Wave-domain Transforms for Irregular Loudspeaker Array Topologies

Christian Hofmann, Walter Kellermann
Friedrich-Alexander University Erlangen-Nürnberg (FAU)
*Multimedia Communications and Signal Processing**
Cauerstraße 7, D-91058 Erlangen, Germany
Email: {hofmann,wk}@LNT.de

Introduction

Today's premium cars are typically equipped with a large number of loudspeakers and microphones to provide high-quality entertainment, hands-free communication via cellphones, and hands-free access to the applications running on the car's head unit relying on a human/machine voice dialog. Especially for multichannel sound rendering, this requires multichannel Acoustic Echo Cancellation (AEC) systems. For lab environments, the concept of wave-domain AEC [1] proved to be an efficient solution for planar loudspeaker arrays by performing the adaptive filtering in a spatial transform domain. While acoustic systems typically show similarly strong couplings between all loudspeakers and all microphones, the wave-domain couplings between loudspeaker modes and microphone modes decrease with increasing difference between mode orders [1]. This prior knowledge can be exploited for improved system identification and reduced computational complexity by modeling only a subset of couplings in the spatial transform domain.

In practice, many environments, such as automotive environments, impose hard constraints on the placement of loudspeakers of the entertainment system, precluding planar loudspeaker arrays. An alternative transform-domain multichannel adaptive filtering approach which does not require geometrical prior knowledge about the transducer placement is the so-called source-domain adaptive filtering [2], for which Singular Value Decompositions (SVDs) of cross-correlation matrices have to be computed and tracked in a reliable way. A similar effect is exploited by the spatial-region AEC described in [3], where independent sources are extracted by beamformers and where transducers may be placed in an arbitrary but known way. An even more sophisticated multichannel AEC concept tailored to the actual characteristics of the Loudspeaker-Enclosure-Microphone System (LEMS) requires the adaptive estimation of the system's eigenspace to obtain a spatial transform before adapting what is considered the actual filter coefficients in the transform domain [4].

However, as these algorithms have only been simulated with a relatively low number of channels and do not exploit prior knowledge about the transducer distributions sufficiently, we focus on Wave-Domain Adaptive Filtering (WDAF) and investigate the consequences of such irregular loudspeaker arrays on wave-domain transforms according to [5].

*This work was supported by the Fraunhofer Institute for Digital Media Technology (IDMT) in Ilmenau, Germany.

Wave-Domain Transforms Based on Circular Harmonics

The core idea of wave-domain analysis emerges from the orthogonal sets of basis functions resulting from solving the acoustic wave equation in a particular coordinate system. Employing them, the sound field is decomposed into components which are elementary solutions of the acoustic wave equation, so-called modes, and is processed in the decomposed domain. The spherical, cylindrical, and circular decompositions [6, 7] are of particular interest, because the decompositions result in countably infinitely many decomposition coefficients also under free-field assumptions and thereby allow for the approximation of the wave field by a few coefficients only—a representation suitable for digital hardware. Initially, wave-domain processing has been proposed in [8] for AEC. Wave-domain transforms based on circular harmonics have been introduced in [9] for active listening room compensation and have been shown to provide approximate eigen functions of the acoustic transfer function matrices—also in reverberant environments. Since then, the concept of WDAF based on circular harmonics has been extended [1, 5] and employed for reduced-complexity AEC [1] as well as Listening Room Equalization (LRE) [10] assuming Uniform Circular Concentric Arrays (UCCAs). Contrasting the spatial transform, denoted as wave domain, the original domain will be denoted as transducer domain from now on.

Notations and Prerequisites

In the following, coordinate vectors are indicated by an arrow, e.g. $\vec{x}$, and $\vec{x}^{\mathrm{T}}$ denotes transposition of $\vec{x}$. A point in space is equivalently described by the triples $\vec{x} = [x, y, z]^{\mathrm{T}}$ in Cartesian coordinates, $[\varrho, \varphi, z]^{\mathrm{T}}$ in cylindrical coordinates, and $[r, \varphi, \vartheta]^{\mathrm{T}}$ in spherical coordinates. Therein, φ is the azimuthal angle in the x-y-plane, ϑ is the elevation angle w.r.t. the x-y-plane, and ϱ and r are cylindrical and spherical radial coordinates of $\vec{x}$, respectively.

Furthermore, let $\vec{x}_{\mathrm{L},l}$ and $\vec{x}_{\mathrm{M},m}$ denote the coordinates of the l^{th} loudspeaker and the m^{th} microphone of an N_{L}-element loudspeaker array and an N_{M}-element microphone array, respectively. Besides, let t be time, f be a continuous-time frequency, and $\omega = 2\pi f$ be the corresponding angular frequency. Additionally, assume that the microphones are in a plane parallel to the x-y-plane and the microphones are centered around the z-axis with uniform angular spacing $2\pi/N_{\mathrm{M}}$ and radius R_{M}. Each of the loudspeakers emits a real-valued signal with a Fourier transform $P_{\mathrm{L}}(l, \omega)$ and each microphone captures the signal $P_{\mathrm{M}}(m, \omega) = P(\vec{x}_{\mathrm{M},m}, \omega)$. The speed of sound is denoted as c_{air} and the wave-number of a monofrequent wave is written as $k(\omega)$.

Circular Harmonics

For uniform circular apertures, transforms based on a circular harmonics decomposition [6] seem intuitive, as they are the orthogonal set of basis functions resulting from the solution of the acoustic wave equation in cylindrical coordinates. This is also justified by the findings in [9], where the transfer function matrices' generalized SVDs lead to a spatial basis strongly resembling circular harmonics. A spatially sampled continuous-time frequency-domain signal $g(\varrho, \varphi_i, \omega)$, sampled equidistantly at N points with angles $\varphi_i = \frac{2\pi}{N} i$ at radius $\varrho = \varrho_0$, where $i = 0, \ldots, N-1$ is the index of the sample point, has the circular harmonics decomposition

$$\mathring{g}\left(\mathring{i}, \omega\right) = \frac{1}{N \mathcal{B}_{\mathring{i}}(\varrho_0 k(\omega))} \sum_{i=0}^{N-1} g(\varrho_0, \varphi_i, \omega) \mathrm{e}^{-\frac{2\pi}{N} \mathrm{j} \mathring{i} i}, \quad (1)$$

where $\mathring{i} \in \mathbb{Z}$ is termed mode index. The term $\mathcal{B}_{\mathring{i}}(\varrho_0 k(\omega))$ is known as baffle function [11] and depends on the boundary conditions at the radius ϱ_0, where the sound field is sampled. In the free-field case, it results in

$$\mathcal{B}_{\mathring{i}}(\varrho_0 k(\omega)) = \mathcal{J}_{\mathring{i}}(\varrho_0 k(\omega)), \quad (2)$$

where $\mathcal{J}_{\mathring{i}}(x)$ denotes the Bessel function of the first kind and order $\mathring{i}$ [11]. Comparing Eq. (1) to a Discrete Fourier Transform (DFT) reveals that $\mathring{g}(\mathring{i}, \omega)$ is N-periodic in $\mathring{i}$: only N non-redundant modes can be determined and mode $\mathring{i}$ is identical to mode $\mathring{i} + N \; \forall \mathring{i}$.

The Microphone Array Signal Transform

The microphone signals of a uniform circular array can straight-forwardly be transformed by a cylindrical harmonics decomposition of the wave field spatially sampled by the microphones [5]. These forward and backward transforms can be represented as linear Multiple-Input/Multiple-Output (MIMO) systems with the transfer functions

$$T_2^{\mathrm{f}}(\mathring{m}, m, \omega) = \frac{1}{N_{\mathrm{M}} \mathcal{B}_{\mathring{m}}(k(\omega) R_{\mathrm{M}})} \mathrm{e}^{-\mathrm{j}\mathring{m}\varphi_{\mathrm{M},m}} \quad (3)$$

$$T_2^{\mathrm{b}}(m, \mathring{m}, \omega) = \mathcal{B}_{\mathring{m}}(k(\omega) R_{\mathrm{M}}) \cdot \mathrm{e}^{\mathrm{j}\mathring{m}\varphi_{\mathrm{M},m}} \quad (4)$$

between microphone $m \in 0, 1, \ldots, N_{\mathrm{M}} - 1$ and microphone mode $\mathring{m} \in \mathbb{Z}$, where $\varphi_{\mathrm{M},m} = \frac{2\pi}{N_{\mathrm{M}}} m$ denotes the azimuthal coordinate of the m^{th} microphone and where the superscripts "f" and "b" denote "forward" and "backward" transform, respectively. For a discrete-time implementation, the normalization to $\mathcal{B}_{\mathring{m}}$ can also be neglected, which shifts its contribution into the wave-domain LEMS [5].

2D Model Loudspeaker Array Signal Transform

For the loudspeaker signal transform in [5], the N_{L} loudspeakers are considered to form a Uniform Circular Array (UCA) with the l^{th} speaker located at $[\varrho_{\mathrm{L},l}, \varphi_{\mathrm{L},l}, z_{\mathrm{L},l}] = [R_{\mathrm{L}}, \frac{2\pi}{N_{\mathrm{L}}} l, 0]$ (loudspeakers and microphones are arranged as UCCAs).

In this case, the loudspeakers' sound field in the vicinity of the microphone array is approximated at $N_{\mathrm{M,v}} = N_{\mathrm{L}}$ points (denoted here as virtual microphone positions $\vec{x}_{\mathrm{M,v}}$) and this approximated sound field $P_{\mathrm{M,v}}(m_{\mathrm{v}}, \omega)$ is transformed by a circular harmonics decomposition. In particular, loudspeakers are considered as point sources to describe the propagation of their signals to the center of the microphone array at $\vec{x}_{\mathrm{M,cent}} = \vec{0}$ and the propagation from the array center to the virtual microphones is modeled by a plane wave approximation. This yields

$$P_{\mathrm{M,v}}(m_{\mathrm{v}}, \omega) := \sum_{l=0}^{N_{\mathrm{L}}-1} P_{\mathrm{L}}(l, \omega) G(\vec{x}_{\mathrm{M,cent}} \,|\, \vec{x}_{\mathrm{L},l}, \omega) \cdot \mathrm{e}^{\mathrm{j}k(\omega) R_{\mathrm{M}} \cos\left(\varphi_{\mathrm{M,v},m_{\mathrm{v}}} - \varphi_{\mathrm{L},l}\right)} \quad (5)$$

with the free-field Green's function [12] $G(\vec{x}_{\mathrm{M,cent}} \,|\, \vec{x}_{\mathrm{L},l}, \omega)$ of the l^{th} loudspeaker, evaluated for the microphone array center, and the azimuthal angle $\varphi_{\mathrm{M,v},m_{\mathrm{v}}}$ of the $m_{\mathrm{v}}^{\mathrm{th}}$ virtual microphone position. Formulating the circular harmonics decomposition for Eq. (5), and exploiting the Jacobi-Anger identity [13] leads to a wave-domain representation of the loudspeaker signals as

$$\tilde{P}_{\mathrm{L}}(\mathring{l}, \omega) := \sum_{l=0}^{N_{\mathrm{L}}-1} \frac{\mathrm{e}^{-\mathrm{j}k(\omega)\varrho_{\mathrm{L},l}}}{\varrho_{\mathrm{L},l}} P_{\mathrm{L}}(l, \omega) \mathrm{j}^{\mathring{l}} \mathrm{e}^{-\mathrm{j}\mathring{l}\varphi_{\mathrm{L},l}}, \quad (6)$$

where $\mathring{l}$ is the loudspeaker mode. From this representation, the loudspeaker signals can be reconstructed by

$$P_{\mathrm{L}}(l, \omega) = \frac{\varrho_{\mathrm{L},l} \cdot \mathrm{e}^{\mathrm{j}k(\omega)\left(\varrho_{\mathrm{L},l} - \max\limits_{l}\{\varrho_{\mathrm{L},l}\}\right)}}{N_{\mathrm{L}}} \cdot \sum_{\mathring{l}=0}^{N_{\mathrm{L}}-1} \tilde{P}_{\mathrm{L}}(\mathring{l}, \omega) \mathrm{j}^{-\mathring{l}} \mathrm{e}^{\mathrm{j}\mathring{l}\varphi_{\mathrm{L},l}}, \quad (7)$$

up to a delay $\max\limits_{l}\{\varrho_{\mathrm{L},l}\}$, which has been introduced to ensure causality [5]. These forward and backward transforms can be represented as MIMO systems with the respective transfer functions

$$T_1^{\mathrm{f,2D}}(\mathring{l}, l, \omega) = \mathrm{j}^{\mathring{l}} \frac{\mathrm{e}^{-\mathrm{j}k(\omega)\varrho_{\mathrm{L},l}}}{\varrho_{\mathrm{L},l}} \mathrm{e}^{-\mathrm{j}\mathring{l}\varphi_{\mathrm{L},l}} \quad (8)$$

$$T_1^{\mathrm{b,2D}}(l, \mathring{l}, \omega) = \frac{\varrho_{\mathrm{L},l} \cdot \mathrm{e}^{\mathrm{j}k(\omega)\left(\varrho_{\mathrm{L},l} - \max\limits_{l}\{\varrho_{\mathrm{L},l}\}\right)}}{N_{\mathrm{L}}} \mathrm{j}^{-\mathring{l}} \mathrm{e}^{\mathrm{j}\mathring{l}\varphi_{\mathrm{L},l}}. \quad (9)$$

3D Model Loudspeaker Array Signal Transform

Different from [5], three-dimensional loudspeaker arrays will be considered in the following. This means that each loudspeaker may appear under an individual elevation angle $\vartheta_{\mathrm{L},l}$ with respect to the microphone array center. Taking this into account when describing the virtual microphones, Eq. (5) changes to

$$P_{\mathrm{M,v}}(m_{\mathrm{v}}, \omega) = \sum_{l=0}^{N_{\mathrm{L}}-1} P_{\mathrm{L}}(l, \omega) G(\vec{x}_{\mathrm{M,cent}} \,|\, \vec{x}_{\mathrm{L},l}, \omega) \cdot \mathrm{e}^{\mathrm{j}k_{\mathrm{M}}(\omega, l) R_{\mathrm{M}} \cos\left(\varphi - \varphi_{\mathrm{L},l}\right)} \quad (10)$$

with r_l being the l^{th} loudspeaker's radial coordinate in spherical coordinates and with

$$k_{\mathrm{M}}(\omega, l) = \frac{\omega}{v_{\mathrm{p}}(l)} = \cos\left(\vartheta_{\mathrm{L},l}\right) \frac{\omega}{c_{\mathrm{air}}} \quad (11)$$

being the loudspeaker's effective wave number in the microphone-array plane at the microphone array center. Performing a circular harmonics decomposition of Eq. (10) results in the wave-domain loudspeaker representation

$$\tilde{P}_{\mathrm{L}}^{(3\mathrm{D})}(\mathring{l},\omega) := \sum_{l=0}^{N_{\mathrm{L}}-1} P_{\mathrm{L}}(l,\omega) \cdot \underbrace{\frac{\mathcal{J}_{\mathring{l}}\big(k(\omega)R_{\mathrm{M}}\cdot\cos\vartheta_{\mathrm{L},l}\big)}{\mathcal{J}_{\mathring{l}}\big(k(\omega)R_{\mathrm{M}}\big)}}_{\substack{\text{depending on}\\ \text{loudspeaker}\\ \text{elevation, mode, and}\\ \text{frequency}}} \cdot \underbrace{\frac{\mathrm{e}^{-\mathrm{j}k(\omega)r_{\mathrm{L},l}}}{r_{\mathrm{L},l}}}_{\substack{\text{mode-}\\ \text{independent}}} \cdot \underbrace{\mathrm{j}^{\mathring{l}}\mathrm{e}^{-\mathrm{j}\mathring{l}\varphi_{\mathrm{L},l}}}_{\substack{\text{frequency-}\\ \text{independent}}}. \tag{12}$$

After approximating the ratio of Bessel functions in Eq. (12) by one[1], a simplified, 3D-modified loudspeaker signal transform can be expressed as MIMO system with the transfer functions

$$T_1^{\mathrm{f,3D}}(\mathring{l},l,\omega) = \frac{\mathrm{e}^{-\mathrm{j}k(\omega)r_{\mathrm{L},l}}}{r_{\mathrm{L},l}}\mathrm{j}^{\mathring{l}}\mathrm{e}^{-\mathrm{j}\mathring{l}\varphi_{\mathrm{L},l}} \tag{13}$$

$$T_1^{\mathrm{b,3D}}(l,\mathring{l},\omega) = \frac{r_{\mathrm{L},l}\cdot\mathrm{e}^{\mathrm{j}k(\omega)\left(r_{\mathrm{L},l}-\max\limits_{l}\{\varrho_{\mathrm{L},l}\}\right)}}{N_{\mathrm{L}}}\mathrm{j}^{-\mathring{l}}\mathrm{e}^{\mathrm{j}\mathring{l}\varphi_{\mathrm{L},l}}. \tag{14}$$

This transform pair contains Eq. (8) and Eq. (9) as special cases for zero elevation angles ($r_{\mathrm{L},l} = \varrho_{\mathrm{L},l}$) of the loudspeakers and, thereby, constitutes a generalization of the 2D transform of the previous section. Note that both the 2D model and the 3D model lead to transforms as linear MIMO systems consisting of fractional delay filters (mode-independent part) and the subsequent mixing of the delayed signals (frequency-independent part). Thereby, the discrete-time realizations of both transform variants can be implemented with the same computational complexity.

Evaluation Method

In this section, we will evaluate the properties of the LEMS in the transform domain. Therein, each of the transfer functions $\tilde{H}(\mathring{m},\mathring{l},\omega)$ describes the coupling from a loudspeaker mode $\mathring{l}$ to a microphone mode $\mathring{m}$. For such a MIMO system, the energy couplings between modes $\mathring{m}$ and $\mathring{l}$ and can be defined as

$$E_{\mathrm{cou}}(\mathring{m},\mathring{l}) = \int_0^{\omega_{\max}} |\tilde{H}(\mathring{m},\mathring{l},\omega)|^2 \mathrm{d}\omega. \tag{15}$$

A concentration of the coupling energy to a few coefficients allows for a good approximation of the respective MIMO system by the subset of paths for strongly coupled modes or exploitation of prior knowledge about typical wave-domain systems during adaptive system identification. For the wave-domain transforms based on circular harmonics, a diagonal dominance with most prominent couplings at $\mathring{l} = \mathring{m}$ has been observed [1].

In order to quantify this compactness, consider the energy compaction measure

$$E_{\mathrm{com}}(N_{\mathrm{D}}) = 10\log_{10}\frac{\sum_{\mathring{m}}\sum_{\mathring{l}=\mathring{m}-\lfloor N_{\mathrm{D}}/2\rfloor}^{\mathring{m}+\lfloor N_{\mathrm{D}}/2\rfloor} E_{\mathrm{cou}}(\mathring{m},\mathring{l})}{\sum_{\mathring{m}}\sum_{\mathring{l}} E_{\mathrm{cou}}(\mathring{m},\mathring{l})}\mathrm{dB}, \tag{16}$$

[1] As for the microphone array transform, this allows for a more efficient discrete-time implementation.

where summation without any limits means summation over all non-redundant modes. This measure corresponds to the logarithmic ratio between the coupling energy of the subset of the N_{D} diagonals centered around $\mathring{m} = \mathring{l}$ to the coupling energy of the complete system[2]. The energy-compaction measure increases monotonically with increasing N_{D} towards $E_{\mathrm{com}}(N_{\mathrm{L}}) = 0\,\mathrm{dB}$. Ideally, $E_{\mathrm{com}}(N_{\mathrm{D}}) = 0\,\mathrm{dB}$ for $N_{\mathrm{D}} = 1$, which means that each microphone mode $\mathring{m}$ is determined only by the corresponding loudspeaker mode $\mathring{l}$. In this case, the wave-domain system could be modeled by N_{M} filters instead of $N_{\mathrm{M}}N_{\mathrm{L}}$ filters in the transducer domain.

Evaluation

The following analyses will be conducted in image-source environments of varying orders N_{im}, where the rooms' dimensions ($1.6\,\mathrm{m} \times 2.9\,\mathrm{m}$ base area) and the transducer geometry is inspired by an automotive environment. Due to the significant absorption of the ceiling, the floor, and seats in an automobile, only four reflective vertical walls are considered. A circular microphone array with $N_{\mathrm{M}} = 48$ elements and radius $R_{\mathrm{M}} = 0.25\,\mathrm{m}$ is placed in the image-source room and is centered between virtual front seats. The loudspeakers are placed symmetrically w.r.t. the y-axis and with equidistant azimuthal spacing along the contour of the room and, on the average, about 0.5 m below the microphones. This represents microphones in the ceiling.

The remaining degrees of freedom are the reflection coefficient R commonly shared by the four walls and the order N_{im} of the image-source model. The top view of

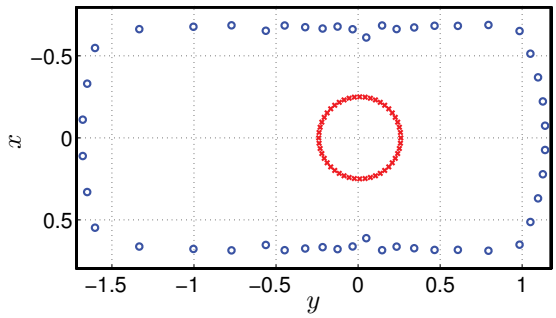

Figure 1: Top view on employed image source room with 48 loudspeakers (blue) and 48 microphones (red).

the complete setup is shown in Fig. 1. The energy compaction measure resulting for a 10^{th}-order image-source model with coefficients or $R = 0.5$ is depicted in Fig. 3 for the wave-domain system employing the 2D model (**cyan**) and the 3D model (**violet**). Obviously, the 3D-modified transform according to Equations (13) and (14) leads to a much more compact wave-domain system, as can be seen by the higher $E_{\mathrm{com}}(N_{\mathrm{D}})$ for low N_{D} of the **violet** curve (3D model). The next analysis, shown in Fig. 3, focuses on the wave-domain system's compactness for varying the orders of the image-source model. To this end, the energy compaction measure for $N_{\mathrm{D}} = 5$ diagonals is visualized for image-source model orders $0 \le N_{\mathrm{im}} \le 10$. As can be seen, the 3D model (**violet**) consistently shows a better energy compaction, although

[2] When considering the i^{th}-strongest coupling instead of all couplings of the i^{th} diagonal, this measure changes to the measure "EC(i)" in [9].

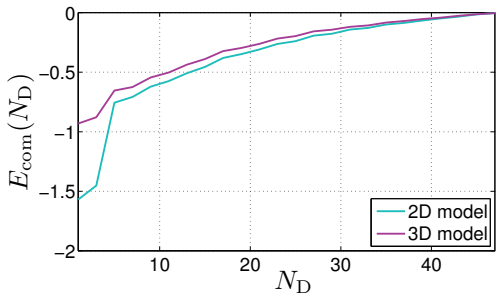

Figure 2: Energy compaction $E_{\mathrm{com}}(N_{\mathrm{D}})$ for different numbers of diagonals N_{D} for a 10^{th}-order image-source model with reflection coefficients of $R = 0.5$.

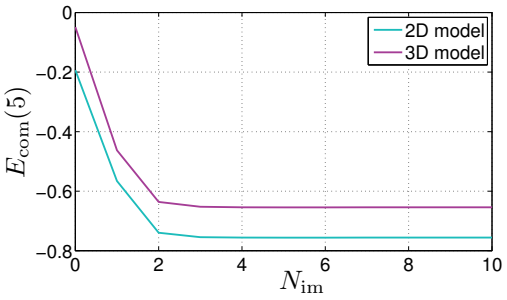

Figure 3: Energy compaction $E_{\mathrm{com}}(N_{\mathrm{D}})$ for $N_{\mathrm{D}} = 5$ diagonals for a varying order N_{im} of the image-source model with reflection coefficients of $R = 0.5$.

best compaction is achieved for fewer reflections—which reflects the fact that the wave-domain transforms have been developed using free-field assumptions and therefore, show least couplings. Note that there is no perfect diagonalization for a 0^{th}-order image-source model, due to the approximations during the loudspeaker signal transform. In case of reflections, each image source signal (capturing a particular $1^{\mathrm{st}}-$ or higher-order reflection) may impinge from a different direction than the original source.

The same effect can be observed in Fig. 4 for a 10^{th}-order image-source environment while varying the reflection coefficient R in the interval $[0.1, 0.9]$. As can be seen, the 3D model (**violet**) leads to a more diagonally dominant system than the 2D model (**cyan**). The more dominant the direct sound (low R), the better the energy compaction of both transforms and the stronger the gain by the 3D model (**violet**).

To sum up, the 3D model increases the diagonal dominance of the wave-domain systems. The improved diagonal dominance stems from early parts of the impulse-responses describing the LEMS. For high-order image source models and highly reflective surfaces, diagonal dominance generally reduces, but a gain by the 3D model in terms of the energy compaction measure is still visible.

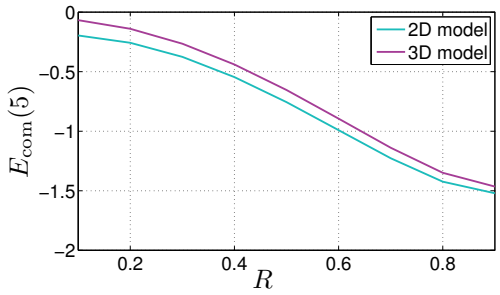

Figure 4: Energy compaction $E_{\mathrm{com}}(N_{\mathrm{D}})$ for $N_{\mathrm{D}} = 5$ diagonals as a function of the reflection coefficient R of the walls.

Conclusion

Wave-domain transforms based on circular harmonics lead to the desired diagonal dominance of the wave-domain systems also for three-dimensional loudspeaker arrays. While the computational effort for a signal processing application is not increased by the 3D-modified transform presented in this manuscript, the diagonal dominance of the wave-domain system can be improved by employing the 3D-modified transform.

References

[1] M. Schneider and W. Kellermann, "A wave-domain model for acoustic mimo systems with reduced complexity," in *Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)*, Edinburgh, UK, May 2011, pp. 133 – 138.

[2] K. Helwani, H. Buchner, and S. Spors, "Source-domain adaptive filtering for MIMO systems with application to acoustic echo cancellation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2010, pp. 321–324.

[3] K.-A. Lee, W.-S. Gan, J. Yang, and F. Sattar, "Multichannel teleconferencing system with multi spatial region acoustic echo cancellation," in *International Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2003, pp. 51–54.

[4] K. Helwani and H. Buchner, "On the eigenspace estimation for supervised multichannel system identification," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013, pp. 630–634.

[5] M. Schneider and W. Kellermann, "A direct derivation of transforms for wave-domain adaptive filtering based on circular harmonics," in *European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, August 2012, pp. 1034–1038.

[6] A. Kuntz, "Wave field analysis using virtual circular and microphone arrays," Ph.D. dissertation, University of Erlangen-Nuremberg, June 2008.

[7] S. Spors, "Active listening room compensation for spatial sound reproduction systems," Ph.D. dissertation, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2005.

[8] H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering: Acoustic echo cancellation for full-duplex systems based on wave-field synthesis," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Montreal, Canada, May 2004.

[9] S. Spors, H. Buchner, R. Rabenstein, and W. Herbordt, "Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering," *Journal of the Acoustical Society of America*, vol. 122, no. 1, pp. 354–369, July 2007.

[10] M. Schneider and W. Kellermann, "Adaptive listening room equalization using a scalable filtering structure in the wave domain," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2012.

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

[12] E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 1999.

[13] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th ed. New York (NY), USA: Dover Publications, 1972.