

---

**Acoustic Array Systems: Paper ICA2016-158****Experimental investigation of multiple-input multiple-output systems for sound-field analysis****Hai Morgenstern<sup>(a)</sup>, Johannes Klein<sup>(b)</sup>, Boaz Rafaely<sup>(a)</sup>, Markus Noisternig<sup>(c)</sup>**

<sup>(a)</sup>Dept. of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Israel,  
haimorg@post.bgu.ac.il, br@bgu.ac.il

<sup>(b)</sup>Institute of Technical Acoustics, RWTH Aachen University, Germany,  
johannes.klein@akustik.rwth-aachen.de

<sup>(c)</sup>Acoustic and Cognitive Spaces Group, IRCAM, CNRS, Sorbonne University, UPMC Paris 6,  
France, markus.noisternig@ircam.fr

**Abstract**

Spherical microphone and loudspeaker arrays have been widely studied for the acquisition of spatial sound-field information. Recently, a theoretical framework, based on systems that combine both arrays, was presented for the spatial analysis of enclosed sound fields. Such systems are referred to as multiple-input multiple-output (MIMO) systems, and they provide means for an enhanced spatial analysis. However, their performance is limited by errors due to spatial sampling and system model mismatch. The effects of these errors on the system performance were studied recently in theory, without experimental validation. Therefore, the practical usefulness of MIMO systems for room-acoustics analysis has yet to be determined. This paper presents an initial investigation in this direction. MIMO system performance and limitations are first evaluated in a simulation study. The system is then studied experimentally, through the analysis of room impulse responses (RIRs). Experimental validation is achieved in several aspects. First, system properties are studied and compared to previous theoretical results. Then, MIMO processing methods are applied for a spatial analysis of early reflections in the RIR, showing that early room reflections can be identified experimentally. The results of this investigation suggest that MIMO systems can be employed in practice for various applications of room acoustics.

**Keywords:** microphone arrays, loudspeaker arrays, MIMO systems, room acoustics, spherical harmonics

---

# Experimental investigation of multiple-input multiple-output systems for sound-field analysis

## 1 Introduction

Spherical arrays, comprising either microphones or loudspeakers, have lately been studied for a spatial analysis of room acoustics. Spherical microphone arrays (SMAs) facilitate a spatial analysis of the sound field [?,?]; spherical loudspeaker arrays (SLAs) have the ability to radiate acoustic energy to specific directions in space, which can potentially improve the separability of room reflections in a room impulse response (RIR) [?,?]. More recently, systems that combine both arrays, typically referred to as multiple-input multiple-output (MIMO) systems, have also been proposed. In particular, a theoretical framework for a spatial analysis using these systems has been developed [?], and the potential of using such systems for spatial analysis has been partially demonstrated in several applications [?, ?, ?].

In practice, the performance of MIMO systems is limited by errors, such as model-mismatch errors and errors due to spatial sampling. Based on these errors, a design framework has been proposed for the spherical arrays that constitute a MIMO system [?]. In particular, guidelines for system design have been derived to ensure high system performance, with the total system error bounded by a given acceptable error level. The performance of MIMO systems in practice, however, has yet to be investigated for determining their practical usefulness.

In this paper an initial experimental investigation of such an acoustic MIMO system is presented. First, a MIMO system model is presented and its properties are briefly discussed. Then, system performance is evaluated in a simulation study, using the design framework proposed in Ref. [?]. Finally, the system is investigated experimentally, and its properties are compared to theoretical results.

## 2 System model

Consider an acoustic MIMO system comprising of  $L$  loudspeakers, mounted on a rigid sphere with a radius of  $r_L$ , and  $M$  microphones, distributed on a rigid sphere with a radius of  $r_M$ . A spherical harmonics (SHs) representation is employed in the analysis of the MIMO system, because the system is based on spherical arrays. Therefore, room transfer functions (RTFs) are derived for  $(N_L + 1)^2 \leq L$  and  $(N_M + 1)^2 \leq M$  SHs coefficients of the SLA and SMA, respectively, where  $N_L$  and  $N_M$  are the SHs order of the SLA and SMA, respectively. These coefficients are rearranged in a matrix  $\mathbf{H}(k)$ , referred to as the MIMO RTF matrix, where  $k$  is the wavenumber. For a system in free-field,  $\mathbf{H}(k)$  is given by [?]:

$$\mathbf{H}(k) = \mathbf{h}_L(k)\mathbf{h}_M^H(k), \quad (1)$$

where  $\mathbf{h}_L(k)$  is the  $(N_L + 1)^2 \times 1$  SLA RTF vector,  $\mathbf{h}_M(k)$  is the  $(N_M + 1)^2 \times 1$  SMA RTF vector, and  $(\cdot)^H$  is the conjugate transpose. The elements of  $\mathbf{h}_L(k)$  and  $\mathbf{h}_M(k)$  are given by [?]:

$$[\mathbf{h}_L(k)]_{nm} = \frac{e^{ikr_0}}{r_0} g_n(kr_L) Y_n^m(\boldsymbol{\theta}_0) \text{ and} \quad (2)$$

$$[\mathbf{h}_M(k)]_{n'm'} = b_{n'}(kr_M) Y_{n'}^{m'}(\boldsymbol{\xi}_0), \quad (3)$$

respectively, with indices  $(n, m)$  and  $(n', m')$  denoting the corresponding SHs coefficients.  $r_0$  is the distance between the arrays,  $g_n(kr_L)$  and  $b_{n'}(kr_M)$  are coefficients defined for the SLA and SMA array types, respectively,  $(\cdot)^*$  is the complex conjugate, and  $Y_n^m(\boldsymbol{\theta}_0)$  is the SH function of order  $n$  and degree  $m$ , evaluated at elevation and azimuth angles,  $\boldsymbol{\theta}_0 = (\vartheta_0, \phi_0)$ .  $\boldsymbol{\theta}_0$  is the direction of radiation (DOR), which points at the SMA position, with respect to the SLA center. Similarly, for the SMA,  $Y_{n'}^{m'}(\boldsymbol{\xi}_0)$  is evaluated at  $\boldsymbol{\xi}_0$ , the direction of arrival (DOA), which points at the SLA position, with respect to the SMA center. The MIMO RTF matrix from eq. (??) shows an inherent unit rank since it is presented as an outer product of the SLA and SMA RTF vectors. Moreover,  $\mathbf{h}_L(k)$  and  $\mathbf{h}_M(k)$  are proportional to the left and right singular vectors of  $\mathbf{H}(k)$ , respectively, and they therefore hold information regarding the DOR and DOA, respectively. For a system positioned in a room, an RTF MIMO matrix,  $\tilde{\mathbf{H}}(k)$ , is formulated as a summation of MIMO RTF matrices for different room reflections [?], by:

$$\tilde{\mathbf{H}}(k) = \sum_q a_q(k) \mathbf{H}_q(k), \quad (4)$$

where  $q$  is the index of the room reflection and  $a_q(k)$  accounts for attenuation due to absorption of the walls and the physical distance. When positioned in a room, the rank of  $\tilde{\mathbf{H}}(k)$  is bounded by the minimum of the significant number of reflections,  $I$ , and the dimensions of the system, i.e.,  $\min(I, (N_L + 1)^2, (N_M + 1)^2)$ . Moreover, the subspaces defined by the left and right singular vectors of  $\tilde{\mathbf{H}}(k)$ , hold information regarding the DORs and DOAs of the significant reflections.

### 3 Simulation-based evaluation of system performance

In practice, the system model presented in the last section includes errors. At high frequencies, errors due to spatial sampling at the arrays are dominant. Model-mismatch errors are typically significant at high frequencies as well, however they become dominant at low frequencies if normalization is applied for a potential enhanced resolution [?] due to the ill-conditioning that this introduces.

For evaluating the system performance, the different errors are incorporated in the system model in order to calculate a total error. The total error is computed as in Ref. [?] for the following system parameters; a SLA with radius  $r_L = 0.2$  m and 672 loudspeakers distributed as in Ref. [?], employed with SHs order  $N_L = 5$ , and a SMA with radius  $r_M = 0.042$  m and 32 microphones distributed as in Ref. [?], employed with SHs order  $N_M = 2$ . Moreover, errors that account for model mismatch at the arrays were generated using a standard normal distribution, rearranged in matrix form, and added to  $\mathbf{h}_L(k)$  and  $\mathbf{h}_M(k)$ , for the SLA and SMA, respectively. The variance of the errors was set constant throughout frequency, 40 dB lower than  $\|\mathbf{h}_L(k)\|$  and

$\|h_M(k)\|$  at frequency  $f = 1$  kHz. Finally, normalization with the coefficients  $g_n(kr_L)$  and  $b_n(kr_M)$  was applied to the SLA and SMA, respectively. Fig. ?? shows the total error of the MIMO system. High levels of the total error are seen at low and at high frequencies, and the lowest

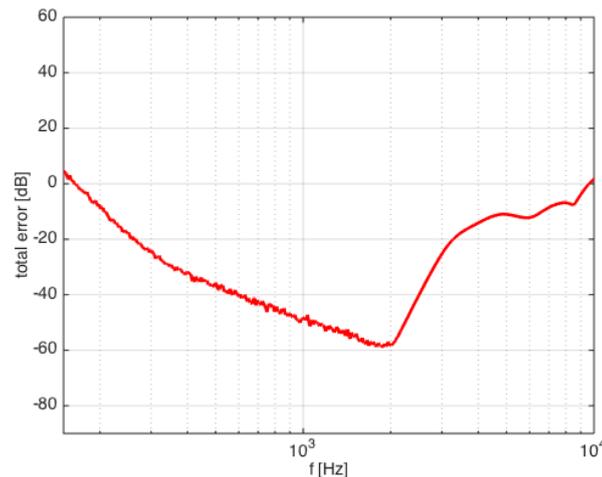


Figure 1: **Total error of a MIMO system comprised of a SLA with radius  $r_L = 0.2$  m and 672 loudspeakers and a SMA with radius  $r_M = 0.042$  m and 32 microphones. SHs orders  $N_L = 5$  and  $N_M = 2$  are employed at the SLA and SMA, respectively.**

level of the total error is evident at approximately  $f = 2$  kHz. The system is therefore expected to perform well at this frequency, with a relatively low error.

## 4 Experimental study

In this section, a MIMO system is investigated experimentally and its properties are compared to the theoretical ones from Sec. ?? to validate the applicability of MIMO systems in practice.

### 4.1 Setup

To evaluate the performance of the MIMO system in a reverberant sound field, a set of room impulse responses were measured in the opera hall, the Großes Festspielhaus of the Salzburg Festival [?]. The hall has an approximate volume of  $15500\text{m}^3$ , accommodating 2158 seats, and a reverberation time of 1.8s (unoccupied) [?]. The MIMO system was comprised of a SLA and a SMA with the same parameters as detailed in the previous section; in particular, a SLA composed of 28 transducers of three sizes (12x2", 12x3" and 4x5") distributed on an  $r_L = 0.2$  sphere [?] was employed. A sequential measurement using all transducers at 24 different azimuth rotations of the sphere results in 672 single speaker excitations of the room. The positions of the SLA transducers were determined by a Gaussian sampling strategy of the order 11 for the 2" and 3" inch transducers, as well as an oversampled order 3 for the 5" transducers. The mh acoustics em32 Eignmike<sup>®</sup> [?] was used. The SLA was positioned

slightly left to the center of the stage, and the SMA in the auditorium (row 10, seat 15). The auditorium was unoccupied. The room temperature was tracked during the measurements and can be considered as stable (with a variance of +/- 0.1 degrees during a measurement cycle). Refer to Fig. ?? for a sketch of the building plan of the Großes Festspielhaus including the measurement setup, i.e. the positions of the arrays. For obtaining the MIMO RTF matrix,

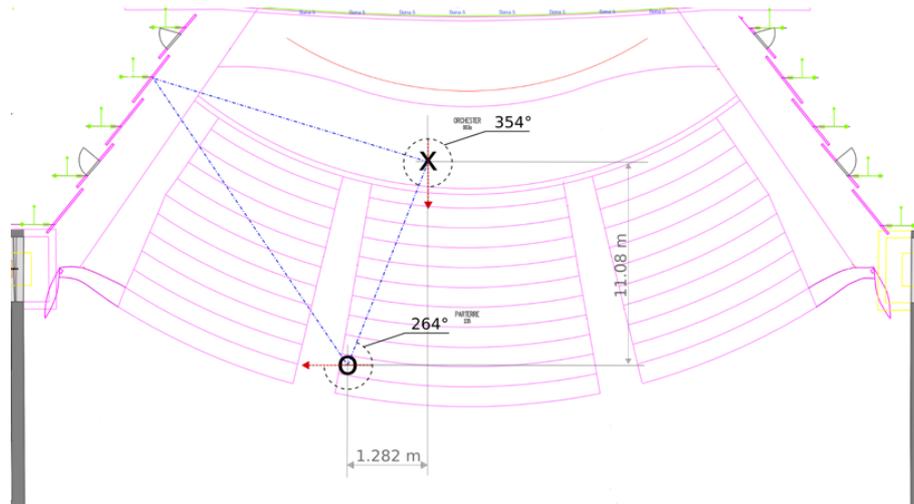


Figure 2: The sketch of the Großes Festspielhaus building plan (top view), including the measurement setup; the SLA and SMA positions are indicated by 'X'- and 'O'-marks, respectively, and their orientation is denoted using an arrow (which shows the corresponding x-axis). The DOR and DOA azimuthal angles are presented for the direct sound. The acoustic paths of the direct sound and a first-order reflection (left wall) are illustrated using dashed lines.

exponential sine-sweep signals were played-back for each loudspeaker sequentially using a sampling frequency of  $f_s = 48\text{kHz}$ , and simultaneously measured by the 32 microphones of the SMA. RTFs were then calculated for the SHs channels of the arrays by applying deconvolution in the time domain and suitable sampling weights in the space domain, with the resulting MIMO RTF rearranged in matrix form as in Eq. (??).

#### 4.2 Processing methods

The properties of the measured system are studied for specific time segments from the MIMO RTF matrix so as to limit the number of reflections. In particular, two time segments were selected within the early-reflections part of the RIRs by applying time-windowing.

Fig. ?? shows the early-reflections part of an RIR between a single loudspeaker and a single microphone, and also shows two time windows with a duration of 25 ms. The two time windows are centered at the time delay of the direct sound and the first reflection, at  $t_0 = 77.8\text{ms}$  and  $t_1 = 109.7\text{ms}$ , respectively. Each one of the time windows is applied on the inverse DFT of the MIMO RTF matrix, and then the DFT is calculated for frequency 2kHz, a frequency which was

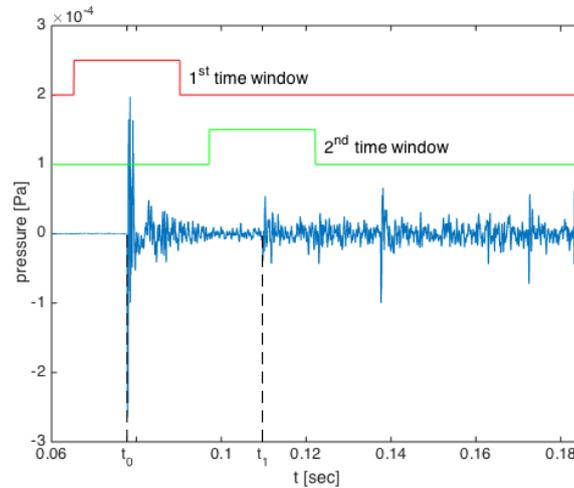


Figure 3: **Early-reflections part of an RIR between a single loudspeaker and a single microphone. Two time windows with a 20 ms duration are added on the plot, centered around  $t_0$  and  $t_1$ , the time delays of the direct sound and first reflection, respectively.**

shown to have a low error in the analysis at section ???. The resulting MIMO RTF matrices are denoted as  $\mathbf{H}_0(k)$  and  $\mathbf{H}_1(k)$  for the first and second time windows respectively.

### 4.3 Results

Since the MIMO RTF matrices from the previous section include a single room reflection, they are expected to have one significant singular value, with the right and left singular vectors holding information regarding the corresponding DOA and DOR, respectively.

Singular values were calculated for  $\mathbf{H}_0(k)$  and  $\mathbf{H}_1(k)$  and are presented in Fig. ???. For both  $\mathbf{H}_0(k)$  and  $\mathbf{H}_1(k)$ , one significant singular value is evident. This implies that one significant reflection is evident in the corresponding time window. In particular, a drop of 16.34dB is seen between the first and the second singular value in the case of  $\mathbf{H}_0(k)$  compared to a drop of 7.41dB in the case of  $\mathbf{H}_1(k)$ . This can be explained by the additional attenuation of the first reflection compared to the direct sound due to a longer acoustic path length and higher absorption at the walls of the room.

The right and left singular vectors are calculated for  $\mathbf{H}_0(k)$  and  $\mathbf{H}_1(k)$ . For estimating the DOA and the DOR for each system, the multiple signal classification (MUSIC) algorithm [?] is employed. MUSIC spectra are calculated for the SLA and SMA, for a linear span of angles, and are presented in Figs. ?? and ?? for  $\mathbf{H}_0(k)$  and  $\mathbf{H}_1(k)$ , respectively. For  $\mathbf{H}_0(k)$ , the DOAs and DORs estimated using MUSIC, indicated using 'X'-marks in Fig. ??, and are in agreement with the angles estimated based on the the arrays' positions, indicated using 'O'-marks in the figure. Similarly, for  $\mathbf{H}_1(k)$ , the angles estimated using the MUSIC spectra are in agreement with the measurement setup and room geometry, in Fig. ???. For both  $\mathbf{H}_0(k)$  and  $\mathbf{H}_1(k)$ , the spread of the peaks in the SLA spectrums (cf. Figs. ?? and ??) is smaller than that

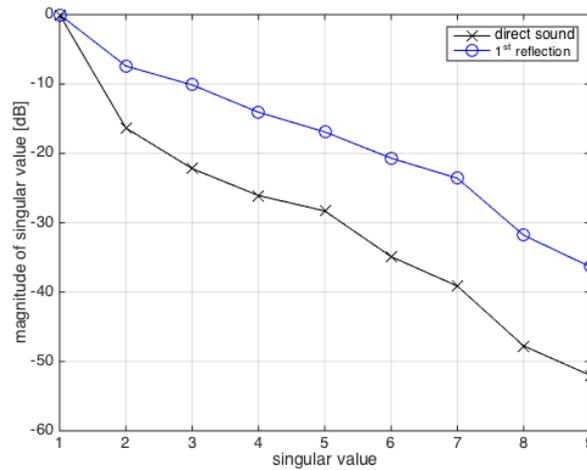


Figure 4: Singular value distribution for  $H_0(k)$  and  $H_1(k)$ .

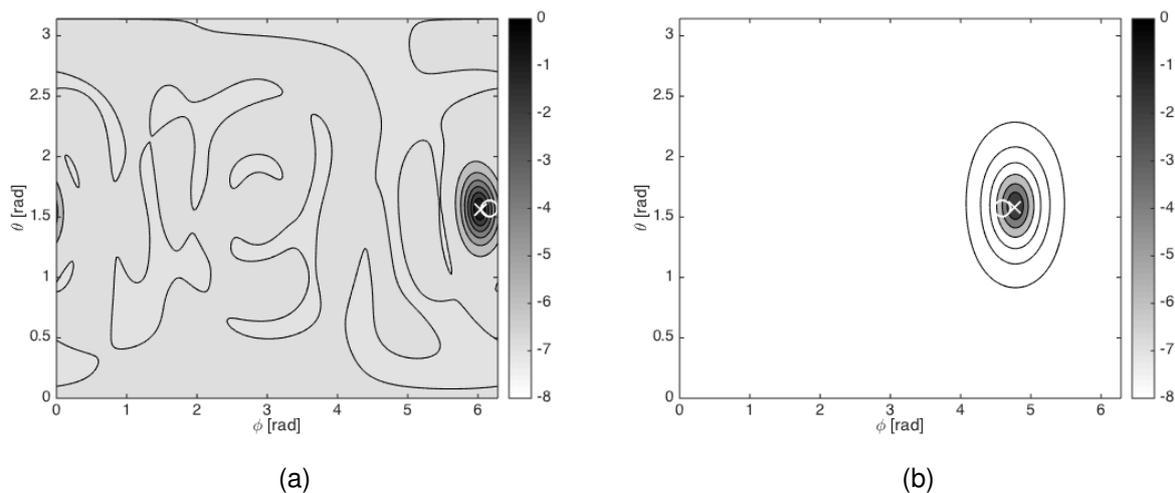


Figure 5: MUSIC spectra for (a) the SLA and (b) the SMA, calculated using the singular vectors of  $H_0(k)$ . An 'X'-mark indicates the DOAs and DORs estimated with MUSIC, and an 'O'-mark indicates those estimated based on the arrays' positions in the room.

of the SMA (cf. Figs. ?? and ??), due to the higher SHs order that is employed at the SLA. Moreover, the ambient noise floor in Figs. ?? and ??, representing the SLA, is higher than that in Figs. ?? and ??, representing the SMA. This can be explained by the higher SHs order employed at the SLA, which introduces excessive ill-conditioning when applying normalization by the radial functions to the SLA. In conclusion, this study showed that early room reflections can be identified experimentally using a spherical MIMO system.

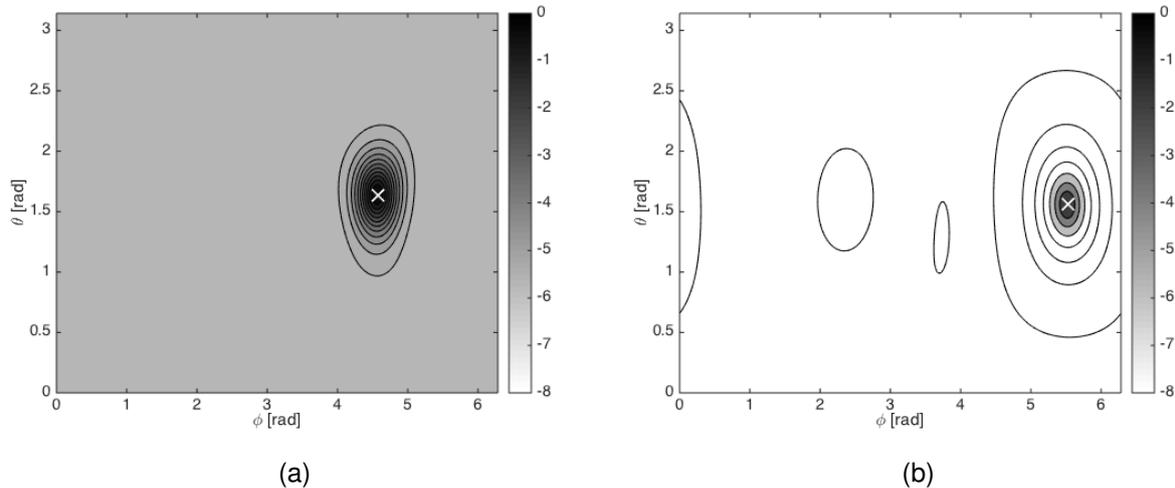


Figure 6: **MUSIC spectra for (a) the SLA and (b) the SMA, calculated using the singular vectors of  $H_1(k)$ . An 'X'-mark indicates the DOAs and DORs estimated with MUSIC.**

## 5 Conclusions

In this paper, an initial experimental investigation was presented for a MIMO system, based on spherical arrays. The analysis of the experimental system supports previous theoretical results, suggesting that MIMO systems can be employed in practice for various applications of room acoustics.

### Acknowledgements

This research was supported by The Israel Science Foundation (Grant No. 146/13) and the French FUI-AAP14 project “Binaural Listening” (BiLi). The authors would like to thank the Salzburg Festival for providing access to the opera hall for several days, and Edwin Pfanzagl-Cordone and his audio engineering team for the technical support during the measurements. We are grateful to Olivier Warusfel, Marco Berzborn, and Michael Vorländer for the many fruitful discussions, as well as to Thibaut Carpentier, Djellal Chalabi, Camille Dianoux, Emmanuel Flety, Arnaud Recher, and Alain Terrier for the help with software and hardware developments.

### References

- [1] B. N. Gover, J. G. Ryan, M. R. Stinson, Microphone array measurement system for analysis of directional and spatial variations of sound fields, *The Journal of the Acoustical Society of America* 112 (2002) 1980.
- [2] B. Rafaely, Plane-wave decomposition of the sound field on a sphere by spherical convolution, *The Journal of the Acoustical Society of America* 116 (2004) 2149.
- [3] S. Tervo, J. Patynen, T. Lokki, Acoustic reflection path tracing using a highly directional

- loudspeaker, in: IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'09), 2009, pp. 245–248.
- [4] A. M. Pasqual, A. de França, J. Roberto, P. Herzog, Application of acoustic radiation modes in the directivity control by a spherical loudspeaker array, *Acta acustica united with Acustica* 96 (1) (2010) 32–42.
- [5] H. Morgenstern, B. Rafaely, F. Zotter, Theory and investigation of acoustic multiple-input multiple-output systems based on spherical arrays in a room, *The Journal of the Acoustical Society of America* 138 (5) (2015) 2998–3009.
- [6] H. Morgenstern and B. Rafaely, “Analysis of acoustic mimo systems in enclosed sound fields,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 209–212, IEEE.
- [7] H. Morgenstern, F. Zotter, and B. Rafaely, “Joint spherical beam forming for directional analysis of reflections in rooms,” *The Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 3207–3207 (2012).
- [8] H. Morgenstern and B. Rafaely, “Enhanced spatial analysis of room acoustics using acoustic multiple-input multiple-output (mimo) systems,” in *Proceedings of Meetings on Acoustics*, vol. 19, 015018, Acoustical Society of America (2013).
- [9] H. Morgenstern, B. Rafaely, M. Noisternig, Design framework for spherical microphone and loudspeaker arrays in a multiple-input multiple-output system, in preparation.
- [10] M. Pollow, J. Klein, P. Dietrich, and M. Vorländer, “Including directivity patterns in room acoustical measurements,” in *Proceedings of Meetings on Acoustics*, vol. 19, 015008, Acoustical Society of America (2013).
- [11] J. Meyer, G. Elko, A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield, in: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 2, IEEE, 2002, pp. II–1781.
- [12] M. Noisternig, J. Klein, M. Berzborn, A. Recher, and O. Warusfel, High-resolution MIMO DRIR measurements in an opera hall, in: *Proc. of the 42nd Annual German Congress on Acoustics (DAGA)*, Aachen, Germany (2016).
- [13] T. Hidaka L. L. Beranek, and T. Okano, Interaural cross-correlation, lateral fraction, and low-and high-frequency sound levels as measures of acoustical quality in concert halls, *The Journal of the Acoustical Society* 98 (2) (1995) 988–1007.
- [14] J. Klein, M. Pollow, and M. Vorländer, Optimized Spherical Sound Source for Auralization with Arbitrary Source Directivity, In *Proc. EAA Joint Symposium on Auralization and Ambisonics*, Berlin, Germany (2014).
- [15] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*, John Wiley & Sons, 2004.