

# MULTI-CHANNEL ROOM IMPULSE RESPONSE SHAPING – A STUDY

Markus Kallinger and Alfred Mertins

University of Oldenburg, Department of Physics, Signal Processing Group  
D-26111 Oldenburg, Germany

{markus.kallinger, alfred.mertins}@uni-oldenburg.de

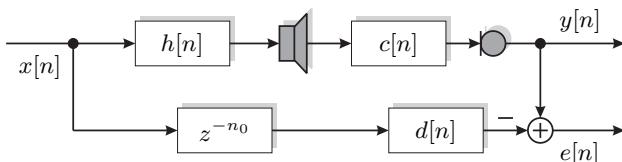
## ABSTRACT

This paper addresses the usability of channel shortening equalizers known for data transmission systems for the equalization of acoustic systems. In multicarrier systems, equalization filters are used to shorten the channel's effective length to the size of a cyclic prefix or the guard interval. In most applications the equalizer succeeds the channel. In acoustic systems, an equalizer is placed in front of a playback loudspeaker to generate a desired impulse response for the concatenation of the equalizer, a loudspeaker, a room impulse response, and a reference microphone. In this paper, we show that shaping the desired impulse response to a shorter reverberation time is more appropriate for acoustical systems than trying to exactly truncate it to a maximum length. Investigations are carried out using a multi-loudspeaker-multi-microphone system.

## 1. INTRODUCTION

Equalization in a listening room is usually carried out on the basis of the following setup: a filter for listening room compensation (LRC) is placed in the signal path in front of a loudspeaker. The goal is to reduce the influence of the succeeding room impulse response (RIR) in order to achieve a signal  $y[n]$  at the position of a reference microphone that is only inaudibly different from the signal  $x[n]$  in front of the equalizer [1]. Let  $c[n]$  be the finite-length RIR and let  $h[n]$  denote the finite-length equalizer. Their  $z$ -transforms are given by  $C(z)$  and  $H(z)$ , respectively. In general,  $C(z)$  is a mixed-phase system, having zeros inside and outside the unit circle. Therefore, only its minimum-phase component can be inverted by a standard causal IIR filter [2]. More recent proposals [3] stress the importance of equalizing the remaining allpass component, too.

Fig. 1 illustrates another traditional setup, where a delayed target system  $\tilde{d}[n] = d[n - n_0]$  is approximated by the concatenated system  $h[n] * c[n]$  in the least squares (LS) sense [4, 5]. Usually, a bandpass-filtered version of a discrete impulse serves as such a system  $d[n]$ .



**Fig. 1.** Single-channel setup for listening room compensation.  $c[n]$  is the room impulse response and  $h[n]$  denotes the equalizer preceding the loudspeaker.

A more relaxed requirement can be found in psychoacoustics: here one uses, for example, the D50-measure for intelligibility of speech, which is defined as the ratio of the energy within 50 ms after the first peak of a RIR versus the complete impulse response's energy [6]. Thus, by choosing a target system with an optimized impulse response of 50 ms duration, we can directly maximize the D50-measure. The appropriate procedure to maximize the energy in a certain region of a desired impulse response has been proposed by Melsa et al. [7] for the application with a discrete multitone transceiver (DMT). If no additional noise sources are assumed, this method is equivalent to Falconer's and Magee's classic approach to channel shortening [8]. In [9] it was generally shown that methods of equalizing channels can be used for RIRs, as well.

In the following sections we shortly summarize the proposal by Melsa et al. and illustrate two major modifications, which become necessary when acoustic impulse responses are to be shortened or shaped. First, we describe why acoustic impulse responses need to be shaped rather than shortened. Finally, we address the spectral properties of an equalized system  $h[n] * c[n]$  and show that the frequency characteristics can be improved by introducing a succeeding short equalizer. A generalization to the multi-input-multi-output (MIMO) case is presented in Section 3. Simulation results are given in Section 4, and some conclusions are drawn in Section 5.

**Notation** Vectors and matrices are printed in boldface. The superscripts  $T$ ,  $*$ , and  $H$  denote transposition, complex conjugation, and Hermitian transposition, respectively. The asterisk  $*$  denotes convolution. The operator  $\text{diag}\{\cdot\}$  turns a vector into a diagonal matrix.

## 2. ACOUSTIC IMPULSE RESPONSE SHAPING

First, let us switch to a formulation used in [10]: a desired concatenated impulse response of equalizer and RIR can be expressed by

$$\mathbf{d}_d = \text{diag}\{\mathbf{w}_d\} \mathbf{C} \mathbf{h} \quad (1)$$

in vector form.  $\mathbf{w}_d$  is a vector that contains ones in the desired region and zeros outside.  $\mathbf{C}$  is a  $(L_c + L_h - 1 \times L_h)$  convolution matrix of the RIR  $c[n]$ .  $L_c$  and  $L_h$  are the lengths of the RIR and the equalizer's coefficient vector  $\mathbf{h}$ , respectively. Accordingly,

$$\mathbf{d}_u = \text{diag}\{\mathbf{w}_u\} \mathbf{C} \mathbf{h} \quad (2)$$

with

$$\mathbf{w}_u = \mathbf{1}_{[L_c + L_h - 1]} - \mathbf{w}_d \quad (3)$$

represents the undesired part of the concatenated impulse response.  $\mathbf{1}_{[L_c+L_h-1]}$  is a vector containing the indicated number of ones. In the following, the matrices  $\mathbf{A}$  and  $\mathbf{B}$  are made up in the same way as in [7]:

$$\mathbf{d}_u^H \mathbf{d}_u = \mathbf{h}^H \mathbf{C}^H \text{diag}\{\mathbf{w}_u\}^H \text{diag}\{\mathbf{w}_u\} \mathbf{C} \mathbf{h} = \mathbf{h}^H \mathbf{A} \mathbf{h}, \quad (4)$$

$$\mathbf{d}_d^H \mathbf{d}_d = \mathbf{h}^H \mathbf{C}^H \text{diag}\{\mathbf{w}_d\}^H \text{diag}\{\mathbf{w}_d\} \mathbf{C} \mathbf{h} = \mathbf{h}^H \mathbf{B} \mathbf{h}. \quad (5)$$

Because most loudspeaker's frequency responses are limited at very low and very high frequencies, it can be advantageous to constrain the maximization to a broad bandpass area. Therefore, we filter the RIR  $c[n]$  with a bandpass filter  $g[n]$  before we set up the accordant convolution matrix  $\mathbf{C}_{\text{BP}}$ . Hence,

$$\mathbf{c}_{\text{BP}}[n] = c[n] * g[n], \quad (6)$$

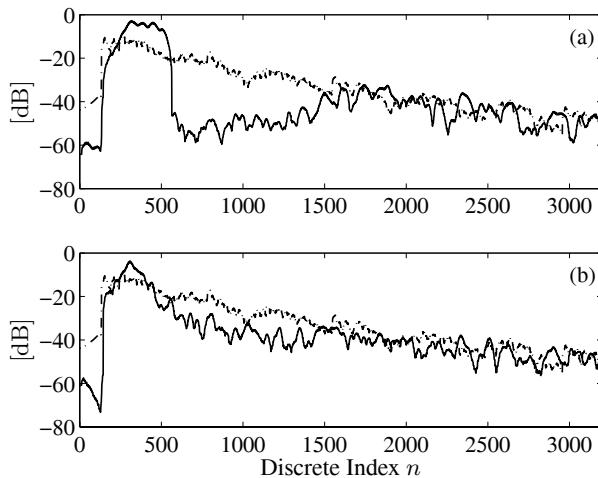
$$\mathbf{B}_{\text{BP}} = \mathbf{C}_{\text{BP}}^H \text{diag}\{\mathbf{w}_d\}^H \text{diag}\{\mathbf{w}_d\} \mathbf{C}_{\text{BP}}. \quad (7)$$

Finally, the optimum equalizer,  $\mathbf{h}_{\text{opt}}$ , for maximizing the energy in a certain region is the solution of the generalized eigenvalue problem [7, 11]

$$\mathbf{B}_{\text{BP}} \mathbf{h}_{\text{opt}} = \mathbf{A} \mathbf{h}_{\text{opt}} \lambda_{\max} \quad (8)$$

with  $\lambda_{\max}$  being the largest eigenvalue and  $\mathbf{h}_{\text{opt}}$  being the corresponding eigenvector.

Fig. 2 shows a comparison of two equalizer design strategies: In chart (a), the solid line shows the squared shortened impulse response, where we directly tried to maximize the D50-measure. Accordingly,  $\mathbf{w}_d$  contains ones between taps 155 and 554; a delay of 154 taps is introduced at the beginning. One major drawback of this procedure is an observable and audible echo with a maximum at tap 1800. The original impulse response (dash-dotted line) exposes a reverberation time of  $\tau_{60} = 400$  ms and its length is 3200. The equalizer has 1200 taps.



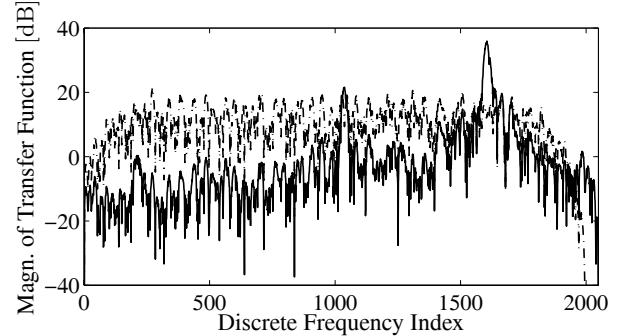
**Fig. 2.** Three squared impulse responses: the dash-dotted line in both plots illustrates the original RIR. In chart (a), the shortened response (solid) results from energy maximization between taps 155 and 554 with a rectangular window. An exponentially decreasing maximization window was used for designing  $\mathbf{h}_{\text{opt}}$ , which produced the shaped impulse response denoted by the solid line in chart (b).

Therefore, we propose to apply an exponentially decreasing window to define the region, whose energy will be maximized. The decreasing envelope is calculated by

$$w_d[n] = \begin{cases} 0 & \text{for } 0 \leq n \leq n_0 - 1 \\ 10^{q(n-n_0)} & \text{for } n_0 \leq n \end{cases} \quad (9)$$

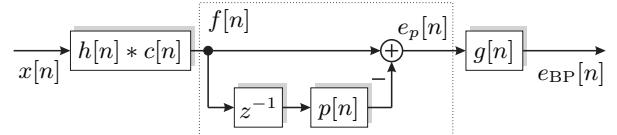
where the factor  $q$  has been chosen heuristically as  $q = -3 \cdot 10^{-5}$ .  $\mathbf{w}_u$  is designed according to equation (3). The initial delay is set to  $n_0 = 155$ . The solid line in Fig. 2(b) shows the resulting impulse response. It displays a slower decay. However, we cannot observe any late echoes.

**Spectral Aspects** In Fig. 3 we investigate the magnitude of the transfer function of the concatenated system  $h_{\text{opt}}[n] * c[n]$ . One characteristic observation of Fig. 3 is the very peaky response of the equalized system using the impulse response shaping approach (solid line).



**Fig. 3.** Magnitude of the transfer function of the concatenated system  $h_{\text{opt}}[n] * c[n]_{\text{BP}}$  (solid line). The dashed-dotted line marks the response after linear predictive post-equalization.

To solve this problem we propose to apply a short post-equalizer to the shaped impulse response  $h_{\text{opt}}[n] * c[n]$ . This is carried out by prediction error filter  $f[n]$  that is based on a one-step linear predictor  $p[n]$  with a relatively short impulse response. Fig. 4 shows its setup.



**Fig. 4.** Signal model of a linear predictive post-equalizer: a band-pass filter  $g[n]$  is used to spectrally weight the initial error signal  $e[n]$ .

The error signal

$$e_p[n] = x[n] * h[n] * c[n] * f[n] \quad (10)$$

is weighted by a bandpass filter  $g[n]$ , the same that is used during the shaping-filter design procedure, which we described at the beginning of this section:

$$\begin{aligned} e_{\text{BP}}[n] &= g[n] * e_p[n] \\ &= \mathbf{x}^T[n] (\mathbf{G} \mathbf{c}_{\text{EQed}} - \mathbf{G} \mathbf{C}_{\text{EQed},-1} \mathbf{p}) \end{aligned} \quad (11)$$

with

$$\mathbf{x}[n] = [x[n], \dots, x[n - L_g - L_c - L_h - L_p + 3]]^T \quad (12)$$

$$\mathbf{g} = [g[0], \dots, g[L_g - 1]]^T \quad (13)$$

$$\mathbf{c}_{\text{EQed}}[n] = h_{\text{opt}}[n] * c[n] \quad (14)$$

$$\mathbf{c}_{\text{EQed}} = \quad (15)$$

$$\underbrace{[c_{\text{EQed}}[0], \dots, c_{\text{EQed}}[L_c + L_h - 2], 0, \dots, 0]}_{L_h + L_c + L_p - 1}^T.$$

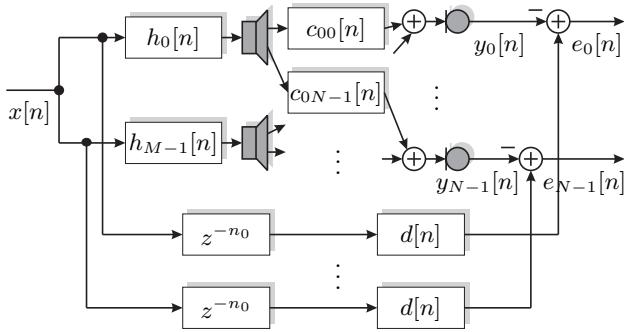
$\mathbf{C}_{\text{EQed},-1}$  is a  $L_h + L_c + L_p - 1 \times L_p$  convolution matrix made of the preceding non-zero part of  $\mathbf{c}_{\text{EQed}}$  with an additional first row of zeros to take into account the prediction delay of one sample (see Fig. 4).  $L_p$  and  $L_g$  are the lengths of the prediction filter and the bandpass, respectively. The calculation of the vector  $\mathbf{p}$  that minimizes the target function  $E\{e_{\text{BP}}^2[n]\}$  leads to

$$\mathbf{p} = \left( \mathbf{C}_{\text{EQed},-1}^H \mathbf{G}^H E \left\{ \mathbf{x}[n] \mathbf{x}^H[n] \right\} \mathbf{G} \mathbf{C}_{\text{EQed},-1} \right)^{-1} \mathbf{C}_{\text{EQed},-1}^H \mathbf{G}^H E \left\{ \mathbf{x}[n] \mathbf{x}^H[n] \right\} \mathbf{G} \mathbf{c}_{\text{EQed}}. \quad (16)$$

The design is usually carried out under the assumption of a white and stationary excitation signal  $x[n]$ . For the predictor design, the bandpass causes a “don’t care” region outside of its passband. This results in a “bathtub-like” spectral shape of the signal  $e[k]$ ;  $e_{\text{BP}}[k]$  is spectrally flat. One further bandpass can be applied to  $e_{\text{BP}}[k]$  in order to achieve a bandpass-weighted signal at the loudspeaker.

### 3. MIMO ACOUSTIC IMPULSE RESPONSE SHAPING

So far, the equalizer design is restricted to a fixed spatial setup with one loudspeaker and one microphone, which the listener has to stay close to. By designing filters with respect to  $N$  microphones arranged on a circle, the zone within the circle can be equalized [12]. The degree of equalization can be increased by using more than one loudspeaker. In the following we extend the shaping approach of Section 2 to the MIMO case. We examine the setup shown in Fig. 5: the  $M$  identical target systems  $d[n]$  are shown for comparability with the traditional MIMO LS solution as introduced in [5]. The  $M$  equalization filters  $h_0[n]$  to  $h_{M-1}[n]$  for each of the



**Fig. 5.** MIMO setup for listening room compensation.  $c_{ik}[n]$  are room impulse responses and  $h_k[n]$  denotes one of the  $M$  equalizers, one preceding each loudspeaker.

$M$  loudspeakers are stacked into a single vector

$$\check{\mathbf{h}} = [h_0[0], \dots, h_0[L_h - 1], \dots, h_{M-1}[0], \dots, h_{M-1}[L_h - 1]]^T. \quad (17)$$

Accordingly, matrices  $\check{\mathbf{A}}_{\text{BP}}$ ,  $\check{\mathbf{B}}_{\text{BP}}$ , and  $\check{\mathbf{C}}_{\text{BP}}$  are now set up with respect to all  $MN$  loudspeaker-microphone pairs. We define each submatrix of  $\check{\mathbf{C}}_{\text{BP}}$ :

$$[\check{\mathbf{C}}_{\text{BP}}]_{ik} = \text{diag}\{\mathbf{w}_d\} \mathbf{C}_{\text{BP},ik}, \quad (18)$$

$$\mathbf{c}_{\text{BP},ik}[n] = g[n] * c_{ik}[n] \quad (19)$$

with  $0 \leq i \leq N - 1$  and  $0 \leq k \leq M - 1$ .  $\mathbf{C}_{\text{BP},ik}$  is a  $(L_h + L_c + L_g - 2) \times L_h$  convolution matrix made up of  $c_{\text{BP},ik}[n]$ . Now we can calculate

$$\check{\mathbf{B}}_{\text{BP}} = \check{\mathbf{C}}_{\text{BP}}^H \check{\mathbf{C}}_{\text{BP}}. \quad (20)$$

$\check{\mathbf{A}}_{\text{BP}}$  is calculated in a similar way without a bandpass filter and using  $\mathbf{w}_u$  instead of  $\mathbf{w}_d$ . Finally, we can extract the stacked coefficient vector by solving the eigenvalue problem

$$\check{\mathbf{B}}_{\text{BP}} \check{\mathbf{h}}_{\text{opt}} = \check{\mathbf{A}}_{\text{BP}} \check{\mathbf{h}}_{\text{opt}} \lambda_{\max}. \quad (21)$$

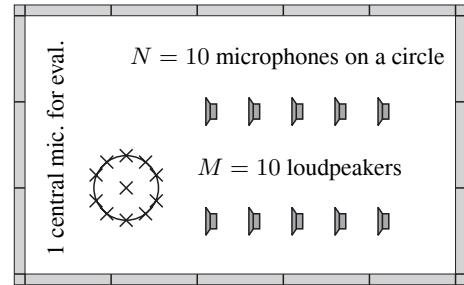
Finally, the resulting impulse response from the loudspeakers to the  $i$ th microphone,

$$\sum_{k=0}^{M-1} h_k[n] * c_{ik}[n],$$

is fed through a separate bandpass-weighted predictor as shown in Fig. 4.

### 4. SIMULATION RESULTS

Fig. 6 shows the spatial setup for the following studies: there are ten loudspeakers in a room of size  $6 \text{ m} \times 10 \text{ m} \times 3 \text{ m}$  (width, depth, and height). Ten microphones are arranged on a circle with a radius of 6 cm. An eleventh microphone is placed in the center of the circle – it is used only for the evaluation of the equalizer design. The reverberation time amounts to 400 ms and all RIRs were simulated using the well-known image method [13]. The RIRs are truncated to a length of 3200 samples. All systems are sampled with a frequency of  $f_s = 8 \text{ kHz}$ .

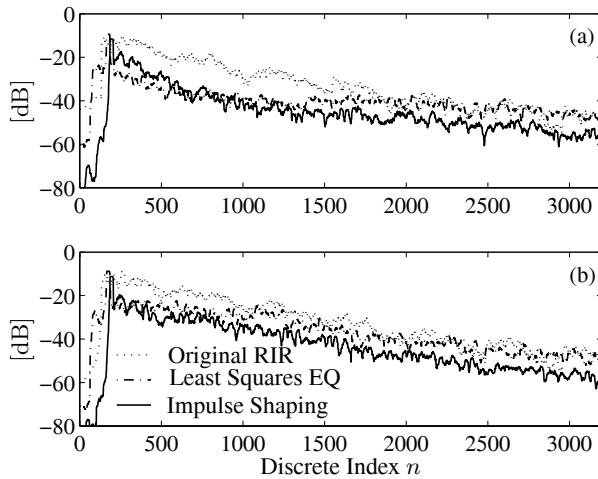


**Fig. 6.** Drawing of the spatial setup used for all MIMO-investigations. The radius for the circular array is 6 cm – a spacing of 3.8 cm between adjacent microphones.

For comparison with the novel shaping equalizer according to Section 3 we used a MIMO LS equalizer as shown in [5]. The

parameters for the equalizer design are:  $N = 10$ ,  $L_h = 1200$ ,  $L_c = 3200$ ,  $L_p = 40$ , and  $n_0 = 154$ . The linear phase bandpass was designed using a Hamming window. We chose  $L_g = 41$ ; -6 dB frequencies lie at 200 Hz and 3600 Hz, respectively.

We examined two cases for the design:  $M = 1$  and  $M = 8$ . Fig. 7 shows the squared impulse responses of the resulting responses at the first microphone of the array (highest in Fig. 6, chart (a) in Fig. 7) and the central microphone (chart (b)). Eight loudspeaker are in operation, the column of two loudspeakers being closest to the microphones is not used: at the position of the design microphone we can observe a fast initial decay of the LS equalized response. However, the slight rise at sample 2000 causes a noticeable echo. After the initial peak, the shaped impulse response decays more slowly but constantly. It stays below the original RIR. Note that the first peak of all three responses are adjusted to the same level. If only one loudspeaker is operated with ten



**Fig. 7.** Six squared impulse responses: all ten microphones in the circular array are used in the design procedure; eight loudspeakers are operated. The dotted line indicates the original RIR, the dash-dotted least squares equalized responses. The solid line is created by the novel impulse response shaping approach. Chart (a) shows the original and equalized responses with respect to microphone 1, which is part of the design procedure. Chart (b) shows equalization results with respect to the eleventh microphone in the center of the array which is not used during the design.

microphones, we cannot observe any important influence of the equalizer: the equalized responses are very similar to the original RIR. Ten loudspeakers produce more rapidly decaying equalized responses compared to eight loudspeakers. Larger equalizer lengths  $L_h$  produce lower tails of the squared impulse responses.

## 5. CONCLUSIONS

In this contribution we have modified the impulse shortening concept for discrete multitone transceivers to the application with acoustic room impulse responses. One major aspect has been the modification of the weighting window for the desired temporal shape of the equalized impulse response. Another aspect was the design of a short succeeding equalizer on the basis of a linear predictive filter. All derivations have been generalized to the MIMO

case with  $M$  loudspeakers and  $N$  microphones. Informal listening tests have confirmed the advantages of the novel equalizer compared to the widely used least squares-equalizer.

## 6. REFERENCES

- [1] J. N. Mourjopoulos, "Digital Equalization of Room Acoustics," *Journal of the Audio Engineering Society*, vol. 42, no. 11, pp. 884–900, Nov. 1994.
- [2] S. T. Neely and J. B. Allen, "Invertibility of a Room Impulse Response," *Journal of the Acoustical Society of America (JASA)*, vol. 66, pp. 165–169, July 1979.
- [3] B. D. Radlović and R. A. Kennedy, "Nonminimum-Phase Equalization and its Subjective Importance in Room Acoustics," *IEEE Trans. on Speech and Audio Processing*, vol. 8, no. 6, pp. 728–737, Nov. 2000.
- [4] S. J. Elliott and P. A. Nelson, "Multiple-Point Equalization in a Room Using Adaptive Digital Filters," *Journal of the Audio Engineering Society*, vol. 37, no. 11, pp. 899–907, Nov. 1989.
- [5] O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast Deconvolution of Multichannel Systems using Regularization," *IEEE Trans. on Speech and Audio Processing*, vol. 6, no. 2, pp. 189–194, Mar. 1998.
- [6] International Organization for Standardization (ISO), "ISO Norm 3382: Acoustics – Measurement of the Reverberation Time of Rooms with Reference to other Acoustical Parameters."
- [7] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse Response Shortening for Discrete Multitone Transceivers," *IEEE Trans. on Communications*, vol. 44, no. 12, pp. 1662–1672, Dec. 1996.
- [8] D. D. Falconer and F. R. Magee, "Adaptive Channel Memory Truncation for Maximum Likelihood Sequence Estimation," *The Bell System Technical Journal*, vol. 52, no. 9, pp. 1541–1562, Nov. 1973.
- [9] M. Kallinger and A. Mertins, "Room Impulse Response Shortening by Channel Shortening Concepts," in *Asilomar Conference on Signals, Systems, and Computers*, Monterey, CA, USA, Oct. 2005.
- [10] G. Arslan, B. L. Evans, and S. Kiaei, "Equalization for discrete multitone transceivers to maximize bit rate," *IEEE Trans. on Signal Processing*, vol. 49, no. 12, pp. 3123–3135, Dec. 2001.
- [11] R. K. Martin, D. Ming, B. L. Evans, and C. R. Johnson Jr., "Efficient Channel Shortening Equalizer Design," *Journal on Applied Signal Processing*, vol. 13, pp. 1279–1290, Dec. 2003.
- [12] T. Ajdler, L. Sbaiz, and M. Vetterli, "Plenacoustic Function on the Circle with Application to HRTF Interpolation," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Philadelphia, PA, USA, Mar. 2005, pp. 273–276.
- [13] J. B. Allen and D. A. Berkley, "Image Method for Efficiently Simulating Small-Room Acoustics," *J. Acoust. Soc. Amer.*, vol. 65, pp. 943–950, 1979.