

# Bestimmung des Reflexionsfaktors mit einem sphärischen Mikrofonarray

Sebastian Lauster, Robert Hupke, Marcel Nophut, Stephan Preihs, Jürgen Peissig

*Institut für Kommunikationstechnik, Leibniz Universität Hannover, Deutschland*

*Email: hupke@ikt.uni-hannover.de*

## Einleitung

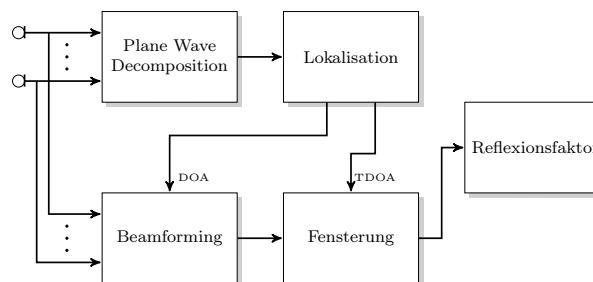
Für eine plausiblere, immersive Verbindung von Räumen in zukünftigen Networked Music Performances (NMP), ist es von großem Interesse, die akustischen Eigenschaften der verbundenen Räume an den verschiedenen Standorten zu bestimmen. Da die Reflexions- und Absorptionseigenschaften der Wände eines Raumes wichtige Parameter für die Auralisierung von virtuellen Umgebungen sind, gibt es eine Vielzahl von etablierten Verfahren zur Bestimmung des Reflexionsfaktors. Standardisierte Laborverfahren verwenden eine Materialprobe in einem Impedanzmessrohr [1, 2] oder einen Hallraum [3]. Ein weiterer Ansatz ist es, die akustische Beschaffenheit von Räumen durch die Analyse von Raumimpulsantworten zu bestimmen. Durch raumakustische Messungen können Parameter aus der Raumimpulsantwort ermittelt werden, die eine allgemeine Aussage über die Akustik an der Messposition erlauben [4]. Diese Messmethoden erlauben es jedoch nicht, das Schallfeld richtungsabhängig zu untersuchen. Die Nutzung eines sphärischen Mikrofonarrays ermöglicht hingegen eine richtungsabhängige Analyse des aufgenommenen Schallfeldes. Durch die Verteilung der Mikrofonkapseln, z.B. auf einer Kugel, lässt sich das Schallfeld im gesamten dreidimensionalen Raum erfassen und anschließend kann, durch die Analyse der räumlich-zeitlichen Struktur der „Plane Wave Decomposition“ (PWD), das Schallfeld in seiner Zusammensetzung untersucht werden. Anschließend können die Reflexionsfaktoren der Wände geschätzt werden, welche dann zur Auralisierung in ein virtuelles Raummodell [5] übertragen werden können. Es gibt mehrere Methoden zur Schätzung von Reflexionsfaktoren mittels sphärischen Mikrofonarrays. In den Arbeiten von Spors [6] und Rafaley [7] werden Raumimpulsantworten mit Sphärischen Harmonischen analysiert und der Reflexionsfaktor über das Verhältnis der frühen Reflexionen zum Direktschall ermittelt.

Im Rahmen dieser Arbeit wird eine Signalverarbeitungskette zur Schätzung von Reflexionsfaktoren auf Grundlage von [6] implementiert. Es soll untersucht werden, inwieweit sich die Methode für die Schätzung von Reflexionsfaktoren in realen Umgebungen eignet. Hierbei wird der Fokus auf die Nutzung des Eigenmike® [8] gelegt. Zuerst werden numerisch simulierte Raumimpulsantworten mit dem Spiegelschallquellenmodell [9] erzeugt, um das Verfahren zu evaluieren. Hierbei können frequenzabhängige Reflexionsfaktoren den Wänden zugeordnet werden. Anschließend wird im Immersive Media Lab [10], einem akustisch präpariertem Abhörraum mit einer mittleren Nachhallzeit von  $T_m = 0,233$  s, ei-

ne Messung zur Schätzung des Absorptionsfaktors einer akustischen Stellwand durchgeführt, dessen Absorptionsfaktor im Vorfeld mit einem Impedanzrohr [11] bestimmt wurde.

## Implementierung

Die implementierte Signalverarbeitungskette ist in Abbildung 1 dargestellt. Das Verfahren lässt sich in drei Hauptteile aufteilen: in die Lokalisation, das Beamforming und die Schätzung der Reflexionsfaktoren. Zur Detektion früher Reflexionen und des Direktschalls wird zuerst die PWD der gemessenen bzw. simulierten Impulsantworten ermittelt. Mit der Anwendung eines geeigneten Lokalisationsverfahrens (Blob-Detection-Methode) auf den Betrag der aus der PWD rekonstruierter Impulsantworten, wird der geschätzte Einfallswinkel „Direction of Arrival“ (DOA), sowie die zeitlichen Verzögerungen der Reflexionen bezogen auf den Zeitpunkt des eintreffenden Direktschalls „Time Difference of Arrival“ (TDOA), ermittelt. Durch die Anwendung eines adaptiven Beamformers, in Richtung der zuvor detektierten DOAs, werden anschließend störende Interferenzen, sowie Diffusfelder und Rauschen unterdrückt. Abschließend werden die frühen Reflexionen erster Ordnung und der Direktschall im Zeitbereich gefenstert und die frequenzabhängigen Reflexionsfaktoren durch das Verhältnis beider Effektivwerte ermittelt.



**Abbildung 1:** Schematisches Blockschaltbild der Signalverarbeitung zur Schätzung des Reflexionsfaktors.

## Rekonstruktion der Raumimpulsantwort

Für die Rekonstruktion des Schallfeldes auf einer Kugel werden Kugelflächenfunktionen verwendet, welche durch folgende Formel beschrieben sind:

$$Y_n^m(\theta, \phi) \equiv \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}, \quad (1)$$

wobei  $P_n^m(\cos\theta)$  die assoziierenden Legendrefunktionen mit der Ordnung  $n \in \mathbb{N}$  und dem Grad  $m \in -n, \dots, n$

darstellt. Durch die inverse sphärische Fouriertransformation kann das Schallfeld in den sphärisch harmonischen Bereich, die „Spherical Harmonic Domain“ (SHD) transformiert werden:

$$P_{nm}(k) = \int_0^{2\pi} \int_0^\pi P(\Omega, k, R) [Y_n^m(\Omega)]^* \sin \theta \, d\theta \, d\phi. \quad (2)$$

Hierbei sind  $P_{nm}(k)$  die sphärischen Fourierkoeffizienten und  $P(\Omega, k, R)$  die Mikrofonsignale in Abhängigkeit von  $k = \omega/c$  die Wellenzahl,  $R$  des Radius der Sphäre sowie dem Winkel  $\Omega = (\theta, \phi)$  entsprechend der Mikrofonposition auf der Sphäre. Die PWD des Signals wird durch die sphärische Fouriertransformation bestimmt, gegeben durch

$$\bar{P}(\hat{\Omega}, k, R) = \sum_{n=0}^N \sum_{m=-n}^n P_{nm}(k) Y_n^m(\hat{\Omega}), \quad (3)$$

wobei  $\bar{P}(\hat{\Omega}, k, R)$  das rekonstruierte Schallfeld am frei wählbaren Punkt  $\hat{\Omega} = (\hat{\theta}, \hat{\phi})$  darstellt. Durch die inverse Fouriertransformation kann das Schallfeld in den Zeitbereich  $\bar{p}(\hat{\Omega}, t, R)$  transformiert werden. Eine fehlerfreie Auflösung des Schallfeldes wird durch die Mikrofonanzahl und Radius des Arrays begrenzt, womit sich eine maximale Schallfeldordnung  $N$  ergibt.

### Lokalisation

Zur Detektion früher Reflexionen wird die Blob-Detection-Methode aus der digitalen Bildverarbeitung verwendet [12]. Hierbei wird ein Laplacian-of-Gaussian-Filter (LoG) genutzt. Der Ansatz wurde modifiziert, um die DOA und TDOAs in Form von lokalen Maxima in der zeitlichen Darstellung der PWD zu detektieren. Das LoG-Filter ist ein zweidimensionales Filter, so dass für die Lokalisation horizontale Impulsantworten mit einem festen Inklinationswinkel  $\hat{\theta} = 90^\circ$  und einem variablen Winkel  $\hat{\phi}$  in dem Bereich  $-180^\circ < \hat{\phi} \leq 180^\circ$  rekonstruiert werden. Um die Methode zu optimieren, wird die Amplitudendämpfung von Schallwellen über den Ausbreitungsweg kompensiert und ein Sum-and-Delay-Beamformer zur Steigerung des SNRs eingesetzt. Das Beamforming erfolgt während der inversen Fouriertransformation durch die Gewichtung des Schallfeldes über

$$\bar{P}(\hat{\Omega}, k, R) = \sum_{n=0}^N \sum_{m=-n}^n b_n(kR) P_{nm}(k) Y_n^m(\hat{\Omega}), \quad (4)$$

wobei als Gewichte die sphärischen Besselfunktionen  $b_n(kR)$  verwendet werden.

### Beamforming

Zur Unterdrückung von störenden Interferenzen und Rauschen wird der Linear-Constraint-Minimum-Variance-Beamformer (LCMV) genutzt. In der SHD werden die Gewichte allgemein wie folgt implementiert:

$$\mathbf{W}_{nm} = \mathbf{c}^H \frac{\mathbf{V}_{nm}^H \mathbf{S}_{nm}^{-1}}{\mathbf{V}_{nm}^H \mathbf{S}_{nm}^{-1} \mathbf{V}_{nm}}, \quad (5)$$

wobei  $\mathbf{S}_{nm}$  das Kreuzleistungsdichtespektrum beschreibt. Der Constraint Vector ist definiert durch  $\mathbf{c}$ ,

während der Steering Vector  $\mathbf{V}_{nm}$  gegeben ist durch

$$\mathbf{V}_{nm}(k) = \mathbf{b}_n(kR) [\mathbf{Y}_n^m(\tilde{\Omega})]^H, \quad (6)$$

wobei  $\tilde{\Omega} = (\tilde{\theta}, \tilde{\phi})$  die detektierten DOAs sind, so dass der Steering Vector in Richtung der einfallenden frühen Reflexionen zeigt.

## Experiment

Raumimpulsantworten für ein Array mit einer starren Oberfläche, 32 Mikrofonen und einem Radius von  $R = 4,2$  cm wurden simuliert. Dies entspricht der Konfiguration des Eigenmikes, welches anschließend für die Messungen verwendet wurde. Das Eigenmike ermöglicht eine maximale Ordnung von  $N = 4$ . Da  $kr < N$  für eine fehlerfreie Rekonstruktion des Schallfeldes gelten muss, ist für diese Konfiguration theoretisch eine Auflösung bis 5,12 kHz ohne räumliches Aliasing möglich. Anschließend wurde eine Messung in einem akustisch präparierten Abhörraum durchgeführt, um den Absorptionsfaktor einer akustischen Stellwand zu bestimmen. Ziel dieser Untersuchungen war es, festzustellen, ob sich das Verfahren auch auf Messungen in realen Umgebungen übertragen lässt.

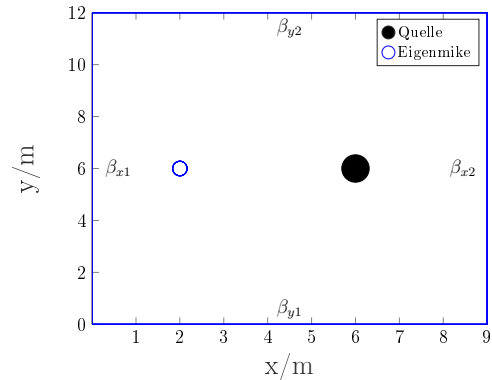


Abbildung 2: Horizontale Darstellung des simulierten Raumes mit der Position des Eigenmikes und der Quelle.

## Simulation

Der simulierte Raum basierte auf einem Rechteckraum ( $9 \times 12 \times 4,4$ ) m und ist dargestellt in Abbildung 2. Den

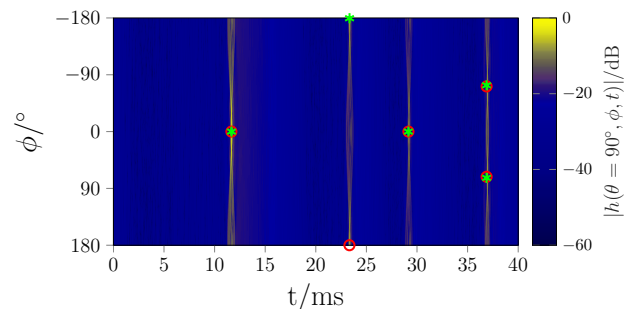


Abbildung 3: PWD der simulierten Impulsantwort mit der erwarteten Einfallsrichtung (roter Kreis) und der detektierten Einfallsrichtung (grüner Stern).

**Tabelle 1:** Tatsächlicher Reflexionsfaktor  $r$  und geschätzter Reflexionsfaktor  $\hat{r}$  mit dem relativen Fehler  $\Delta_r$ .

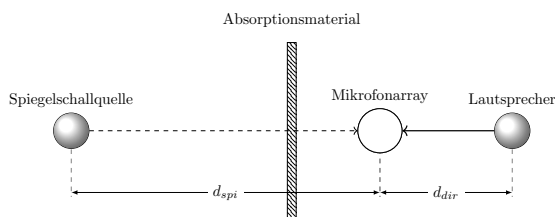
Oktavfrequenz/Hz	$\beta_{x1}$			$\beta_{x2}$			$\beta_{y1}$			$\beta_{y2}$		
	$r$	$\hat{r}$	$\Delta_r/\%$	$r$	$\hat{r}$	$\Delta_r/\%$	$r$	$\hat{r}$	$\Delta_r/\%$	$r$	$\hat{r}$	$\Delta_r/\%$
125	0,82	0,932	13,643	0,98	1,043	6,429	0,82	0,941	14,73	0,75	0,871	16,1
250	0,94	0,946	0,6	0,98	0,995	1,529	0,94	0,95	1,06	0,82	0,871	6,265
500	0,96	0,965	0,555	0,970	0,95	2,027	0,96	0,9602	0,02	0,89	0,892	0,16
1000	0,97	0,969	0,088	0,97	0,962	0,82	0,97	0,973	0,274	0,92	0,924	0,44
2000	0,98	0,977	0,328	0,96	0,953	0,744	0,98	0,985	1,075	0,93	0,924	0,663
4000	0,98	0,985	0,4933	0,95	0,948	0,127	0,98	0,958	1,222	0,94	0,924	1,64
8000	0,98	0,981	0,091	0,95	0,949	0,018	0,98	0,684	30,241	0,94	0,667	28,983

Wänden  $\beta_{x1}$ ,  $\beta_{x2}$ ,  $\beta_{y1}$  und  $\beta_{y2}$  wurden verschiedene frequenzabhängige Reflexionsfaktoren zugewiesen. Eine omnidirektionale Quelle stand im Raum an der Position (6, 6, 1,8) m. Um ein Peak-Signal-to-Noise-Ratio (PSNR) von 60 dB zu erhalten wurde auf die simulierten Mikrofonsignalen zusätzliches Rauschen hinzugefügt. Abbildung 3 zeigt die zeitlich dargestellte PWD der simulierten Raumimpulsantwort mit der erwarteten Einfallsrichtung (roter Kreis) der Reflexionen, sowie den geschätzten Einfallsrichtungen (grüner Stern) auf der Horizontalebene für  $\theta = 90^\circ$ . Die Detektion für den Winkel  $\phi$  weist eine Abweichung von 2,44 % auf, wohingegen das geschätzte TDOA um 0,042 % abweicht.

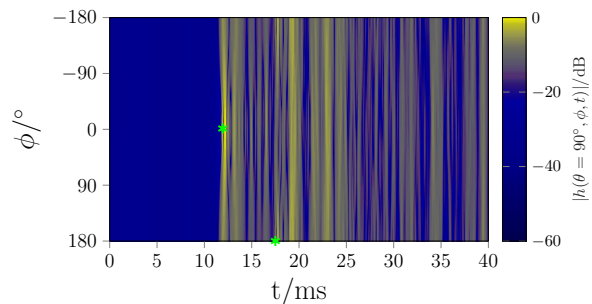
Die geschätzten Reflexionsfaktoren  $\hat{r}$  und die tatsächlichen Reflexionsfaktoren  $r$ , der vier äußeren Wände, sind in Tabelle 1 dargestellt. Ebenfalls wird für jede Oktavfrequenz der relative Fehler  $\Delta_r/\%$  abgebildet. Es ist zu erkennen, dass bei tieferen Frequenzen ein deutlich höherer Fehler in der Schätzung entsteht. Dies lässt sich durch die frequenzabhängige Begrenzung der Größe des simulierten Mikrofonarrays erklären, welche in einer geringen Richtcharakteristik des Beamformer bei tieferen Frequenzen resultiert. Im Frequenzbereich zwischen 500 Hz und 4 kHz liegt der relative Fehler der Schätzung bei maximal 2,03 %. Es ist festzustellen, dass manche Schätzungen große Abweichungen im 8 kHz Oktavband aufweisen, wohingegen die Schätzung bei anderen Wänden nur einen sehr geringen Fehler haben. Dies ist nicht weiter untersucht worden und muss in Zukunft genauer betrachtet werden.

## Messungen

Der zuvor durch ein Impedanzrohr ermittelte Reflexionsfaktor einer Stellwand wird als Referenz verwendet um mit den geschätzten Ergebnissen der Messung verglichen

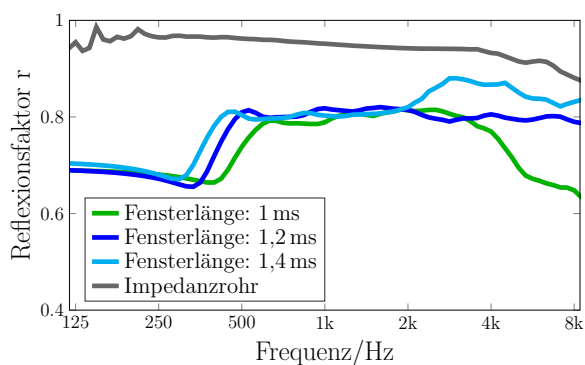
**Abbildung 4:** Schematischer Messbau zur Schätzung des frequenzabhängigen Reflexionsfaktors einer Stellwand.

zu werden. Zur Anregung des Raumes wurde ein Dodekaederlautsprecher genutzt, welcher eine Monopolschallquelle nachbildet. Der Messaufbau ist in Abbildung 4 dargestellt. Die Stellwand wurde, mit einem Abstand von 1 m, hinter das Eigenmikro gestellt, so dass eine Reflexion bei  $(\theta, \phi) = (90^\circ, -180^\circ)$  zu erwarten war. Der Abstand  $d_{dir}$  zwischen Lautsprecher und Mikrofonarray betrug 3 m, sodass sich eine Entfernung der Spiegelschallquelle zum Array von  $d_{spi} = 5$  m ergab. Die Messung wurde mit einem logarithmischen Sinusweep im Frequenzbereich von 50 Hz bis 20 kHz durchgeführt.

**Abbildung 5:** PWD der gemessenen Impulsantwort mit der detektierten Einfallsrichtung (grüner Stern).

Die zeitliche PWD für  $\theta = 90^\circ$  in Abbildung 5 zeigt, dass obwohl eine hohe Anzahl an zusätzlichen Reflexionen in der Messung auftreten, sowohl Direktschall, als auch die Reflexion, hervorgerufen durch die Stellwand, detektiert werden konnten. In Abbildung 6 sind die geschätzten Reflexionsfaktoren für unterschiedliche Fensterlängen abgebildet. Zur Glättung der Kurven wurde ein 1/3-Oktavbandfilter verwendet. Betrachtet man den Verlauf der Kurven, so ist eine deutlich größere Abweichung im Frequenzbereich von 125 Hz bis 500 Hz mit einem gemittelten Fehler von 27,51 % festzustellen. Im Frequenzbereich von 1 kHz bis 8 kHz hingegen ist die Abweichung zu der Referenzmessung geringer.

Die untersuchte Fensterlänge hat erst ab einer Frequenz von 2 kHz einen Einfluss auf das Ergebnis. So ergibt sich ein gemittelter Fehler von 14,69 % im Frequenzbereich von 500 Hz bis 4 kHz. Je größer die Fensterlänge gewählt wird, desto genauer wird die Schätzung im Frequenzbereich oberhalb von 2 kHz. Bei einer Fensterlänge von 1 ms ergibt sich ein Fehler von 16,11 %, bei 1,2 ms ein Fehler von 14,86 % und bei einer Fensterlänge von 1,4 ms ein Fehler von 13,11 %. Bei tieferen Frequenzen unterhalb



**Abbildung 6:** Ergebnis der Messung bei verschiedenen Fensterlängen im Vergleich zur Impedanzrohrmessung.

von 2 kHz hat die Fensterlänge nur geringen Einfluss auf das Ergebnis. Die größere Abweichung zu tiefen Frequenzen kann auf die unterschiedlichen Anregungsarten beider Verfahren zurückzuführen sein. So wird bei der Nutzung des Impedanzrohrs nur ein kleiner Teil der Materialprobe angeregt, wohingegen bei der Messung im realen Raum es nicht vermeidbar ist, die komplette Stellwand anzuregen. Zusätzlich kann nicht ausgeschlossen werden, dass aufgrund des Messaufbaus, eine Art Plattenresonator entstanden ist, der zu einer zusätzlich Schallabsorption im niedrigen Frequenzbereich führte. Den Einfluss der Position der Stellwand auf die Messung wurde nicht weiter betrachtet und ist Bestandteil zukünftiger Untersuchungen.

## Zusammenfassung

In dieser Arbeit wurde untersucht, inwieweit sich mit einem sphärischen Mikrofonarray Reflexionsfaktoren von Raumwänden abschätzen lassen. Dafür wurde zuerst eine Signalverarbeitungskette auf Grundlage der Publikation von Spors [6] implementiert und simulativ evaluiert. Um zu überprüfen, ob das Verfahren auf real gemessene Impulsantworten anwendbar ist, wurden mit einem sphärischen Mikrofonarray Messungen einer Materialprobe durchgeführt. Die geschätzten Reflexionsfaktoren wurden mit einer Referenzmessung eines Impedanzmessrohrs verglichen.

Die Ergebnisse aus der Simulation zeigen, dass die Schätzung einen relativen Fehler von maximal 2,03 % im Frequenzbereich zwischen 500 Hz und 4 kHz aufweist. Dies sollte für eine Auralisation ein hinreichend kleiner Fehler sein. Zu tieferen Frequenzen wird dieser Fehler jedoch größer. Die reale Messung hingegen weist eine deutlich größere Abweichung zu den Reflexionskoeffizienten, gemessen mit dem Impedanzrohr, auf. Hierbei wurde gezeigt, dass die Wahl der Fenstergröße in der Signalverarbeitung einen Einfluss auf das Ergebnis ab einer Frequenz von 2 kHz hat. Nichtsdestotrotz werden Ergebnisse mit einem gemittelten Fehler von 14,69 % im Frequenzbereich von 500 Hz bis 4 kHz erreicht. Der Einfluss des Fehlers auf eine Auralisation des simulierten Raumes wurde bisher noch nicht betrachtet und ist Bestandteil zukünftiger Untersuchungen. Des Weiteren sollte untersucht werden, inwiefern die vorgestellte Methode der automatischen De-

tektion der ersten Reflexionen auf komplexere Schallszenarien anzuwenden ist. Eine Erweiterung der Schätzung von anderen raumakustischen Parametern wie Raumvolumen sowie Raumgeometrie kann ebenfalls Bestandteil von zukünftigen Untersuchungen sein.

## Förderung

Diese Arbeit ist gefördert vom Bundesministerium für Wirtschaft und Energie unter dem Förderkennzeichen 01MD18010G.

## Literatur

- [1] DIN EN ISO 10534-1:2001-10, *Acoustics - Determination of sound absorption coefficient and impedance in impedance tubes - Part 1: Method using standing wave ratio.*
- [2] DIN EN ISO 10534-2:2001-10, *Acoustics - Determination of sound absorption coefficient and impedance in impedance tubes - Part 2: Transfer-function method.*
- [3] ISO 354:2003-05, *Acoustics - Measurement of sound absorption in a reverberation room.*
- [4] Wolfgang Ahnert and Hans-Peter Tennhardt, "Raumakustik," in *Handbuch der Audiotechnik*, pp. 181–266. Springer, 2008.
- [5] Henning Piper, Marcel Nophut, Robert Hupke, Stephan Preihs, and Jürgen Peissig, "Investigations on loudspeaker-based auralization of immersively connected rooms," in *Fortschritte der Akustik: DAGA 2020*. 2020.
- [6] Sascha Spors and Till Rettberg, "On the Estimation of Acoustic Reflection Coefficients from In-Situ Measurements using a Spherical Microphone Array," in *Fortschritte der Akustik: DAGA 2018*. 2018.
- [7] Boaz Rafaely, *Fundamentals of Spherical Array Processing*, Springer, 8 edition, 2015.
- [8] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2002.
- [9] Emanuël Habets, "Spherical Microphone Array Impulse Response generator (SMIRgen)," 2016.
- [10] Robert Hupke, James Ordner, Jakob Bergner, et al., "Towards a virtual audiovisual environment for interactive 3D audio productions," in *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*, March, 2019.
- [11] Roman Schlieper, Song Li, and Jürgen Peissig, "Development and validation of a full range acoustic impedance tube," in *Audio Engineering Society Convention 144*, May 2018.
- [12] Akshay Chopra, "Implementation of a Laplacian Blob Detector," 2018.