 17 Panning-Positionen. Die Simulation wurde exemplarisch für alle $\mathbf{r_E}$-Alg. durchgeführt. Dabei wurde ein Dekoder 3. Ordnung mit max-$\mathbf{r_E}$-Gewichtung und ohne Delay-Kompensation verwendet. Es ist zu erkennen, dass mit dem $\mathbf{r_E}$-Alg. 1 keine sinnvolle Prädiktion der Lokalisation an dezentralen Hörpositionen durchgeführt werden kann. Wie bereits erwähnt bewirkt die Berücksichtigung der Ausbreitungsdämpfung eine erhebliche Verbesserung der Prädiktionsqualität. $\mathbf{r_E}$-Alg. 3, 4 und 5 erzielen eine sehr ähnliche Gestalt des Hörbereichs, wobei Alg. 3 den kleinsten Hörbereich vorhersagt.



(a) (b)

Abbildung 18 – Einfluss der Delay-Kompensation auf den absoluten Winkelfehler $|\Delta\varphi|$ für zentrale (a) und dezentrale Hörposition (b). Die Daten der Wiedergabe/Simulation mit basic- und max-$\mathbf{r_E}$-Gewichtung wurden verwendet.

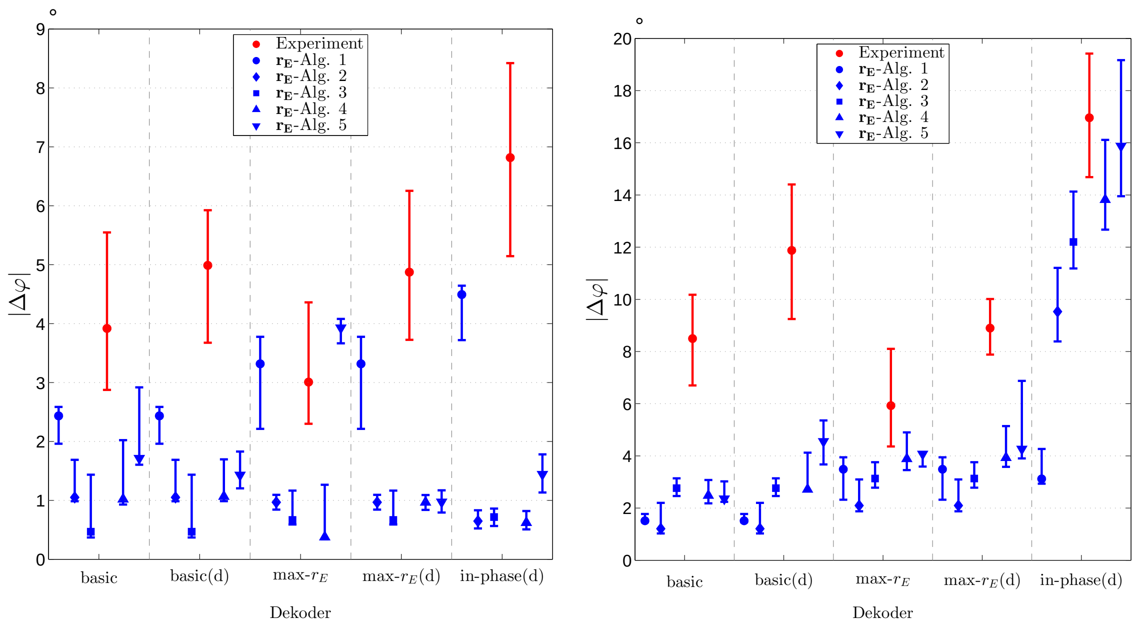
Abbildung 18 zeigt den Einfluss der Delay-Kompensation auf den absoluten Winkelfehler $|\Delta\varphi|$ für die Hörposition inner- und außerhalb des Sweetspots. Zur Auswertung wurden die Versuchsdaten der Wiedergabe mit basic- und max-$\mathbf{r_E}$-Gewichtung verwendet. Zunächst ist zu erkennen, dass die $\mathbf{r_E}$-Algorithmen 1 und 2 für ein- und ausgeschaltete Delay-Kompensation exakt die selben Winkelfehler liefern, da diese die Laufzeiten der Schallwellen bei der Simulation nicht berücksichtigen. Weiterhin ist festzustellen das die gemessenen und prädizierten $|\Delta\varphi|$ denen der Abb. 15 stark ähneln.

Die Simulationsergebnisse der $\mathbf{r_E}$-Alg. 3 und 4 an der zentralen Hörposition unterscheiden sich für Wiedergabe mit und ohne Delay-Kompensation nur marginal. Interessant ist, dass der $\mathbf{r_E}$-Alg. 5 für die Wiedergabe mit Delay-Kompensation ein geringeres $|\Delta\varphi|$ prädiziert, speziell für die Wiedergabe mit 3. und 5. Ordnung Ambisonics. Dieses Verhalten ist gegensätzlich zu bisherigen Versuchsergebnissen [13].

Der Einfluss der Delay-Kompensation ist an der Hörposition 2 wiederum nur marginal

(Abb. 18 (b)). Sowohl die Ergebnisse des Hörversuchs, als auch der $\mathbf{r_E}$-Algorithmen für aktive Delay-Kompensation sind denen ohne Delay-Kompensation fast identisch. Dieses Verhalten ist auf den geringen Einfluss einer Sweetspot-bezogenen Delay-Kompensation an dezentralen Hörpositionen zurückzuführen. An diesen Hörpositionen sind die Laufzeitunterschiede der einzelnen Lautsprechersignale ohnehin so groß, dass die im Vergleich minimale Kompensation der Laufzeitverzögerung perzeptiv keinen Einfluss hat. Es ist zu erkennen, dass die Berücksichtigung der Ausbreitungsdämpfung und das Einbringen einer zeitlichen Gewichtung sehr gute Prädiktionsergebnisse liefert. Bei Wiedergabe mit erster Ordnung Ambisonics sind die prädizierten Winkelfehler des $\mathbf{r_E}$-Alg. 3 und 4 nahezu identisch zu den Hörversuchsergebnissen. Die Berücksichtigung des Präzendenzeffekts und des Cone-of-Confusion führt zu einer signifikanten Überschätzung von $|\Delta\varphi|$. Für die Wiedergabe mit 3. und 5. Ordnung Ambisonics unterschätzen alle $\mathbf{r_E}$-Algorithmen den Winkelfehler der Lokalisation.

Bei der ambisonischen Wiedergabe sind laut Frank [13] Dekoder ohne Delay-Kompensation zu bevorzugen. Diese Aussage kann durch die Simulation mit den $\mathbf{r_E}$-Alg. leider nicht untermauert werden.



(a) (b)

Abbildung 19 – Absoluter Winkelfehler $|\Delta\varphi|$ zum Vergleich der Dekoder für zentrale (a) und dezentrale Hörposition (b). Dekoder mit basic-, max-$\mathbf{r_E}$- und in-phase-Gewichtung. Die Delay-Kompensation ist durch „(d)“ gekennzeichnet.

Abbildung 19 zeigt den Einfluss des verwendeten Dekoders auf den Winkelfehler $|\Delta\varphi|$ für beide Hörpositionen. Es wurden die Versuchsdaten der Wiedergabe mit 5. Ordnung Ambisonics verwendet. Es sind die Ergebnisse des Hörversuchs und der Simulationen für den basic- und den max-$\mathbf{r_E}$-gewichteten Dekoder mit und ohne Delay-Kompensation sowie für den in-phase-gewichteten Dekoder dargestellt. Wie bereits im vorigen Abschnitt festgestellt wurde, sind die Winkelfehler für die $\mathbf{r_E}$-Algorithmen 1 und 2 für die Wiedergabe mit gleichem Dekoder mit und ohne Delay-Kompensation identisch, da die unterschiedlichen Schalllaufzeiten in der Berechnung nicht berücksichtigt werden. Ähnlich zu den Abbildungen 15 (a) und 18 (a) prädizieren die $\mathbf{r_E}$-Alg. 2 bis 4 für die zentrale Hörposition sehr kleine $|\Delta\varphi|$. $\mathbf{r_E}$-Alg. 1 prädiziert bei allen 3 Gewichtungen mit Delay-Kompensation den größten Winkelfehler, liegt jedoch immer noch signifikant unter den Versuchsergebnissen (ausgenommen max-$\mathbf{r_E}$(d)-Gewichtung). Für die einfache max-$\mathbf{r_E}$-Gewichtung scheint Alg. 1 ein guter Prädiktor zu sein. Gleiches gilt für $\mathbf{r_E}$-Alg. 5. Dieser Algorithmus zeigt wieder ein interessantes Verhalten. Er prädiziert wieder für aktive Delay-Kompensation eine signifikant kleineres $|\Delta\varphi|$, vor allem für max-$\mathbf{r_E}$-Gewichtung (vgl. Abb. 18). Es kann festgehalten werden, dass die $\mathbf{r_E}$-Alg. für verschiedene Gewichtungen nur unzureichende Vorhersagen von $|\Delta\varphi|$ an der zentralen Hörposition treffen. Auch der Einfluss der Delay-Kompensation kann nicht abgebildet werden.

In der Abbildung 19 (b) ergeben sich positive Winkelfehler $|\Delta\varphi|$. Die virtuellen Schallquellen werden an der dezentralen Hörposition in Blickrichtung nach links verschoben wahrgenommen. Bei Wiedergabe mit in-phase-Gewichtung ergibt sich im Hörversuch der größte Winkelfehler $|\Delta\varphi|$. Die Delay-Kompensation für basic- und max-$\mathbf{r_E}$-Gewichtung verschlechtert die Lokalisation. Das geringste $|\Delta\varphi|$ wird durch die max-$\mathbf{r_E}$-Gewichtung verursacht. Die $\mathbf{r_E}$-Algorithmen unterschätzen für basic-Gewichtung $|\Delta\varphi|$ signifikant. Die prädizierten Winkelfehler für die Simulation mit und max-$\mathbf{r_E}$-Gewichtung mit und ohne Delay-Kompensation ähneln einander stark. Wieder kann der Einfluss der Delay-Kompensation nicht modelliert werden. Die besten Prädiktionsergebnisse können bei der in-phase-Gewichtung erzielt werden. Der $\mathbf{r_E}$-Alg. 1 prädiziert für diese Gewichtung ein sehr kleines $|\Delta\varphi|$ und weicht signifikant vom Versuchsergebnis ab. Die Berücksichtigung der Ausbreitungsdämpfung ($\mathbf{r_E}$-Alg. 2) sorgt für eine starke Verbesserung des prädizierten Winkelfehlers. Die Simulation mit zusätzlicher zeitlicher Gewichtung ($\mathbf{r_E}$-Alg. 3) bringt eine weitere Verbesserung der Prädiktion. Die Simulationen mit den $\mathbf{r_E}$-Algorithmen 4 bis 5 geben eine gute Schätzung des Ergebnisses des Hörversuchs ab. Tendenziell sind die Konfidenzintervalle der Simulationen mit basic und max-$\mathbf{r_E}$-Gewichtung kleiner als die der Simulation mit in-phase-Gewichtung. Das bedeutet, dass $|\Delta\varphi|$ für die in-phase-Gewichtung stärker über die Panning-Richtungen variiert. Die Analyse der Abbildung 19 (b) bestätigt die Feststellung von Frank in [13], dass die in-phase-Gewichtung für die Lokalisation bei ambisonischer Wiedergabe eher ungeeignet ist. Im Allgemeinen neigen die $\mathbf{r_E}$-Alg. dazu $|\Delta\varphi|$ zu unterschätzen. Dieses Verhalten könnte durch die Wiedergabe mit Ambisonics 5. Ordnung in Zusammenhang mit der gewählten dezentralen Hörpositionen liegen. Diese hätte wahrscheinlich näher am Rand des zirkulären Arrays liegen sollen. Es ist somit sinnvoll die Simulation für das gesamte Auditorium des IEM CUBE durchzuführen.

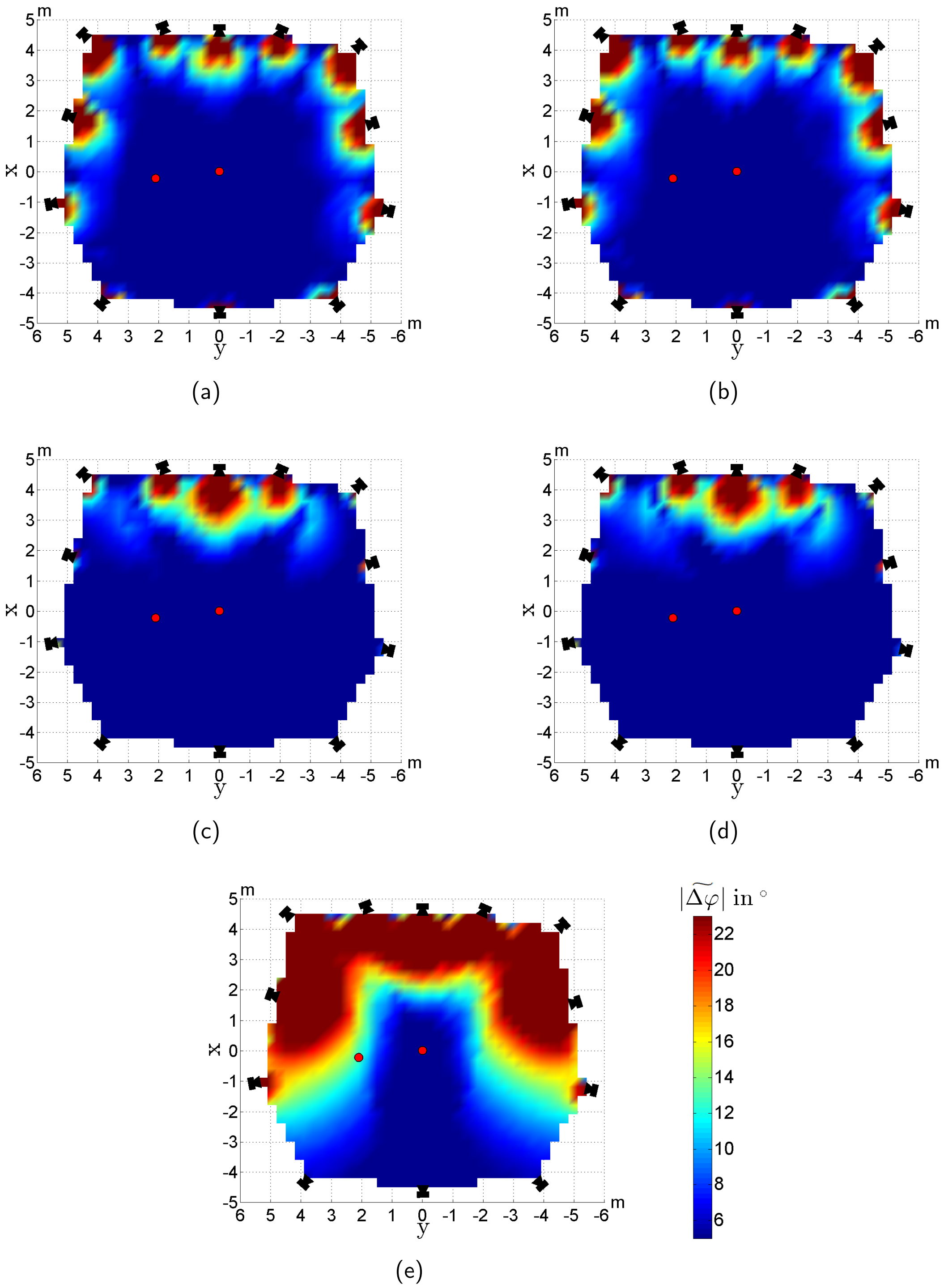


Abbildung 20 – Auditorium des IEM CUBE für die ambisonische Wiedergabe mit 5. Ordnung. Ausgewertet mit $\mathbf{r_E}$-Alg. 4. (a) basic-Gewichtung, (b) basic(d)-Gewichtung, (c) max-$\mathbf{r_E}$-Gewichtung, (d) max-$\mathbf{r_E}$(d)-Gewichtung, (e) in-phase(d)-Gewichtung. Zentrale und dezentrale Hörposition sind durch die roten Punkte gekennzeichnet.

Abbildung 20 zeigt den Medianwert des absoluten Winkelfehlers $|\widetilde{\Delta\varphi}|$ im IEM CUBE ermittelt über die 17 Panning-Positionen. Die Simulation für die verschiedenen Dekodergewichtungen wurde exemplarisch mit dem $\mathbf{r_E}$-Alg. 4 durchgeführt. Wie schon angenommen befinden sich die Hörpositionen für die Wiedergabe mit basic- und max-$\mathbf{r_E}$-Gewichtung zu weit im Zentrum der Lautsprecheranordnung (Abb. 20 (a) bis (d)). An diesen Positionen ist $|\widetilde{\Delta\varphi}|$ minimal. Trotzdem können in der Simulation Unterschiede zwischen den Gewichtung ersichtlich gemacht werden. Die Bedämpfung der Nebenkeulen bei max-$\mathbf{r_E}$-Gewichtung ist sehr gut zu erkennen (vgl. Abb. 20 (a) mit (c)). Bei der basic-Gewichtung spielen im Mittel alle Lautsprecher mit Ausnahme der hinteren drei (Abb. 20 (a), Lautsprecher 6, 7 und 8). Die $\cos$-Fensterung der zirkulären Harmonischen bewirkt, dass im Mittel nur mehr die vorderen 3 Lautsprecher (Abb. 20 (c), Lautsprecher 1, 2 und 12) spielen. Bereits Lautsprecher 3 und 11, mit einem $\varphi \approx \pm 45°$, werden stark bedämpft. Wie schon in Abb. 19 zu erkennen war hat Delay-Kompensation bei der Simulation kaum einen Einfluss auf das prädizierte $|\widetilde{\Delta\varphi}|$ (vgl. in Abb. 20 (a) mit (b) und (c) mit (d)). Die Bereiche mit sehr guter Lokalisation sind für basic- und max-$\mathbf{r_E}$-Gewichtung bei Delay-Kompensation geringfügig kleiner bzw. die Bereiche schlechterer Lokalisation unmittelbar vor den Lautsprechern sind minimal stärker ausgeprägt. Abb. 20 (e) zeigt $|\widetilde{\Delta\varphi}|$ für die Simulation mit in-phase-Gewichtung. Es ist deutlich zu erkennen, dass sich der Hörbereich stark verkleinert. Nahezu der gesamte vordere Bereich mit einem Radius $r \geq 2$ m weist unzureichende Lokalisation auf.

6 Zusammenfassung und Ausblick

In dieser Projektarbeit wurden vier Erweiterungen des $\mathbf{r_E}$-Vektormodells vorgestellt und miteinander verglichen. Die Erweiterungen nehmen schrittweise in ihrer Komplexität zu. Zunächst wird nur die Ausbreitungsdämpfung berücksichtigt ($\mathbf{r_E}$-Alg. 2). Danach folgen die Einbindung von zeitabhängigen Gewichten ($\mathbf{r_E}$-Alg. 3), des Präzedenzeffekts ($\mathbf{r_E}$-Alg. 4) und schließlich des „Cone-of-Confusion" im erweiterten Vektormodell von P. Stitt ($\mathbf{r_E}$-Alg. 5 [38]). Der Überblendparameter α, welcher von P. Stitt in [38] vorgeschlagen wurde, wurde auch im $\mathbf{r_E}$-Alg. 4 integriert. Er sorgt für ein stufenloses Überblenden zwischen $\mathbf{r_E}$-Alg. 2 ($\alpha = 1$) und $\mathbf{r_E}$-Alg. 4 bzw. 5 ($\alpha = 0$). Mit den Versuchsdaten von [16] wurde ein optimales $\alpha = 0.62$ für $\mathbf{r_E}$-Alg. 4 und $\alpha = 0.76$ für $\mathbf{r_E}$-Alg. 5 gefunden. Des weiteren wurde eine Laufzeitanalyse der Berechnungsdauer der Erweiterung des $\mathbf{r_E}$-Vektormodells durchgeführt. Diese kam zu dem Ergebnis, dass die rechenaufwändigeren $\mathbf{r_E}$-Alg. 4 und 5 ca. die 4-fache Berechnungsdauer des einfachen $\mathbf{r_E}$-Vektormodells benötigen. $\mathbf{r_E}$-Alg. 2 und 3 hingegen benötigen nur unwesentlich mehr Zeit.

Die Vorhersage von experimentell bestimmten Hörflächen aus [16] kann schon mit dem recht einfachen $\mathbf{r_E}$-Alg. 2 gut durchgeführt werden. Die Einbindung der zeitabhängigen Gewichte ($\mathbf{r_E}$-Alg. 3) kann die Vorhersage für Wiedergabe mit höheren Ordnungen weiter verbessern. Die aufwändigeren $\mathbf{r_E}$-Alg. 4 und 5 erzielen nur geringfügig bessere Ergebnisse. Der Mehraufwand an Rechenleistung kann somit nicht gerechtfertigt werden. Alle Algorithmen neigen dazu den rückseitigen Teil der Hörfläche größer zu schätzen. Die lateralen Bereiche werden hingegen unterschätzt. Dieser Effekt ist vor allem bei der Wiedergabe mit 1. Ordnung Ambisonics zu erkennen. Der breitere Hörbereich beim Hörversuch könnte aus Raumreflexionen und damit verbundener Verdeckung lateraler Signale resultieren. Es könnte sein, dass lateral einfallende Lautsprechersignale durch Reflexionen maskiert wurden und die Lokalisation somit weniger gestört wurde. Weiterhin wurde ein Vergleich mit den Daten aus dem Hörversuch von [10] durchgeführt. Es konnte festgestellt werden, dass die $\mathbf{r_E}$-Alg. im Allgemeinen dazu neigen den absoluten Winkelfehler $|\Delta\varphi|$ zu unterschätzen. Für höhere Ordnungen der Wiedergabe prädizieren die Algorithmen jedoch genauer und mit einer geringeren Streuung. Der negative Einfluss der Delay-Kompensation auf die Wiedergabequalität kann durch die Simulationen mit den $\mathbf{r_E}$-Alg. an den Hörpositionen des Hörversuchs von [10] nicht bestätigt werden. Jedoch kann man für die Randbereiche der Hörfläche eine Verschlechterung der Lokalisation für die Delay-kompensierte Wiedergabe erkennen (vgl. Abb. 20). Dieses Verhalten bestätigt die Tendenz der Versuchsergebnisse von [10] wiederum. Vergleicht man die Plots für den Median des absoluten Winkelfehlers $|\widetilde{\Delta\varphi}|$ sind die Vorteile einer Wiedergabe mit max-$\mathbf{r_E}$-Gewichtung sehr gut zu erkennen. Die in-phase-Gewichtung erzeugt in großen Bereichen des Auditoriums größere $|\widetilde{\Delta\varphi}|$. Somit bestätigt die Simulation die von Frank getroffene Aussage, dass die max-$\mathbf{r_E}$-Gewichtung bei ambisonischer Wiedergabe zu bevorzugen ist [13].

Um die Vorhersage des Hörbereichs weiter zu verbessern, wäre es möglich, die Richtcharakteristik der im sphärischen Array verwendeten Lautsprecher in die $\mathbf{r_E}$-Algorithmen zu integrieren. Da die bisherigen Simulationen von Freifeldbedingungen ausgehen, sich sphärische Lautsprecher-Arrays aber für gewöhnlich in Räumen befinden, wäre es sinnvoll auch raumakustische Gütemaße wie z.b. die Nachhallzeit T_{60} in die Simulationen einfließen zu lassen. Um die ungenaue Prädiktion der $\mathbf{r_E}$-Alg. im Zentrum eines Lautsprecher-Arrays zu verbessern, wäre es möglich $|\mathbf{r_E}|$ als Maß für die Unsicherheit der Lokalisation in die Algorithmen einfließen zu lassen. Weiterhin könnte auch der Source-Splitting-Effekt in die $\mathbf{r_E}$-Algorithmen integriert werden um die Simulation größerer Wiedergabeumgebungen zu ermöglichen. Unter Source-Splitting ist hierbei die Aufspaltung der virtuellen Schallquelle auf die einzelnen Lautsprecher eines Arrays mit wachsender Array-Größe und damit verbundenen größeren Lautsprecherabständen gemeint. Die bisherigen Simulationen dieser Projektarbeit beschränken sich leider nur auf Lokalisationsuntersuchungen unverhallter virt. Schallquellen. In der Aufführungspraxis werden jedoch meist verhallte Signale verwendet. Es ist somit notwendig die Verhallung, z.B. durch die Spatial Decomposition Method (SDM) [40], von Signalen in die erweiterten $\mathbf{r_E}$-Vektormodelle einfließen zu lassen.

Um die Performance und Robustheit der implementierten Algorithmen weiter validieren zu können, wäre es sinnvoll Simulationen für weitere Lautsprecher-Arrays durchzuführen. Hierfür kämen das mobile Ambisonics-Array des IEM (mAmbA) [17] oder das Lautsprecher-Array des György-Ligeti-Saals der Universität für Musik und darstellende Kunst Graz in Frage. Für diese Lautsprecheranordnungen könnten auch Hörversuche zur Evaluierung der erweiterten $\mathbf{r_E}$-Vektormodelle vorgenommen werden.

7 Referenzen

[1] A. J. Berkhout, D. de Vries und P. Vogel: „Acoustic control by wave field synthesis," *Journal of the Acoustical Society of America*, vol. 93, pp. 2764-2778, 1993

[2] A. J. Berkhout, M. M. Boone, D. de Vries und P. Vogel: „Generation of sound fields using wave field synthesis, an overview," *Proceedings of Active 95*, 1995

[3] D. H. Cooper und T. Shiga: „Discrete-matrix multichannel stereo," *J. Audio Eng. Soc.*, vol. 20, no. 5, pp. 346-360, 1972

[4] P. Craven und M. A. Gerzon: „Coincident microphone simulation covering three dimensional space and yielding various directional outputs," *U.S. Patent*, no. 4,042,779, 1977

[5] L. Cremer: „Die wissenschaftlichen Grundlagen der Raumakustik," Band 1, Stuttgart, S. Hirzel Verlag, 1948

[6] 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*, University of Paris, 2000.

[7] M. Dietz, S. D. Ewert und V. Hohmann: „Auditory model based direction estimation of concurrent speakers from binaural signals," *Speech Communication*, vol. 53, no. 5, pp. 592-605, Mai 2011

[8] „EBU SQAM CD: Sound Quality Assessment Material recordings for subjective tests," 2008

[9] P. Felgett: „Ambisonic reproduction of directionality in surround-sound systems," *Nature*, vil. 252, pp. 534-538, 1974

[10] M. Frank, F. Zotter und A. Sontacchi: „Localization Experiments Using Different 2D Ambisonics Decoders," *25th Tonmeistertagung*,Leipzig, Deutschland, November 2008

[11] M. Frank und F. Zotter: „All-Round Ambisonic Panning and Decoding," *Journal of the Audio Engineering Society*, vol. 60, no. 10, pp. 807-820, Oktober 2012

[12] M. Frank: „Phantom Sources using Multiple Loudspeakers in the Horizontal Plane," *Ph.D. Dissertation*, Universität für Musik und darstellende Kunst Graz, Juni 2013

[13] M. Frank: „Source Width of Frontal Phantom Sources: Perception, Measurement, and Modeling," *Archives of Acoustics*, vol. 38, no. 3, pp. 311-319, Januar 2013

[14] M. Frank: „Localization Using Different Amplitude-panning Methods in the Frontal Horizontal Plane," *Proc. of EAA Joint Symposium on Auralization and Ambisonics, 3-5 April*, Berlin, April 2014, pp. 41-47

[15] M. Frank: „Simple Uncertainty Prediction for Phantom Source Localization," *DAGA Tagungsbericht*, Nürnberg, pp. 1630-1633, 2015

[16] M. Frank und F. Zotter: „Exploring the perceptual sweet area in Ambisonics," *142nd AES Convention*, Berlin, Deutschland, Mai 2017

[17] M. Frank und A. Sontacchi: „Case Study on Ambisonics for Multi-Venue and Multi-Target Concerts and Broadcasts," submitted to *Journal of the Audio Engineering Society*, 2017

[18] M. Frank und F. Zotter: „Extension of the generalized tangent law for multiple loeudspeakers," *DAGA Tagungsbericht*, Kiel, pp. 1081-1084, 2017

[19] M. A. Gerzon: „The design of precisely coincident microphone arrays for stereo and surround sound," *prepr. L-20 of 50th Audio Eng. Soc. Conv.*, 1975

[20] M. Gerzon: „Design of Ambisonics Decoders for for Multispeaker Surround Sound," *58th Convention of the Audio Engineering Society*, New York, NY, Nov. 1977.

[21] M. Gerzon: „General metatheory of auditory localisation," *92nd Convention of the Audio Engineering Society*, Wien, März 1992.

[22] H. Haas: „Über den Einfluss eines Einfachechosauf die Hörsamkeit von Sprache," *Acustica 1*, 1951, pp. 49-58

[23] Recommendation ITU-R BS.2051-0, „Advanced sound system for programme production," Genf, 2014

[24] L. A. Jeffress: „A place theory of sound localization," *Journal of comparative and physiological psychology*, vol. 41, pp. 35-39, 1948.

[25] M. Kronlachner: „Räumliche Transformationen zur Veränderung von Ambisonics Aufnahmen," *Masterarbeit*, Universität für Musik und darstellende Kunst Graz, 2014

[26] D. M. Leakey: „Some measurements on the effects of interchannel intensity and time differences in two channel sound systems," *The Journal of the Acoustical Society of America*, vol. 31, no. 7,pp. 977-986,1959

[27] H. Lee und F. Rumsey: „Level and time panning of phantom images for musical sources," *Journal of the Audio Engineering Society*, vol. 61, no. 12, pp. 978-988, 2013

[28] W. Lindemann: „Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals," *The Journal of the Acoustical Society of America*, vol. 80, no. 6, pp. 1608-1622, 1986

[29] W. Lindemann: „Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals," *The Journal of the Acoustical Society of America*, vol. 80, no. 6, pp. 1623-1630, 1986

[30] R. Y. Litovsky, H. S. Colburn, W. A. Yost und S. J. Guzman: „The Precedence Effect," *The Journal of the Acoustical Society of America*, vol. 106, no. 4 Pt. 1, pp. 1633-1654, Oktober 1999

[31] D. G. Malham: „Experience with large area 3-D ambisonic sound system," *Proceedings of the Institute of Acoustics*, vol. 14, no. 5, pp. 209-216, 1992.

[32] C. Moldryzk, A. Goertz, M. Makarski, S. Feistel, W. Ahnert und S. Weinzierl: „Wellenfeldsynthese für einen großen Hörsaal," *DAGA Tagungsbericht*, Stuttgart, 2007

[33] V. Pulkki: „Virtual sound source positioning using vector base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456-466, 1997

[34] J. Ootomo, K. Tanno, A. Saji, J. Huang und W. Hatano: „The precedence effect of sound from a side direction," *30th AES International Conference*, Saariselkä, Finnland, März 2007

[35] B. Rakerd, W. M. Hartmann und J. Hsu: „Echo suppression in the horizontal and median sagittal planes," *The Journal of the Acoustical Society of America*, Vol. 107, pp. 1061 - 1064, 2000

[36] P. W. Robinson, A. Walther, C. Faller und J. Braasch: „Echo thresholds for reflections from acoustically diffusive architectural surfaces," *The Journal of the Acoustical Society of America*, Vol. 134, pp. 2755 - 2764, 2013

[37] B. G. Shinn-Cunningham, P. M. Zurek und N. I. Durlach: „Adjustment and discrimination measurements of the precedence effect," *The Journal of the Acoustical Society of America*, vol. 93, no. 5, pp. 2923-2932, 1993

[38] P. Stitt: „Ambisonics and Higher-Order Ambisonics for Off-Centre Listeners: Evaluation of Perceived and Predicted Image Direction," *Doctoral Thesis*, Queen's University Belfast, April 2015

[39] P. Stitt, S. Bertet, M. van Walstijn: „Off-Center Listening with Third-Order Ambisonics: Dependence of Perceived Source Direction on Signal Type," *Journal of the Audio Engineering Society*, vol. 65, no. 3, March 2017

[40] S. Tervo, J. Pätynen, A. Kuusinen und T. Lokki: „Spatial decomposition method for room impulse responses," *Journal of the Audio Engineering Society*, vol. 61, no. 1/2, pp. 17 - 28, 2013

[41] H. Wallach, E. B. Newman und M. R. Rosenzweig: „The precedence effect in sound localization," *The American Journal of Psychology*, vol. 62, no. 3, pp. 315-336, April 1949

[42] D. Ward and T. Abhayapala: „Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Transactions on Speech and Audio Processing*, 9(6), pp. 697-707, 2001.

[43] J. Zmölnig: „Entwurf und Implementierung einer Mehrkanal-Beschallungsanlage," *Diplomarbeit*, Universität für Musik und darstellende Kunst Graz, 2002

[44] F. Zotter: „Analysis and Synthesis of Sound-Radiation with Spherical Arrays," *Doktorarbeit*, Institut für Elektronische Musik und Akustik, Graz, 2009

[45] F. Zotter: „Sampling strategies for acoustic holography/holophony on the sphere," in *NAG-DAGA*, Rotterdam, 2009

[46] F. Zotter, M. Frank, A. Fuchs und D. Rudrich: „Preliminary study on the perception of orientation-changing directional sound sources in rooms," *Proceedings of Forum Acusticum*, Krakau, September 2014