

Bachelorarbeit im Studiengang Audiovisuelle Medien

mit dem Titel

**Vergleich des automatischen Downmixes von 3D-  
Hauptmikrofonarrays bei NGA-Systemen in Hinblick auf  
die Entstehung von Kammfiltereffekten**

Vorgelegt von: Klara Schulmeyer

Matrikelnummer: 39096

an der Hochschule der Medien Stuttgart

am 21.07.2023

zur Erlangung des akademischen Grades eines Bachelor of Engineering

Erst-Prüfer: Prof. Dr. Frank Melchior

Zweit-Prüfer: Prof. Oliver Curdt

## **Ehrenwörtliche Erklärung**

Hiermit versichere ich, Klara Schulmeyer, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel „Vergleich des automatischen Downmixes von 3D-Hauptmikrofonarrays bei NGA-Systemen in Hinblick auf die Entstehung von Kammfiltereffekten“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 26 Abs. 2 Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. § 19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

  
\_\_\_\_\_

Stuttgart, 20.07.2023

## **Danksagung**

Besonderer Dank gilt meinem Erstprüfer Prof. Dr. Frank Melchior, welcher mir während der Bearbeitungszeit für fachlichen Fragen zu Verfügung stand und sich dafür stets viel Zeit nahm. Weiterer Dank gilt meinem Zweitprüfer Prof. Oliver Curdt. Beide haben meinen Werdegang durch ihre Lehrveranstaltungen und Gespräche abseits der Hochschulveranstaltungen entscheidend geprägt.

Zudem möchte ich meiner Familie und meinem Partner für ihre emotionale Unterstützung danken.

## **Kurzfassung**

Die vorliegende Arbeit beschäftigt sich mit dem Vergleich verschiedener 3D-Downmixverfahren in Hinblick auf Produktionen von akustischer Musik. Ziel der Arbeit ist es, den Einfluss unterschiedlicher 3D-Mikrofonarrays und der bei Dolby Atmos und MPEG-H integrierten Downmixalgorithmen auf das Downmixergebnis von akustischen 3D-Musikaufnahmen festzustellen.

Dazu wurden mit einem PCMA-3D, einem Decca Cuboid und einem ORTF-3D angefertigte Mehrkanalaufnahmen mithilfe eines Testsignals nachgebildet. Dabei wurde angenommen, dass ideale Mikrofone mit idealer Richtcharakteristik zum Einsatz kommen sowie Freifeld-Bedingungen herrschen. Anschließend wurden von diesen im Ausgangsformat 5.0.4 vorliegenden Simulationen sowie von einer akustischen 3D-Orchesteraufnahme jeweils mit beiden Algorithmen ein 5.0- und ein Stereodownmix erstellt und die Ergebnisse bezüglich auftretender Kammfiltereffekte verglichen.

Es konnte Folgendes festgestellt werden: Je größer die Abstände und damit die Laufzeitdifferenzen zwischen den Mikrofonkapseln eines Arrays und je kleiner die Pegeldifferenzen zwischen den jeweiligen Signalen, desto ausgeprägtere Kammfilterstrukturen ergeben sich im Downmix. Diese waren in den mit dem MPEG-H zugrundeliegenden Codec erstellten Downmixes durch den aktiven Algorithmus und die damit verbundene Kammfilterkompensation weniger ausgebildet.

## **Abstract**

This thesis deals with the comparison of different 3D downmix methods regarding acoustic music productions. The purpose of this work is to examine the influence of different 3D microphone arrays and the downmix algorithms integrated in Dolby Atmos and MPEG-H on the downmix result of acoustic 3D music recordings.

Therefore, 5.0.4 multichannel recordings made with a PCMA-3D, a Decca Cuboid and an ORTF-3D were simulated using a test signal. It was assumed that ideal microphones with ideal polar patterns were used and that free-field conditions applied. Subsequently, a 5.0- and a stereo downmix were created from these simulations as well as from an acoustic 3D recording of an orchestra using both algorithms. The results were compared with regard to occurring comb filters.

The following results were determined: The greater the distances and thus the delay differences between the microphone capsules of an array and the smaller the level differences between the signals, the more prominent comb filter structures arise in the downmix. These were less pronounced in the downmixes created with the MPEG-H codec due to the active algorithm and the associated comb filter compensation.

## Abkürzungsverzeichnis

C	Center
DAW	Digital Audio Workstation
FC	Front Center
FL	Front Left
FR	Front Right
L	Left
LFE	Low Frequency Effects
Lh	Left Height
Lhs	Left Height Surround
Lrs	Left Rear Surround
Ls	Left Surround
Lss	Left Side Surround
Ltr	Left Top Rear
Ltf	Left Top Front
MHAPi	MPEG-H Authoring Plug-in
ML	Middle Layer
MPEG	Moving Pictures Expert Group
ORTF	Office de Radiodiffusion Télévision Française
PCMA	Perspective Control Microphone Array
UL	Upper Layer
R	Right
Rh	Right Height
Rhs	Right Height Surround
Rrs	Right Rear Surround
RL	Rear Left
RR	Rear Right
Rs	Right Surround
Rss	Right Side Surround
Rtf	Right Top Front
Rtr	Right Top Rear
TL	Top Left
Tr	Top Rear

## Inhaltsverzeichnis

1.	Einleitung .....	1
2.	Erläuterung verschiedener Downmixverfahren .....	2
2.1.	Aktiver vs. Passiver Downmix .....	2
2.2.	Dolby Atmos .....	2
2.3.	MPEG-H .....	7
3.	Forschungsfrage .....	13
4.	Methodik .....	14
5.	Aufbau der DAW-Projekte .....	16
5.1.	Dolby Atmos mit Avid Pro Tools .....	16
5.2.	MPEG-H mit Steinberg Nuendo .....	17
6.	Verwendete Arrays .....	18
6.1.	PCMA-3D .....	18
6.2.	Decca Cuboid .....	20
6.3.	ORTF 3D .....	23
7.	Vergleich verschiedener Downmixe .....	26
7.1.	Downmix Dolby Atmos Bed .....	26
7.2.	Downmix Dolby Atmos Objekte .....	45
7.3.	Downmix MPEG-H Bed .....	53
7.4.	Downmix MPEG-H Objekte .....	60
7.5.	Zusammenfassung der Ergebnisse .....	66
8.	Grenzen der Methodik .....	74
9.	Praktische Anwendung .....	75
10.	Ausblick .....	76

11.	Literaturverzeichnis .....	77
12.	Abbildungsverzeichnis .....	79
13.	Tabellenverzeichnis .....	83
14.	Anhang A .....	84
	A.1 Pegel- und Delaywerte PCMA-3D 0° .....	84
	A.2 Pegel- und Delaywerte PCMA 135° .....	84
	A.3 Pegel- und Delaywerte Decca Cuboid 0° .....	85
	A.4 Pegel- und Delaywerte Decca Cuboid 135° .....	85
	A.5 Pegel- und Delaywerte ORTF-3D 0° .....	86
	A.6 Pegel- und Delaywerte ORTF-3D 135° .....	86
15.	Anhang B Routing in Studio U48 .....	87
	B.1 Routing für die Arbeit mit Dolby Atmos .....	87
	B.2 Routing für die Arbeit mit MPEG-H .....	87
16.	Anhang C elektronischer Anhang.....	88

## **1. Einleitung**

Aus der Einführung von 3D-Audio in den Musikbereich resultieren neue Herausforderungen an die Produktion und Distribution der Inhalte. Durch die fehlende Kontrolle über die Wiedergabebedingungen beim Konsumenten ergeben sich mit steigender Kanalanzahl zunehmend Probleme. Es ist davon auszugehen, dass nicht jeder Konsument die Anzahl an Lautsprechern des Produktionsformates zur Verfügung hat. Um das Hörerlebnis des Zuhörers einschätzen zu können und soweit möglich auf allen Lautsprecheranordnungen gleichbleibende Qualität sicherzustellen, müssen daher bereits während der Produktion Vorkehrungen für die Wiedergabe auf einem kleineren Lautsprechersetup getroffen werden. Eine Möglichkeit ist, für jedes übliche Format, beispielsweise 7.1.4, 5.1 und 2.0, eine separate Mischung anzufertigen. Dies ist in den meisten Fällen wirtschaftlich unattraktiv, weswegen sog. automatische Downmixe zum Einsatz kommen.

Dabei werden bei verschiedenen Genres und Aufnahmetechniken unterschiedliche Anforderungen an den Downmix gestellt. So ist es beispielsweise bei elektronisch produzierter Musik mit vielen im Raum verteilten Elementen und eventuellen Panning-Automatiken wichtig, wie sich die Ortung der Inhalte auf kleinere Lautsprecher setups übertragen lässt. Bei klassischer Musik ist es hingegen u.a. von Bedeutung, inwiefern sich der natürliche Raumeindruck und damit zusammenhängend die Balance aus Direktschall, Reflexionen und Diffusschall mit weniger Lautsprechern erzielen lässt.

## **2. Erläuterung verschiedener Downmixverfahren**

Beim Downmix werden Signale, die für M Lautsprecher vorgesehen sind in ein Format mit N Lautsprechern überführt. Dabei gilt:  $N < M$

Zur Einschätzung der Ergebnisse verschiedener Downmixer, ist es nötig einen Überblick über die verwendeten Algorithmen zu haben. In dieser Arbeit werden aufgrund ihrer wirtschaftlichen Bedeutung die in Dolby Atmos und MPEG-H integrierten Downmixalgorithmen verglichen.

### **2.1. Aktiver vs. Passiver Downmix**

Beim Downmix von Mehrkanalsignalen wird zwischen aktivem und passivem Downmix unterschieden. Bei einem passiven Algorithmus werden die Input-Signale mit feststehenden Koeffizienten gewichtet und entsprechend einer statischen Matrix miteinander summiert, um die Output-Signale zu erhalten. Dieses Verfahren führt durch die Summierung ähnlicher, aber in der Phase verschobener Signale unter anderem zu Kammfiltern. (Vilkamo et al., 2014) Diese Effekte sollen durch aktive Algorithmen kompensiert werden. Dies kann entweder durch eine Bearbeitung des Downmixsignals, beispielsweise durch eine Anhebung oder Absenkung bestimmter Frequenzbänder durch einen Equalizer (Faller & Baumgarte, 2003), oder durch eine Angleichung der Phase der Ursprungssignale vor dem Downmixvorgang erfolgen. (Adami et al., 2014)

Dolby Atmos arbeitet mit einem passiven Algorithmus, während MPEG-H einen aktiven Algorithmus mit Phasen Anpassung vor dem Downmixvorgang nutzt (Vilkamo et al., 2014).

### **2.2. Dolby Atmos**

Dolby Atmos wurde 2012 im Kino eingeführt und findet seit 2014 auch im Heimkinobereich und seit 2017 auf der Videostreaming-Plattform Netflix Verwendung. 2019 wurde mit der Einführung von Dolby Atmos Music auch der Musikstreamingmarkt erschlossen. Mittlerweile können unter anderem auf den Plattformen Apple Music, Amazon Music und Tidal Dolby Atmos Inhalte abgerufen werden.

Dolby Atmos bezeichnet eine geschützte Marke von Dolby und steht für ein Audioerlebnis, welches über die verschiedenen Codecs Dolby True HD, Dolby

Digital Plus und Dolby AC-4 übertragen werden kann (Dolby, 2023). Dolby True HD ist der einzige verlustfreie Codec, der Dolby Atmos unterstützt (Dolby, 2022b). Dolby Digital Plus wurde für die Übertragung von Dolby Atmos um die Komponenten JOC (Joint Object Coding) und OAMD (Object Audio Metadata) erweitert. Diese ergänzen den Dolby Digital zugrundeliegenden Codec AC-3 und sind vom Europäischen Institut für Telekommunikationsnormen in ETSI TS 103 420 standardisiert. (Dolby, 2022d) AC-4 ist der neuere und effizientere der beiden verlustbehafteten Codecs (Dolby, 2022c) und in ETSI TS 103 190-1 und ETSI TS 103 190-2 standardisiert. Zudem ist Dolby AC-4 Teil des ATSC 3.0 Systems, eine Reihe aus Standards für next generation broadcasting (Advanced Television Systems Committee). Für eine tiefergehende Erläuterung der Codecs sei auf die entsprechenden Standards verwiesen. Während Dolby TrueHD für den Einsatz auf Blu-Ray Discs konzipiert ist, werden DD+ und AC-4 für Streaminganwendungen und die Übertragung von Dolby Atmos Music verwendet.

Dolby Atmos gehört sowohl zu den kanalbasierten als auch den objektbasierten Produktionsformaten. Die kanalbasierte Komponente ist das sog. Bed, welches bis zu zehn Kanäle im Format 7.1.2 unterstützt. Es ist möglich mehrere Beds gleichzeitig zu verwenden. (Dolby, 2022f) Je nach Samplingfrequenz stehen 128 (48 kHz) oder 54 (96 kHz) Input Channel zur Verfügung. Die übrigen Input Channel können für Objekte verwendet werden. Hier sind sowohl Mono- als auch Stereoobjekte möglich (Dolby, 2022e). Wie in ETSI TS 103 190-2 V1.2.1 und ETSI TS 103 420 V1.2.1, beide aus dem Jahr 2018, beschrieben, werden mithilfe von Metadaten Informationen für das Rendering von Objekten übertragen. Diese beinhalten u.a. deren Position und die Objekte werden dementsprechend (unter Anpassung der Pegel) auf die Lautsprecher im verwendeten Setup verteilt. Bei kanalbasierten Komponenten hingegen sind Spezifikationen für den Down- oder Upmix nötig, falls das verwendete Lautsprecher-setup nicht mit der kodierten Konfiguration übereinstimmt.

Das von Dolby empfohlene Lautsprecher-setup für das in dieser Arbeit verwendete Ausgangsformat Format 5.0.4 wird den Abbildungen eins und zwei sowie Tabelle eins dargestellt (Dolby, 2023b).

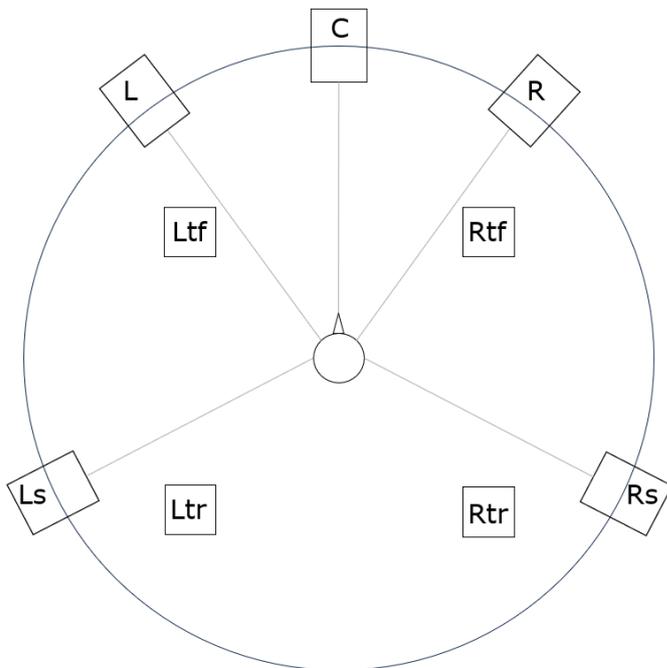
**Tabelle 1**

*Lautsprecherkonfigurationen bei Dolby Atmos im Format 5.1.4*

#	Name		Azimuth	Elevation
1	Left	L	+22° ... +30°	0°
2	Right	R	-22° ... -30°	0°
3	Center	C	0°	0°
4	LFE			
5	Left Surround	Ls	+110° ... +120°	0°
6	Right Surround	Rs	-110° ... -120°	0°
7	Left Top Front	Ltf		+30° ... +55°
8	Right Top Front	Rtf		+30° ... +55°
9	Left Top Rear	Ltr		+125° ... +150°
10	Right Top Rear	Rtr		+125° ... +150°

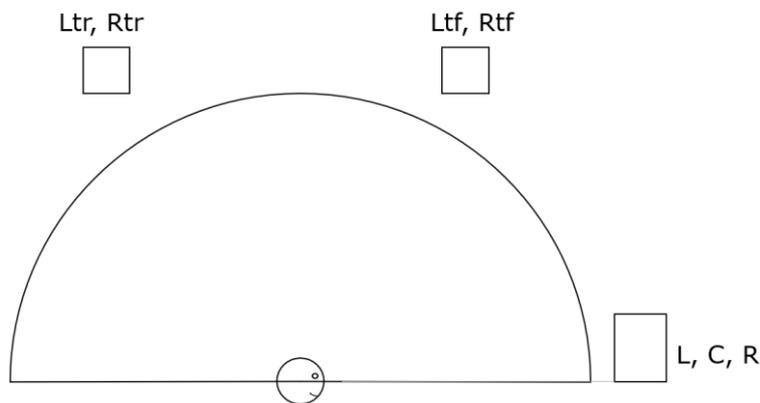
**Abbildung 1**

*Lautsprecheranordnung bei Dolby Atmos im Format 5.0.4, Draufsicht*



**Abbildung 2**

*Lautsprecheranordnung bei Dolby Atmos im Format 5.0.4, Seitenansicht*



Da im Bed die größtmögliche Kanalanzahl im Upper Layer zwei beträgt, werden hier Ltf und Ltr zu TI (Top Left) und Rtf und Rtr zu Tr (Top Right) zusammengefasst.

Der Downmix des Beds bei Dolby Atmos erfolgt immer stufenweise: Zunächst wird ein 7.1-Render des Contents erstellt, indem jeweils die linke (Kanal TI) bzw. rechte Seite (Tr) des Upper Layers auf den Kanal Left Side Surround (Lss) bzw. Right Side Surround (Rss) angerechnet werden. Der Koeffizient hierfür wurde durch den Downmix eines Mono-Signals ermittelt und beträgt -1,5 dB. Die Lautsprecher der Kanäle Lss und Rss bei der Wiedergabe im Format 7.1 entsprechen dabei den Lautsprechern der Kanäle Ls und Rs im Format 5.1. Aus dem 7.1-Render wird ein 5.1-Downmix berechnet. Dieser ist anschließend die Grundlage für einen 2.0-Downmix. Im Dolby Atmos Renderer kann unter folgenden Downmixsettings für das Format 5.1 gewählt werden:

Standard (Lo/Ro) – default:

$$Ls = 0 \text{ dB} \times Lss + 0 \text{ dB} \times Lrs$$

$$Rs = 0 \text{ dB} \times Rss + 0 \text{ dB} \times Rrs$$

Dolby Pro Logic IIx:

$$L_s = L_{ss} + (-1.2 \text{ dB} \times L_{rs}) + (-6.2 \text{ dB} \times R_{rs})$$

$$R_s = R_{ss} + (-6.2 \text{ dB} \times L_{rs}) + (-1.2 \text{ dB} \times R_{rs})$$

Zusätzlich gibt es die Möglichkeit einen *direct render* auszuwählen. Hier werden die Spuren ohne die vorherige Erstellung eines 7.1-Renders zu 5.1 downgemixt. Zusätzlich steht ein *direct render with room balance* zur Verfügung, bei welchem Kammfiltereffekte, welche durch Phantomschallquellen zwischen den Front und Rear Channels entstehen, reduziert werden. (Dolby, 2022a)

Die Wahl der 5.1-Downmixeinstellungen beeinflusst durch den stufenweisen Downmix das Ergebnis des Stereodownmixes. Für den Downmix von 5.1 zu 2.0 können folgende Einstellungen verwendet werden:

Lo/Ro - default:

$$L_o = L + (-3 \text{ dB} \times C) + (-3 \text{ dB} \times L_s)$$

$$R_o = R + (-3 \text{ dB} \times C) + (-3 \text{ dB} \times R_s)$$

Lt/Rt (Pro Logic II):

$$L_t = L + (-3 \text{ dB} \times C) - (-1.2 \text{ dB} \times L_s) - (-6.2 \text{ dB} \times R_s)$$

$$R_t = R + (-3 \text{ dB} \times C) + (-6.2 \text{ dB} \times L_s) + (-1.2 \text{ dB} \times R_s)$$

Lt/Rt (Pro Logic II) w/Phase 90:

$$L_t = L + (-3 \text{ dB} \times C) - (-1.2 \text{ dB} \times L_s) - (-6.2 \text{ dB} \times R_s)$$

$$R_t = R + (-3 \text{ dB} \times C) + (-6.2 \text{ dB} \times L_s) + (-1.2 \text{ dB} \times R_s)$$

Bei der letzten Einstellung werden die Kanäle  $L_s$  und  $R_s$  um  $90^\circ$  in der Phase verschoben. Dies reduziert ungewollte Signalauslöschungen, verbessert die Abbildung von Schallereignissen und ermöglicht eine Matrix Dekodierung. Daher wird dieser für Lt/Rt-Downmixes empfohlen. (Dolby, 2022a)

### **2.3. MPEG-H**

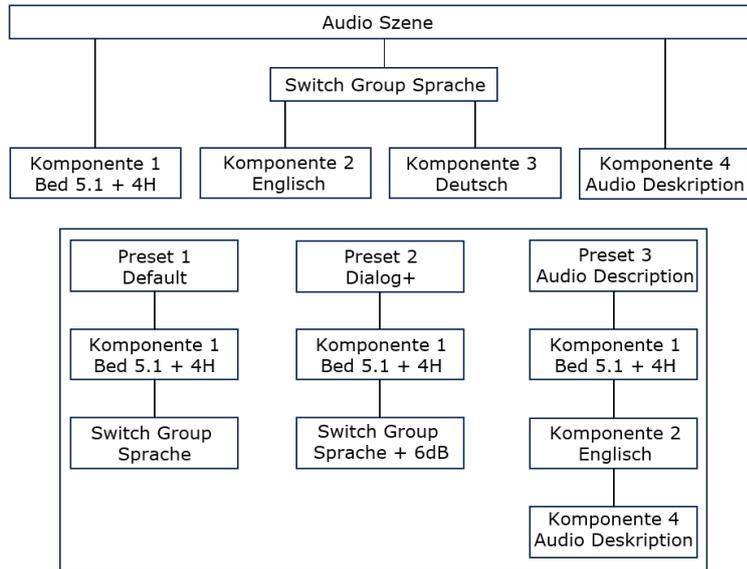
MPEG-H ist ein als ISO/IEC 23008-3 spezifizierter Audiocodec und wurde 2015 für den Broadcastbereich entwickelt. 2019 wurde durch die Einführung des 3D-Audioformats Sony 360°, welches auf MPEG-H basiert, die Nutzung des Codecs auf die Musikproduktion ausgeweitet. Sony 360° kann u.a. über Amazon Music HD, Deezer und Tidal gestreamt werden. (Fraunhofer-Institut für Integrierte Schaltungen IIS, 2023)

MPEG-H unterstützt kanalbasierte und objektbasierte Inhalte sowie Higher Order Ambisonics (Herre & Quackenbush, 2022). Dabei können bis zu 128 Signale auf bis zu 64 Lautsprecher gerendert werden. Um die Komplexität der Implementierung gering zu halten, wurde ein sog. Low-Complexity Profile (LC-Profile) mit verschiedenen Levels definiert. Das LC-Profile Level 3 unterstützt die Wiedergabe von höchstens 32 Signalen auf bis zu 24 Lautsprechern und ist Teil des ATSC 3.0 Systems. (Simon et al., 2018)

Audio-Objekte werden meist in Kombination mit kanal- oder szenenbasiertem Content verwendet. Durch die Konzipierung von MPEG-H für den Broadcastbereich sind Objekte für die Übertragung alternativer Audiospuren vorgesehen, welche dem Endkonsumenten Möglichkeit zur Interaktion bzw. Personalisierung geben, während die Mischung im kanalbasierten Bed stattfindet. Wie viel Interaktion erlaubt ist, wird bei der Produktion festgelegt. Dafür kann eine Organisation der Objekte in sog. Switch Groups vorgenommen werden. Elemente einer Switch Group lassen sich nie gleichzeitig abspielen. Dies bietet sich z.B. für die Übertragung mehrerer Sprachen an. Zusätzlich können verschiedene Presets (beispielsweise für Dialogue Enhancement) bereitgestellt werden. Der Endkonsument kann diese über ein Menü abrufen (Bleidt et al., 2017; Grewe et al., 2021). Abbildung drei zeigt ein Beispiel für die Struktur einer aus Bed und Objekten bestehenden Szene:

**Abbildung 3**

Beispiel für eine MPEG-H Audio Szene



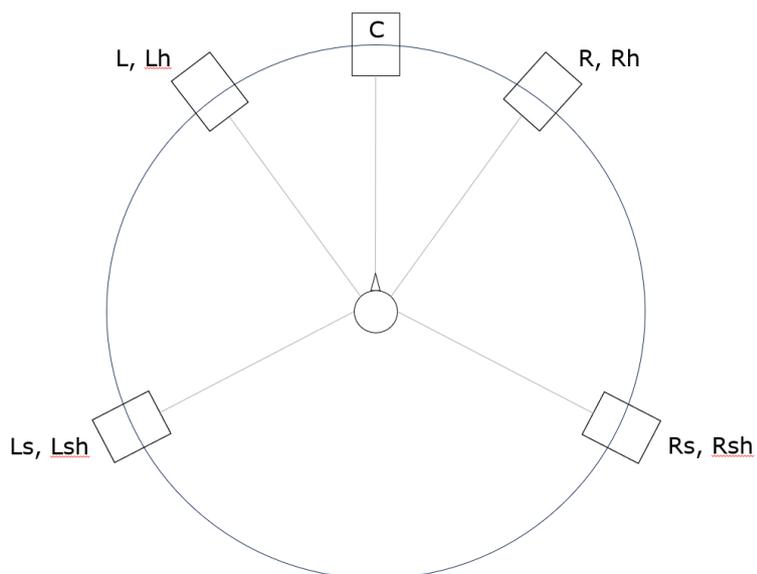
Quelle: In Anlehnung an (Grewe et al., 2021)

Für die Erstellung einer Szene bzw. die gesamte Organisation des Contents stehen zwei Werkzeuge zur Verfügung: das MPEG-H Authoring Tool und das MPGE-H Authoring Plug-in (MHAPI), mit welchem direkt in der DAW gearbeitet werden kann.

MPEG-H unterstützt für die kanalbasierte Komponente Lautsprecheranordnungen bis zum Format 22.2 (Bleidt et al., 2017). Das empfohlene Lautsprechersetup für das in dieser Arbeit verwendete Ausgangsformat 5.1+4H entspricht dem ISO/IEC 23091-3 [3] Standard und wird in der folgenden Tabelle sowie den Abbildungen vier und fünf dargestellt.

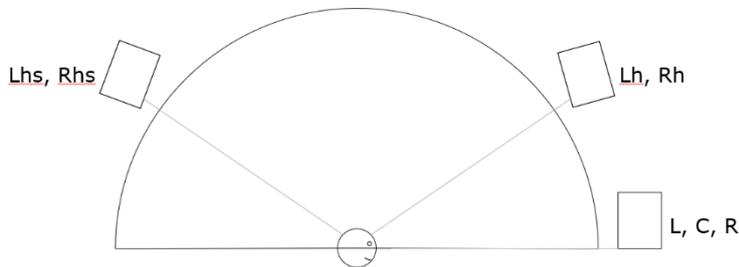
**Tabelle 2***Lautsprecherkonfigurationen bei MPEG-H im Format 5.1 + 4H*

#	Name		Azimuth	Elevation
1	Left	L	+30°	0°
2	Right	R	-30°	0°
3	Center	C	0°	0°
4	LFE			
5	Left Surround	Ls	+100° ... +120°	0°
6	Right Surround	Rs	-100° ... -120°	0°
7	Front Left Height	Lh	+30° ... + 45°	+30° ... + 55°
8	Front Right Height	Rh	-30° ... - 45°	+30° ... +55°
9	Left Surround Height	Lsh	+100° ... +135°	+30° ... + 55°
10	Right Surround Height	Rsh	-100° ... -135°	+30° ... + 55°

*Quelle:* (Grewe et al., 2020)**Abbildung 4***Lautsprecheranordnung bei MPEG-H im Format 5.0+4H, Draufsicht*

**Abbildung 5**

*Lautsprecheranordnungen bei MPEG-H im Format 5.0+4H, Seitenansicht*



Während HOA-Inhalte mittels eines HOA Renderers und Objekte mittels VBAP (Vector Base Amplitude Panning) auf das Wiedergabe-Lautsprechersetup gerendert werden, ist für kanalbasierte Inhalte ein Downmix nötig, falls das Zielformat weniger Lautsprecher verwendet als das Produktionsformat (Herre & Quackenbush, 2022).

Dabei kommt der sog. Format Converter zum Einsatz. Als erster Schritt des Downmixes wird eine Matrix, welche die Signale des Ausgangsformats -durch Koeffizienten gewichtet- in das Zielformat überführt, ermittelt. Für übliche Zielformate sind diese vordefiniert. Für unsymmetrische oder unübliche Lautsprecherkonfigurationen wird iterativ für jedes Signal des Ausgangsformats aus einer Liste an festgelegten Regeln eine sog. Mapping Rule ausgewählt, mit welcher das Signal auf einen oder mehrere Lautsprecher des Zielformats gelegt wird. So wird für jeden Downmix von einem Produktionsformat zu einem kleineren Zielformat eine Downmixmatrix adaptiert. Bevor diese Downmixkoeffizienten angewendet werden wird ein aktiver Downmixalgorithmus angewendet, um aus passivem Downmix entstehende Artefakte wie Kammfilter zu vermeiden. Dazu wird die Korrelation zwischen denjenigen Signalen gemessen, die im Verlauf des Downmixvorgangs summiert werden. Falls nötig, werden diese jeweils zueinander in ihrer Phase angepasst. Außerdem wird eine frequenzabhängige Anpassung der Downmixgains vorgenommen, um einen Leistungsverlust der Input Signale zu verhindern. Bei diesem aktiven Algorithmus bleiben dekorrelierte Inputsignale unverändert und somit werden durch den passiven Downmixvorgang entstehende ungewollte Effekte mit minimalen Anpassungen eliminiert. (Herre & al, 2015; Herre & Quackenbush, 2022)

Für den Downmix der kanalbasierten Komponente kann im MHAPi-Plugin und dem MPEG-H Authoring Tool für den 5.1- und 2.0-Downmix zwischen verschiedenen Einstellungen gewählt werden. Diese sind nach verschiedenen Anwendungsfällen bzw. Genres kategorisiert (Tabelle 3-7):

**Tabelle 3***Broadcast Legacy*

Center → L/R	Surround → Front	Upper Front → Down	Upper Surround → Down	LFE → L/R
-3,00 dB	-4,50 dB	-4,50 dB	-4,50 dB	Muted

**Tabelle 4***Dynamic*

Center → L/R	Surround → Front	Upper Front → Down	Upper Surround → Down	LFE → L/R
-3,00 dB	-3,00 dB	-3,00 dB	-3,00 dB	-3,00 dB

**Tabelle 5***Classics*

Center → L/R	Surround → Front	Upper Front → Down	Upper Surround → Down	LFE → L/R
-3,00 dB	-6,00 dB	-3,00 dB	-6,00 dB	Muted

**Tabelle 6***Loud*

Center → L/R	Surround → Front	Upper Front → Down	Upper Surround → Down	LFE → L/R
-3,00 dB	-6,00 dB	-3,00 dB	-3,00 dB	-3,00 dB

**Tabelle 7***Sports*

Center → L/R	Surround → Front	Upper Front → Down	Upper Surround → Down	LFE → L/R
0,00 dB	-4,50 dB	-3,00 dB	-4,50 dB	Muted

Wird keine dieser Optionen gewählt, kommt eine Default Einstellung zum Einsatz. Die Koeffizienten hierfür wurden durch den Downmix von Monosignalen ermittelt und werden in Tabelle 8 dargestellt:

**Tabelle 8***Default*

Center → L/R	Surround → Front	Upper Front → Down	Upper Surround → Down	LFE → L/R
-3,00 dB	-2,00 dB	-1,00 dB	-1,50 dB	muted

### **3. Forschungsfrage**

Die verschiedenen Downmixverfahren sollen in der vorliegenden Arbeit bezüglich des Auftretens von Kammfiltereffekten beim Downmixing akustischer Musikaufnahmen verglichen werden.

Als akustische Aufnahmen werden im Rahmen dieser Arbeit Aufnahmen bezeichnet, bei denen mit einem Hauptmikrofon-Array gearbeitet wird, anstatt jedes Instrument einzeln nah zu mikrofonieren. Das hat zur Folge, dass die Aufnahme den Gesamtklang in einem Raum abbildet. Das gewünschte Verhältnis zwischen Direktschall und Raumanteil wird dabei über den Abstand des Arrays zur Schallquelle eingestellt.

Der Abstand der Mikrofone eines Arrays zueinander sowie deren Richtcharakteristika verursachen Pegel- und Laufzeitunterschiede zwischen den einzelnen Spuren der Mehrkanalaufnahme. Inwiefern sich diese Unterschiede zusätzlich auf im Downmix auftretende Kammfiltereffekte auswirken, soll in dieser Arbeit ebenfalls untersucht werden.

Die dieser Arbeit zugrundeliegende Forschungsfrage lautet dementsprechend: Welchen Einfluss haben unterschiedliche 3D-Hauptmikrofonarrays sowie die bei Dolby Atmos und MPEG-H integrierten Downmixalgorithmen hinsichtlich auftretender Kammfiltereffekte auf das Downmixergebnis akustischer 3D-Musikaufnahmen?

#### 4. Methodik

Zunächst wurden mithilfe eines Testsignals verschiedene 3D-Mikrofonarrays nachgebildet. Um Rückschlüsse darüber ziehen zu können, inwiefern sich Pegel- und Laufzeitunterschiede zwischen den verschiedenen Mikrofonen des Arrays auf das Downmixergebnis auswirken, war es nötig, Arrays mit möglichst großen Unterschieden betreffend Mikrofonrichtcharakteristika sowie -Abständen zu wählen. Folgende drei Setups, welche in den nächsten Kapiteln näher beschrieben werden, wurden verglichen: PCMA-3D, Decca Cuboid und ORTF-3D.

Um mit den Arrays erzeugte Signale nachzubilden, wurde jeweils für eine oder mehrere Schalleinfallrichtungen die Laufzeit des Schalls zu jeder Mikrofonkapsel des Arrays sowie der aus der jeweiligen Richtcharakteristik des Mikrofons resultierende Pegelabfall mithilfe von dafür angelegten Matlab-Skripten (s. elektronischer Anhang) ausgerechnet. Der entfernungsabhängige Pegelabfall wurde nicht in das Modell mit einbezogen. Unter der Annahme, dass sich eine Schallwelle ab einem Abstand von  $r > 2\lambda$ , was bei dem gewählten Abstand der Schallquelle zum Mittelpunkt des Arrays von fünf Metern ab einer Frequenz von ca. 150 Hz zutrifft, als Linienwelle ausbreitet und der entfernungsabhängige Pegelabfall damit (-)3 dB pro Entfernungsverdopplung beträgt, liegt dieser beim größten verglichenen Array unter 2 dB. Damit sind die Werte für die Ausprägung der Kammfilterstruktur im Downmix nicht relevant. Zudem wird durch das Außer-Acht-Lassen des entfernungsabhängigen Pegelabfalls die Entfernung der Schallquelle zum Array aus der Betrachtung ausgeschlossen. Um dies auch für die Delaywerte zu erreichen, wurde hier jeweils mit den Laufzeitunterschieden zwischen den Mikrofonen anstatt den absoluten Werte von jedem Mikrofon zur Schallquelle gearbeitet.

Außerdem wurden die Frequenzabhängigkeit der Richtcharakteristika und Reflexionen im Raum nicht in die Betrachtung inkludiert. So wird der größtmögliche Pegelunterschied zwischen den einzelnen Kanälen erreicht.

Diese Berechnungen wurden für jedes Hauptmikrofonarray für eine Schalleinfallrichtung von  $0^\circ$  und für eine Schalleinfallrichtung von  $135^\circ$  links, um seitlichen und schräg von hinten kommenden Schall zu simulieren, durchgeführt. Während mit der Schalleinfallrichtung von  $0^\circ$  eine Schallquelle simuliert werden sollte, wurden im Kontext von akustischen Aufnahmen aus einer Schalleinfallrichtung von  $135^\circ$  Reflexionen modelliert. Bei allen Modellen lag das

simulierte Schallereignis dabei auf einer Höhe mit dem Middle Layer des Arrays. Dies führt zu maximalen Laufzeiten zwischen der Quelle und den einzelnen Mikrofonkapseln des Middle Layers.

Da die vorliegende Arbeit das Verhalten der Algorithmen in Hinblick auf akustische Musikaufnahmen untersuchen soll, wird das Downmixverhalten des LFE-Kanals außer Acht gelassen. Deswegen und da alle verwendeten Hauptmikrofonarrays aus höchstens neun Mikrofonen bestehen, handelt es sich beim Ausgangsformat um 5.0.4.

Um die Mehrkanalaufnahmen zu simulieren, wurde zuerst mit dem Pro Tools internen Signal Generator in einer separaten DAW-Session ein Testsignal generiert und aufgenommen. Als Testsignal diente weißes Rauschen. Als ein Signal, in dem alle hörbaren Frequenzen mit gleicher Leistung vertreten sind, ist es geeignet, um damit Kammfiltereffekte zu detektieren. Anschließend wurde pro Downmixverfahren für jedes Array eine DAW-Session angelegt, in welche das Audiofile auf 9 Spuren gelegt wurde. Jede Spur war dabei jeweils für einen Kanal des zu vergleichenden Arrays vorgesehen und wurde in Pegel und Delay dementsprechend angepasst. Für den genauen Aufbau der DAW-Sessions sei an dieser Stelle auf das nächste Kapitel verwiesen.

Die nachgebildeten Array-Signale wurden jeweils mit den beiden Verfahren einmal zu Surround (5.0) und Stereo (2.0) downgemixt. Verglichen wurden dabei jeweils die Resultate von einem Downmix des Dolby Beds, einem Downmix von Dolby Objekten, einem Downmix des MPEG-H Beds sowie einem Downmix von MPEG-H Objekten. Alle Downmixvarianten wurden mit den Default-Einstellungen benutzt.

## **5. Aufbau der DAW-Projekte**

Als DAW wurde bei der Arbeit mit Dolby Atmos Pro Tools Studio (Version 2022.5.0) von Avid verwendet und mit MPEG-H Nuendo 11 von Steinberg. Durch die gute Integration von Dolby Atmos in Pro Tools wurde sich hier für die DAW von Avid entschieden. Beim Routing des bei MPEG-H erforderlichen Signalwegs traten in Pro Tools Probleme auf, weshalb hier auf Nuendo ausgewichen wurde. Da es für die Downmixergebnisse unerheblich sein sollte, welche DAW verwendet wird, war die Nutzung zwei verschiedener DAWs unproblematisch.

Aus Gründen der Übersichtlichkeit wurde für jedes Array und jede Schallquellenrichtung jeweils eine neue Session angelegt (s. elektronischer Anhang).

### **5.1. Dolby Atmos mit Avid Pro Tools**

Zunächst wurden neun Spuren angelegt und entsprechend der bei Dolby Atmos üblichen Konventionen (s. Tabelle eins) benannt. Die jeweiligen Pegel- und Laufzeitanpassungen für die verschiedenen Kanäle der Arrays wurden über die Pro Tools nativen Plug-ins „Time Adjuster“ und „Trim“ vorgenommen.

Die Downmixe des Dolby Beds und der Dolby Objects wurden jeweils in einer Session angefertigt. Dafür wurde für jede Spur jeweils auf das Bed geroutet sowie einem Objekt zugewiesen. Durch das Halten der Alt-Taste und Klicken auf den Object-Button einer beliebigen Spur konnten so alle Spuren gleichzeitig zwischen Bed und Objekten umgeschaltet werden. Die Spuren wurden alle entsprechend gepant, sodass bei der Wiedergabe im Ausgangsformat 5.0.4 jedes Signal nur von dem zugehörigen Lautsprecher wiedergegeben wird.

Für die Produktion mit Dolby Atmos ist ein Dolby Atmos Renderer von Nöten. Da es unter Windows nicht möglich ist die DAW-Session und den Renderer auf einem Endgerät zu nutzen, wurde der Renderer auf einem externen Rechner genutzt. Genaue Informationen bezüglich des Routings zwischen DAW und Renderer können dem Diagramm in Anhang B.1 entnommen werden.

Um die Downmixe aufzeichnen und exportieren zu können, müssen der Renderer und die DAW über einen Timecode synchronisiert werden. Da der LTC Timecode Generator von Dolby für Windows nicht verfügbar ist, wurde mithilfe einer Software ein Timecode generiert und als Audiofile exportiert. Dieses wurde auf einer

separaten Spur abgelegt und dieser wurde ein Objekt zugewiesen. Anschließend wurde dieses Objekt in den Einstellungen des Dolby Atmos Renderers als Timecode-Quelle definiert.

## **5.2. MPEG-H mit Steinberg Nuendo**

Auch in der Nuendo Session wurden zunächst neun Spuren angelegt, welche entsprechend der MPEG-H Konvektionen (s. Tabelle zwei) benannt und auf einen Bus geroutet wurden, in welchem das MHAPi Plug-in insertiert war. Für eine genaue Anleitung zur Einbindung des MHAPi Plug-ins sei auf die vom Fraunhofer IIS herausgegebene Bedienungsanleitung zur MPEG-H Authoring Suite verwiesen.

Die für die einzelnen Arrays nötigen Pegelanpassungen wurden in dieser Session über den Fader des jeweiligen Kanals geregelt. Für die Delaywerte wurde ein im Track-Inspector integriertes Tool verwendet.

Das Routing der Signale über das Bed bzw. als Objekte wird im MHAPi Plug-in vorgenommen. Das Panning der Objekte erfolgt ebenso im Plug-in. Diese wurden dabei so gepant, dass im Ausgangsformat 5.0.4 jeder Kanal des Arrays direkt auf dem entsprechenden Lautsprecher liegt. Um die Downmixe anzuhören, muss in der Monitoring-Einheit des Plug-ins das entsprechende Setup ausgewählt werden. Da bei der Anwahl des Formats 5.0 Routingfehler auftraten, wurde an dieser Stelle Format 5.1 gewählt.

Auch der Export der Dateien kann über das Plug-in erfolgen. Anders als bei Dolby Atmos gibt es hier nicht die Möglichkeit die Downmixe zu exportieren. Daher wurde der Ausgang der DAW über Dante auf einen weiteren Rechner geroutet und dort die Einzelspuren aufgezeichnet. Dies war essenziell, um die Frequenzspektren der einzelnen Spuren eines Downmixes analysieren zu können. In Anhang B.2 befindet sich ein Blockdiagramm des vorgenommenen Routings.

## 6. Verwendete Arrays

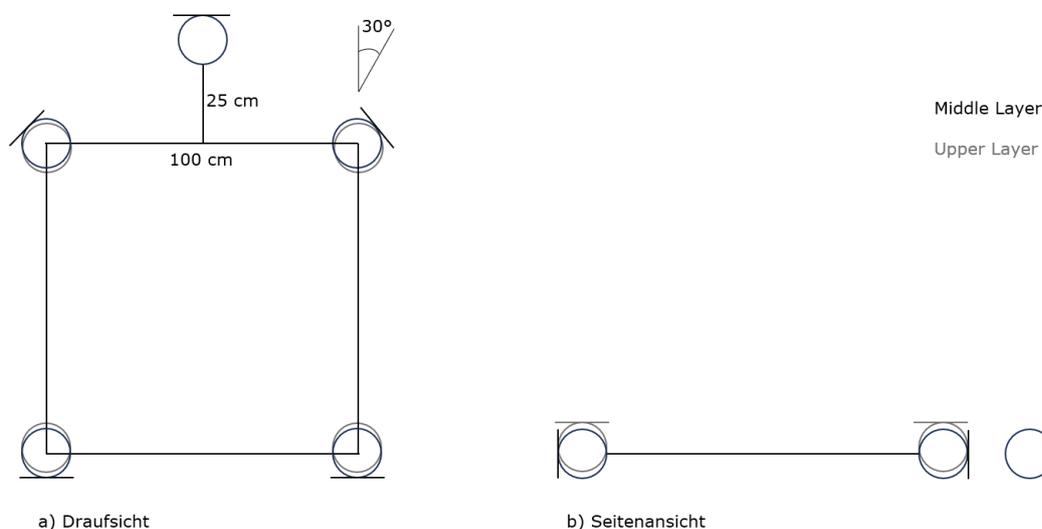
Im Folgenden werden die verglichenen Arrays bezüglich der Anordnung und Richtcharakteristika ihrer Mikrofone beschrieben.

### 6.1. PCMA-3D

Das PCMA-3D ist eine Erweiterung des von Lee entwickelten Perspective Control Microphone Arrays und besteht aus vier Paaren jeweils koinzident angeordneter Mikrofone sowie einem Center-Mikrofon. Die vier Paare bilden dabei ein Quadrat mit der Seitenlänge eines Meters. Jedes Paar setzt sich zusammen aus einem Mikrofon mit Nierencharakteristik und einem Mikrofon mit der Richtcharakteristik einer Superniere. Die vier Supernieren werden dabei nach oben gerichtet und bilden den Upper Layer. Der Middle Layer besteht aus den übrigen vier Mikrofonen und dem Center. Das Center-Mikrofon befindet sich mittig zwischen dem linken und rechten Mikrofon, wobei der Abstand von der Basis des Arrays zum Mikrofon 25 cm beträgt. Das linke und rechte Mikrofon (FL und FR) sind jeweils um  $30^\circ$  nach außen gedreht. Die jeweils gegenüberliegenden Mikrofone mit Nierencharakteristik (RL und RR) zeigen in die entgegengesetzte Richtung. (H. Lee & Johnson, 2019) Bei einer vor dem Hauptmikrofonarray liegenden Schallquelle sind diese Mikrofone also nach hinten gerichtet, während FL, FR und FC auf die Schallquelle gerichtet werden. Abbildungen sechs bis acht veranschaulichen die Anordnung der Mikrofone im Middle und Upper Layer.

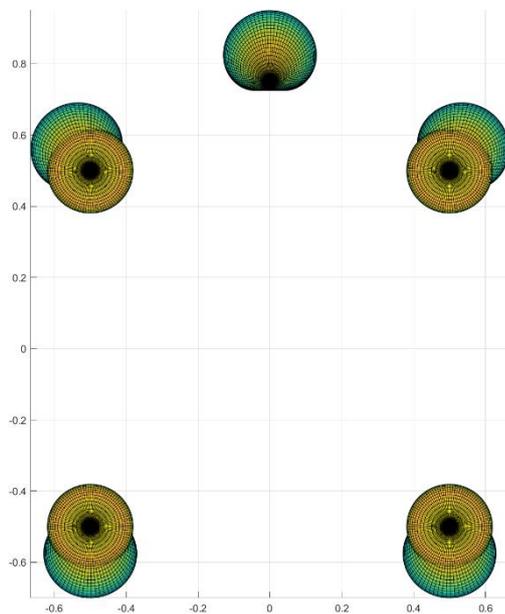
#### Abbildung 6

Mikrofonanordnung im PCMA-3D



**Abbildung 7**

*PCMA-3D: Darstellung der Richtcharakteristika, Draufsicht*

**Abbildung 8**

*PCMA-3D: Darstellung der Richtcharakteristika, Seitenansicht*



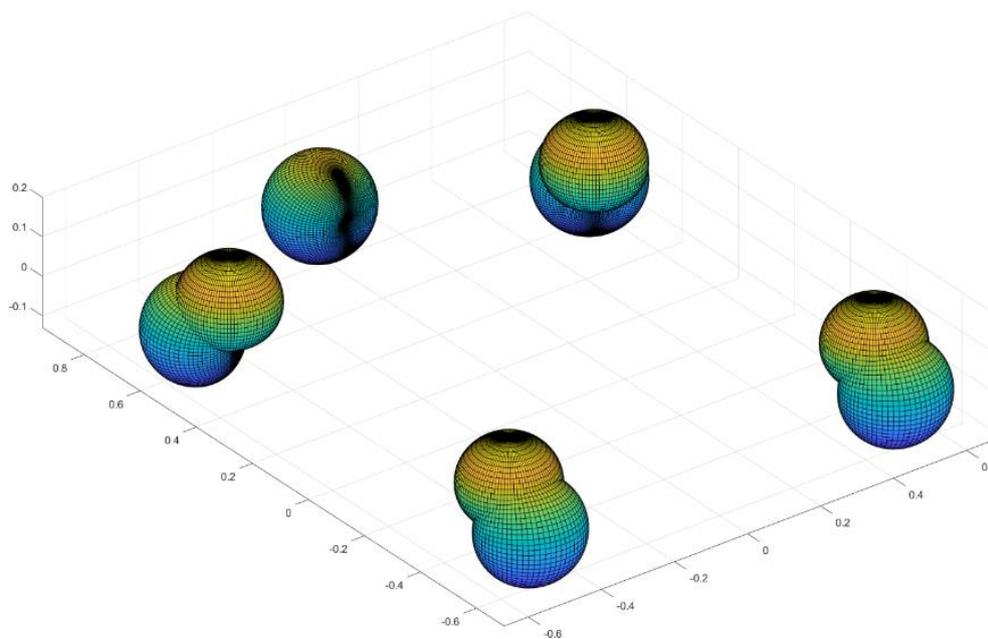
Durch die Anordnung der Mikrofone ergibt sich ein vertikal koinzidentes Array. Dieser Ansatz beruht auf folgenden Annahmen: Zum einen trägt ein Abstand zwischen dem Middle Layer und dem Upper Layer nicht maßgeblich zum wahrgenommenen räumlichen Eindruck bei einer 3D-Wiedergabe bei (H. K. Lee & Gibben, 2015). Zum anderen wurde bei einem Versuch mit bandbegrenzten Testsignalen festgestellt, dass die Lokalisation einer Schallquelle in der vertikalen Ebene weniger durch Laufzeitdifferenz zwischen den Kanälen als dem sog. pitch-height-Effekt bestimmt wird (Wallis & Lee, 2015). Um eine Verlagerung der Phantomschallquelle nach oben zu verhindern, ist sicherzustellen, dass das Level des vom Mikrophon aufgenommenen Direktschalls im Upper Layer jeweils um

mindestens 9,5 dB geringer ist als das beim entsprechenden Mikrofon des Middle Layers (Wallis & Lee, 2017). Diese Prämisse ist erfüllt, indem die Mikrofone des Upper Layers nach oben zeigen und hier Supernieren zur Verwendung kommen (H. Lee, 2021). Durch den Abstand von einem Meter zwischen den vorderen und hinteren Mikrofonen des Arrays und dadurch, dass die hinteren Mikrofone ebenfalls von der Schallquelle weg zeigen, erzeugen außerdem FL und RL sowie FR und RR voneinander dekorrelierte Signale (H. Lee & Johnson, 2019).

Abbildung neun zeigt das Array mit den entsprechenden Richtcharakteristika im Raum.

### Abbildung 9

*PCMA-3D: 3D-Plot des Arrays*



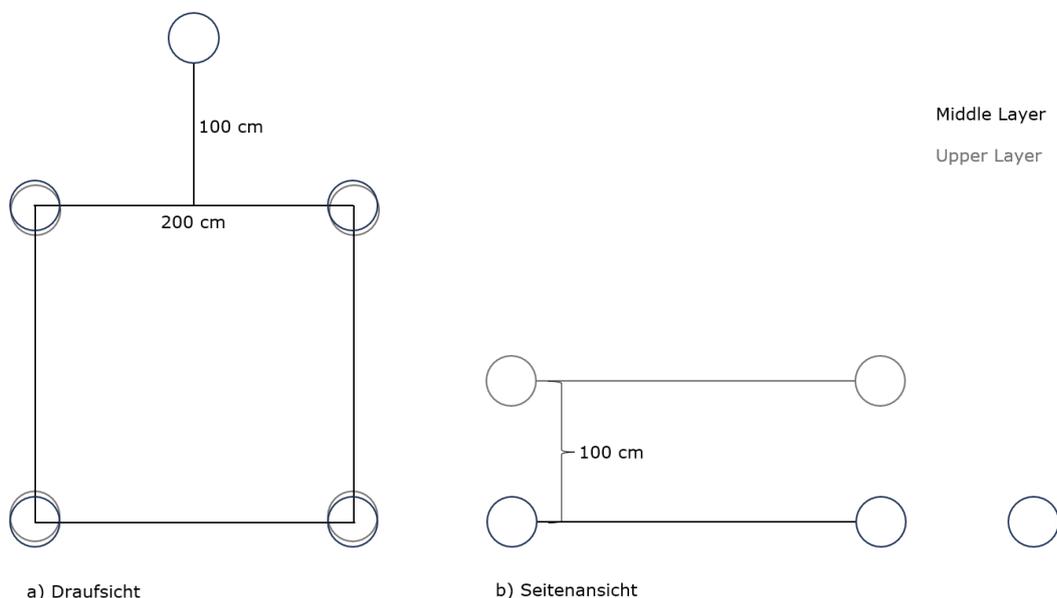
## 6.2. Decca Cuboid

Als zweites Array wurde der Decca Cuboid nachgestellt. Dabei handelt es sich um eine Erweiterung des Decca Trees. Dieser wurde in den 1950er Jahren bei Decca Records entwickelt und besteht aus drei ungerichteten Mikrofonen. Dadurch entsteht das Stereobild hauptsächlich durch Laufzeitdifferenzen zwischen den Kanälen. Der Abstand zwischen dem linken und dem rechten Mikrofon beträgt

dabei 1,5 m bis 2 m. Der Center befindet sich in der Mitte und ist zwischen einem und anderthalb Meter von der Basis entfernt. Durch die großen Abstände zwischen den Mikrofonen wird eine hohe Dekorrelation der Signale bis in tiefe Frequenzen erreicht. Dieser Decca Tree wird um sechs weitere Mikrofone ergänzt. Zwei davon befinden sich auf einer Ebene mit dem ursprünglichen Tree und ergeben zusammen mit diesem den Middle Layer des 3D-Arrays. Dabei befindet sich jeweils in einer Linie mit dem linken und rechten Mikrofon ein weiteres Mikrofon mit Kugelcharakteristik. Diese sind jeweils 2m von den Front-Mikrofonen entfernt und spannen so ein Rechteck auf. Der Upper Layer befindet sich einen Meter über dem Middle Layer. Auch dieser besteht ausschließlich aus ungerichteten Mikrofonen, welche ein Rechteck derselben Größe wie die des Middle Layers aufspannen. (H. Lee & Johnson, 2019) Die folgenden Abbildungen veranschaulichen das beschriebene Array:

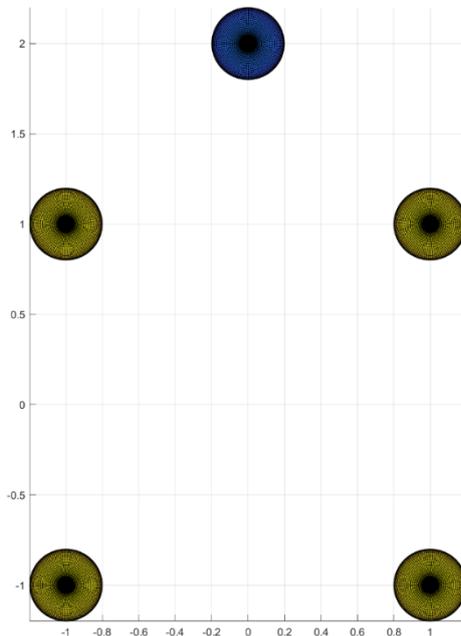
### Abbildung 10

#### Mikrofonanordnungen im Decca Cuboid

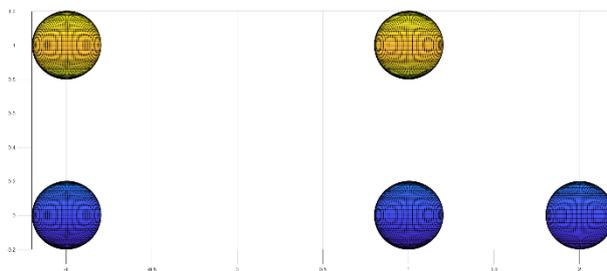


**Abbildung 11**

*Decca Cuboid: Darstellung der Richtcharakteristika, Draufsicht*

**Abbildung 12**

*Decca Cuboid: Darstellung der Richtcharakteristika, Seitenansicht*

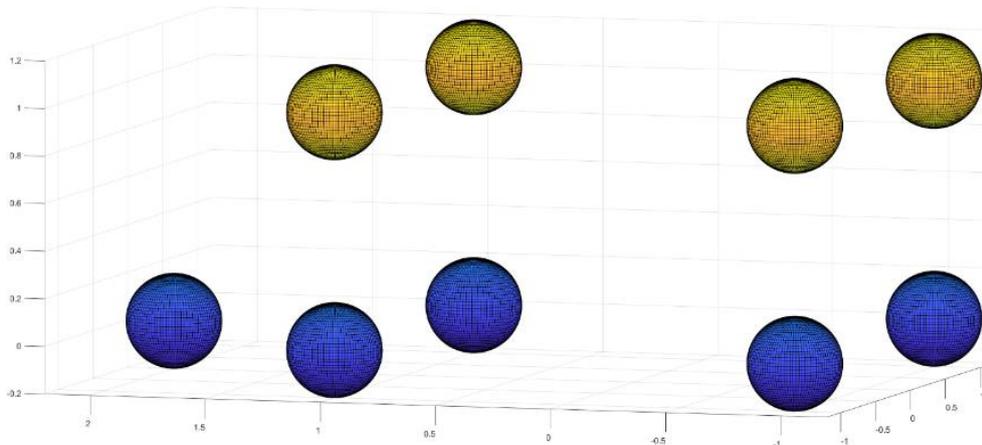


Der Decca Cube beschreibt, bezüglich Laufzeit- und Pegeldifferenzen zwischen den Mikrofonen, jeweils ein Extrem in der Vergleichsgruppe. Mit dem großen Abstand zwischen den Mikrofonkapseln kann hier der größte Laufzeitunterschied unter den verglichenen Arrays erreicht werden. Durch die Verwendung von ungerichteten Mikrofonen sind die Pegelunterschiede hier am geringsten.

Abbildung 13 veranschaulicht das Array und die Richtcharakteristika im Raum.

**Abbildung 13**

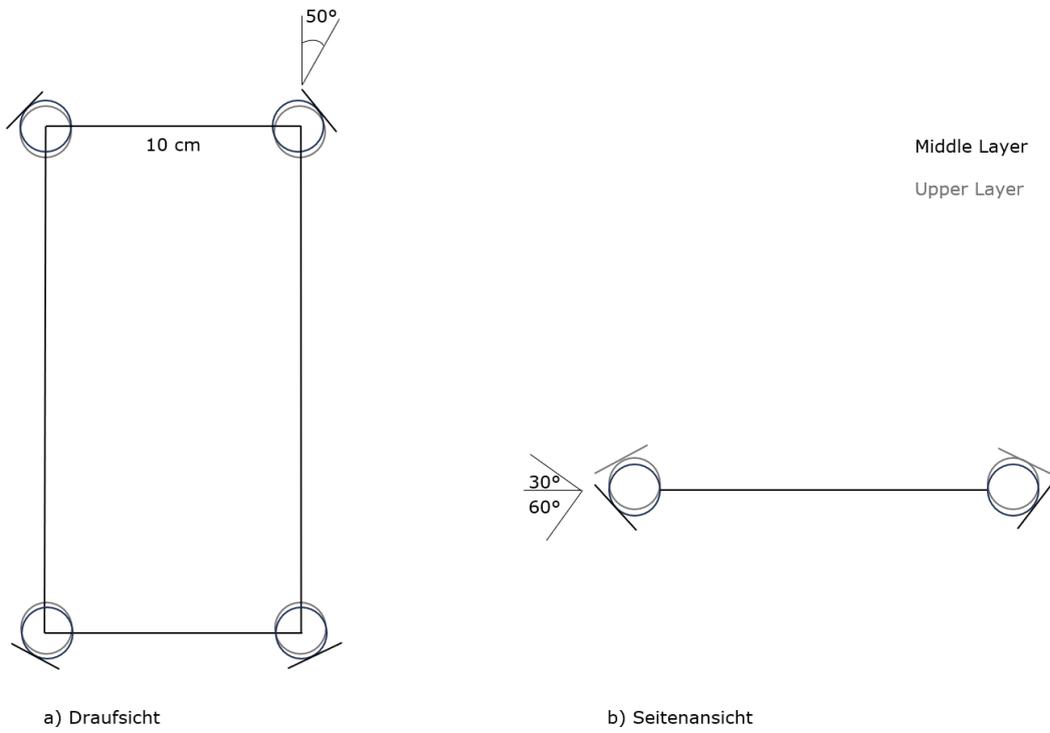
*Decca Cuboid: 3D-Plot des Arrays*

**6.3. ORTF 3D**

Beim letzten verwendeten Setup handelt es sich um das ORTF-3D. Diesem liegt ein Prinzip zugrunde, welches dem der gleichnamigen Stereo-Mikrofonierungstechnik ähnlich ist. Beim ORTF-3D werden acht Mikrofone mit der Richtcharakteristik einer Superniere verwendet. Dabei bilden jeweils zwei Mikrofone ein koinzident angeordnetes Paar. Die Mikrofone des Middle Layers sind jeweils um  $50^\circ$  nach außen gedreht und der Winkel zwischen den oberen und unteren Mikrofonen beträgt jeweils  $90^\circ$ . Dadurch und durch die Verwendung von Supernieren, wird eine gute Separation der Kanäle erreicht. Es gibt verschiedene Ausführungen des Arrays, welche sich jeweils in den Abständen zwischen den Mikrofonen voneinander unterscheiden. (H. Lee, 2021; Wittek & Theile, 2017) In Abb. 14-16 wird das für diese Arbeit verwendete Setup dargestellt.

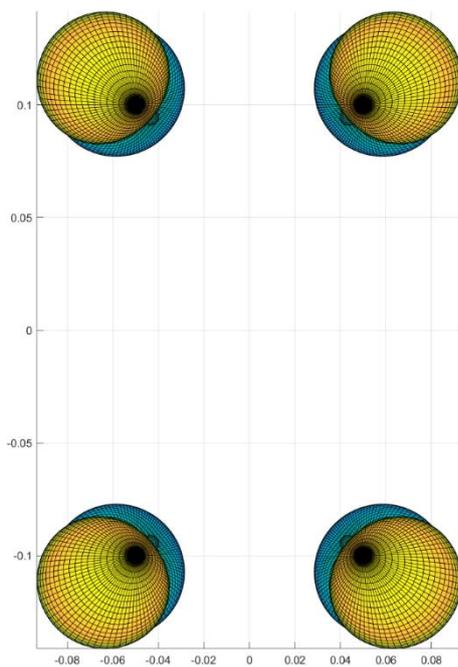
**Abbildung 14**

*Mikrofonanordnungen im ORTF-3D*



**Abbildung 15**

*ORTF-3D: Darstellung der Richtcharakteristika, Draufsicht*



**Abbildung 16**

*ORTF-3D: Darstellung der Richtcharakteristika, Seitenansicht*

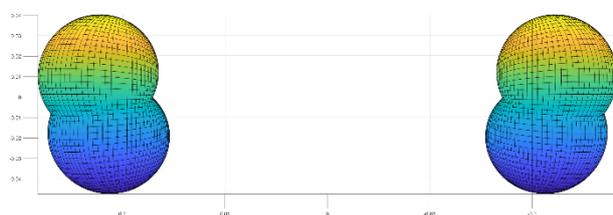
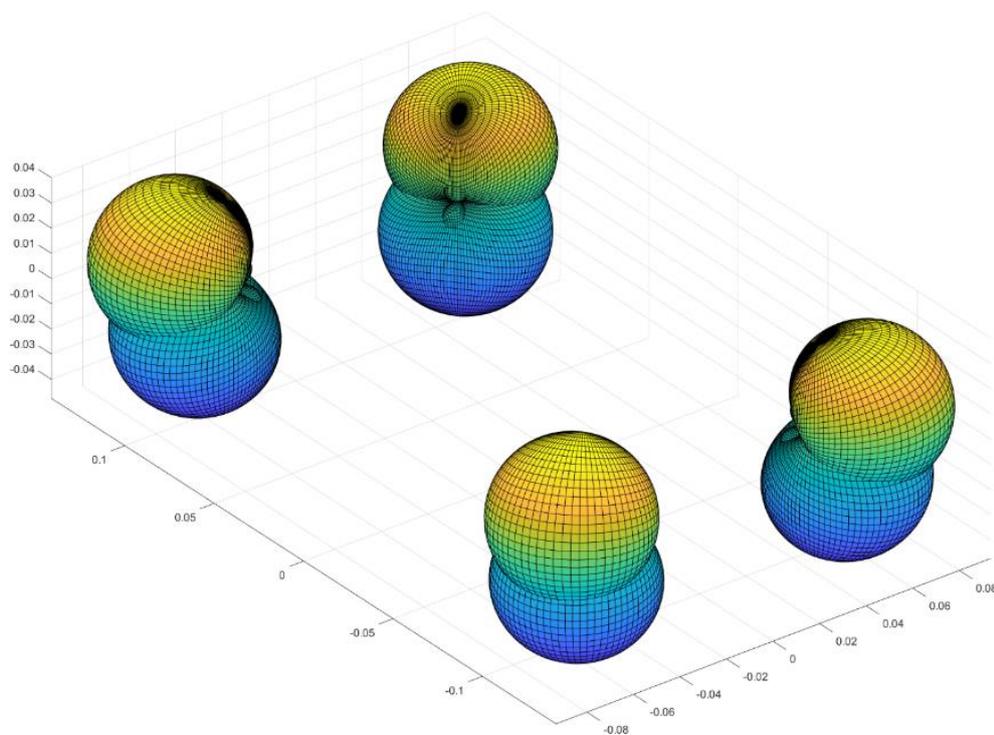


Abbildung 17 illustriert das Array mit den entsprechenden Richtcharakteristika der Mikrofone im Raum.

**Abbildung 17**

*ORTF-3D: 3D-Plot des Arrays*



## 7. Vergleich verschiedener Downmixe

Beim Vergleich der verschiedenen Downmix-Varianten wurden die Spuren der Arrays zunächst über das Dolby Atmos Bed geroutet und downgemixt. Anschließend wurde jeder Spur ein Dolby Atmos Object zugeordnet und das Downmixverhalten dieser Objekte untersucht. Um eine bessere Abgrenzung der Ergebnisse zu erhalten, wurde auf eine Mischung zwischen Dolby Atmos Bed und Dolby Atmos Objects verzichtet. Als letzter Schritt wurden dieselben Spuren mit MPEG-H jeweils als Channel-Bed und im Anschluss als Objekte downgemixt.

Ausgewählte Hörbeispiele sind im elektronischen Anhang zu finden.

### 7.1. Downmix Dolby Atmos Bed

Beim Downmix des Dolby Atmos Beds wurde aufgrund der Tatsache, dass hier im Upper Layer (UL) lediglich ein Stereopaar vorgesehen ist, jeweils noch zwischen verschiedenen Signalkombinationen der oberen vier Signale unterschieden. Dabei wurden einmal alle vier Signale verwendet und jeweils einmal die vorderen (Ltf und Rtf) bzw. hinteren (Ltr und Rtr) zwei Signale. Hier ist anzumerken, dass bei einer Verwendung aller vorhandenen Signale diese jeweils auf der linken und rechten Seite summiert werden. So ergeben sich die Signale  $TL = Ltf + Ltr$  und  $Tr = Rtf + Rtr$ . Mit überkreuzten Signalen (Ltf und Rtr oder Ltr und Rtf) wurde nicht gearbeitet. Da diese sich in jedem untersuchten Fall in ihrem Delaywert unterscheiden, ist hier bei der Wiedergabe im Format 5.0.4 im UL und damit in jedem Downmix mit Kammfiltereffekten zu rechnen, welche sich zudem auf der linken und rechten Seite unterscheiden.

#### PCMA-3D

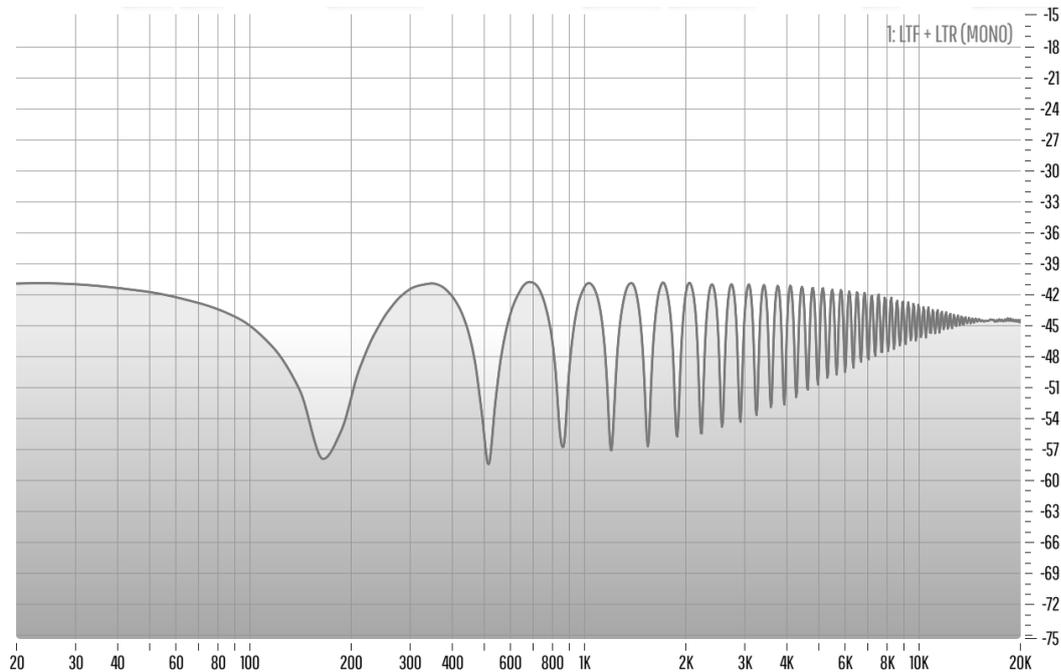
Als erstes wurde die Simulation des PCMA-3D mit einer Schalleinfallrichtung von  $0^\circ$  downgemixt. Bei der Verwendung aller vier Signale im Upper Layer kommt es bereits vor dem Downmix im Ausgangsformat 5.0.4 zu Kammfiltereffekten im Upper Layer. Diese resultieren aus der Summierung von Ltf und Ltr sowie Rtf und Rtr und der Verzögerung zwischen diesen beiden Kanälen um jeweils 140 Samples.

Abbildung 18 zeigt die spektrale Analyse des summierten Signals auf der linken Seite des Upper Layers. Die rechte Seite weist aufgrund der Übereinstimmung von Pegel- und Delaywerten eine identische Kammfilterstruktur auf. Um diese und folgende Abbildungen bezüglich der Kammfilterstruktur miteinander vergleichen

zu können, wurde der Offset der Kurven jeweils so eingestellt, dass diese in der Mitte des Graphen liegen.

### Abbildung 18

*Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim PCMA-3D bei einem Schalleinfallswinkel von 0°*

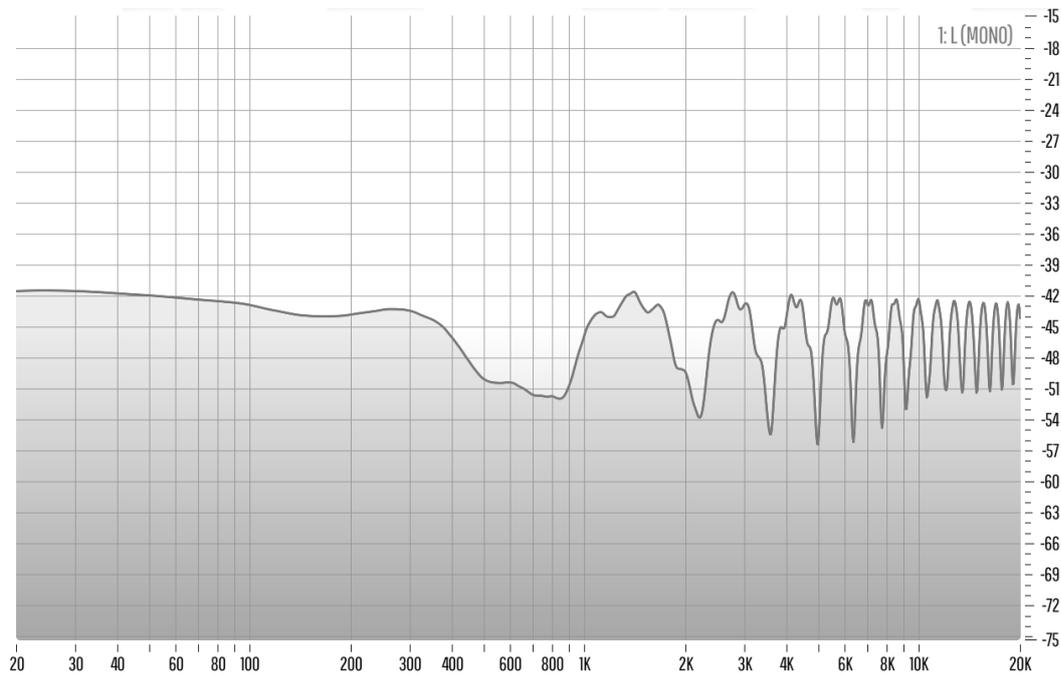


Dieser Kammfilter wird im 5.0-Downmix auf die Surround-Kanäle verlagert. Wie in Kapitel 2.2 erläutert, werden die Kanäle des Upper Layers zunächst mit reduziertem Pegel zu Lss und Rss summiert. Da Lss und Rss aufgrund der Simulation im Freifeld einen geringen Pegel von -53,75 dB aufweisen, ändert sich die Kammfilterstruktur nicht. Im Stereodownmix werden die Surroundkanäle sowie der Centerkanal jeweils mit verringertem Pegel zu L und R addiert. Abbildung 19 zeigt das resultierende Frequenzspektrum.

Der Unterschied, zwischen den in den Abbildungen 18 und 19 dargestellten Frequenzspektren lässt sich durch den Laufzeitunterschied zwischen C und L sowie C und R von jeweils 39 Samples erklären.

**Abbildung 19**

*Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung des Dolby Atmos Bed*



Bereits vor dem Downmix entstehende Kammfilter im UL können vermieden werden, indem nur zwei der vier vorhandenen Kanäle im Upper Layer verwendet werden. Im betrachteten Model weisen alle vier Signale denselben Pegel auf. Hinsichtlich des Delaywertes stimmen jeweils Ltf und Rtf sowie Ltr und Rtr überein. Der Laufzeitunterschied zwischen vorne und hinten beträgt jeweils 140 Samples. Aufgrund des geringen Pegels der Surroundkanäle im Ausgangsformat kommt es im 5.0-Downmix weder bei Verwendung der vorderen Signale (Ltf und Rtf) noch bei Verwendung der hinteren Signale des UL (Ltr und Rtr) zu Kammfiltereffekten auf den Surround Kanälen.

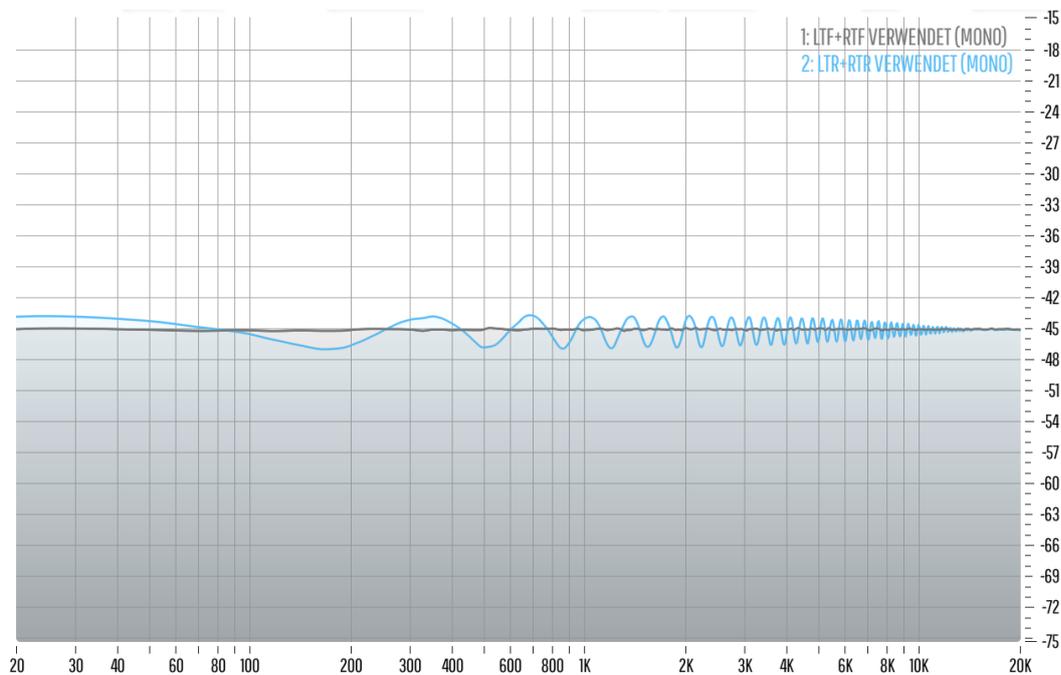
Im Stereodownmix ergeben sich bei Verwendung der hinteren Kanäle des Upper Layers im Ausgangsformat ein Kammfilter durch das Delay zwischen den vorderen Kanälen des Middle Layer (ML) und Ltr und Rtr. Durch die koinzidente Anordnung von L und Ltf bzw. R und Rtf führt die Verwendung der vorderen Kanäle des UL nicht zu einem Kammfilter. In beiden Fällen hat die Addition des Centers zum linken sowie rechten Kanal Auswirkungen auf das endgültige Frequenzspektrum. Abbildung 20 zeigt die spektrale Analyse des linken Kanals im Stereodownmix unter Verwendung der vorderen (graue Linie) bzw. hinteren Kanäle des UL (blaue

Linie) im Ausgangsformat, jeweils mit gemutetem Centerkanal. Abbildung 21 zeigt beide Szenarien mit hinzuaddiertem Center. Durch die mittig vor dem Array simulierte Schallquelle und die daraus resultierenden übereinstimmenden Delaywerte für die linke und rechte Seite des Arrays, verhält sich R analog zu L.

Der Vergleich von Abb.20 und Abb.21 zeigt, dass, unabhängig von den für den UL gewählten Signalen, die Kammfilterstruktur auf den Stereosignalen durch die Addition des Centers auf Links und Rechts bestimmt wird.

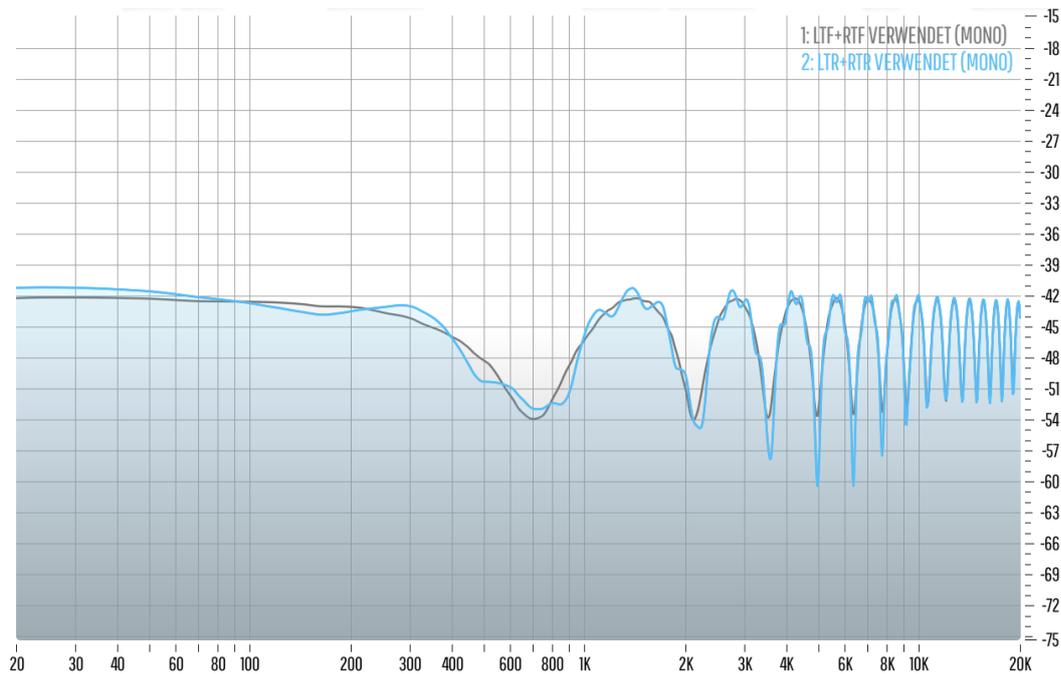
### Abbildung 20

*Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von unterschiedlichen Ausgangssignalen im Dolby Atmos Bed (ohne C)*



**Abbildung 21**

*Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von  $0^\circ$  unter Verwendung unterschiedlicher Ausgangssignalen im Dolby Atmos Bed (mit C)*



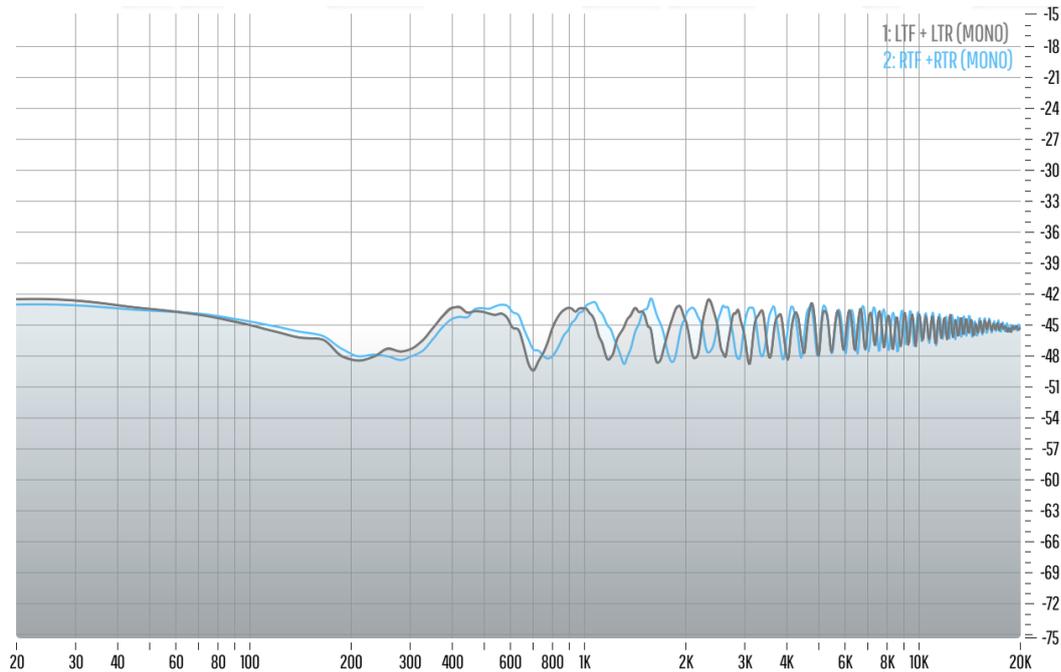
Bei der zweiten simulierten Schalleinfallrichtung ( $135^\circ$  links) führt die Verwendung aller vier Signale des Upper Layers erneut zu kammfilterähnlichen Strukturen auf den Kanälen Tl und Tr. Durch die seitliche Schallquellenposition und die daraus resultierenden unterschiedlichen Laufzeiten von der Quelle bis zu den Mikrofonkapseln des Arrays ergeben sich auf der linken und rechten Seite unterschiedliche Kammfilterstrukturen.

Abbildung 22 zeigt die spektrale Analyse der Kanäle Tl (graue Kurve) und Tr (blaue Kurve).

Wie beim vorherigen Modell übertragen sich diese Kammfilterstrukturen in die Downmixe. Die Ausbildung der Kammfilter im UL kann verhindert werden, indem nur die vorderen oder nur die hinteren Kanäle des UL im Ausgangsformat genutzt werden.

**Abbildung 22**

*Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim PCMA-3D bei einem Schalleinfallswinkel von 135°*

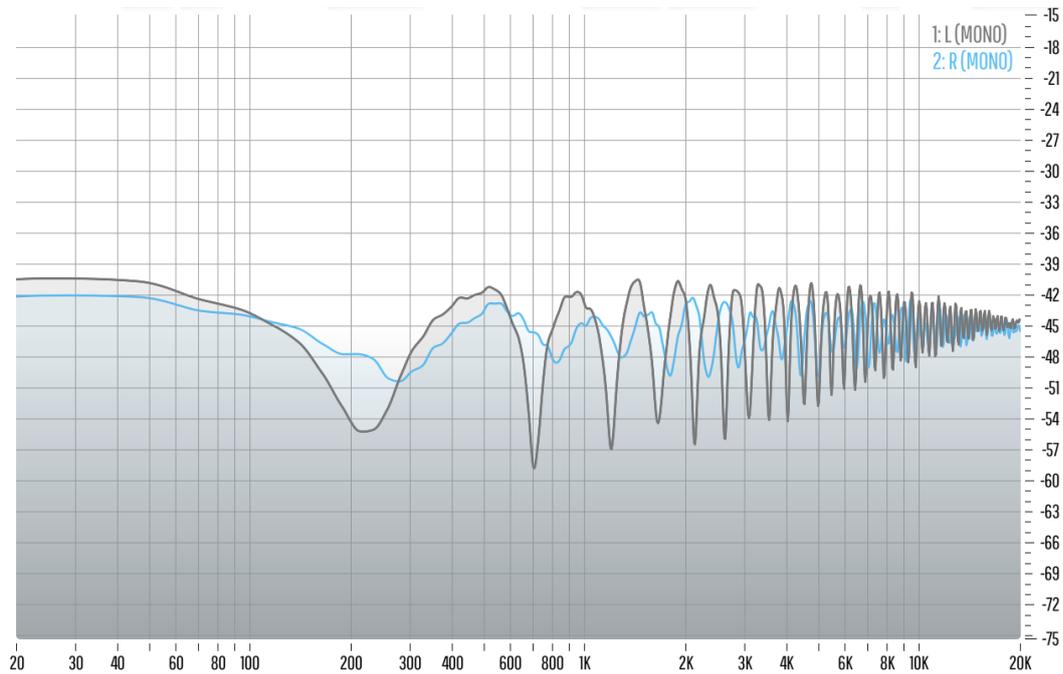


Im Gegensatz zur Simulation mit mittlerer Schallquellenposition weisen bei seitlichem Schalleinfallswinkel die Kanäle Lss und Rss einen höheren Pegel auf. Dadurch bilden sich im 5.0-Downmix Kammfiltereffekte auf Ls und Rs, falls diese mit einem Signal mit anderem Delaywert summiert werden. Dies ist der Fall, wenn im UL im Ausgangsformat die vorderen Signale (Ltf und Ltr) verwendet werden. Die sich ergebenden Kammfiltereffekte stimmen aufgrund der vertikal koinzidenten Anordnung der Mikrofone in ihrer Struktur mit den in Abb. 22 gezeigten Spektren überein. Bei der Verwendung der hinteren Signale des UL kommt es im 5.0-Downmix auf keinem Kanal zu Kammfiltereffekten. An dieser Stelle soll angemerkt werden, dass auch keine anderen Änderungen des Frequenzspektrums auftreten, d.h. die spektrale Analyse der Kanäle zeigt das Frequenzspektrum des Testsignals. Abweichungen von diesem Spektrum werden in jedem Fall beleuchtet. Dies gilt für das Gesamte Kapitel 7.

Die Ergebnisse des 5.0-Downmix wirken sich durch den stufenweisen Downmixalgorithmus auf den Stereodownmix aus. Zusätzlich wird das resultierende Frequenzspektrum der Stereokanäle auch hier durch die Addition des Centers auf L und R bestimmt. Die Abbildungen 23, 24 und 25 zeigen jeweils den linken und rechten Kanal des Stereodownmixes nach Verwendung unterschiedlicher Signale des UL im Ausgangsformat.

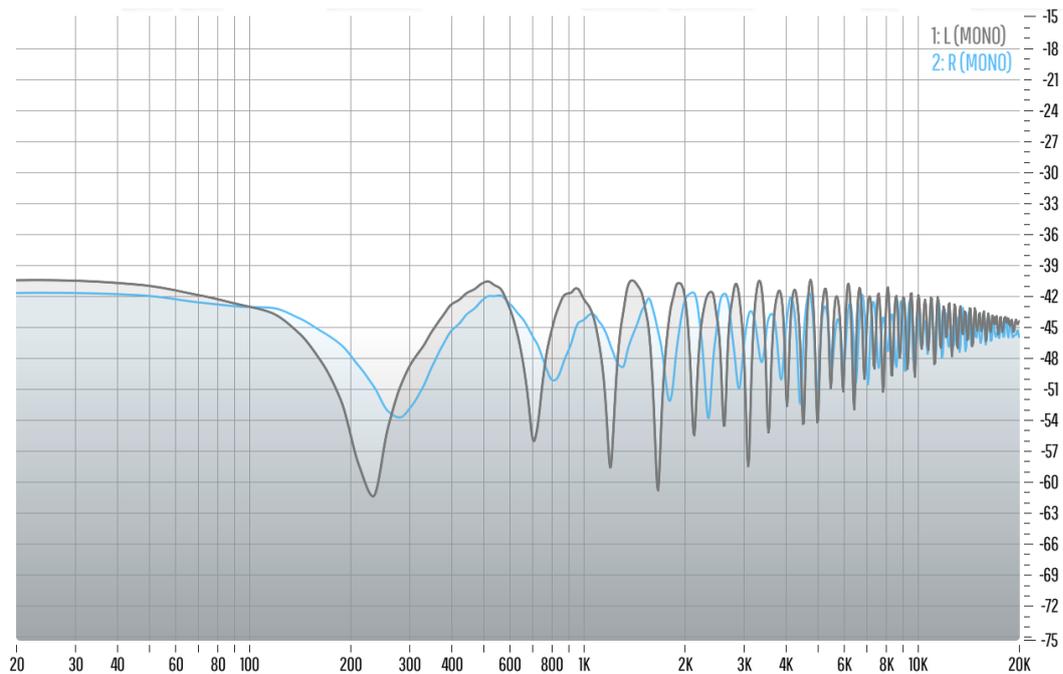
**Abbildung 23**

*Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung aller Signale des UL im Dolby Atmos Bed*



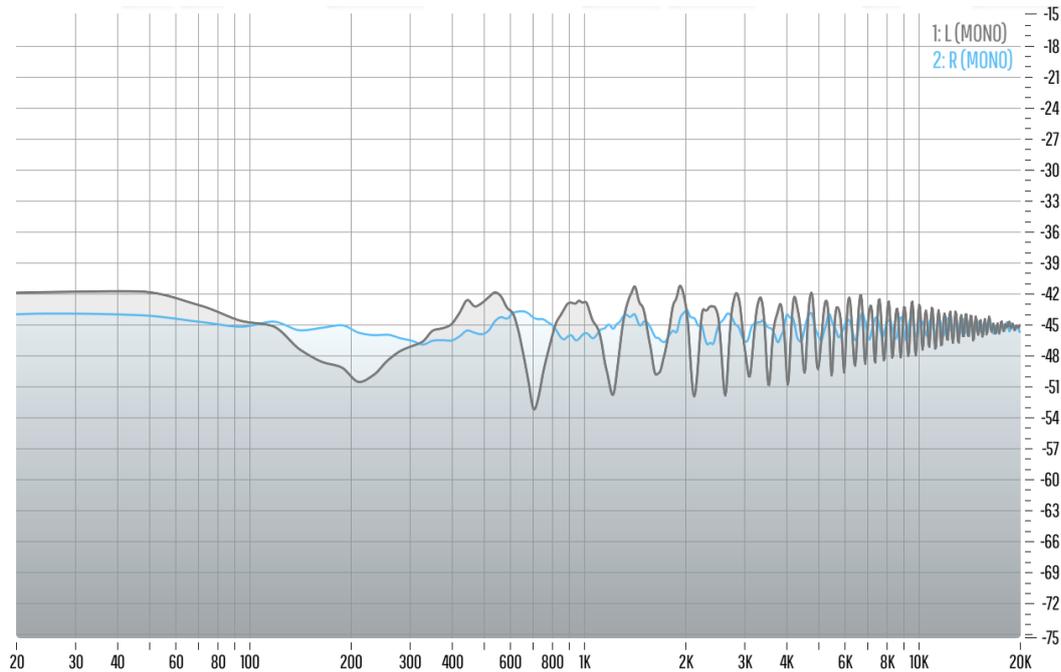
**Abbildung 24**

*Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung der Signale Ltf und Rtf des UL im Dolby Atmos Bed*



**Abbildung 25**

*Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von  $135^\circ$  unter Verwendung der Signale Ltr und Rtr des UL im Dolby Atmos Bed*



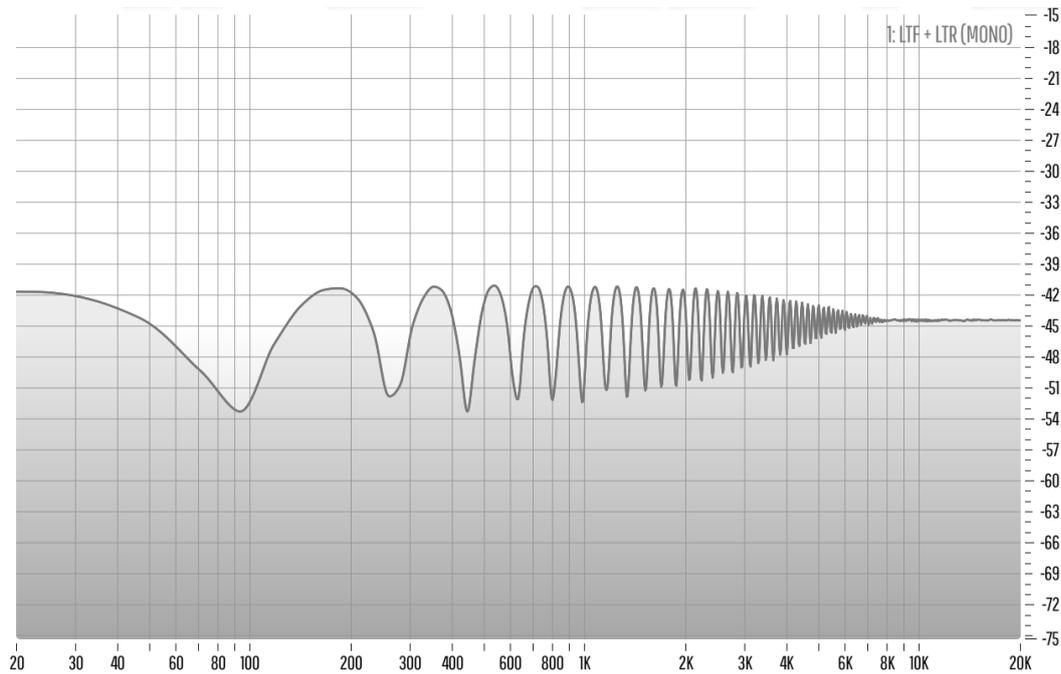
Beim Vergleich der Abbildungen fällt auf, dass sich die kammfilterähnlichen Strukturen hauptsächlich bezüglich des Pegelunterschieds zwischen von Interferenzen angehobenen und abgeschwächten Frequenzen unterscheiden.

**Decca**

Beim Downmix der Simulation des Decca Cuboids und einem Schalleinfallswinkel von  $0^\circ$  wiederholen sich einige Beobachtungen aus dem vorangegangenen Kapitel. Auch hier kommt es bei Verwendung aller vier Signale im Upper Layers bereits im Ausgangsformat zu Kammfiltereffekten im Upper Layer. Abbildung 26 zeigt den Kanal TI vor dem Downmix. Die dargestellte Kammfilterstruktur sowie alle weiteren Analysen des linken Kanals für dieses Modell gelten ebenso für rechts.

**Abbildung 26**

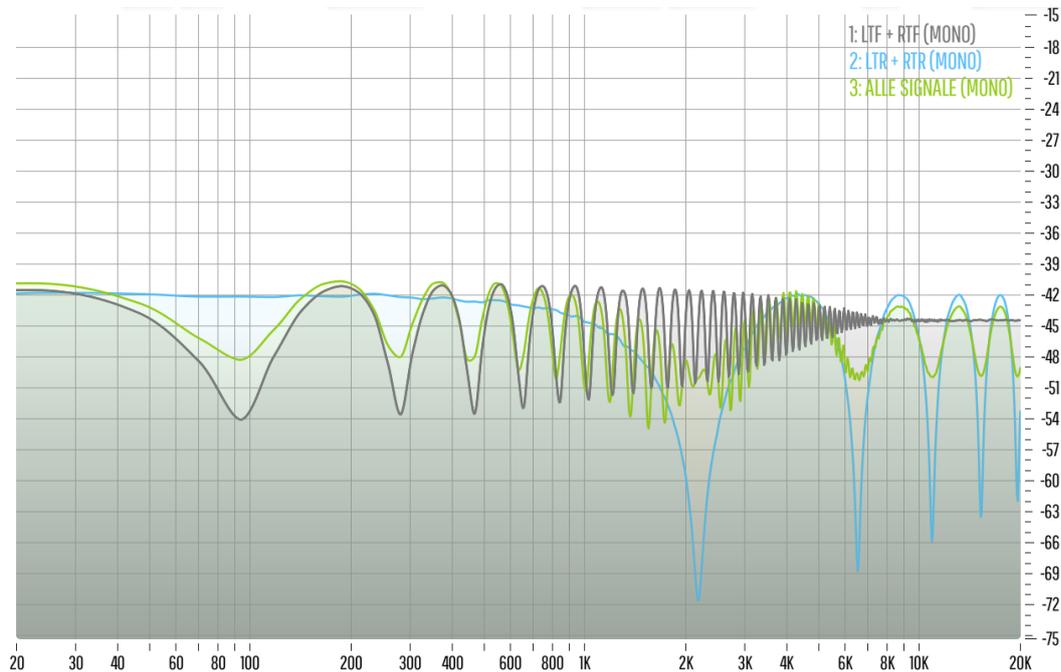
*Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim Decca Cuboid bei einem Schalleinfallswinkel von 0°*



Diese Kammfilterstruktur überträgt sich im 5.0-Downmix auf die Surround Kanäle. Da aufgrund der getroffenen Annahmen für die Simulationen und aufgrund der Verwendung ungerichteter Mikrofone alle Kanäle des Decca Cuboids denselben Pegel aufweisen, kommt es anders als beim PCMA-3D zu einer Veränderung des im UL entstehenden Kammfilters durch die Summation mit Ls bzw. Rs. Das heißt, anders als beim PCMA kann durch die Betrachtung des im UL entstehenden Kammfilters kein Rückschluss auf auftretende Kammfilter auf den Surroundsignalen im 5.0-Downmix gezogen werden. Außerdem kann durch die vertikale Trennung von Middle und Upper Layer und die dadurch -bei der gewählten Schallquellenposition- unterschiedlichen Laufzeiten ein Kammfilter auf den Surround Kanälen im 5.0-Downmix nicht durch die Verwendung bestimmter Signale des UL vermieden werden. Abbildung 27 zeigt den linken Surround Kanal im 5.0-Downmix nach Verwendung aller UL-Signale (grün), der vorderen UL-Signale (grau) und der hinteren UL-Signale (blau).

**Abbildung 27**

*Frequenzspektrum des Kanals Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 0° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds*



In Abb. 27 ist zu sehen, dass unter Beteiligung der vorderen Kanäle des Upper Layers der Kammfilter in tieferen Frequenzen einsetzt. Dies ist auf den Laufzeitunterschied zwischen den Surround Kanälen und den vorderen bzw. hinteren Kanälen des Upper Layers zurückzuführen. Die jeweiligen Delaywerte für alle Kanäle könne für diese und alle anderen Simulationen Anhang A entnommen werden.

Das Downmixergebnis des 5.0-Downmixes wirkt sich durch die Summierung von Ls und L sowie Rs und R auf den Stereodownmix aus. Zudem hat die Addition des Centerkanals auf den linken sowie rechten Kanal Auswirkungen auf die entstehenden Frequenzspektren. Abbildung 28 zeigt den linken Kanal des Stereodownmixes jeweils unter Verwendung aller Signale des UL (grün), der vorderen UL-Signale (grau) und der hinteren UL-Signale (blau).

**Abbildung 28**

*Frequenzspektrum des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 0° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds*

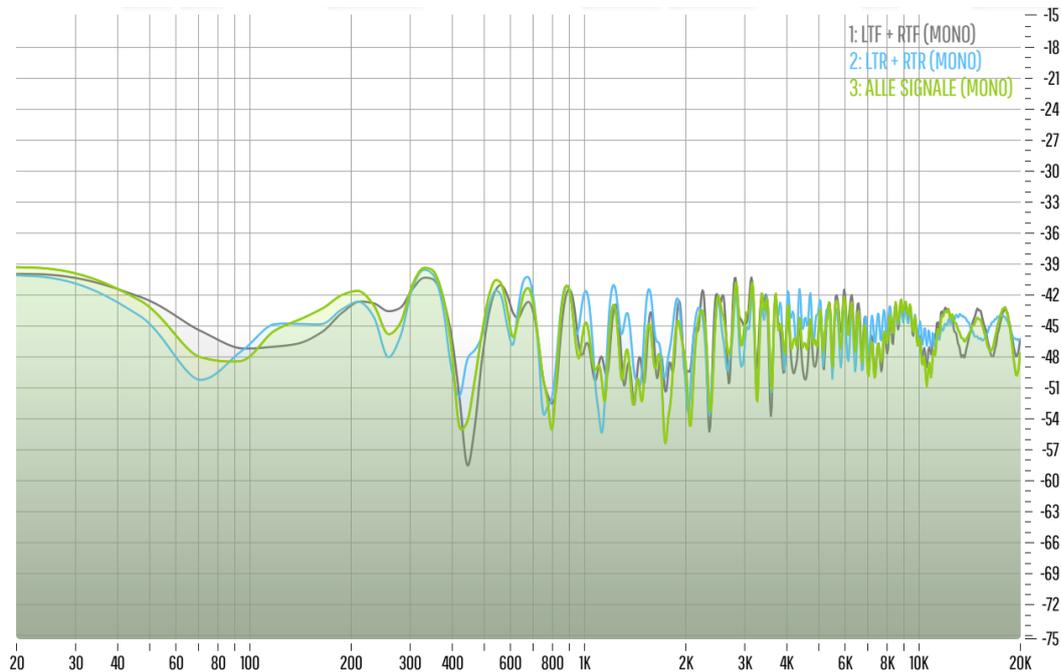
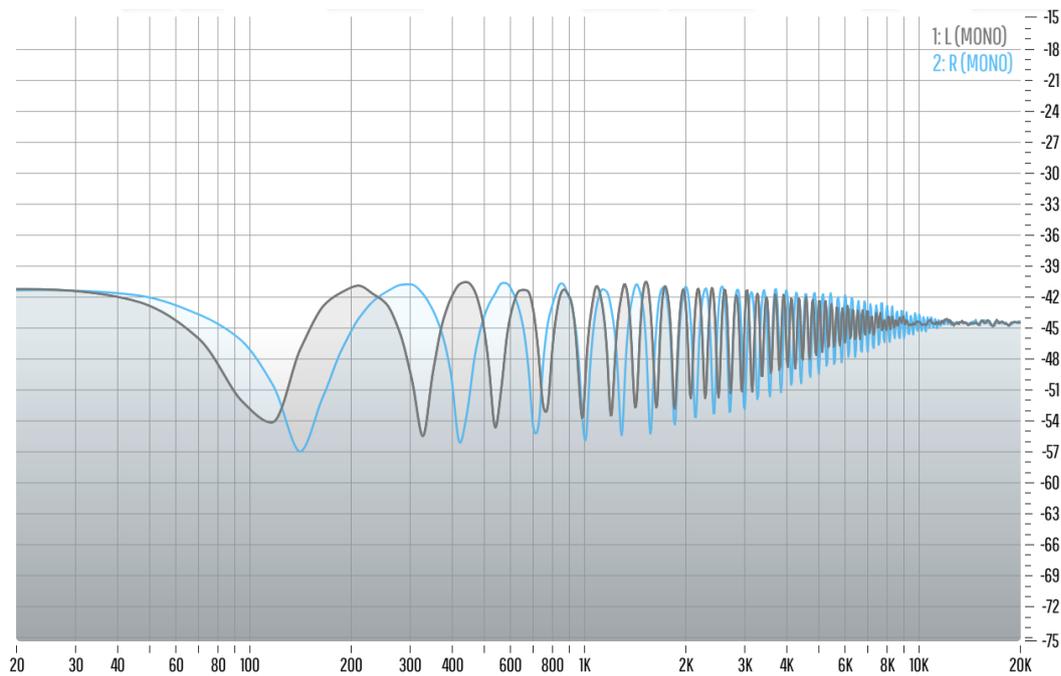


Abbildung 28 verbildlicht die durch den Downmix entstehenden destruktiven Interferenzen. Diese nehmen erst in hohen Frequenzen (ab 10 kHz) die für eine Kammfilter typische Struktur an. Genau wie aus einem Kammfilter resultieren aus diesem Frequenzspektrum klangliche Veränderungen. Diese in ihrer Qualität miteinander zu vergleichen ist nicht Gegenstand dieser Arbeit. Nichtsdestotrotz sind alle dargestellten Frequenzspektren in Bezug auf deren Auswirkungen auf den Gesamtklang zu betrachten.

Einige der beobachteten Ergebnisse wiederholen sich bei der Betrachtung der zweiten Schallquellenposition. Auch hier ergeben sich bei der Verwendung aller vier Signale im UL des Ausgangsformats Kammfiltereffekte auf den Kanälen Tl und Tr. Diese unterscheiden sich aufgrund der unterschiedlichen Delaywerte der Signale Ltf und Ltr sowie Rtf und Rtr auf der linken und rechten Seite. Abbildung 29 zeigt die Frequenzspektren der Kanäle Tl und Tr im Format 5.0.4.

**Abbildung 29**

*Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim Decca Cuboid bei einem Schalleinfallswinkel von 135°*



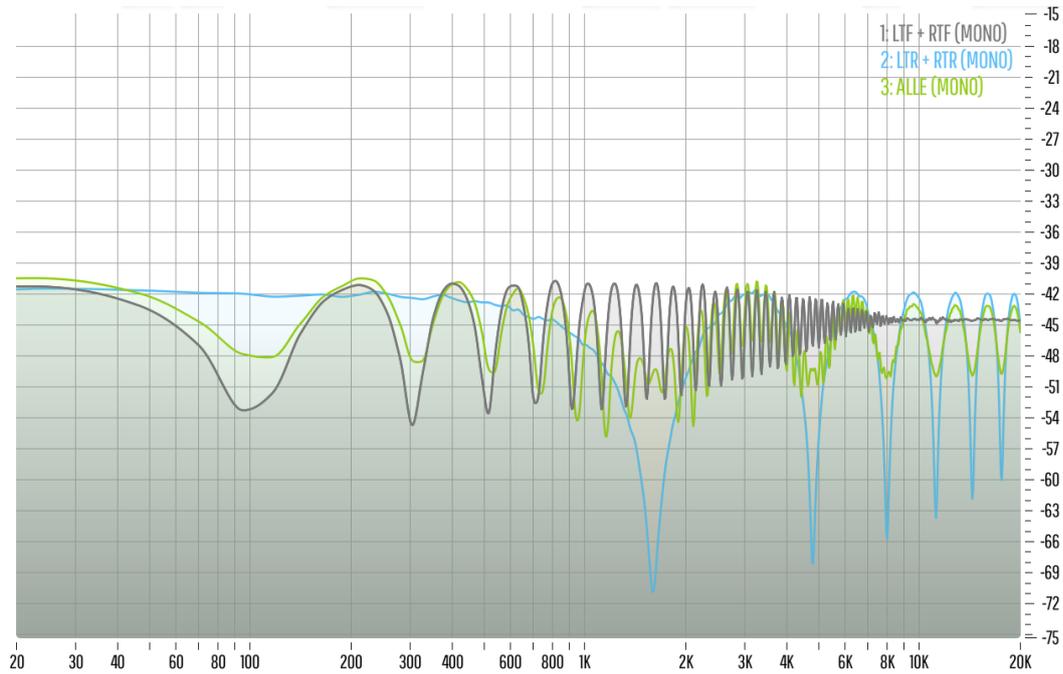
Wie im zuvor betrachteten Modell kann die Bildung der in Abbildung 29 dargestellten Kammfilter durch das Weglassen von zwei Signalen des Upper Layers verhindert werden. Anderenfalls hat diese Kammfilterstruktur Einfluss auf die Downmixergebnisse.

In jedem Fall ergeben sich bei der Summierung der Kanäle des UL auf Ls und Rs durch die räumliche Trennung des Middle und Upper Layers Kammfiltereffekte bzw. kammfilterähnliche Strukturen im 5.0-Downmix.

Die Abbildungen 30 und 31 zeigen den linken bzw. rechten Surround-Kanal, jeweils nach Verwendung unterschiedlicher Signale im Upper Layer des Ausgangsformats.

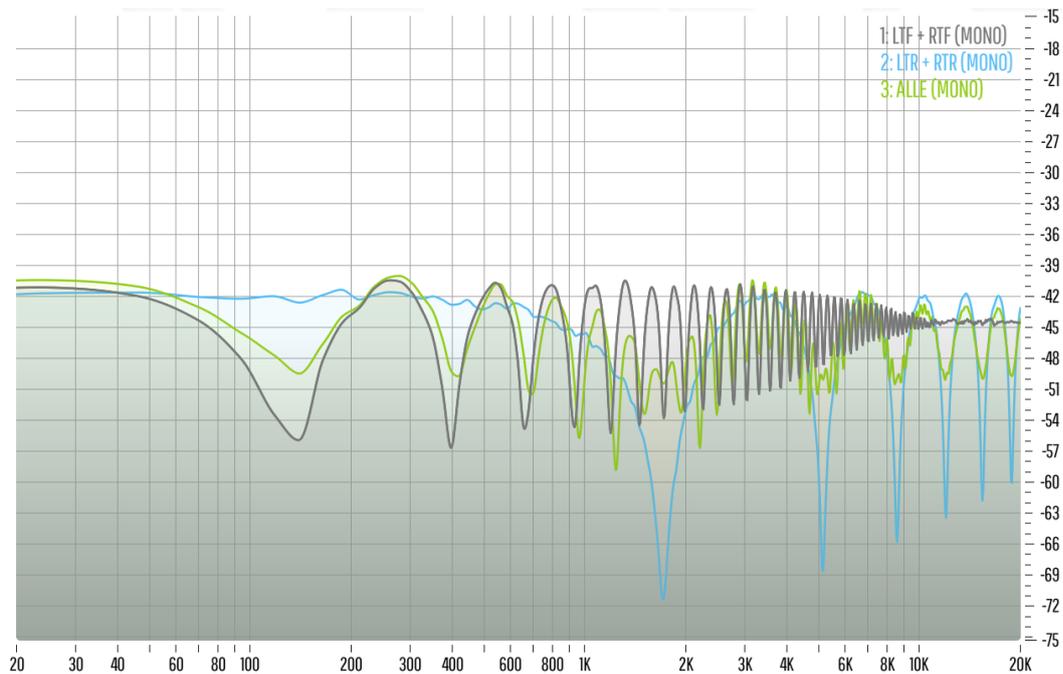
**Abbildung 30**

Frequenzspektrum des Kanals Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 135° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds



**Abbildung 31**

Frequenzspektrum des Kanals Rs im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 135° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds



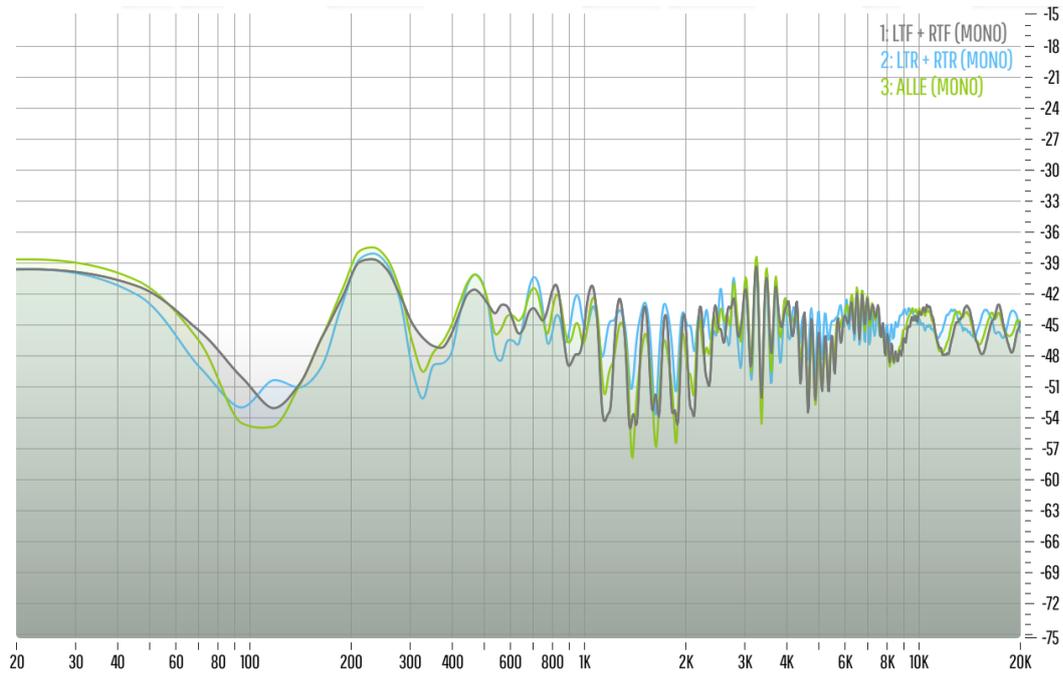
Ein Vergleich von Abb. 31 mit Abb. 30 zeigt, dass die Kurven jeweils um einen bestimmten Wert im Frequenzspektrum nach rechts verschoben sind. Dies resultiert aus dem Delayunterschied zwischen den Kanälen Lss und Rss im Ausgangsformat. Die Lage des ersten Minimums bzw. des ersten Maximums lässt sich dabei jeweils wie folgt berechnen: Bei der Frequenz, deren Wellenlänge dem Verzögerungsweg entspricht, liegt das erste Maximum des Kammfilters. Weitere Maxima folgen bei ganzzahligen Vielfachen dieser Frequenz. Das erste Minimum liegt bei der Frequenz, deren Wellenlänge dem halben Verzögerungsweg entspricht und weitere Minima liegen jeweils mittig zwischen zwei benachbarten Maxima. Zudem ist zu beobachten, dass sich keine Unterschiede bezüglich der Pegeldifferenz zwischen den Maxima und Minima der Kammfilter ergibt. Dies ist auf die Verwendung von ungerichteten Mikrofonen und die Vernachlässigung des entfernungsabhängigen Pegelabfalls zurückzuführen, wodurch jeweils Signale von identischem Pegel summiert werden.

Diese Kammfilterstruktur hat durch die Summation der Kanäle Ls und Rs auf L bzw. R Einfluss auf das Frequenzspektrum der Kanäle im 2.0-Downmix. Dieses wird außerdem durch die Addition des Centers zu den Kanälen L und R beeinflusst. Abb. 32 zeigt den linken Stereokanal unter Verwendung verschiedener Signalkombinationen im UL, Abb. 33 zeigt den rechten Stereokanal.

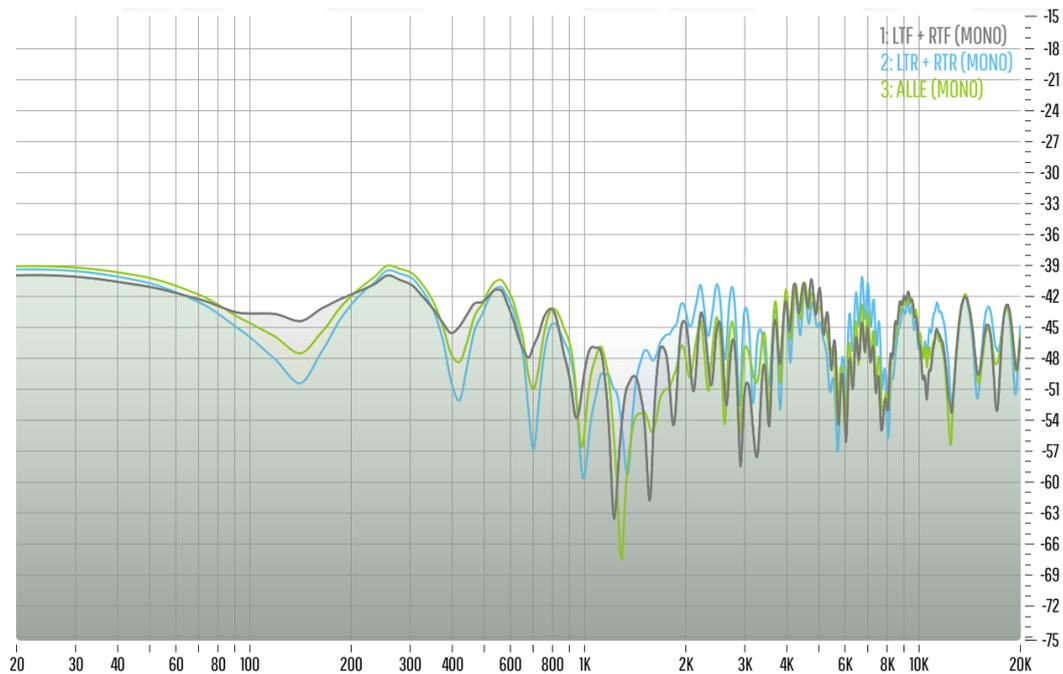
In beiden Abbildungen kann keine typischen Kammfilterstruktur erkannt werden.

**Abbildung 32**

Frequenzspektrum des linken Kanals im Stereodownmix des Decca Cuboids bei einem Schalleinfallswinkel von  $135^\circ$  nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds

**Abbildung 33**

Frequenzspektrum des rechten Kanals im Stereodownmix des Decca Cuboids bei einem Schalleinfallswinkel von  $135^\circ$  nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds

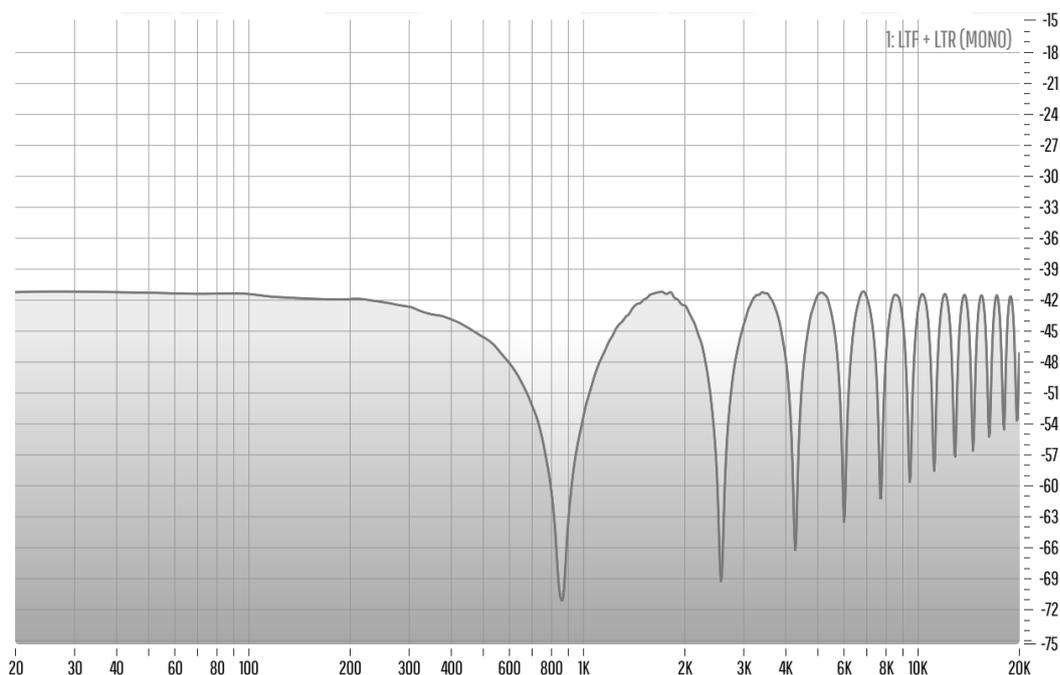


## ORTF-3D

Auch für das ORTF-3D wurde bezüglich der aus dem Upper Layer verwendeten Signale eine Unterscheidung getroffen. Für die erste Schallquellenposition (mittig vor dem Array, auf einer Höhe mit dem Middle Layer) ergibt sich im Format 5.0.4 bei Verwendung aller vier Signale die in Abbildung 34 dargestellte Kammfilterstruktur auf den Kanälen Tl und Tr.

### Abbildung 34

*Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim ORTF-3D bei einem Schalleinfallswinkel von 0°*



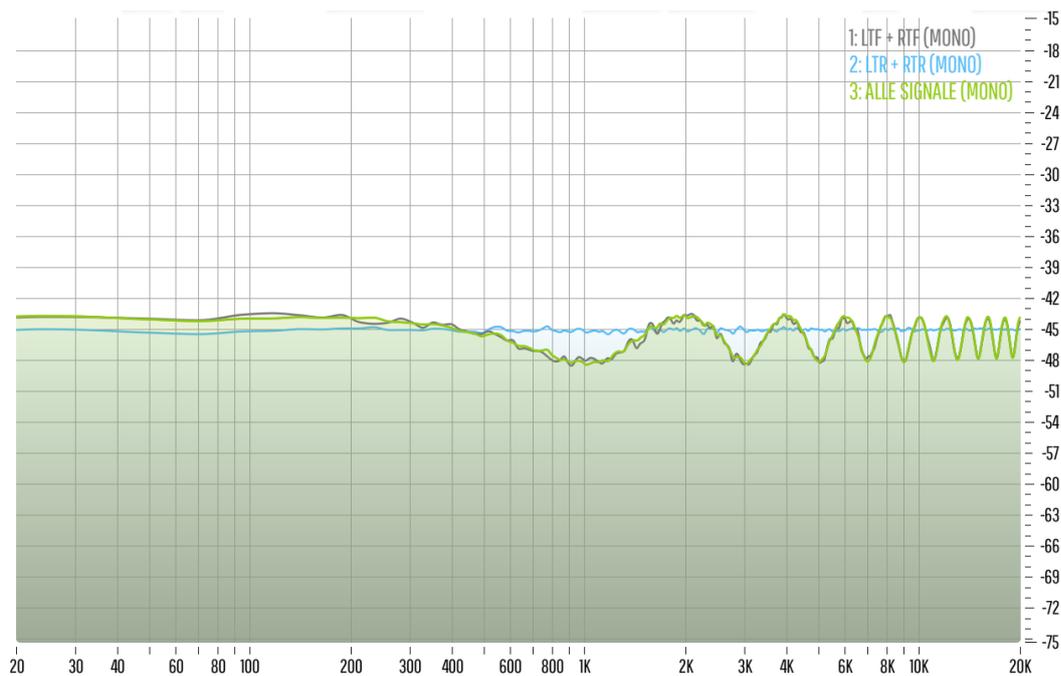
Dieser Kammfilter beeinflusst durch Addition des UL zu den Surround Kanälen den 5.0- und damit den Stereodownmix. Durch die Nutzung von nur zwei Signalen kann die Entstehung eines Kammfilters im UL verhindert werden. Bei Verwendung von Ltr und Rtr kommt es im 5.0-Downmix durch die koinzidente Anordnung der Mikrofone und den damit übereinstimmenden Delaywerten von Ltr und Ls sowie Ltr und Rs nicht zu Kammfiltereffekten. Bei Verwendung von Ltf und Rtf entstehen im 5.0-Downmix Kammfiltereffekte auf den Surround Lautsprechern. Diese werden im Stereodownmix zum linken bzw. rechten Kanal addiert. Durch das Fehlen eines Centerkanals beim ORTF-3D kommt es in keinem der beschriebenen Fälle zu zusätzlichen Veränderungen des Frequenzspektrums durch die Addition eines Centerkanals zu L und R, wie sie beim Downmix des PCMA-3D und des Decca

Cuboids beobachtet werden konnten. Abbildung 35 zeigt die spektrale Analyse von Ls im 5.0-Downmix und Abbildung 36 zeigt die spektrale Analyse des linken Kanals des Stereodownmixes jeweils nach Verwendung aller Signale des UL (grün) im Ausgangsformat und nach Verwendung von Ltf und Rtf (grau) bzw. Ltr und Rtr (blau) im Ausgangsformat.

Bei der Auswertung der Abbildungen 35 und 36 fällt auf, dass die Kammfilterstruktur des Stereosignals durch die Summierung von vorne und hinten liegenden Kanälen, unabhängig von deren Zugehörigkeit zum Middle oder Upper Layer, bestimmt wird. Dies liegt an der vertikal koinzidenten Anordnung der Mikrofone.

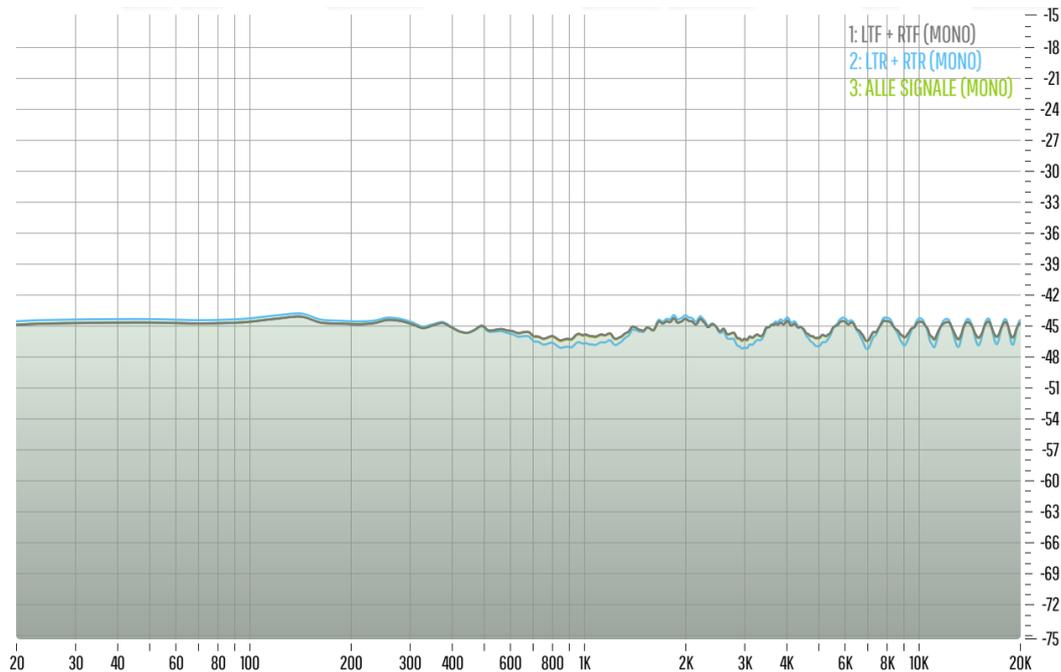
### Abbildung 35

*Frequenzspektrum des Kanals Ls im 5.0-Downmix des ORTF-3D bei einem Schalleinfallswinkel von  $0^\circ$  nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds*



**Abbildung 36**

*Frequenzspektrum des Stereodownmix des ORTF-3D bei einem Schalleinfallswinkel von 0° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds*



*Anmerkung:* Die grüne und blaue Kurve sind deckungsgleich.

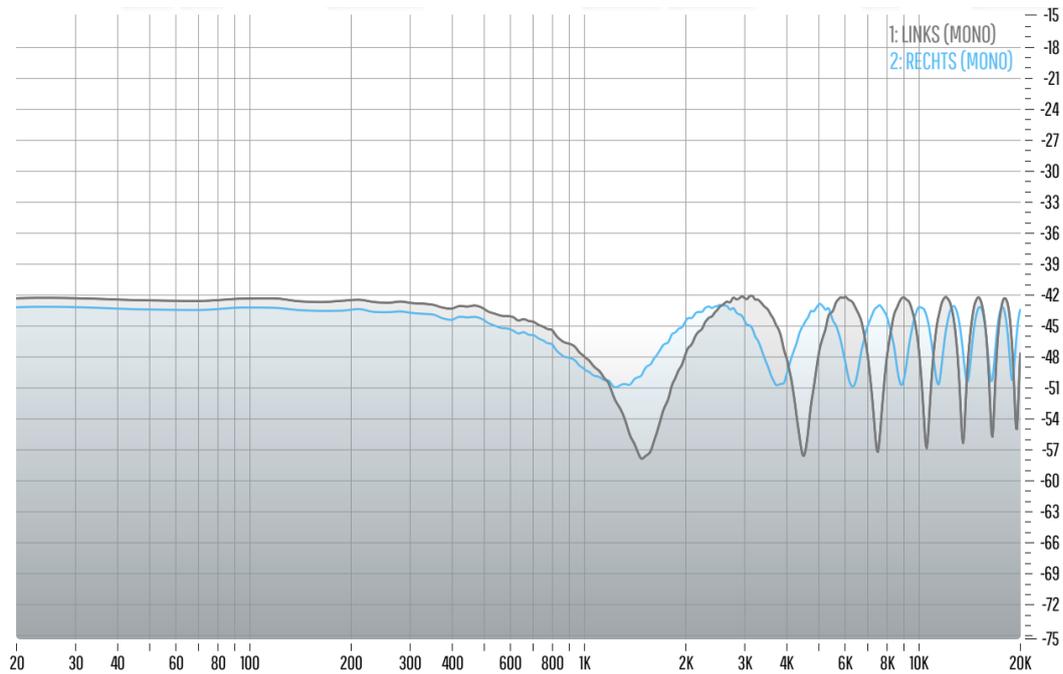
Während im Modell mit mittlerer Schallquellenposition die Pegelwerte für die hinteren bzw. vorderen Mikrofone eines Layers jeweils übereinstimmten, ergeben sich durch den seitlichen Schalleinfallswinkel im zweiten Modell unterschiedliche Pegelwerte für jedes Mikrofon. Dies führt zu unterschiedlichen Strukturen des Frequenzspektrums in den Downmixes auf der linken und rechten Seite.

Kammfiltereffekte entstehen analog zum ersten Modell bei Verwendung aller vier Signale im Upper Layer sowie durch die Summation vorderer und hinterer Signale. Wie im ersten Modell können Kammfiltereffekte im UL durch das Weglassen der vorderen oder hinteren Kanäle des UL und Kammfilter im 5.0-Downmix durch die Verwendung der hinteren Kanäle im UL vermieden werden.

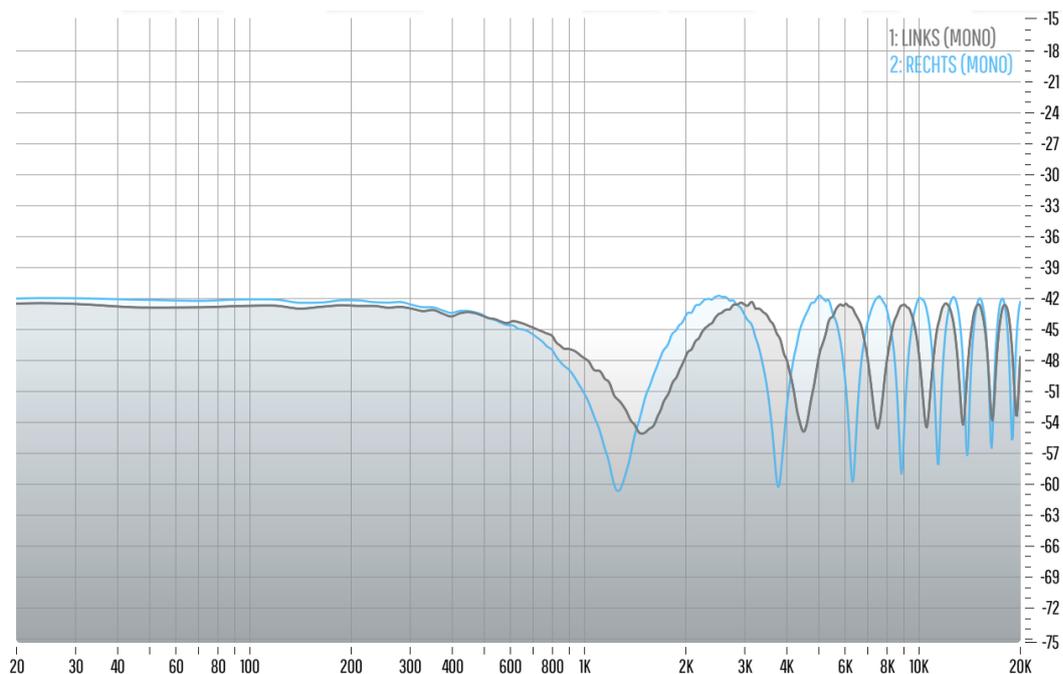
Die Abbildungen 37, 38 und 39 zeigen jeweils den linken (blaue Kurve) und rechten Kanal (graue Kurve) des Stereodownmix unter Verwendung unterschiedlicher Signale des Upper Layers im Ausgangsformat.

**Abbildung 37**

Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von  $135^\circ$  unter Verwendung aller Signale des UL im Dolby Atmos Bed

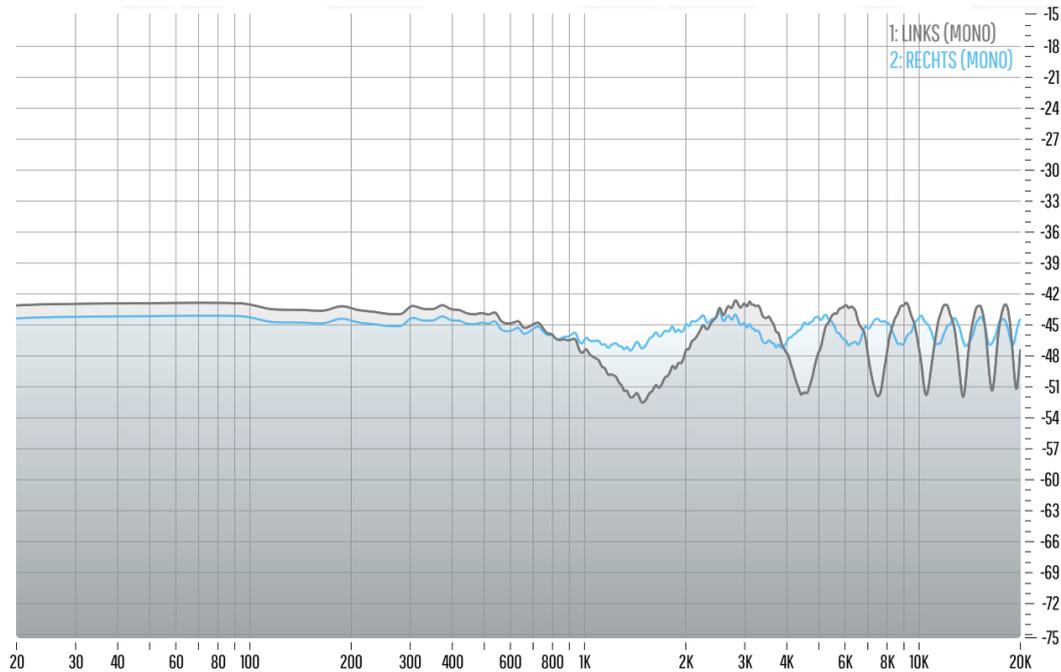
**Abbildung 38**

Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von  $135^\circ$  unter Verwendung der Signale Ltf und Rtf des UL im Dolby Atmos Bed



**Abbildung 39**

*Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 135° unter Verwendung der Signale Ltr und Rtr des UL im Dolby Atmos Bed*



Beim Vergleich der Abb. 37, 38 und 39 zeigt sich, dass der wesentliche Unterschied in der Pegeldifferenz zwischen den Minima und Maxima des Kammfilters liegt.

## 7.2. Downmix Dolby Atmos Objekte

Bei der Verwendung von Dolby Objekten anstelle des Beds ist eine Unterscheidung bezüglich der im Upper Layer verwendeten Signale nicht nötig, da für Objekte eine separate Ansteuerung aller vier Deckenlautsprecher möglich ist. Um jeweils die Downmixergebnisse bei Nutzung aller vom jeweiligen Array generierten Signale zu untersuchen, wurde hier so wie in den Kapiteln 7.3 und 7.4 darauf verzichtet, unterschiedliche Signalkombinationen zu verwenden.

Die Objekte werden nicht anhand einer Matrix in ein anderes Wiedergabeformat überführt, sondern mithilfe von Metadaten gerendert. In den Metadaten wird u.a. die Position der Objekte übertragen. Diese wird für jedes Wiedergabeformat berechnet. Dadurch, dass die Objekte in den Array-Simulationen im Ausgangsformat genau auf den Lautsprecherpositionen liegen, ergibt sich für die Position der Objekte im 5.0- und Stereodownmix Folgendes: Die Signale des Upper Layers werden im 5.0 Downmix auf die darunterliegenden Lautsprecher und damit

die darunterliegenden Array-Kanäle übertragen. Im Stereodownmix werden die auf den Surroundlautsprechern liegenden Signale sowie die auf Ltr und Rtr liegenden Objekte auf L und R gelegt. Zusätzlich wird der Center auf den linken und rechten Lautsprecher aufgeteilt.

### PCMA-3D

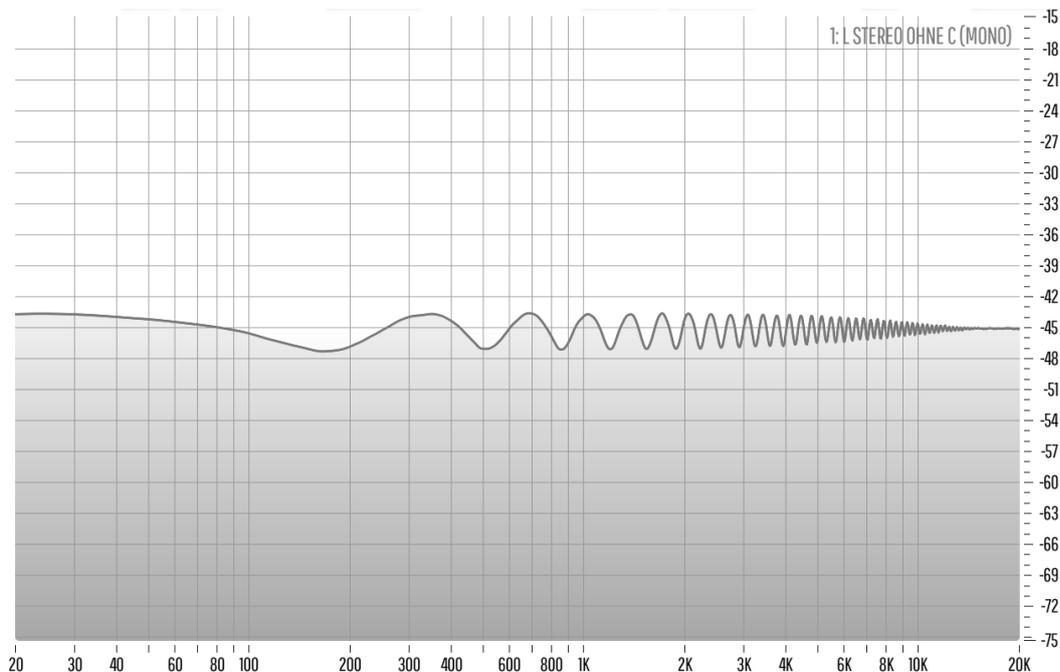
Bei der ersten simulierten Schallquellenposition kommt es bei der Wiedergabe im Format 5.0 nicht zu Kammfiltern, da die summierten Signale (L und Ltf, Ls und Ltr, R und Rtf und Rs und Rtr) jeweils denselben Delaywert aufweisen.

Bei der Stereowiedergabe kommt es durch die Addition der hinteren auf die vorderen Kanäle sowie die Addition des Centers auf Links und Rechts zu Kammfiltereffekten auf den Stereokanälen. Durch die symmetrische Anordnung der Mikrofone des Arrays und die mittig liegende Schallquellenposition stimmen diese links und rechts überein.

Abbildung 40 zeigt den linken Kanal des Stereomixes mit gemutetem Center. Abbildung 41 zeigt den linken Kanal des Stereodownmixes inklusive Center.

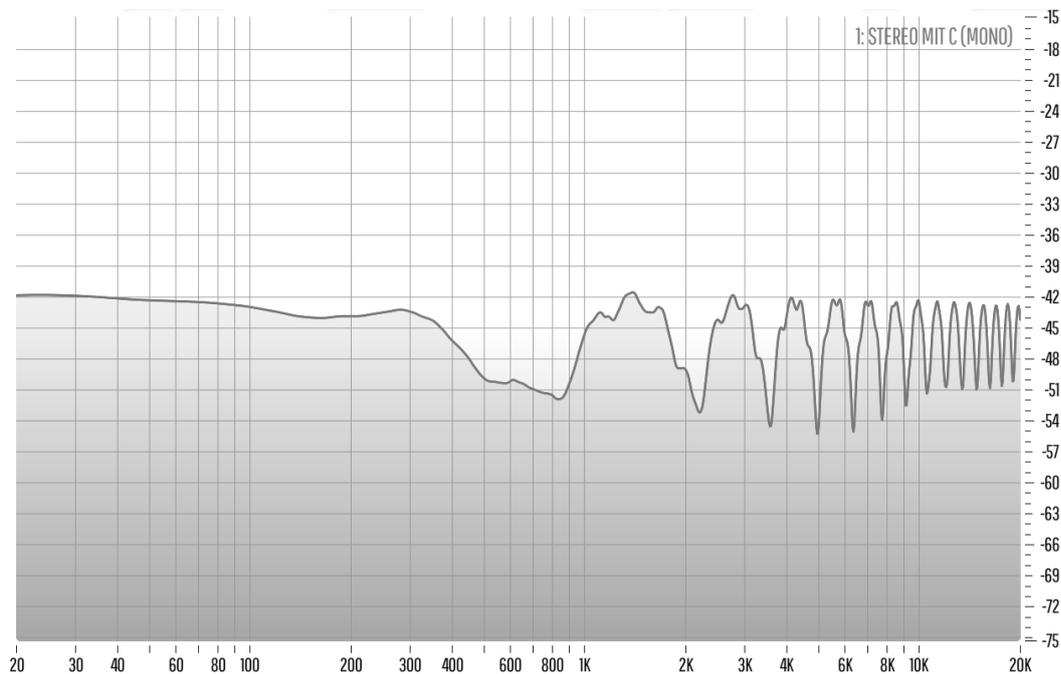
### Abbildung 40

*Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten (ohne C)*



**Abbildung 41**

*Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten (mit C)*

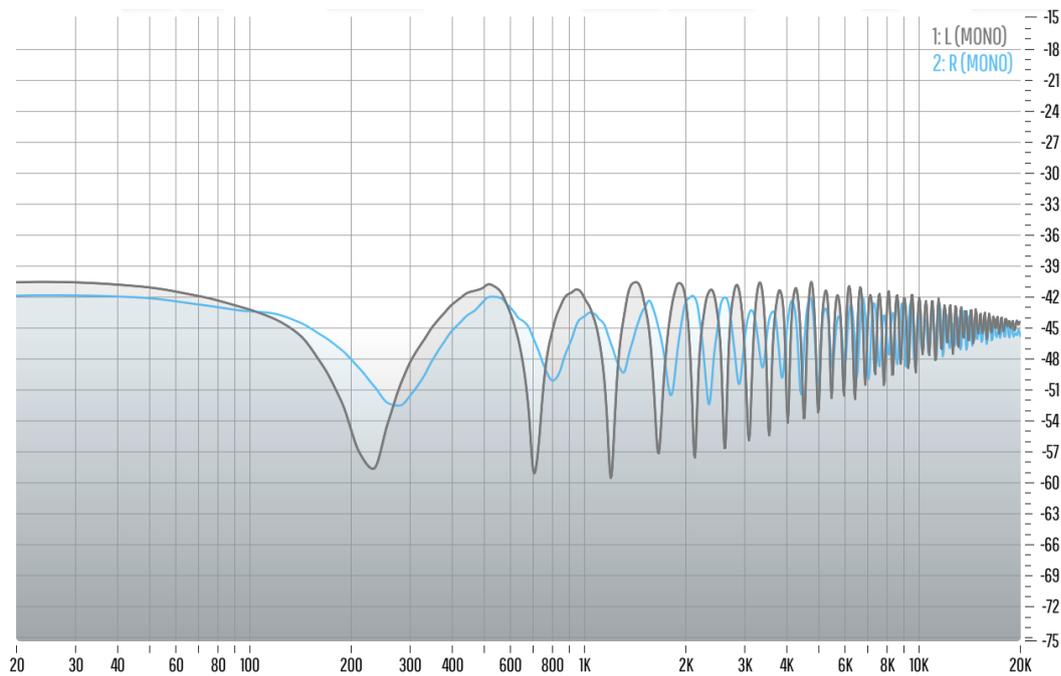


Der Vergleich von Abbildung 40 und 41 macht die Auswirkungen des Centerkanals bzw. des Laufzeitunterschiedes zwischen Center und Links sowie Rechts auf das Frequenzspektrum der Kanäle des Stereodownmixes deutlich.

Auch bei der zweiten Schallquellenposition ergeben sich bei der Wiedergabe im Format 5.0 keine Kammfiltereffekte. Im Stereomix kommt es durch die Summierung von Ls und Rs sowie der Addition des Centerkanals auf L und R zu Kammfiltereffekten. Diese fallen aufgrund des seitlichen Schalleinfalls und der damit unterschiedlichen Laufzeit von der simulierten Schallquelle zu den Mikrofonen auf beiden Seiten unterschiedlich aus. Abbildung 42 zeigt die spektrale Analyse des linken und rechten Stereokanals.

**Abbildung 42**

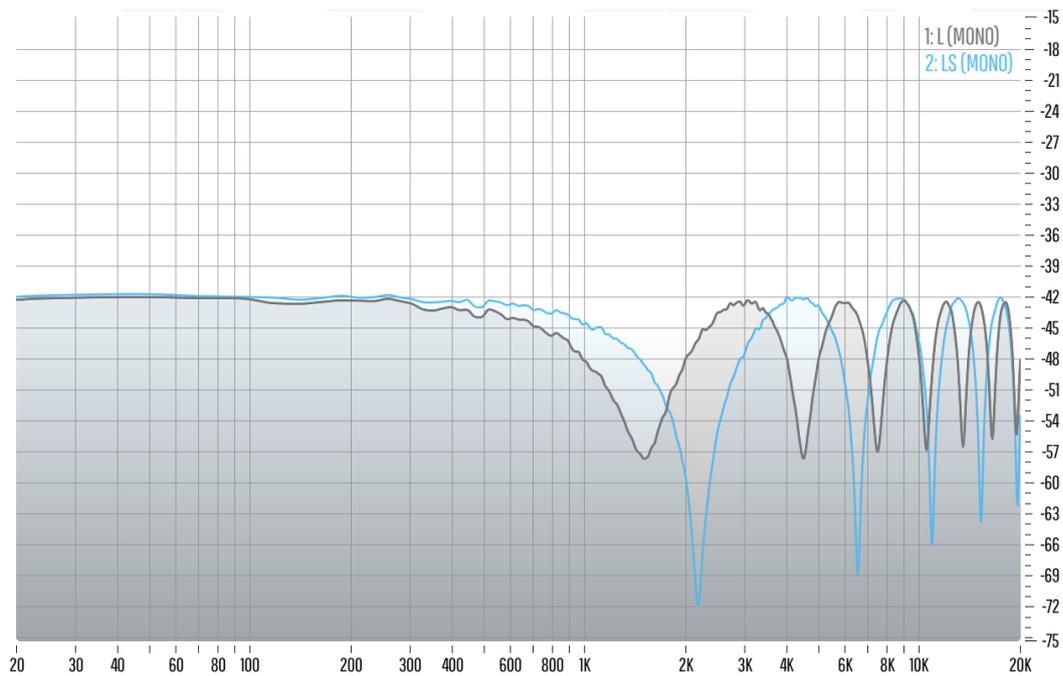
*Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung von Dolby Atmos Objekten*

**Decca Cuboid**

Bei der Simulation des Decca Cuboids mit der ersten Schallquellenposition kommt es bei der Wiedergabe im Format 5.0 auf L, R, Ls und Rs zu Kammfiltereffekten. Diese resultieren aus der Summierung der jeweils übereinanderliegenden Kanäle. Da die Mikrofone des Decca Cuboids nicht vertikal koinzident angeordnet sind, unterschieden sich die betreffenden Kanäle jeweils in ihrem Delaywert, was bei der Summierung zu Kammfiltereffekten führt. Abbildung 43 zeigt die spektrale Analyse des linken Kanals (grau) und des linken Surround-Kanals (blau) im 5.0-Mix. Auch hier stimmen die Frequenzspektren der linken und rechten Seite jeweils überein.

**Abbildung 43**

*Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboid bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten*

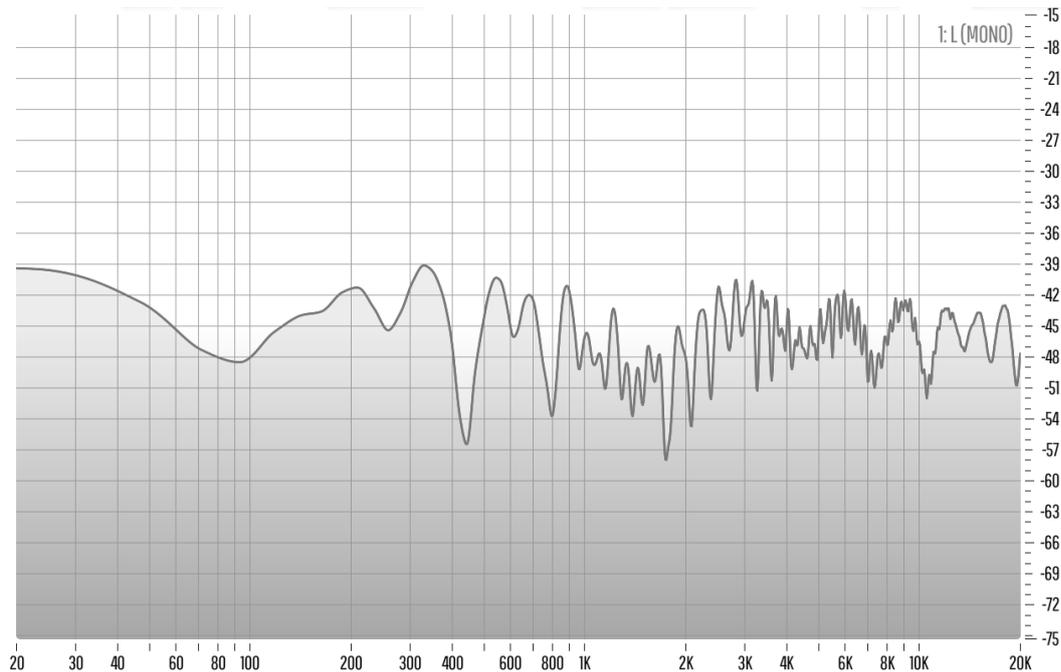


Im Stereomix werden die hinteren Kanäle zu den vorderen addiert sowie der Center auf L und R addiert. Beides hat Auswirkungen auf die Frequenzspektren der Stereokanäle.

Abbildung 44 zeigt den linken Kanal des Stereomixes. Der rechte Kanal weist dieselbe Struktur auf. Im Stereodownmix ist, wie in Abbildung 44 zu sehen, keine klare Kammfilterstruktur mehr zu erkennen.

**Abbildung 44**

*Frequenzspektrum des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten*

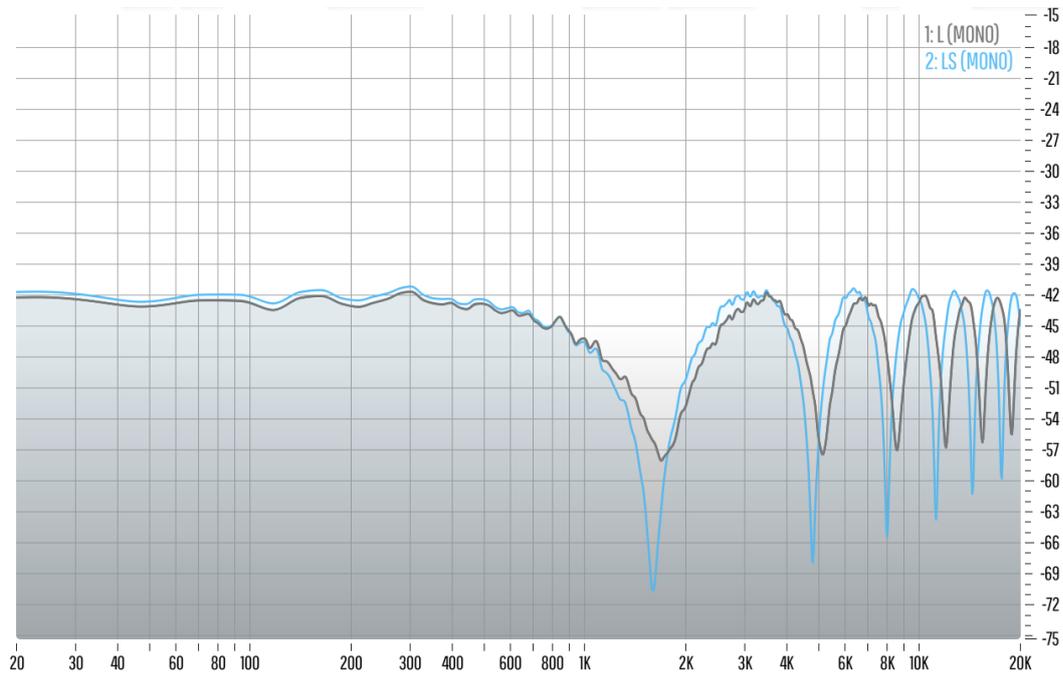


Auch bei der zweiten Schallquellenposition kommt es beim Decca Cuboid im 5.0-Downmix zu Kammfiltereffekten auf den Kanälen L, R, Ls und Rs. Diese unterscheiden sich wie bereits erläutert jeweils auf der linken und rechten Seite. Wie beim Downmix des Dolby Atmos Beds gezeigt, unterscheiden sich die Kammfilterstrukturen von Ls und Rs nur durch eine Verschiebung der Kurve im Frequenzspektrum. Dies konnte auch für die Kanäle L und R beobachtet werden. Daher zeigt Abbildung 45 jeweils die spektrale Analyse von L und Ls im 5.0-Downmix.

Im Frequenzspektrum der Kanäle des 2.0-Downmixes ergeben sich wie bereits bei der anderen Schallquellenposition gezeigt keine typischen Kammfiltereffekte. Die aus destruktiven Interferenzen entstehenden Frequenzspektren der beiden Stereokanäle zeigt Abb. 46

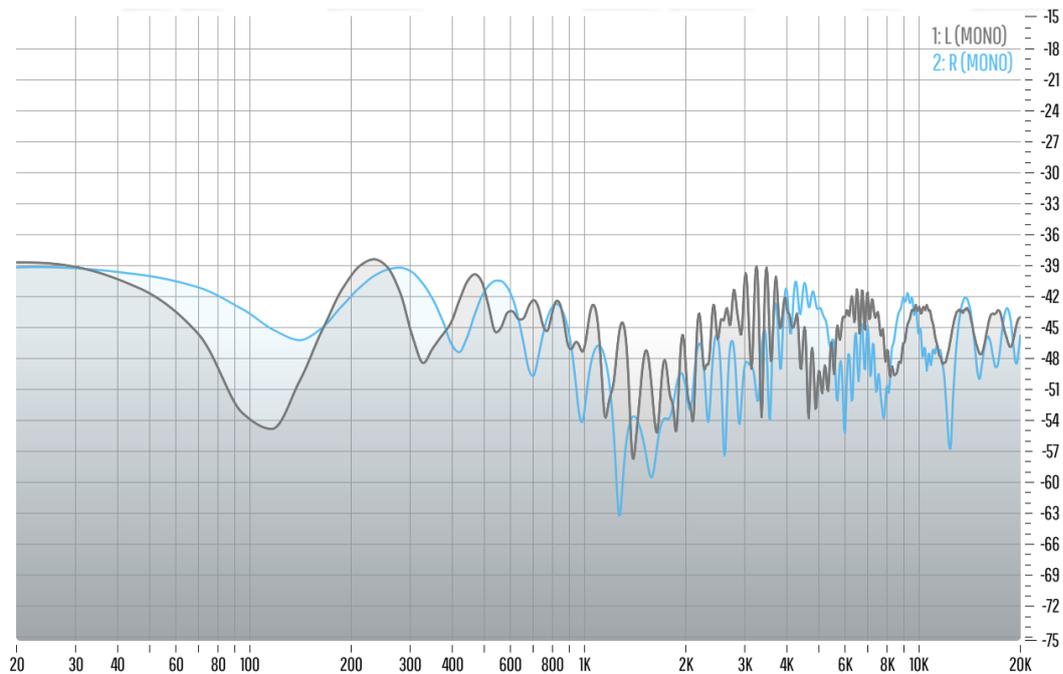
**Abbildung 45**

*Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboid bei einem Schalleinfallswinkel von 135° unter Verwendung von Dolby Atmos Objekten*



**Abbildung 46**

*Frequenzspektrum der Kanäle des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 135° unter Verwendung von Dolby Atmos Objekten*



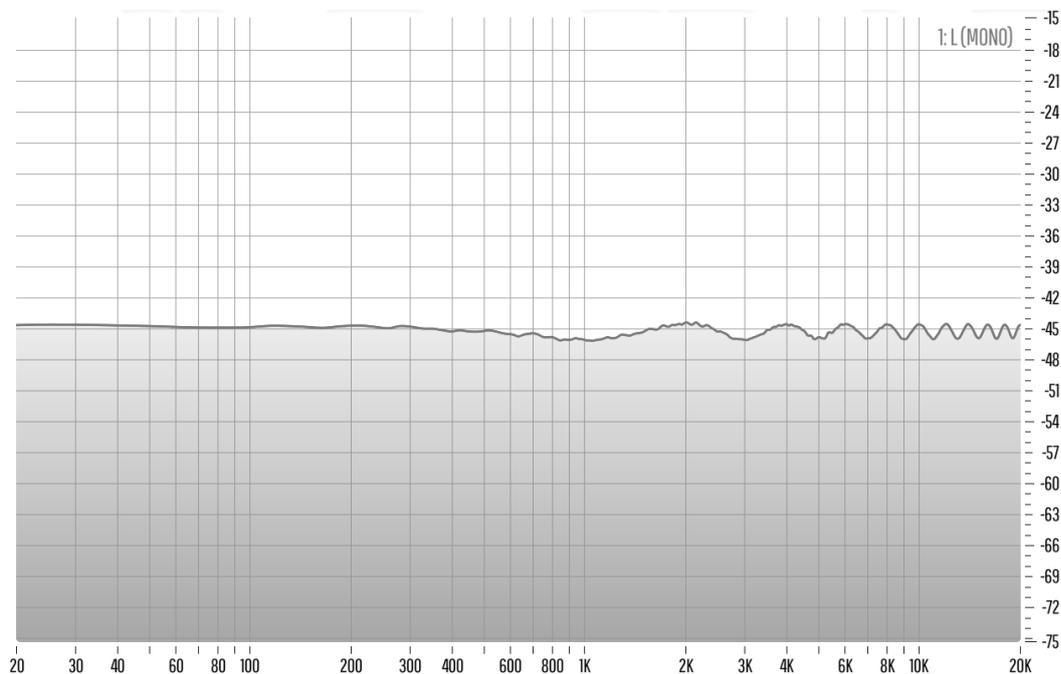
## ORTF-3D

Bei der Simulation des ORTF-3D mit der ersten Schalleinfallrichtung kommt es durch die koinzidente Anordnung des Middle und Upper Layers bei der Wiedergabe im Format 5.0 auf keinem Kanal zu Kammfiltereffekten.

Im Stereomix ergibt sich durch die Überlagerung der hinteren und vorderen Kanäle und die Laufzeitunterschiede zwischen diesen die in Abbildung 47 gezeigte Kammfilterstruktur. Diese stimmt auf beiden Stereokanälen überein.

### Abbildung 47

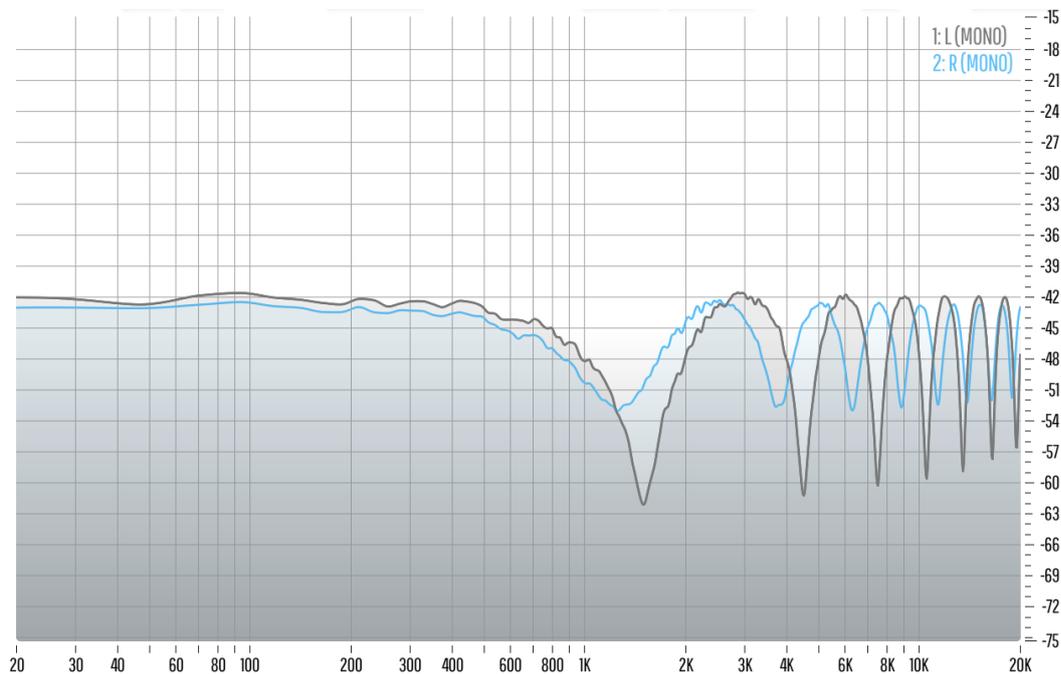
*Frequenzspektrum des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten*



Auch in der zweiten Simulation mit anderer Schallquellenposition entstehen durch die vertikal koinzidente Anordnung der Mikrofone und der ausschließlichen Summation von Kanälen mit selbem Delaywert im 5.0-Mix keine Kammfilter. Im Stereomix ergeben sich durch die Addition der hinteren und vorderen Kanäle Kammfilter. Diese unterscheiden sich durch den seitlichen Schalleinfallswinkel und der daraus resultierenden Unterschiede in der Laufzeit zu den einzelnen Mikrofonen auf der linken und rechten Seite. Abb. 48 zeigt den linken und rechten Stereokanal.

**Abbildung 48**

*Frequenzspektrum der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung von Dolby Atmos Objekten*

**7.3. Downmix MPEG-H Bed**

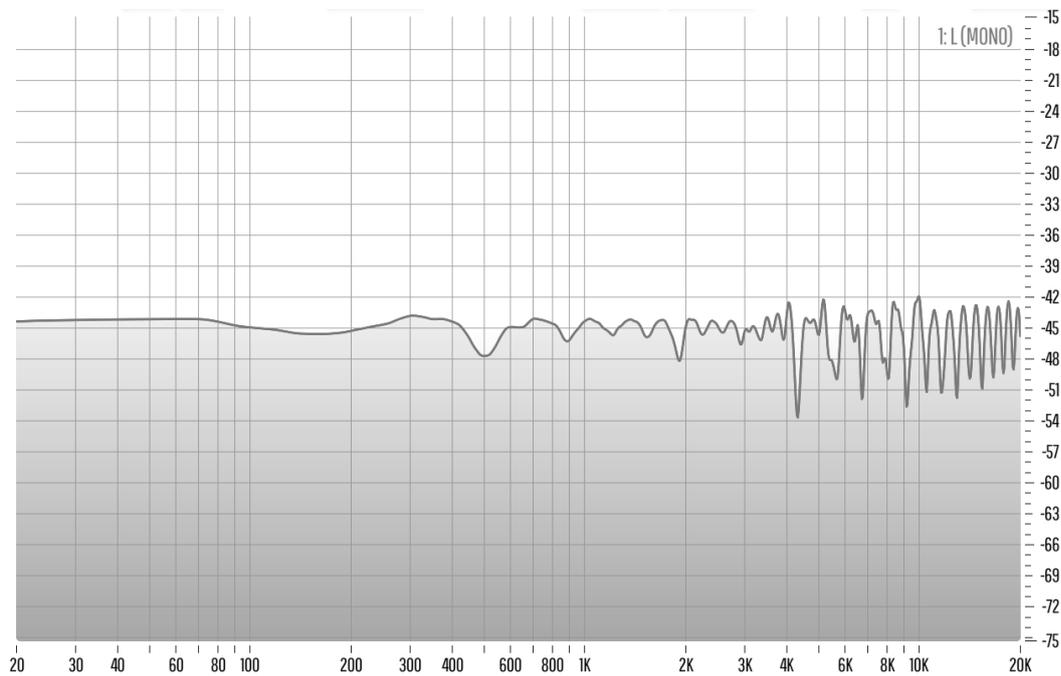
Bei Verwendung des kanalbasierten MPEG-H Beds werden im Downmix zu 5.0 im Unterschied zu Dolby Atmos die Signale des Upper Layers jeweils zu den darunterliegenden Kanälen addiert. Zudem kommt hier, wie in Kapitel 2.3 erläutert, ein aktiver Algorithmus zum Einsatz.

**PCMA-3D**

Beim Downmix des Modells des PCMA-3D mit der ersten Schallquellenposition (Schallquelle auf einer Höhe mit dem Array, mittig vor dem Array), werden im 5.0-Downmix ausschließlich Signale mit demselben Delaywert ( $L+L_h$ ,  $L_s+L_{sh}$ ,  $R+R_h$ ,  $R_s+R_{sh}$ ) summiert. Dadurch entstehen auf keinem Kanal Kammfiltereffekte durch den Downmix zu 5.0. Im Stereomix kommt es zu einer Summierung von jeweils  $L$ ,  $L_h$ ,  $L_s$  und  $L_{sh}$  sowie  $R$ ,  $R_h$ ,  $R_s$  und  $R_{sh}$ . Zusätzlich wird der Center zu Links und Rechts addiert. Nach der Anwendung des aktiven Algorithmus ergibt sich in den beiden Stereokanälen das in Abb. 49 dargestellte Frequenzspektrum.

**Abbildung 49**

*Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung des MPEG-H Beds*

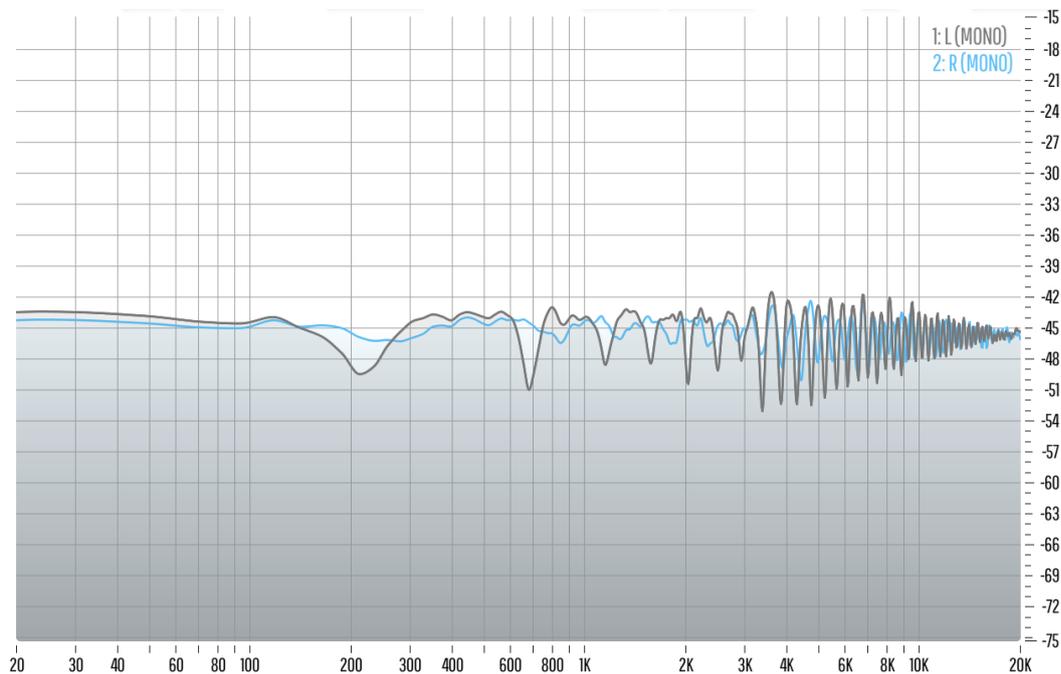


Hier kann keine typische Kammfilterstruktur erkannt werden.

Auch bei der zweiten simulierten Schalleinfallrichtung kommt es im 5.0-Mix aus den erläuterten Gründen nicht zu Kammfiltereffekten. Im Stereomix ergeben sich durch die Summation der Kanäle L, Ls und C sowie R, Rs und C die in Abbildung 50 dargestellten Frequenzspektren des linken und rechten Kanals.

**Abbildung 50**

*Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung des MPEG-H Beds*



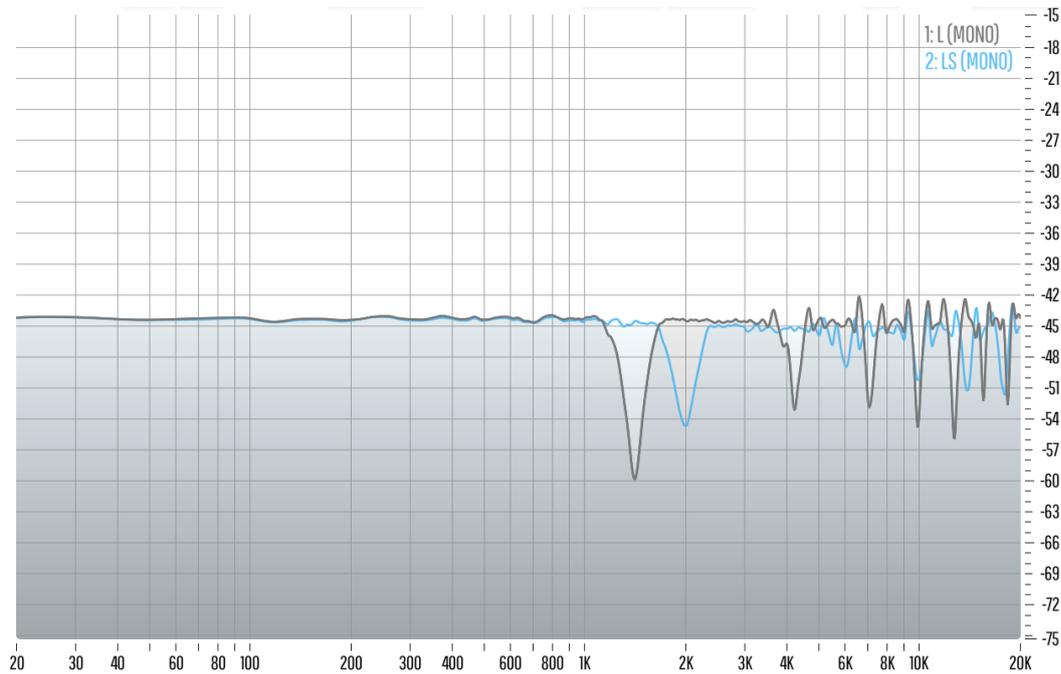
In Abbildung 50 kann ebenso keine typische Kammfilterstruktur erkannt werden.

**Decca Cuboid**

Beim Decca Cuboid kommt es durch den Abstand zwischen Middle und Upper Layer beim 5.0-Mix durch die Addition der im Array jeweils übereinanderliegenden Kanälen zu destruktiven Interferenzen auf L, R, Ls und Rs. Dabei sind die Strukturen auf der linken und rechten Seite jeweils identisch. Abbildung 51 zeigt den linken Kanal und den linken Surround Kanal im 5.0-Mix.

**Abbildung 51**

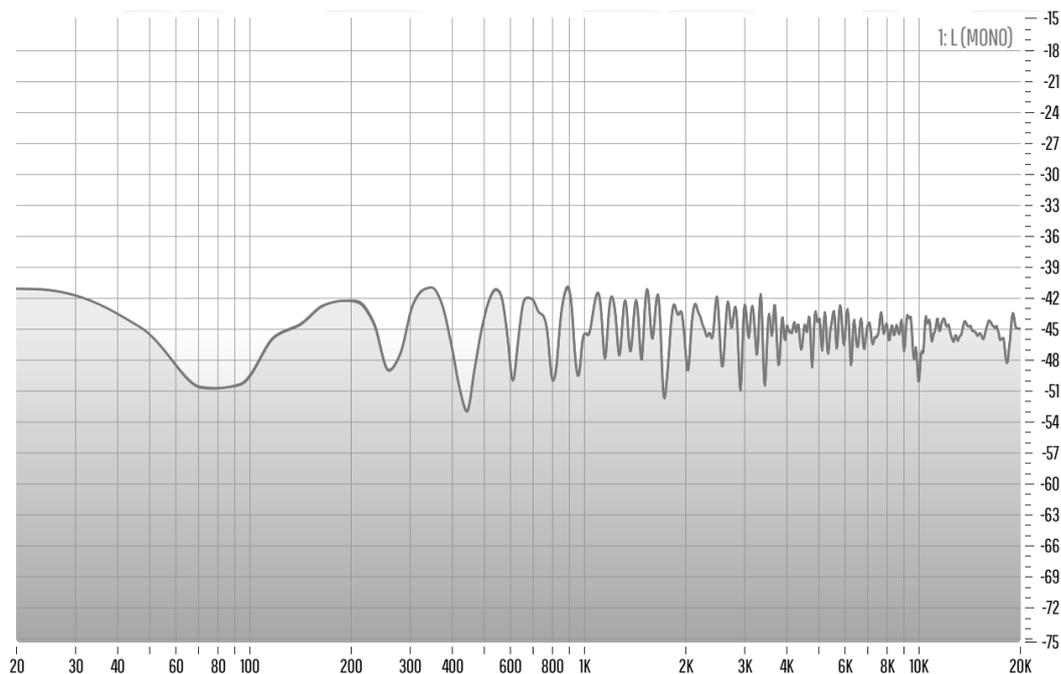
*Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 0° unter Verwendung des MPEG-H Beds*



Auch durch den Stereodownmix kommt es zu Interferenzen. Diese stimmen auf beiden Seiten überein. Das aus dem aktiven Downmix resultierende Frequenzspektrum des linken Kanals wird in Abbildung 52 dargestellt.

**Abbildung 52**

*Frequenzspektrum des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 0° unter Verwendung des MPEG-H Beds*

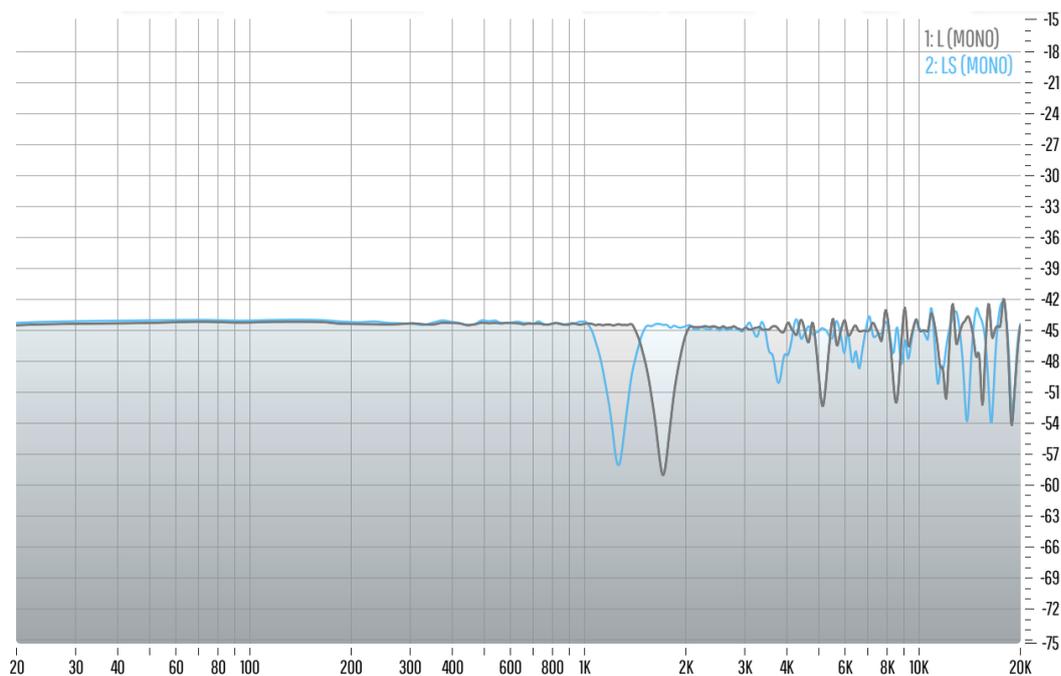


Auch hier nimmt das Spektrum nicht die Form eines Kammfilters an.

Beim Downmix der Simulation des Decca Cuboids mit einer Schalleinfallrichtung von  $135^\circ$  zu 5.0 mit Dolby Atmos konnte sowohl unter der Verwendung des kanalbasierten Beds als auch von Objekten gezeigt werden, dass sich die Frequenzspektren von Ls und Rs bzw. von L und R lediglich durch eine Offset entlang der x-Achse unterscheiden. Da dies durch die Delay- und Pegelwerte und nicht durch den Downmixalgorithmus bzw. den Object-Renderer bedingt ist, gilt dies auch für den Downmix dieser Simulation mit MPEG-H. Um den Einfluss des aktiven Downmixalgorithmus dazustellen, genügt es demnach jeweils eine der beiden Seiten zu analysieren. Abbildung 53 zeigt das Frequenzspektrum der Kanäle L und Ls im 5.0-Downmix.

### Abbildung 53

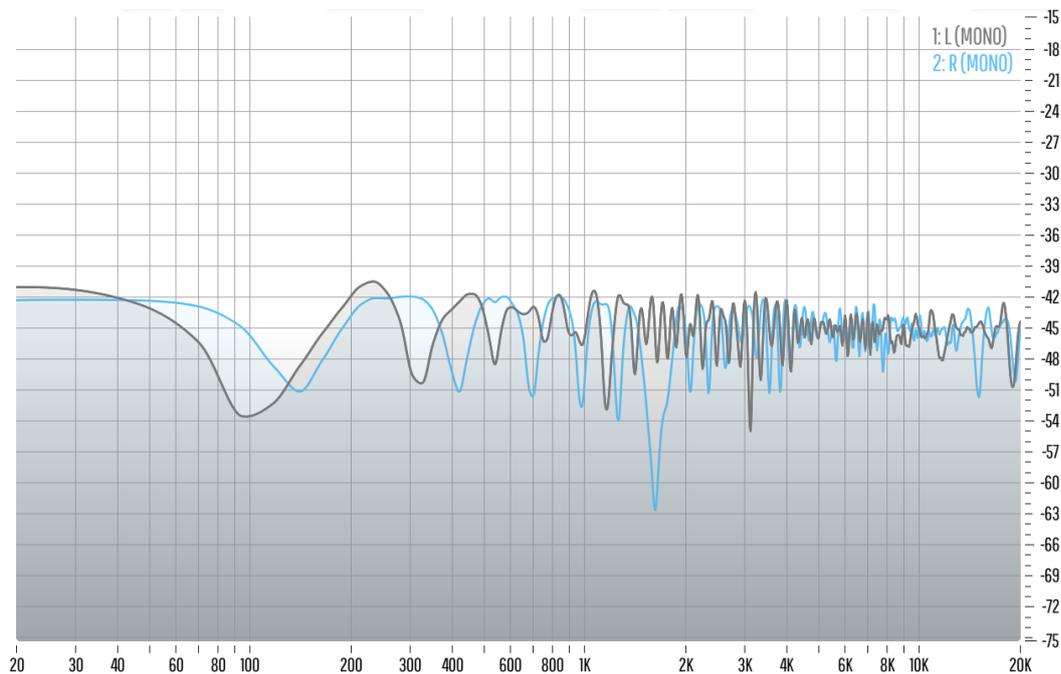
*Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von  $135^\circ$  unter Verwendung des MPEG-H Beds*



Im 2.0 Downmix kommt es durch die Addition der hinteren Kanäle des UL nach vorne sowie der Addition des Centers zu den Kanälen L und R zu den in Abb. 54 dargestellten Frequenzspektren.

**Abbildung 54**

*Frequenzspektren der Kanäle des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 135° unter Verwendung des MPEG-H Beds*



Hier ist keine typische Kammfilterstruktur zu erkennen.

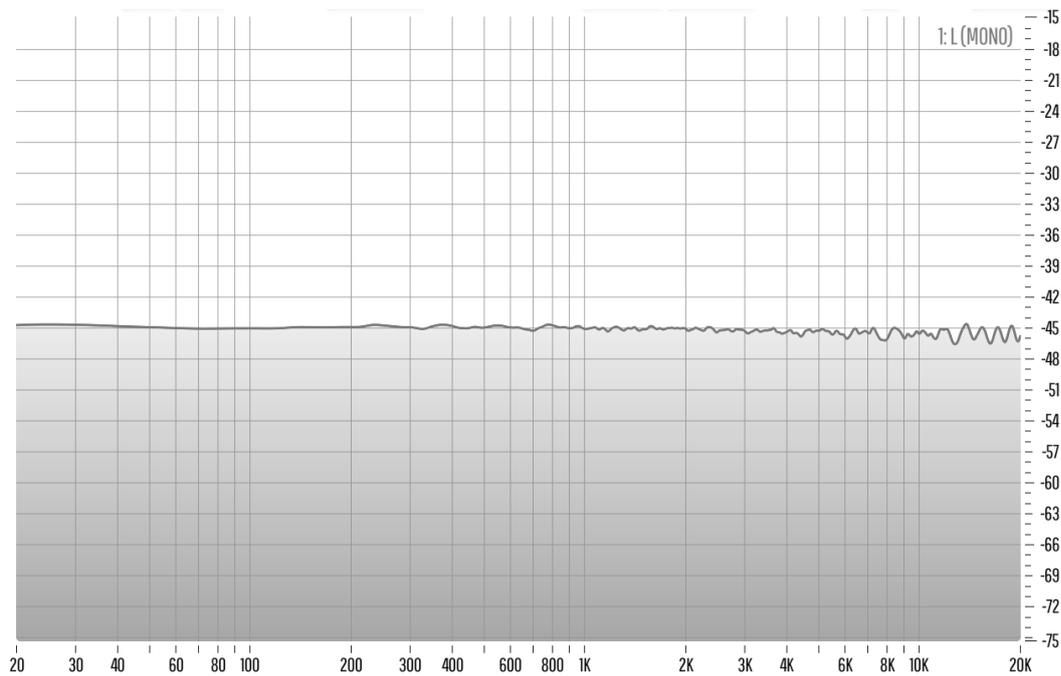
**ORTF-3D**

Beim ORTF-3D kommt es im Modell mit der ersten Schallquellenposition im 5.0-Downmix nicht zu Kammfiltereffekten, da nur Kanäle, welche denselben Delaywert aufweisen, summiert werden. Dies gilt auch für die zweite simulierte Schalleinfallrichtung.

Im Stereodownmix ergeben sich Kammfilter auf dem linken und rechten Kanal. Abbildung 55 zeigt den linken Kanal des Stereodownmixes. Auch hier stimmen L und R aufgrund der Symmetrie des Modells in ihrem Frequenzspektrum überein. Hier ist zu beachten, dass der Kammfiltereffekt erst in hohen Frequenzen (ab ca. 10 kHz) einsetzt.

**Abbildung 55**

*Frequenzspektrum des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 0° unter Verwendung des MPEG-H Beds*

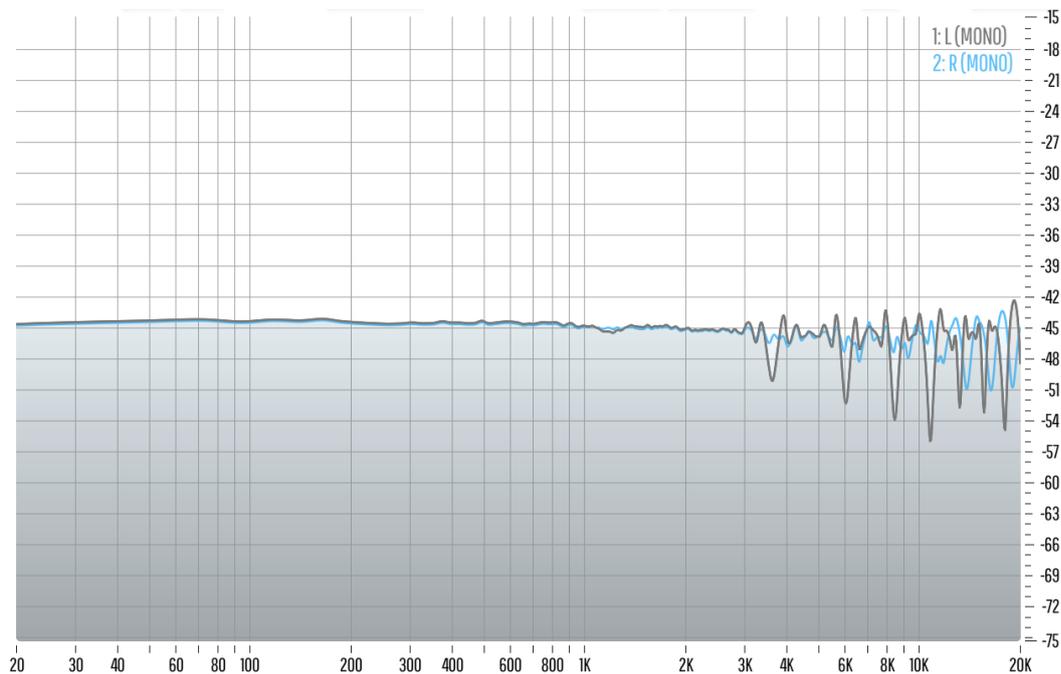


Bei der zweiten Schalleinfallrichtung kommt es durch die Summation von L und Ls sowie R und Rs zu destruktiven Interferenzen, welche sich auf der linken und rechten Seite aufgrund der jeweils verschiedenen Laufzeit vom Schallereignis zu jedem Mikrofon des Arrays unterscheiden.

Abbildung 56 zeigt das aus dem Downmix resultierende Frequenzspektrum des linken und rechten Kanals.

**Abbildung 56**

*Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 135° unter Verwendung des MPEG-H Beds*



Hier können keine typischen Kammfilterstrukturen beobachtet werden.

#### **7.4. Downmix MPEG-H Objekte**

Die Überführung der Signale in ein kleineres Wiedergabeformat erfolgt bei der Verwendung von Objekten nicht über eine Downmixmatrix. Stattdessen werden die Objekte entsprechend der in den Metadaten überlieferten Information auf ihre Positionen gerendert. Dadurch entfällt auch die aktive Komponente des Downmixalgorithmus.

#### **PCMA-3D**

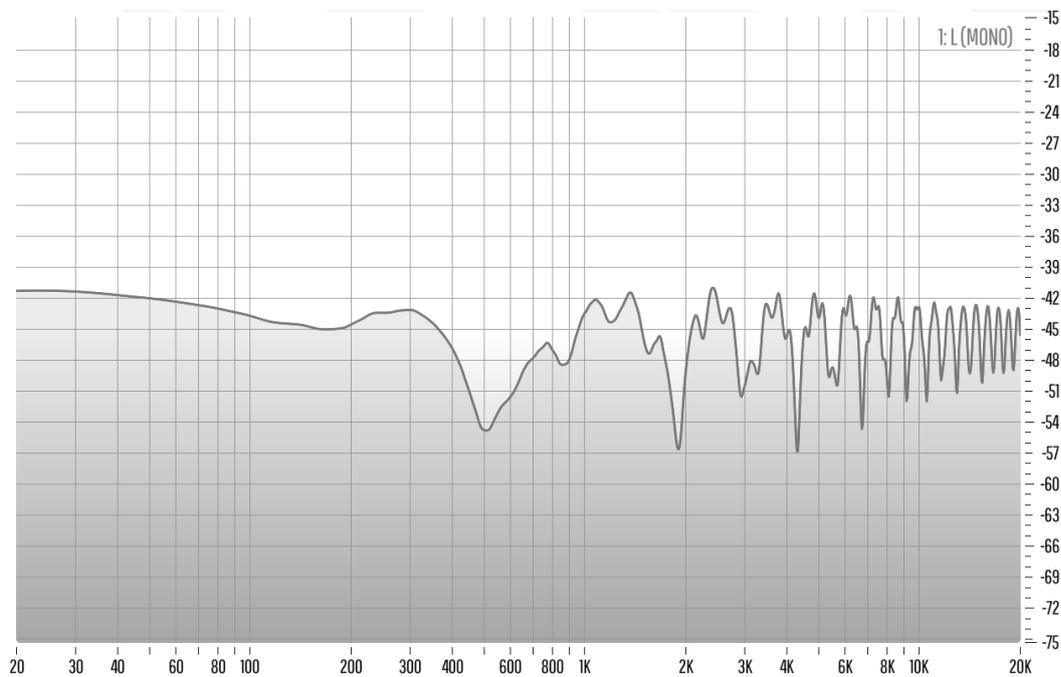
Genau wie beim Downmix von über das MPEG-H Bed gerouteten Signalen, kommt es bei der Simulation des PCMA-3D durch die vertikal koinzidente Anordnung der Mikrofone und den damit übereinstimmenden Delaywerten bei den sich addierenden Signalen in keinem Kanal des 5.0-Mix bei einer der beiden Schallquellenpositionen zu Kammfiltereffekten.

Im Stereomix ergeben sich durch die Summierung der hinten liegenden Signale (Ls, Lsh, Rs und Rsh) zu den Signalen des linken und rechten Signals destruktive Interferenzen. Abbildung 57 zeigt das resultierende Frequenzspektrum des linken

Kanals des Stereomixes. Der linke und rechte Kanal stimmen in ihrem Spektrum überein.

### Abbildung 57

*Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von  $0^\circ$  unter Verwendung von MPEG-H Objekten*

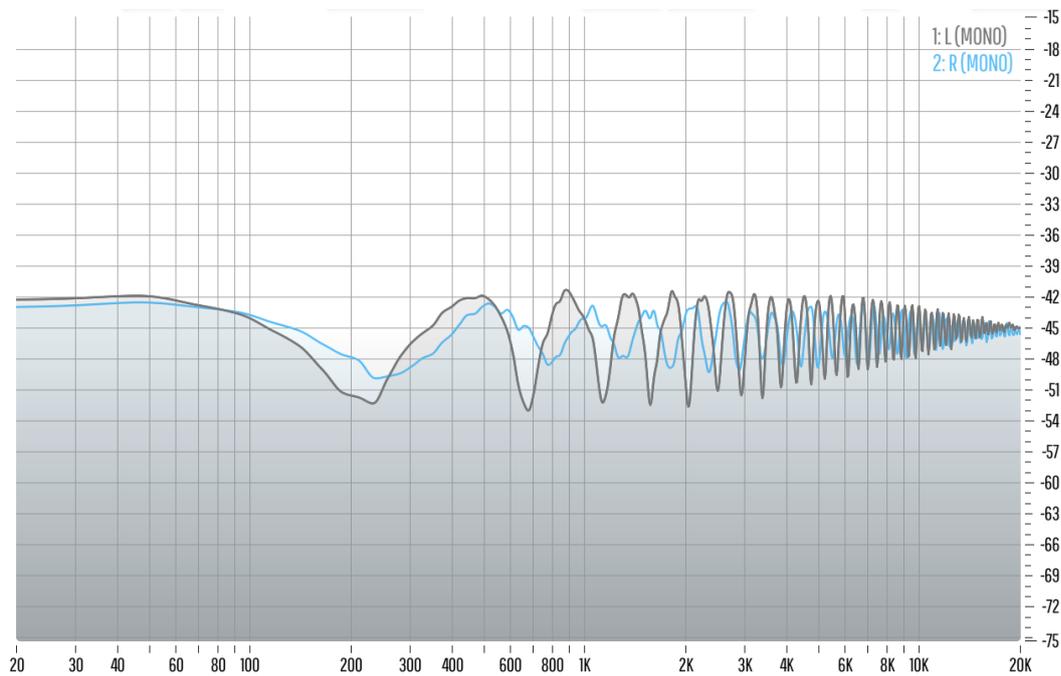


In Abb. 57 sind insbesondere in hohen Frequenzen kammfilterartige Strukturen zu sehen.

Beim zweiten Modell mit einer Schalleinfallrichtung von  $135^\circ$  links kommt es aus den genannten Gründen ebenfalls ausschließlich im 2.0-Mix zu Kammfiltereffekten. Abbildung 58 zeigt die spektrale Analyse des linken und rechten Stereokanals.

**Abbildung 58**

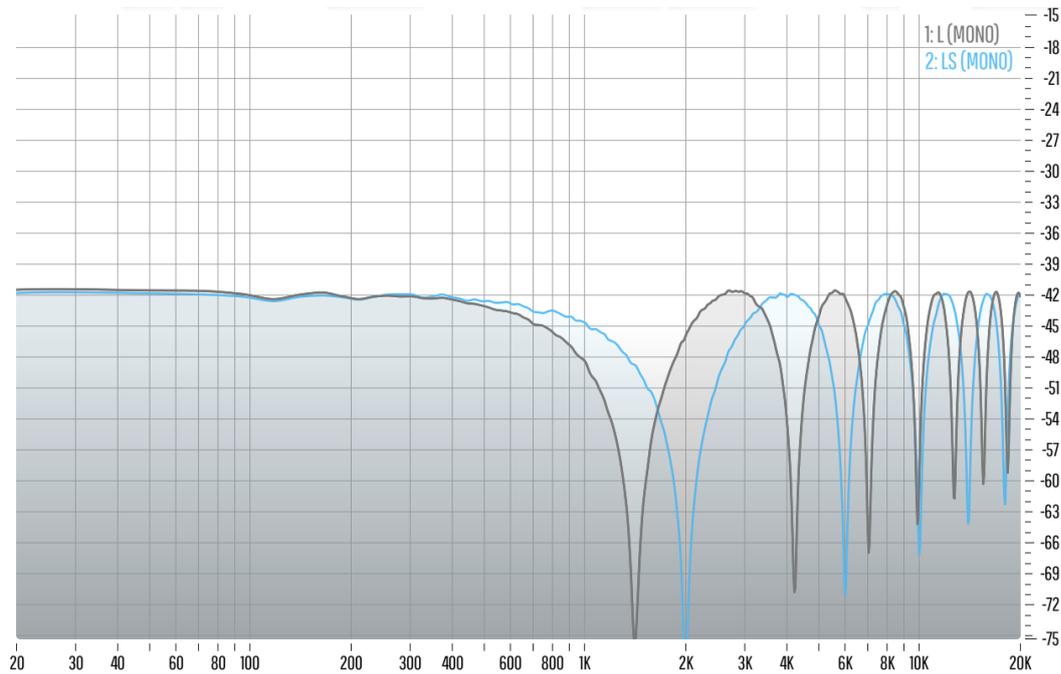
*Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung von MPEG-H Objekten*

**Decca Cuboid**

Bei der Simulation des Decca Cuboids mit einem Schalleinfallswinkel von 0° ergeben sich durch den Abstand des Upper Layers zum Middle Layer und den daraus resultierenden unterschiedlichen Delaywerten bei der Summierung von L und Lh, R und Rh, Ls und Lsh sowie Rs und Rsh im 5.0-Mix Kammfiltereffekte auf den Kanälen L, R, Ls und Rs. Diese sind jeweils auf der linken und rechten Seite identisch. Abbildung 59 zeigt die spektrale Analyse des linken Surround Kanals und den linken Kanals des 5.0-Mixes.

**Abbildung 59**

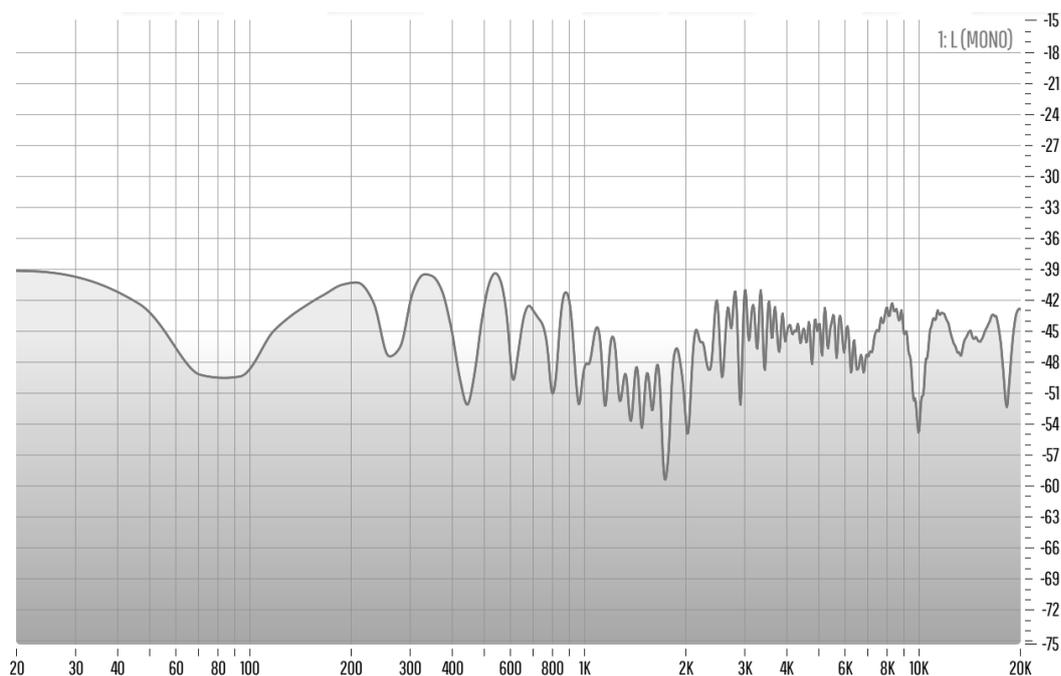
Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von  $0^\circ$  unter Verwendung von MPEG-H Objekten



Bei der Stereowiedergabe ergeben sich erneut Interferenzen auf dem linken und rechten Kanal, welche jedoch nicht die Struktur eines Kammfilters aufweisen. Abbildung 60 zeigt den linken Kanal des Stereomixes.

**Abbildung 60**

Frequenzspektrum des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von  $0^\circ$  unter Verwendung von MPEG-H Objekten

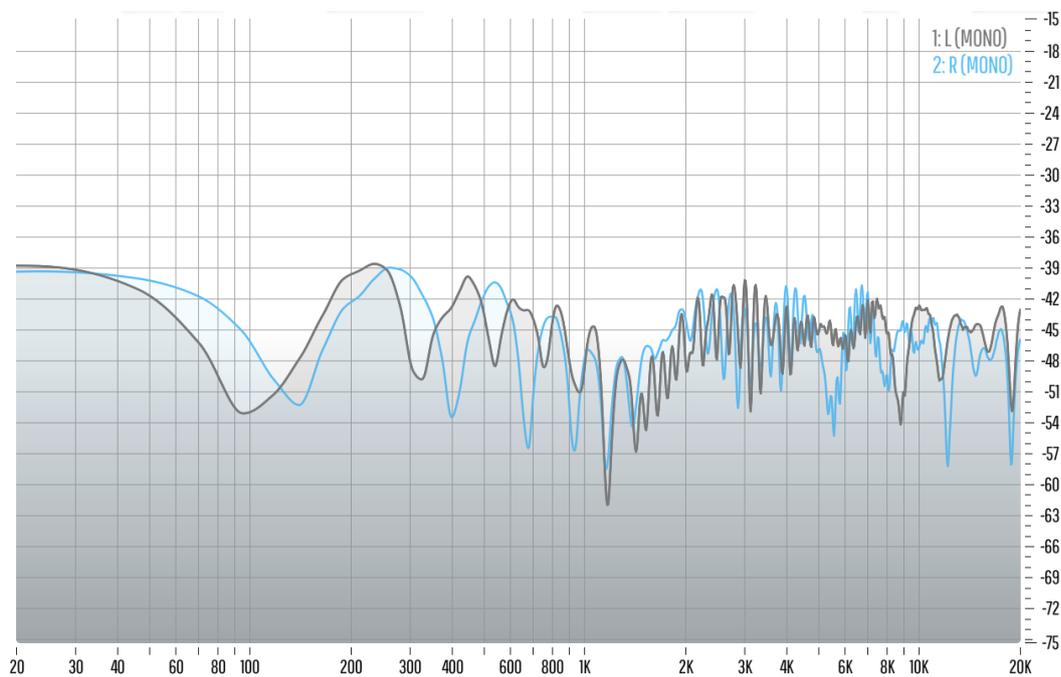


Bei der zweiten Schallquellenposition ergibt sich im 5.0-Mix auf den Kanälen L, R, Ls und Rs jeweils die gleiche Kammfilterstruktur, welche wie in den vorangegangenen Kapiteln abhängig von den Delaywerten der summierten Kanäle im Frequenzspektrum verschoben ist. Die Struktur der Kammfilter ist mit der aus Abb. 59 identisch und die Lage der Maxima und Minima lässt sich wie bereits erläutert berechnen.

Die Frequenzspektren der Kanäle im Stereodownmix sind in Abb. 61 dargestellt.

### Abbildung 61

*Frequenzspektren der Kanäle des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 135° unter Verwendung von MPEG-H Objekten*



Hier ist keine typische Struktur eines Kammfilters zu sehen.

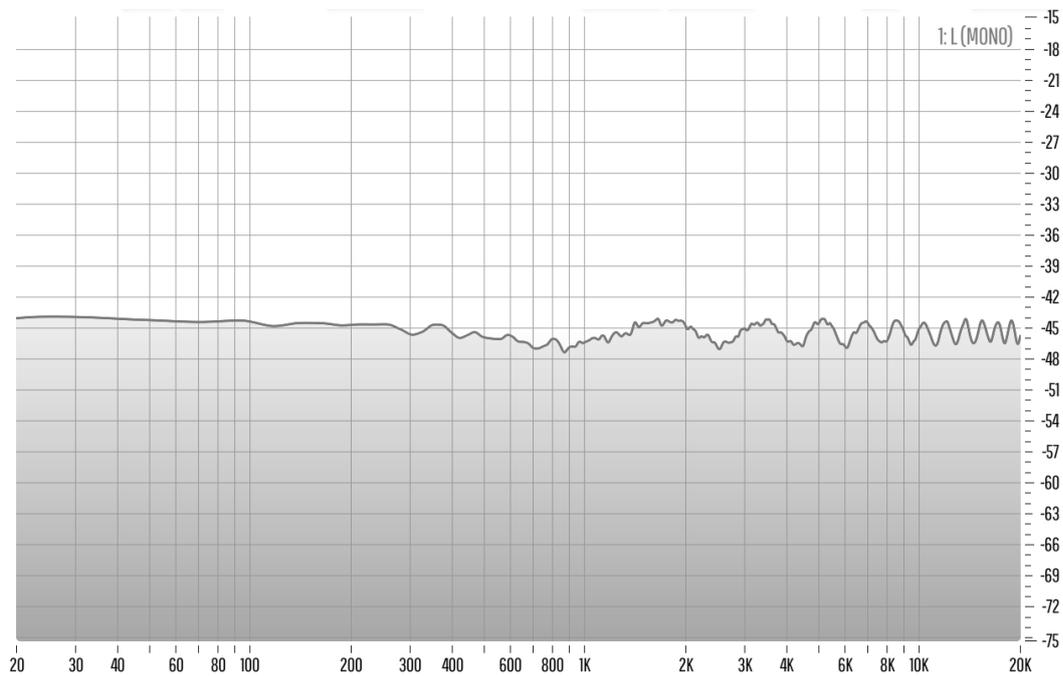
### ORTF-3D

Für das Modell des ORTF-3D gilt aufgrund der vertikal koinzidenten Anordnung der Mikrofone dasselbe wie für das PCMA-3D: Im 5.0-Mix kommt es unabhängig von der Schalleinfallrichtung auf keinem Kanal zu Kammfiltern.

Abbildung 62 zeigt für die erste Schallquellenposition (Schalleinfall von 0°) die spektrale Analyse des linken Kanals im Stereomix.

**Abbildung 62**

*Frequenzspektrum des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von MPEG-H Objekten*

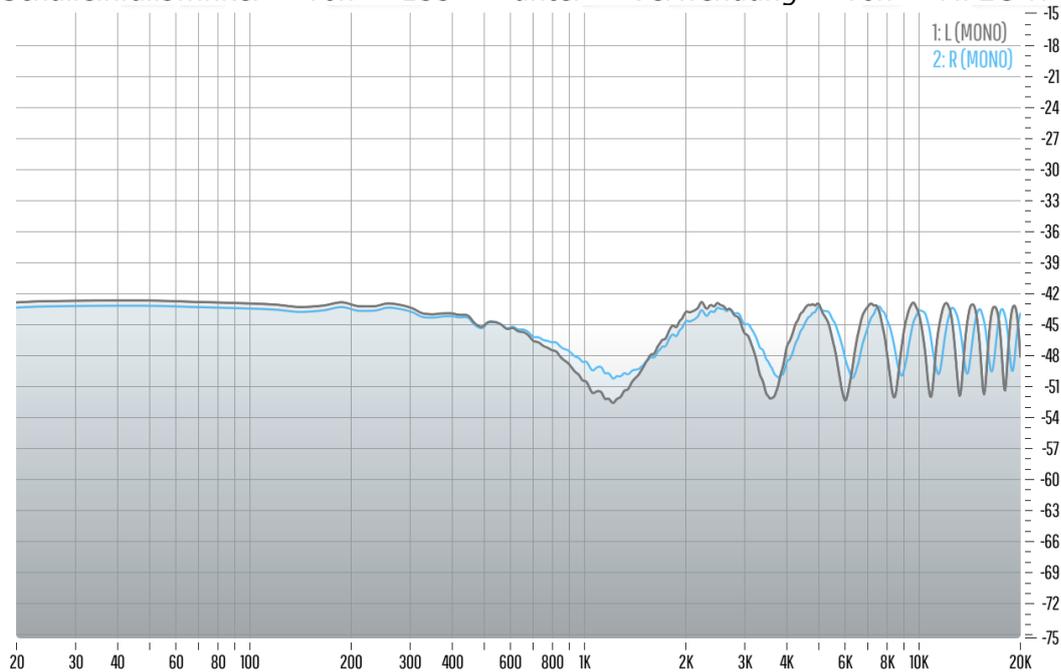


Hier sind ab einer Frequenz von 1 kHz kammfilterartige Strukturen zu erkennen.

Für die zweite Schalleinfallrichtung (135°) ergeben sich im Stereomix auf beiden Kanälen ebenfalls Kammfiltereffekte. Da diese durch den seitlichen Schalleinfall links und rechts nicht übereinstimmen, ist in Abbildung 63 jeweils die spektrale Analyse des linken als auch des rechten Kanals dargestellt.

**Abbildung 63**

*Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 135° unter Verwendung von MPEG-H Objekten*



## 7.5. Zusammenfassung der Ergebnisse

Beim Vergleich der mit den Dolby Atmos bzw. MPEG-H Downmixalgorithmen sowie den jeweiligen Object-Renderern erstellten 5.0- und Stereomixe in Bezug auf Kammfiltereffekte kann Folgendes festgestellt werden:

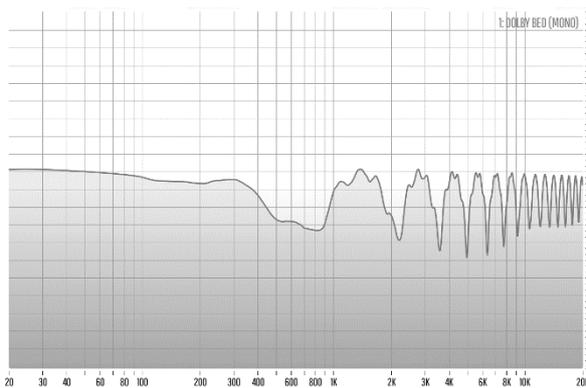
Ein entscheidender Unterschied liegt in der größtmöglichen Kanalanzahl im Upper Layer des Ausgangsformats. Während im MPEG-H Bed und bei Verwendung von MPEG-H Objekten sowie Dolby Objekten 4 Kanäle zur Verfügung stehen und damit jedes simulierte Array-Signal auf einen separaten Lautsprecher geroutet werden konnte, ist im Dolby Bed für den Upper Layer lediglich ein Stereopaar vorgesehen. Die Signale werden jeweils auf der rechten und linken Seite summiert. In den betrachteten Simulationen werden dabei Ltf und Ltr sowie Rtf und Rtr summiert. Dadurch entstehen Kammfilter auf den Kanälen des Upper Layers, TI und Tr. Bei welcher Frequenz dieser Kammfilter beginnt kann aus dem Delayunterschied zwischen Ltf und Ltr bzw. Rtf und Rtr berechnet werden. Wie zuvor bereits erläutert, liegt bei der Frequenz, deren Wellenlänge dem Verzögerungsweg entspricht, das erste Maximum des Kammfilters. Weitere Maxima folgen bei ganzzahligen Vielfachen dieser Frequenz. Das erste Minimum liegt bei der

Frequenz, deren Wellenlänge dem halben Verzögerungsweg entspricht. Um im Ausgangsformat entstehende Kammfilter im Upper Layer und die damit verbundenen Auswirkungen auf die Downmixe zu verhindern, müssen entweder Ltf und Rtf oder Ltr und Rtr im 5.0.4-Mix weggelassen werden.

Der bei Dolby Atmos für kanalbasierte Inhalte zum Einsatz kommende Downmixalgorithmus sowie der in Dolby Atmos integrierte Object-Renderer führen im Stereodownmix bei Verwendung derselben Kanäle des Upper Layers im Ausgangsformat zu denselben Ergebnissen bezüglich der Kammfilterstruktur. Dies wird exemplarisch an den folgenden Abbildungen 64 und 65 gezeigt. Diese zeigen jeweils den linken Kanal des Stereomixes aus der Simulation des PCMA-3D bei Verwendung aller Signale im Ausgangsformat.

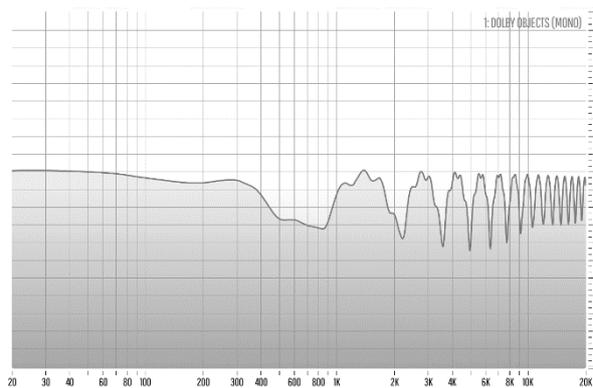
**Abbildung 64**

*Stereodownmix des PCMA-3D unter Verwendung des Dolby Atmos Beds*



**Abbildung 65**

*Stereodownmix des PCMA-3D unter Verwendung von Dolby Atmos Objekten*



Die hier veranschaulichte Übereinstimmung der Frequenzspektren konnte in allen betrachteten Simulationen festgestellt werden.

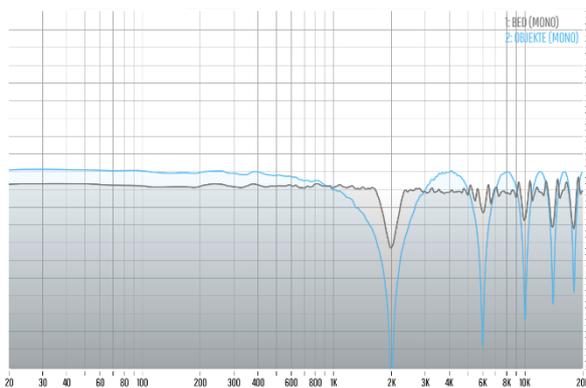
Im 5.0-Mix unterscheiden sich die Ergebnisse aus dem Downmix kanalbasierter Inhalte und Objekte bei Dolby Atmos durch die unterschiedliche Handhabung des Upper Layers. Durch die Positionierung der Objekte anhand der übertragenen Metadaten werden bei objektbasierten Inhalten jeweils übereinanderliegende Kanäle summiert, während im kanalbasierten Downmix alle Signale des Upper Layers auf die Surround-Kanäle addiert werden. Dadurch ergeben sich insbesondere zwei Unterschiede: (i) Beim Decca Cuboid als vertikal nicht koinzidentes Array kommt es bei der Verwendung von Objekten durch den

Unterschied der Delaywerte der jeweils übereinanderliegenden Kanäle zu Kammfiltereffekten auf den Kanälen L und R, während sich im 5.0-Downmix des kanalbasierten Beds in keinem Szenario Kammfiltereffekte auf den vorderen Kanälen ergeben. (ii) Bei vertikal koinzidenten Arrays bilden sich bei der Verwendung von Objekten durch die ausschließliche Summierung von Signalen mit übereinstimmendem Delaywert im 5.0-Mix keine Kammfiltereffekte. Beim Downmix von kanalbasierten Inhalten kann es, wie oben erläutert, abhängig von den im Upper Layer des Ausgangsformats verwendeten Signalen auch bei vertikal koinzidenten Arrays zu Kammfiltern auf den Surround-Kanälen des 5.0-Downmixes kommen. Dies ist unter Beteiligung von Ltf und Rtf im Ausgangsformat der Fall.

Beim Vergleich der aus dem aktiven Downmixalgorithmus resultierenden Ergebnissen des kanalbasierten MPEG-H Beds mit den Ergebnissen nach Verwendung von MPEG-H Objekten konnte folgendes beobachtet werden: Obwohl jeweils dieselben Signale bei der Überführung der Kanäle in das Format 5.0- bzw. 2.0 summiert wurden, kommt es zu Unterschieden in den resultierenden Frequenzspektren. Dies konnte in allen Modellen beobachtet werden und soll beispielhaft in den Abbildungen 66 und 67 dargestellt werden. Abbildung 66 zeigt jeweils den linken Surroundkanal des 5.0-Mixes in der Simulation des Decca Cuboids. Abbildung 67 zeigt jeweils den linken Kanal des Stereomixes im Model des ORTF-3D mit einem Schalleinfallswinkel von 135°.

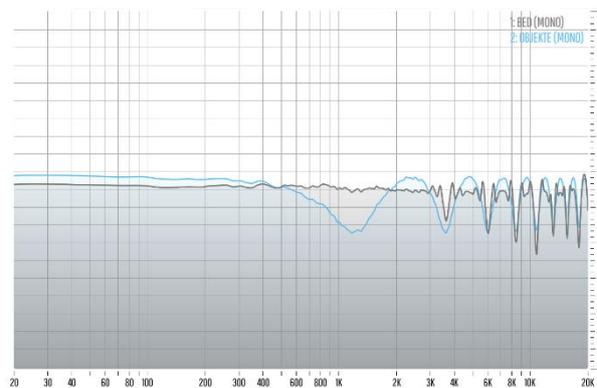
### Abbildung 66

*Gegenüberstellung der Downmixe des MPEG-H Beds und MPEG-H Objekten am Beispiel des Decca Cuboids*



### Abbildung 67

*Gegenüberstellung der Downmixe des MPEG-H Beds und MPEG-H Objekten am Beispiel des ORTF-3D*

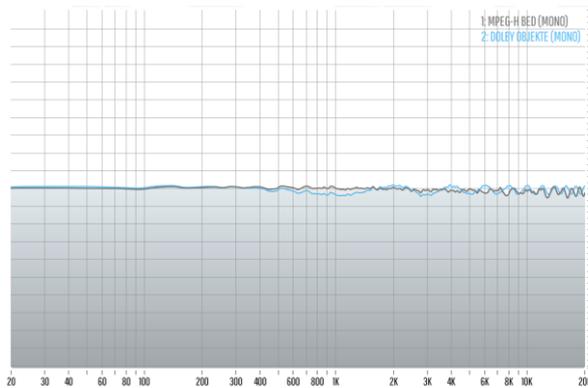


Die Unterschiede sind auf die aktive Komponente des Format Converters, welche beim Rendering von Objekten fehlt, zurückzuführen. Durch die hauptsächlichliche Verwendung von MPEG-H im Broadcastbereich sind Objekte hier weniger zur Positionierung einzelner Audio-Events im Raum als zur Übertragung alternativer Audiospuren vorgesehen. Von dieser Bestimmung abzuweichen ist in Hinblick auf die entstehenden Kammfiltereffekte nicht empfehlenswert.

Um die Unterschiede zwischen den beiden Produktionsumgebungen deutlich zu machen, werden in den folgenden Abbildungen jeweils die Stereodownmixe, welche im Gegensatz zu den 5.0-Downmixen in jedem Modell von Kammfiltereffekten betroffen sind, des MPEG-H Beds mit Ergebnissen des Dolby Object Renderers aller Simulationen gegenübergestellt. Die Ergebnisse des Object Renderers stimmen wie bereits gezeigt für das Format 2.0 mit den Ergebnissen des kanalbasierten Downmixes bei Dolby Atmos überein.

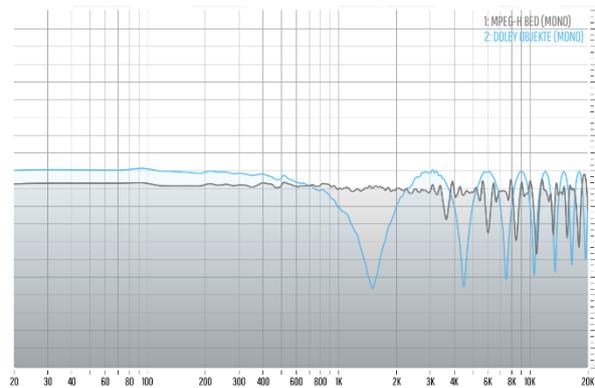
#### Abbildung 68

*Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (ORTF-3D, 0°)*



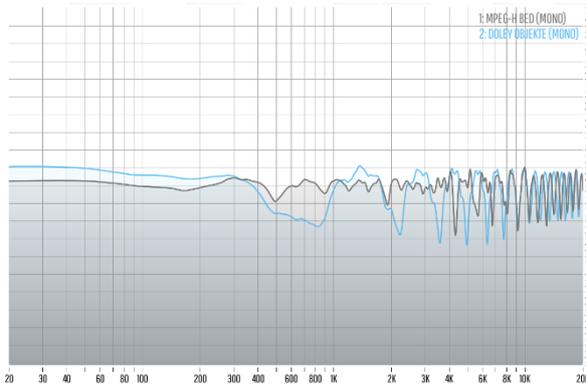
#### Abbildung 69

*Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (ORTF-3D, 135°)*



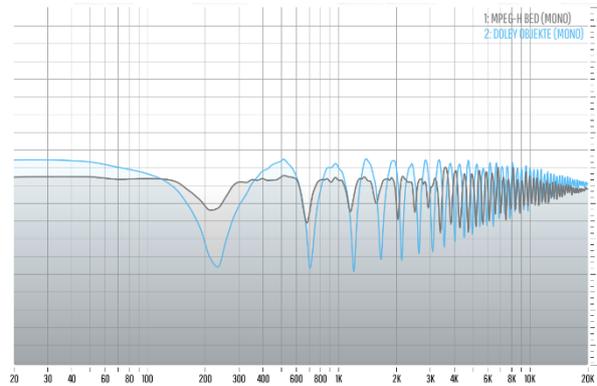
**Abbildung 70**

Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (PCMA-3D, 0°)



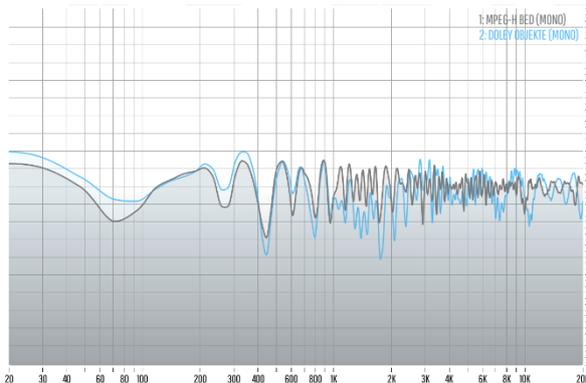
**Abbildung 71**

Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (PCMA-3D, 135°)



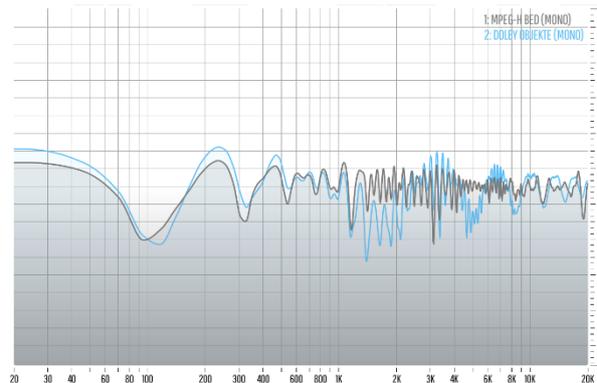
**Abbildung 72**

Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (Decca Cuboid, 0°)



**Abbildung 73**

Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (Decca Cuboid, 135°)

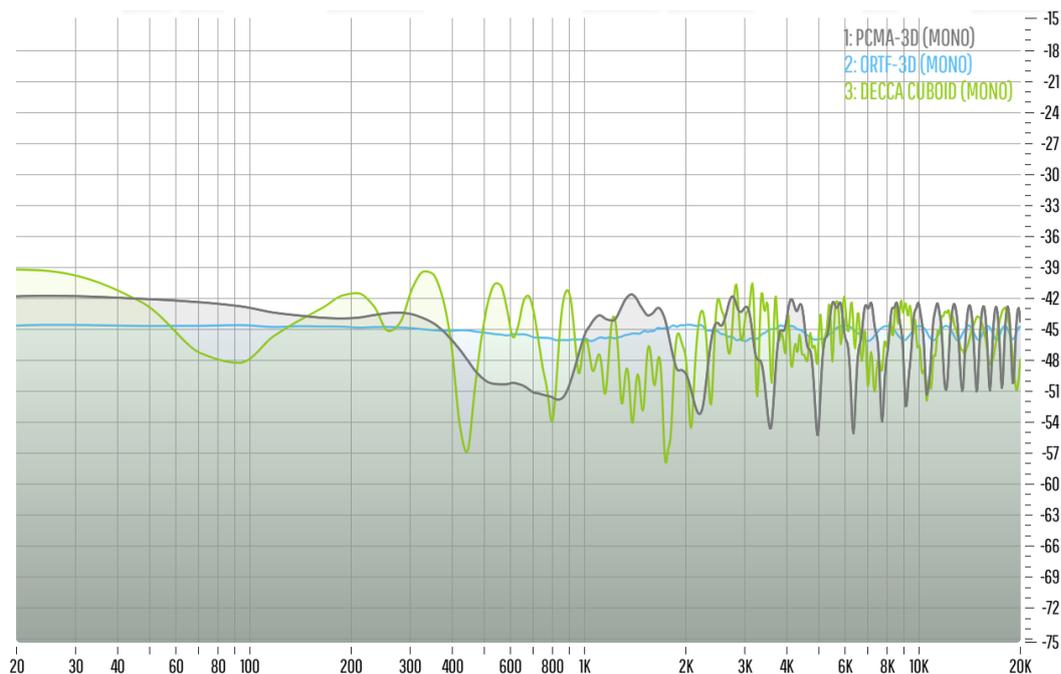


Die Abbildungen 68-71 zeigen, dass durch den aktiven Downmixalgorithmus Kammfilterstrukturen ausgebessert werden. Dies funktioniert umso besser, je kleiner die Abstände zwischen den Mikrofonen eines Arrays sind. Beim Decca Cuboid als dem Array mit den größten räumlichen Abmessungen ist, wie in den Abbildung 72 und 73 zu sehen ist, der Unterschied zwischen Dolby Atmos und MPEG-H am geringsten. An dieser Stelle sei auch darauf hingewiesen, dass auch der Unterschied zwischen den Frequenzspektren nach Verwendung des MPEG-H Beds oder MPEG-H Objekten bei den Modellen des Decca Cuboids sehr gering ausfällt.

Die Auswirkungen der unterschiedlichen 3D-Hauptmikrofonarrays auf das Auftreten von Kammfiltereffekten im Downmix lassen sich demnach folgendermaßen zusammenfassen: Je größer die Abstände zwischen den Mikrofonen und damit die Laufzeitunterschiede zwischen den im Downmix summierten Kanälen, desto tiefer beginnen Kammfilter bzw. desto tiefer liegt die erste sich ergebende destruktive Interferenz im Frequenzspektrum. Dies lässt sich bei allen verglichenen Downmixvarianten feststellen und bereits im Ausgangsformat 5.0.4 bei Verwendung aller Signale des Upper Layers im Dolby Bed beobachten, was ein Vergleich der Abbildungen 18, 26 und 34 zeigt. Zudem soll die folgende Abbildung 74, welche jeweils den linken Kanal des Dolby Object Stereomixes aller Simulationen mit einem Schalleinfallswinkel von  $0^\circ$  zeigt, diesen Zusammenhang nochmals illustrieren.

### Abbildung 74

*Auswirkungen des Mikrofonabstandes auf Frequenzspektrum des Stereodownmixes*

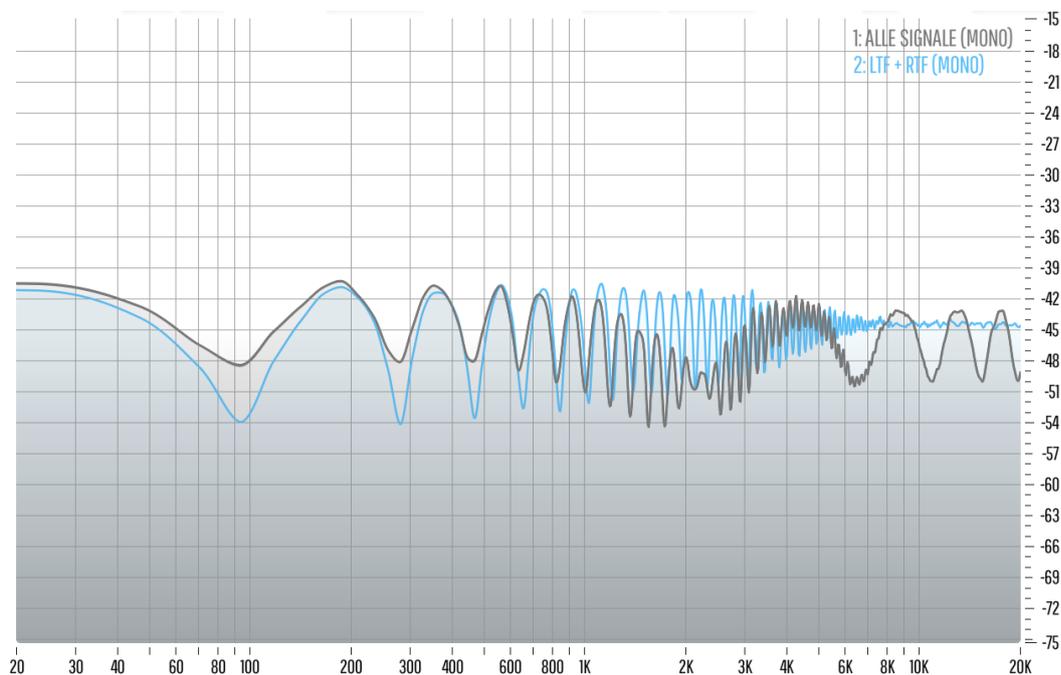


Außerdem ist an dieser Stelle anzumerken, dass es (außer bei Verwendung aller Signale im Upper Layer des Ausgangsformats im Dolby Bed) bei vertikal koinzidenten Mikrofonanordnungen nicht zu Kammfiltern im 5.0 Downmix kommt. Falls es sich beim gewählten Zielformat um 5.0 handelt, ist demnach davon abzuraten mit Dolby Atmos kanalbasiert zu arbeiten.

Der aus den verschiedenen Richtcharakteristika der Mikrofone resultierende Pegel führt bei Summierung der Signale zu unterschiedlichen Pegeldifferenzen zwischen Maxima und Minima des Kammfilters (Dickreiter et al., 2023). Dies ist nach mehreren Downmixschritten und den damit verbundenen Verminderungen des Pegels der ursprünglichen Signale meist nicht deutlich zu erkennen. Abbildung 75 zeigt daher den linken Surround Kanal des kanalbasierten 5.0-Downmixes der Simulation des Decca Cuboids mit einem Schalleinfallswinkel von 0°. Hier wurden im Ausgangsformat einmal alle vier Signale des Upper Layers verwendet (graue Linie) und einmal nur die vorderen beiden (blaue Linie). Durch die Nutzung ungerichteter Mikrofone weisen alle Kanäle des UL denselben Pegel auf, d.h. bei Nutzung aller Signale ist ein um 3 dB höherer Pegel im Upper Layer zu erwarten. Der Einfluss der hinteren Kanäle des UL zeigt sich im 5.0 Mix, wie in Abb.27 zu erkennen, ab einer Frequenz von 1 kHz. Unterhalb dieser Frequenz sind die Unterschiede in den Ausprägungen der Kammfilter also auf die unterschiedlichen Pegel der summierten Signale zurückzuführen.

### Abbildung 75

*Auswirkung des Pegels der Ausgangskanäle auf die Amplitude des im Downmix entstehenden Kammfilters*



Dieser Zusammenhang kann auch beim Vergleich der Abbildungen 23, 24 und 25 sowie der Abbildungen 37, 38 und 39 erkannt werden. Durch die seitliche Schallquellenposition und die Richtcharakteristika ergeben sich hier jeweils unterschiedliche Pegel auf den beiden Seiten, was im Stereodownmix zu einer unterschiedlichen Amplitude der Kammfilter auf den Kanälen L und R führt. Nicht zuletzt können die Auswirkungen des Pegelunterschiedes auf die Kammfilterstruktur am Beispiel des 5.0-Downmixes des PCMA-3D mit einem Schalleinfallswinkel von  $0^\circ$  beobachtet werden: Hier haben die Kanäle Ls und Rs des Ausgangsformats keinen Einfluss auf sich ergebende Kammfilterstrukturen im Downmix, da diese einen sehr geringen Pegel von ca. -53 dB aufweisen. Des Weiteren zeigt der Vergleich der Abbildungen 40 und 41 den großen Einfluss des Centers auf das Frequenzspektrum des 2.0-Downmixes bei einer Schalleinfallrichtung von  $0^\circ$ . Dies ist auf den hohen Pegel des Centerkanals zurückzuführen.

Zuletzt muss erneut auf die grundsätzlichen Unterschiede bezüglich der empfohlenen Lautsprecheraufstellung zwischen den Produktionsformaten Dolby Atmos und MPEG-H hingewiesen werden. Diese haben keine Auswirkung auf das Auftreten von Kammfiltereffekten im Downmix, sollen jedoch aus Gründen der Vollständigkeit nicht unerwähnt bleiben. Die verschiedenen Aufstellungen sind in den Abbildungen eins, und zwei sowie vier und fünf dargestellt.

## **8. Grenzen der Methodik**

Die in dieser Arbeit gewonnenen Erkenntnisse sind stets unter Beachtung der angewandten Methodik zu verstehen. Durch die für die Simulationen getroffenen Bedingungen lassen sich die Ergebnisse nicht direkt auf akustische Musikaufnahmen übertragen. Dadurch, dass die Frequenzabhängigkeit der Richtcharakteristik der Mikrofone und der laufzeitabhängige Pegelabfall nicht in die Modelle mit einbezogen wurden, ergeben sich im Downmix von realen Aufnahmen abweichende Ergebnisse. Dasselbe gilt für die Tatsache, dass Reflexionen außer Acht gelassen wurden, womit Freifeld-Bedingungen simuliert werden.

Zudem wurden mit der Höhe der simulierten Schallquelle auf einer Ebene mit dem Middle Layer des Arrays jeweils Extremwerte für die Delayunterschiede zwischen den einzelnen Kanälen auf dieser Ebene ermittelt. Befindet sich die Schallquelle auf einer anderen Höhe, ist mit geringeren Unterschieden zwischen den jeweiligen Delaywerten zu rechnen, was sich auf die Struktur der resultierenden Kammfiltereffekte auswirkt.

Außerdem lässt die Untersuchung der Downmixergebnisse auf entstehende Kammfilter andere eventuell ungewollte Phänomene unbeleuchtet. Dies kann am Beispiel des aktiven MPEG-H Downmixalgorithmus erläutert werden: Die damit erstellten Downmixe weisen zwar keine typische Kammfilterstruktur auf, dennoch tritt auch hier eine klangliche Veränderung des Ausgangsmaterials auf.

## 9. Praktische Anwendung

Um die mit den Simulationen ermittelten Ergebnisse auf ein praktisches Beispiel auszuweiten, wurde zum Abschluss mit den verglichenen Verfahren Downmixe einer akustischen 3D-Orchesteraufnahme angefertigt. Bei dem Orchester handelt es sich um das Hochschulsinfonieorchester der HMDK Stuttgart. Als Hauptmikrofon-Array wurde ein PCMA-3D verwendet. Die Signale wurden analog zu den Testsignalen aus den Modellen geroutet und nicht in ihrem Pegel verändert. Im elektronischen Anhang befinden sich als Hörbeispiele von jeder der Downmixvarianten jeweils ein Ausschnitt des Stereodownmixes.

Hier werden die Grenzen der Methodik deutlich. Während in der Simulation des PCMA-3D bei einer Schalleinfallrichtung von  $0^\circ$  die hinteren Kanäle des Middle Layers aufgrund der Nierencharakteristik und der Ausrichtung der Mikrofone einen Pegel von -53 dB aufweisen und damit keinen Einfluss auf den Gesamtklang haben, verzeichnen diese Kanäle bei einer realen Aufnahme hörbaren Pegel und haben damit Einfluss auf die Downmixe.

An dieser Stelle musste festgestellt werden, dass die Downmixergebnisse der vorliegenden Orchesteraufnahme keine weiteren Erkenntnisse für die Forschungsfrage dieser Arbeit liefern. Der erste Höreindruck der Downmixe liefert zwar kammfilterähnliche klangliche Veränderungen, diese lassen sich jedoch durch eine bloße Analyse des Frequenzspektrums nicht aussagekräftig nachweisen. Zudem zeigt der erste Höreindruck eine Veränderung des Verhältnisses von Direkt- und Diffussschall mit jeder Verringerung der Kanalanzahl. Dies müsste anhand weiterer Modelle und mit unterschiedlichen Aufnahmen sowie Analysen bezüglich der Korrelation bzw. Kohärenz der Signale verifiziert werden und wurde im Rahmen dieser Arbeit nicht weiter verfolgt.

## **10. Ausblick**

Die vorliegende Arbeit beleuchtete die Auswirkungen verschiedener 3D-Hauptmikrofonarrays sowie unterschiedlicher Downmixverfahren auf die Bildung von Kammfiltereffekten im 5.0- sowie Stereodownmix mithilfe verschiedener Simulationen mit einem Testsignal. Dabei konnte gezeigt werden, dass der Laufzeitunterschied zwischen den Mikrofonkapseln eines Arrays die größte Auswirkung auf die Entstehung von Kammfiltern hat. Außerdem wurde der Nutzen aktiver Downmixalgorithmen gegenüber passiven Downmixalgorithmen herausgestellt. Wie Kapitel 9 zeigt, sind diese Erkenntnisse nicht auf reale Aufnahmen übertragbar. Um die Resultate mit Downmixergebnissen akustischer Aufnahmen in Bezug setzen zu können, sind Ausweitungen der Simulationsbedingungen nötig. Nützlich wären unter anderem eine Translation der Modelle in Diffusfeld-Bedingungen sowie die Einbeziehung der Frequenzabhängigkeit der Richtcharakteristika der Mikrofone. Zudem wurden in dieser Arbeit alle Downmixvarianten mit den Default-Einstellungen verwendet. Die Nutzung anderer Voreinstellungen oder manueller Koeffizienten kann die gewonnenen Ergebnisse ergänzen.

## 11. Literaturverzeichnis

- Adami, A., Habets, E. A. P. & Herre, J [Jurgen] (2014). Down-mixing using coherence suppression. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014): Florence, Italy, 4 - 9 May 2014* (S. 2878–2882). IEEE. <https://doi.org/10.1109/ICASSP.2014.6854126>
- Advanced Television Systems Committee. *ATSC Standard: ATSC 3.0 System (A/300:2023-03)*.
- Bleidt, R. L., Sen, D., Niedermeier, A., Czelhan, B., Fug, S., Disch, S., Herre, J [Jurgen], Hilpert, J., Neuendorf, M., Fuchs, H., Issing, J., Murtaza, A., Kuntz, A., Kratschmer, M., Kuch, F., Fug, R., Schubert, B., Dick, S., Fuchs, G., . . . Kim, M.-Y. (2017). Development of the MPEG-H TV Audio System for ATSC 3.0. *IEEE Transactions on Broadcasting*, 63(1), 202–236. <https://doi.org/10.1109/TBC.2017.2661258>
- Dickreiter, M., Dittel, V., Hoeg, W. & Wöhr, M. (Hrsg.). (2023). *De Gruyter reference. Handbuch der Tonstudioteknik* (9. Aufl.). de Gruyter Saur.
- Dolby. (2022a, 18. Juni). *How do the 5.1 and Stereo downmix settings work?* [https://professionalsupport.dolby.com/s/article/How-do-the-5-1-and-Stereo-downmix-settings-work?language=en\\_US](https://professionalsupport.dolby.com/s/article/How-do-the-5-1-and-Stereo-downmix-settings-work?language=en_US)
- Dolby. (2022b, 23. Juni). *Are there ways to distribute Dolby Atmos Music as lossless audio?* [https://professionalsupport.dolby.com/s/article/Are-there-ways-to-distribute-Dolby-Atmos-Music-as-lossless-audio?language=en\\_US](https://professionalsupport.dolby.com/s/article/Are-there-ways-to-distribute-Dolby-Atmos-Music-as-lossless-audio?language=en_US)
- Dolby. (2022c, 24. Juni). *Is Dolby AC-4 more efficient than Dolby Digital or Dolby Digital Plus?* [https://professionalsupport.dolby.com/s/article/Is-Dolby-AC-4-more-efficient-than-Dolby-Digital-or-Dolby-Digital-Plus?language=en\\_US](https://professionalsupport.dolby.com/s/article/Is-Dolby-AC-4-more-efficient-than-Dolby-Digital-or-Dolby-Digital-Plus?language=en_US)
- Dolby. (2022d, 24. Juni). *Is Dolby Digital Plus with Dolby Atmos standardized?* [https://professionalsupport.dolby.com/s/article/Is-Dolby-Digital-Plus-with-Dolby-Atmos-standardized?language=en\\_US](https://professionalsupport.dolby.com/s/article/Is-Dolby-Digital-Plus-with-Dolby-Atmos-standardized?language=en_US)
- Dolby. (2022e, 25. Juni). *How many objects can I/should I use?* [https://professionalsupport.dolby.com/s/article/How-many-objects-can-I-should-I-use?language=en\\_US](https://professionalsupport.dolby.com/s/article/How-many-objects-can-I-should-I-use?language=en_US)
- Dolby. (2022f, 25. Juni). *What is a bed?* [https://professionalsupport.dolby.com/s/article/What-is-a-bed?language=en\\_US](https://professionalsupport.dolby.com/s/article/What-is-a-bed?language=en_US)
- Dolby. (2023, 24. Juni). *Dolby Atmos | Dolby AC4 UK*. <https://dolbyac4.com/de/atmos>
- European Telecommunications Standards Institute (2018a-02). *Digital Audio Compression (AC-4) Standard: Part 1: Channel based coding* (ETSI TS 103 190-1 V1.3.1).
- European Telecommunications Standards Institute (2018b-02). *Digital Audio Compression (AC-4) Standard: Part 2: Immersive and personalized audio* (ETSI TS 103 190-2 V1.2.1).
- European Telecommunications Standards Institute (2018c-10). *Backwards-compatible object audio carriage using Enhanced AC-3* (ETSI TS 103 420 V1.2.1).
- Faller, C. & Baumgarte, F. (2003). Binaural cue coding-part II: schemes and applications. *IEEE Transactions on Speech and Audio Processing*, 11(6), 520–531. <https://doi.org/10.1109/TSA.2003.818108>
- Fraunhofer-Institut für Integrierte Schaltungen IIS. (2022). *MPEG-H Authoring Suite: User Guide for Version 5.0.0*.
- Fraunhofer-Institut für Integrierte Schaltungen IIS. (2023, 27. Juni). *MPEG-H Audio*. <https://www.iis.fraunhofer.de/de/ff/amm/rundfunk-streaming/mpegh.html>

- Grewe, Y., Eibl, P., Rieger, D., Torcoli, M., Simon, C. & Scuda, U. (Hrsg.) (2021). *MPEG-H Audio Production Workflows for a Next Generation Audio Experience in Broadcast, Streaming, and Music*.
- Grewe, Y., Scuda, U., Murtaza, A., Kaheelin, M. & Duare Nuno. (2020). *STUDIO RECOMMENDATIONS Studio Recommendations for 3D-Audio Productions with MPEG-H Audio*. Fraunhofer IIS.
- Herre, J [Jürgen] & al, e. (2015). MPEG-H Audio—The New Standard for Universal Spatial/3D Audio Coding. *Journal of the Audio Engineering Society*, 62(12), 821–830. <https://doi.org/10.17743/jaes.2014.0049>
- Herre, J [Jürgen] & Quackenbush, S. R. (2022). MPEG-H 3D Audio: Immersive Audio Coding. *Acoustical Science and Technology*, 43(2), 143–148. <https://doi.org/10.1250/ast.43.143>
- Lee, H. K. & Gibben, C. (2015). Effect of Vertical Microphone Layer Spacing for a 3D Microphone Array. *Journal of the Audio Engineering Society*, 62(12), 870–884. <https://doi.org/10.17743/jaes.2014.0045>
- Lee, H. (2021). Multichannel 3D Microphone Arrays: A Review. *Journal of the Audio Engineering Society*, 69(1/2), 5–26. <https://doi.org/10.17743/jaes.2020.0069>
- Lee, H. & Johnson, D. (2019). *3D Microphone Array Recording Comparison (3D-MARCo)*. <https://doi.org/10.5281/ZENODO.3477602>
- Simon, C., Grewe, Y., Faecks, N. & Scuda, U. (2018). Field Tests for Immersive and Interactive Broadcast Audio Production using MPEG-H 3D Audio. *SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING*, 2018(1), 40–46. <https://doi.org/10.18580/setijbe.2018.5>
- Vilkamo, J., Kuntz, A. & Füg, S. (2014). Reduction of Spectral Artifacts in Multichannel Downmixing with Adaptive Phase Alignment. *Journal of the Audio Engineering Society*, 62(7/8), 516–526. <https://doi.org/10.17743/jaes.2014.0027>
- Wallis, R. & Lee, H. (2015). The Effect of Interchannel Time Difference on Localization in Vertical Stereophony. *Journal of the Audio Engineering Society*, 63(10), 767–776. <https://doi.org/10.17743/JAES.2015.0069>
- Wallis, R. & Lee, H. (2017). The Reduction of Vertical Interchannel Crosstalk: The Analysis of Localisation Thresholds for Natural Sound Sources. *Applied Sciences*, 7(3), 278. <https://doi.org/10.3390/app7030278>
- Witteck, H. & Theile, G. (Hrsg.) (2017). *Development and application of a stereophonic multichannel recording technique for 3D Audio and VR*.

## 12. Abbildungsverzeichnis

Abbildung 1	<i>Lautsprecheranordnung bei Dolby Atmos im Format 5.0.4, Draufsicht</i> .....	4
Abbildung 2	<i>Lautsprecheranordnung bei Dolby Atmos im Format 5.0.4, Seitenansicht</i> .....	5
Abbildung 3	<i>Beispiel für eine MPEG-H Audio Szene</i> .....	8
Abbildung 4	<i>Lautsprecheranordnung bei MPEG-H im Format 5.0+4H, Draufsicht</i> .....	9
Abbildung 5	<i>Lautsprecheranordnungen bei MPEG-H im Format 5.0+4H, Seitenansicht</i> .....	10
Abbildung 6	<i>Mikrofonanordnung im PCMA-3D</i> .....	18
Abbildung 7	<i>PCMA-3D: Darstellung der Richtcharakteristika, Draufsicht</i> .....	19
Abbildung 8	<i>PCMA-3D: Darstellung der Richtcharakteristika, Seitenansicht</i> .....	19
Abbildung 9	<i>PCMA-3D: 3D-Plot des Arrays</i> .....	20
Abbildung 10	<i>Mikrofonanordnungen im Decca Cuboid, Draufsicht (links) und Seitenansicht (rechts)</i>	21
Abbildung 11	<i>Decca Cuboid: Darstellung der Richtcharakteristika, Draufsicht</i> .....	22
Abbildung 12	<i>Decca Cuboid: Darstellung der Richtcharakteristika, Seitenansicht</i> .....	22
Abbildung 13	<i>Decca Cuboid: 3D-Plot des Arrays</i> .....	23
Abbildung 14	<i>Mikrofonanordnungen im ORTF-3D, Draufsicht (links) und Seitenansicht (rechts)</i> .....	24
Abbildung 15	<i>ORTF-3D: Darstellung der Richtcharakteristika, Draufsicht</i> .....	24
Abbildung 16	<i>ORTF-3D: Darstellung der Richtcharakteristika, Seitenansicht</i> .....	25
Abbildung 17	<i>ORTF-3D: 3D-Plot des Arrays</i> .....	25
Abbildung 18	<i>Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim PCMA-3D bei einem Schalleinfallswinkel von 0°</i> .....	27
Abbildung 19	<i>Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung des Dolby Atmos Bed</i> .....	28
Abbildung 20	<i>Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von unterschiedlichen Ausgangssignalen im Dolby Atmos Bed (ohne C)</i> .....	29
Abbildung 21	<i>Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung unterschiedlicher Ausgangssignalen im Dolby Atmos Bed (mit C)</i> .....	30
Abbildung 22	<i>Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim PCMA-3D bei einem Schalleinfallswinkel von 135°</i> .....	31
Abbildung 23	<i>Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung aller Signale des UL im Dolby Atmos Bed</i> .....	32
Abbildung 24	<i>Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung der Signale Ltf und Rtf des UL im Dolby Atmos Bed</i> .....	32
Abbildung 25	<i>Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung der Signale Ltr und Rtr des UL im Dolby Atmos Bed</i> .....	33
Abbildung 26	<i>Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim Decca Cuboid bei einem Schalleinfallswinkel von 0°</i> .....	34
Abbildung 27	<i>Frequenzspektrum des Kanals Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 0° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds</i> .....	35
Abbildung 28	<i>Frequenzspektrum des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 0° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds</i> .....	36

---

Abbildung 29	<i>Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim Decca Cuboid bei einem Schalleinfallswinkel von 135°</i> .....	37
Abbildung 30	<i>Frequenzspektrum des Kanals Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 135° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds</i> .....	38
Abbildung 31	<i>Frequenzspektrum des Kanals Rs im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 135° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds</i> .....	38
Abbildung 32	<i>Frequenzspektrum des linken Kanals im Stereodownmix des Decca Cuboids bei einem Schalleinfallswinkel von 135° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds</i> .....	40
Abbildung 33	<i>Frequenzspektrum des rechten Kanals im Stereodownmix des Decca Cuboids bei einem Schalleinfallswinkel von 135° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds</i> .....	40
Abbildung 34	<i>Kammfilter auf den Kanälen Tl und Tr des Dolby Atmos Beds beim ORTF-3D bei einem Schalleinfallswinkel von 0°</i> .....	41
Abbildung 35	<i>Frequenzspektrum des Kanals Ls im 5.0-Downmix des ORTF-3D bei einem Schalleinfallswinkel von 0° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds</i> .....	42
Abbildung 36	<i>Frequenzspektrum des Stereodownmix des ORTF-3D bei einem Schalleinfallswinkel von 0° nach Verwendung unterschiedlicher Signalkombinationen im UL des Dolby Atmos Beds</i> .....	43
Abbildung 37	<i>Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 135° unter Verwendung aller Signale des UL im Dolby Atmos Bed</i> .....	44
Abbildung 38	<i>Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 135° unter Verwendung der Signale Ltf und Rtf des UL im Dolby Atmos Bed</i> .....	44
Abbildung 39	<i>Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 135° unter Verwendung der Signale Ltr und Rtr des UL im Dolby Atmos Bed</i> .....	45
Abbildung 40	<i>Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten (ohne C)</i> .....	46
Abbildung 41	<i>Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten (mit C)</i> .....	47
Abbildung 42	<i>Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung von Dolby Atmos Objekten</i> .....	48
Abbildung 43	<i>Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboid bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten</i> .....	49
Abbildung 44	<i>Frequenzspektrum des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten</i> .....	50
Abbildung 45	<i>Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboid bei einem Schalleinfallswinkel von 135° unter Verwendung von Dolby Atmos Objekten</i> .....	51
Abbildung 46	<i>Frequenzspektrum der Kanäle des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 135° unter Verwendung von Dolby Atmos Objekten</i> .....	51
Abbildung 47	<i>Frequenzspektrum des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von Dolby Atmos Objekten</i> .....	52

---

---

Abbildung 48	<i>Frequenzspektrum der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung von Dolby Atmos Objekten</i> .....	53
Abbildung 49	<i>Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung des MPEG-H Beds</i> .....	54
Abbildung 50	<i>Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung des MPEG-H Beds</i> .....	55
Abbildung 51	<i>Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 0° unter Verwendung des MPEG-H Beds</i> .....	56
Abbildung 52	<i>Frequenzspektrum des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 0° unter Verwendung des MPEG-H Beds</i> .....	56
Abbildung 53	<i>Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 135° unter Verwendung des MPEG-H Beds</i> .....	57
Abbildung 54	<i>Frequenzspektren der Kanäle des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 135° unter Verwendung des MPEG-H Beds</i> .....	58
Abbildung 55	<i>Frequenzspektrum des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 0° unter Verwendung des MPEG-H Beds</i> .....	59
Abbildung 56	<i>Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 135° unter Verwendung des MPEG-H Beds</i> .....	60
Abbildung 57	<i>Frequenzspektrum des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von MPEG-H Objekten</i> .....	61
Abbildung 58	<i>Frequenzspektren der Kanäle des Stereodownmixes des PCMA-3D bei einem Schalleinfallswinkel von 135° unter Verwendung von MPEG-H Objekten</i> .....	62
Abbildung 59	<i>Frequenzspektren der Kanäle L und Ls im 5.0-Downmix des Decca Cuboids bei einem Schalleinfallswinkel von 0° unter Verwendung von MPEGH Objekten</i> .....	63
Abbildung 60	<i>Frequenzspektrum des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 0° unter Verwendung von MPEG-H Objekten</i> .....	63
Abbildung 61	<i>Frequenzspektren der Kanäle des Stereodownmixes des Decca Cuboids bei einem Schalleinfallswinkel von 135° unter Verwendung von MPEG-H Objekten</i> .....	64
Abbildung 62	<i>Frequenzspektrum des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 0° unter Verwendung von MPEG-H Objekten</i> .....	65
Abbildung 63	<i>Frequenzspektren der Kanäle des Stereodownmixes des ORTF-3D bei einem Schalleinfallswinkel von 135° unter Verwendung von MPEG-H Objekten</i> .....	66
Abbildung 64	<i>Stereodownmix des PCMA-3D unter Verwendung des Dolby Atmos Beds</i> .....	67
Abbildung 65	<i>Stereodownmix des PCMA-3D unter Verwendung von Dolby Atmos Objekten</i> .....	67
Abbildung 66	<i>Gegenüberstellung der Downmixe des MPEG-H Beds und MPEG-H Objekten am Beispiel des Decca Cuboids</i> .....	68
Abbildung 67	<i>Gegenüberstellung der Downmixe des MPEG-H Beds und MPEG-H Objekten am Beispiel des ORTF-3D</i> .....	68
Abbildung 68	<i>Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (ORTF-3D, 0°)</i> .....	69
Abbildung 69	<i>Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (ORTF-3D, 135°)</i> .....	69
Abbildung 70	<i>Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (PCMA-3D, 0°)</i> .....	70
Abbildung 71	<i>Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (PCMA-3D, 135°)</i> .....	70
Abbildung 72	<i>Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (Decca Cuboid, 0°)</i> .....	70

---

Abbildung 73 *Gegenüberstellung der mit MPEG-H und Dolby Atmos generierten Downmixe (Decca Cuboid, 135°)*..... 70

Abbildung 74 *Auswirkungen des Mikrofonabstandes auf Frequenzspektrum des Stereodownmixes.* 71

Abbildung 75 *Auswirkung des Pegels der Ausgangskanäle auf die Amplitude des im Downmix entstehenden Kammfilters*..... 72

### 13. Tabellenverzeichnis

Tabelle 1	<i>Lautsprecherkonfigurationen bei Dolby Atmos im Format 5.1.4</i> .....	4
Tabelle 2	<i>Lautsprecherkonfigurationen bei MPEG-H im Format 5.1 + 4H</i> .....	9
Tabelle 3	<i>Broadcast Legacy</i> .....	11
Tabelle 4	<i>Dynamic</i> .....	11
Tabelle 5	<i>Classics</i> .....	11
Tabelle 6	<i>Loud</i> .....	11
Tabelle 7	<i>Sports</i> .....	12
Tabelle 8	<i>Default</i> .....	12

## 14. Anhang A

### A.1 Pegel- und Delaywerte PCMA-3D 0°

	Pegel in dB	Delay in Samples
L	- 0,89	39
R	- 0,89	39
C	0	0
Ls	- 53,75	178
Rs	- 53,75	178
Ltf	- 8,73	39
Ltr	- 8,73	178
Rtf	- 8,73	39
Rtr	- 8,73	178

### A.2 Pegel- und Delaywerte PCMA 135°

	Pegel in dB	Delay in Samples
L	- 10,34	106
R	- 35,37	198
C	- 18,84	177
Ls	- 1,38	0
Rs	- 1,93	106
Ltf	- 8,73	106
Ltr	- 8,73	0
Rtf	- 8,73	198
Rtr	- 8,73	106

## A.3 Pegel- und Delaywerte Decca Cuboid 0°

---

	Pegel in dB	Delay in Samples
L	0	157
R	0	157
C	0	0
Ls	0	431
Rs	0	431
Ltf	0	174
Ltr	0	443
Rtf	0	174
Rtr	0	443

---

## A.4 Pegel- und Delaywerte Decca Cuboid 135°

---

	Pegel in dB	Delay in Samples
L	0	225
R	0	396
C	0	417
Ls	0	0
Rs	0	225
Ltf	0	239
Ltr	0	19
Rtf	0	407
Rtr	0	239

---

## A.5 Pegel- und Delaywerte ORTF-3D 0°

	Pegel in dB	Delay in Samples
L	- 4,92	0
R	- 4,92	0
C	/	/
Ls	- 15,92	28
Rs	- 15,92	28
Ltf	- 2,92	0
Ltr	- 40,91	28
Rtf	- 2,92	0
Rtr	- 40,91	28

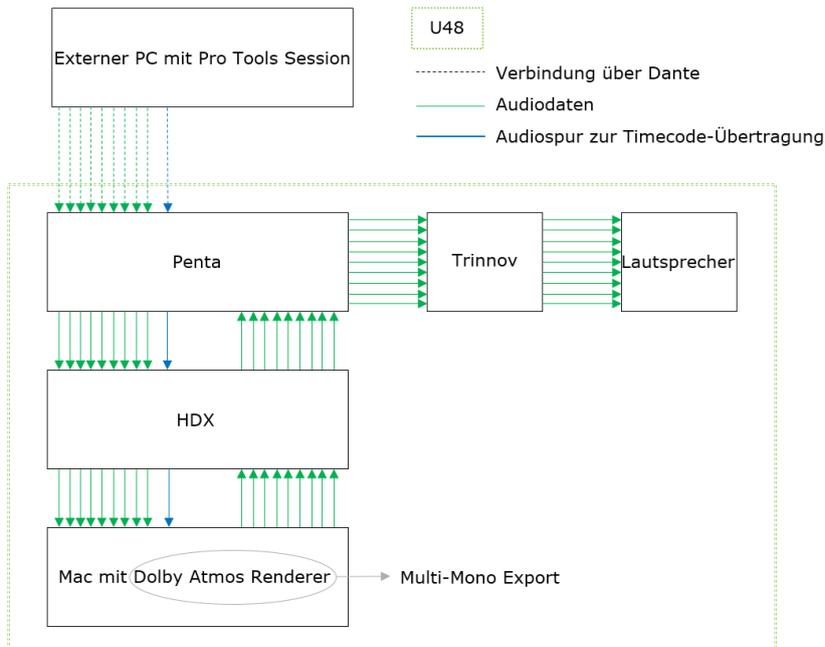
## A.6 Pegel- und Delaywerte ORTF-3D 135°

	Pegel in dB	Delay in Samples
L	- 8,25	20
R	- 25,94	30
C	/	/
Ls	- 3,32	0
Rs	- 9,59	10
Ltf	- 7,90	20
Ltr	- 0,79	0
Rtf	- 14,87	30
Rtr	- 10,27	10

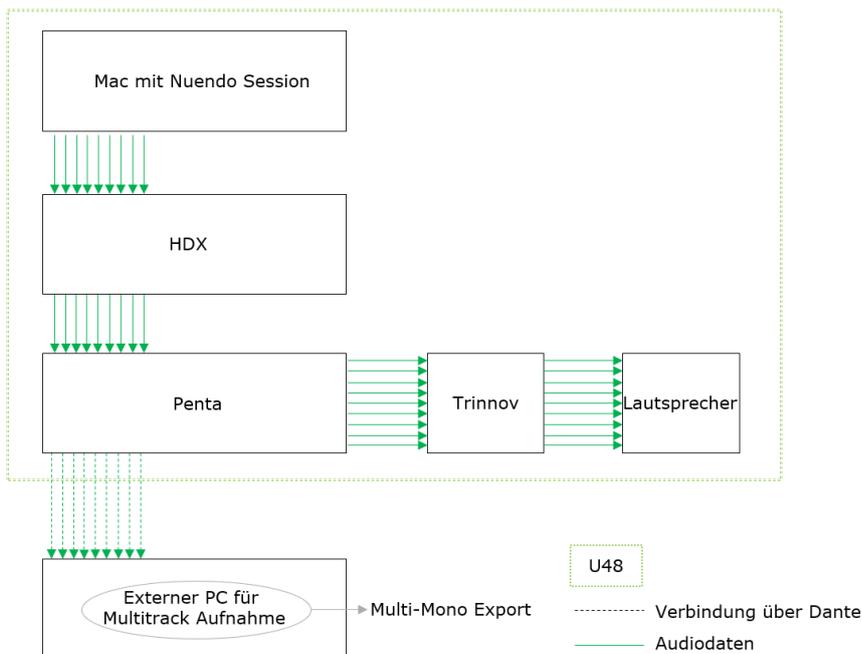
## 15. Anhang B Routing in Studio U48

Alle Downmixvergleiche wurden im Tonstudio U48 der Hochschule der Medien Stuttgart durchgeführt. Die folgenden Abbildungen zeigen den Signalweg inklusive der Einbindung eines externen Rechners.

### B.1 Routing für die Arbeit mit Dolby Atmos



### B.2 Routing für die Arbeit mit MPEG-H



## 16. Anhang C elektronischer Anhang

Der elektronische Anhang enthält folgende Dateien:

1. Alle für die Downmixvergleiche mit Testsignal angelegten DAW-Sessions
2. Alle für die Pegel- und Delayberechnung sowie für die Erstellung der Plots aus den Abbildungen 7, 8, 9, 11, 12, 13, 15, 16 und 17 programmierten MATLAB-Skripte und Funktionen
3. Hörbeispiele aller Stereodownmixe der Orchesteraufnahme
4. Ausgewählte Hörbeispiele aus den Simulationen mit einem Testsignal:
5. Die Dateien sind in zwei Ordner (Stereo- und 5.0-Downmix) einsortiert und die Dateibenennung folgt folgender Struktur:

*Array\_Schalleinfallswinkel\_Produktionsumgebung\_Bed oder Objekt\_Format\_Kanal*

Bei allen Downmixes wurden alle Signale des Ausgangsformats verwendet, d.h. auch unter Verwendung des Dolby Beds wurden alle Signale des UL verwendet

Alle Hörbeispiele aus den Stereodownmixes dienen der Einordnung der durch die in den Abbildungen in Kapitel sieben gezeigten Kammfilter hervorgerufenen klanglichen Verfärbungen. Da die beiden Stereokanäle bei allen enthaltenen Beispielen bezüglich ihrer Frequenzspektren übereinstimmen, ist jeweils nur der linke Kanal als Hörbeispiel angefügt. Da gezeigt wurde, dass die Ergebnisse des Stereodownmixes bei Dolby Atmos unter Verwendung des Beds und Objekten übereinstimmen, wurde darauf verzichtet Hörbeispiele der Stereodownmixe des Dolby Beds anzuhängen.

Folgende Hörbeispiele der Stereodownmixe sind im elektronischen Anhang zu finden:

- *Decca\_0°\_Dolby Atmos\_Objekte\_2.0\_L*
- *Decca\_0°\_MPEG-H\_Bed\_2.0\_L*
- *Decca\_0°\_MPEG-H\_Objekte\_2.0\_L*
- *ORTF\_0°\_Dolby Atmos\_Objekte\_2.0\_L*
- *ORTF\_0°\_MPEG-H\_Bed\_2.0\_L*
- *ORTF\_0°\_MPEG-H\_Objekte\_2.0\_L*
- *PCMA\_0°\_Dolby Atmos\_Objekte\_2.0\_L*
- *PCMA\_0°\_MPEG-H\_Bed\_2.0\_L*
- *PCMA\_0°\_MPEG-H\_Objekte\_2.0\_L*

Die Hörbeispiele der 5.0-Downmixe dienen dem direkten Vergleich zwischen der Verwendung des Dolby Atmos Beds und Dolby Atmos Objekten am Beispiel der Surround Kanäle des PCMA-3D sowie dem Vergleich zwischen MPEG-H und Dolby Atmos am Beispiel von Multimonoexports des Decca Cuboids.

Demnach sind folgende Hörbeispiele der 5.0-Downmixe im elektronischen Anhang zu finden:

- *PCMA\_0°\_Dolby Atmos\_Bed\_5.0\_Ls*
- *PCMA\_0°\_Dolby Atmos\_Objekte\_5.0\_Ls*
  
- *Decca\_0°\_Dolby Atmos\_Objekte\_5.0\_L*
- *Decca\_0°\_Dolby Atmos\_Objekte\_5.0\_R*
- *Decca\_0°\_Dolby Atmos\_Objekte\_5.0\_C*
- *Decca\_0°\_Dolby Atmos\_Objekte\_5.0\_Ls*
- *Decca\_0°\_Dolby Atmos\_Objekte\_5.0\_Rs*
  
- *Decca\_0°\_MPEG-H\_Bed\_5.0\_L*
- *Decca\_0°\_MPEG-H\_Bed\_5.0\_R*
- *Decca\_0°\_MPEG-H\_Bed\_5.0\_C*
- *Decca\_0°\_MPEG-H\_Bed\_5.0\_Ls*
- *Decca\_0°\_MPEG-H\_Bed\_5.0\_Rs*