

ROOM IMPULSE RESPONSE SHORTENING BY CHANNEL SHORTENING CONCEPTS

Markus Kallinger and Alfred Mertins

University of Oldenburg, Institute 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 from 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 data-transmission 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 modify the channel shortening paradigm and show that shaping the desired impulse response to a shorter reverberation time is more appropriate for acoustical systems than exactly truncating it.

1. INTRODUCTION

Equalization of an acoustic system 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 obtain a signal $y[n]$ at the position of a reference microphone that is hardly distinguishable by a human listener from the signal $s[n]$ in front of the equalizer [1]. The basic setup is depicted in Fig. 1. $c[n]$ is the finite-length RIR and $h[n]$ denotes the finite-length equalizer. Their z -transforms are given by $C(z)$ and $H(z)$, respectively. In general, $C(z)$ is

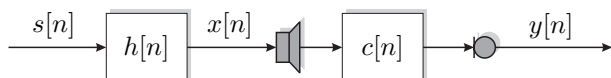


Fig. 1. Single-channel setup for listening room compensation. $c[n]$ is the room impulse response and $h[n]$ denotes the equalizer preceding the loudspeaker.

a 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.

Finite-length equalizers are designed with the aim of minimizing the squared error between the concatenation of $c[n]$, $h[n]$, and a given target system [4, 5]. Usually, a bandpass-filtered version of a delayed impulse serves as such a system. In order to compare the least-squares equalizer with the novel impulse shaping approach, we briefly review its derivation in Section 2.

A more relaxed requirement than choosing a bandpass weighted impulse as the target system 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) and further investigated in [8]. In Section 3, we summarize this method. Section 4 introduces two major modifications that make the technique more appropriate for the shortening of acoustic impulse responses.

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 discrete time index is denoted by n . The operator $\text{diag}\{\cdot\}$ turns a vector into a diagonal matrix. $\delta[n]$ denotes a discrete impulse.

2. LEAST-SQUARES EQUALIZATION

In traditional least-squares equalization for LRC, a finite-length filter $h[n]$ precedes the RIR $c[n]$. The equalizer is designed to minimize the squared error between the concatenation $h[n]*c[n]$ and a target system $g[n]$ that is delayed by n_0 taps. The filter $g[n]$ is typically chosen as a bandpass filter. Fig. 2 shows the according setup.

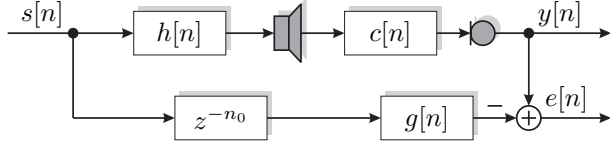


Fig. 2. Setup for listening room compensation using a least-squares equalizer. $c[n]$ is the room impulse response and $h[n]$ denotes the equalizer preceding the loudspeaker.

The error signal $e[n]$ can be expressed as

$$e[n] = \mathbf{s}^T[n] \mathbf{C} \mathbf{h} - \mathbf{s}^T[n] \mathbf{g}_{n_0} \quad (1)$$

with

$$\begin{aligned} \mathbf{s}[n] &= [s[n], \dots, s[n - L_h - L_c + 2]]^T, \\ \mathbf{g}_{n_0} &= [\underbrace{0, \dots, 0}_{n_0}, g[0], \dots, g[L_g - 1], \underbrace{0, \dots, 0}_{L_h + L_c - 1 - L_g - n_0}]^T. \end{aligned} \quad (2)$$

In these expressions, L_h and L_c are the lengths of the equalizer and the RIR, respectively. L_g represents the length of the target system, which is usually a bandpass filter. The $(L_h + L_c - 1 \times L_h)$ -dimensional convolution matrix \mathbf{C} is made up by the RIR $c[n]$. The equalizer that minimizes the error signal's power $E\{e^2[n]\}$ for a white-noise input $s[n]$ is given by

$$\mathbf{h} = (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H \mathbf{g}_{n_0}. \quad (4)$$

3. IMPULSE RESPONSE SHORTENING

We choose a formulation used by Arslan et al. [9]: a desired concatenated impulse response of equalizer and RIR can be expressed by

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

in vector form. \mathbf{w}_d is a vector that contains ones in the desired region and zeros outside. The convolution matrix \mathbf{C} and the equalizer's coefficient vector \mathbf{h} are defined in the same way as in Section 2. Accordingly,

$$\mathbf{d}_u = \text{diag}\{\mathbf{w}_u\} \mathbf{C} \mathbf{h} \quad (6)$$

with

$$\mathbf{w}_u = \mathbf{1}_{[L_c + L_h - 1]} - \mathbf{w}_d \quad (7)$$

represents the undesired part of the concatenated impulse response. As a side condition, this part's energy is kept constant while the energy of \mathbf{d}_d is maximized. $\mathbf{1}_{[L_c + L_h - 1]}$ is a vector containing the indicated number of entries, which

are all ones. In the following, the matrices \mathbf{A} and \mathbf{B} are assembled in the same way as in [7]:

$$\mathbf{d}_u^H \mathbf{d}_u = \mathbf{h}^H \mathbf{C}^H \text{diag}\{\mathbf{w}_u\}^2 \mathbf{C} \mathbf{h} = \mathbf{h}^H \mathbf{A} \mathbf{h}, \quad (8)$$

$$\mathbf{d}_d^H \mathbf{d}_d = \mathbf{h}^H \mathbf{C}^H \text{diag}\{\mathbf{w}_d\}^2 \mathbf{C} \mathbf{h} = \mathbf{h}^H \mathbf{B} \mathbf{h}. \quad (9)$$

One first modification for acoustic systems becomes necessary when we take into account the loudspeaker's limited playback capabilities at very low and very high frequencies. Therefore, we constrain the maximization procedure to a broad bandpass area. First, we apply the bandpass $g[n]$ of Section 2 to the RIR:

$$c_{\text{BP}}[n] = c[n] * g[n]. \quad (10)$$

Consequently, we assemble a $(L_h + L_c + L_g - 2 \times L_h)$ -dimensional convolution matrix \mathbf{C}_{BP} on the basis of $c_{\text{BP}}[n]$. We obtain

$$\mathbf{B}_{\text{BP}} = \mathbf{C}_{\text{BP}}^H \text{diag}\{\mathbf{w}_{\text{BP},d}\}^H \text{diag}\{\mathbf{w}_{\text{BP},d}\} \mathbf{C}_{\text{BP}}. \quad (11)$$

Compared to \mathbf{w}_d , $\mathbf{w}_{\text{BP},d}$ is $L_g - 1$ elements longer; $\lfloor L_g/2 \rfloor$ zeros are prefixed to compensate the bandpass' group delay. \mathbf{A} is not modified. Finally, the optimum equalizer, \mathbf{h}_{opt} , for maximizing the energy in a certain region is the solution of the generalized eigenvalue problem

$$\mathbf{B}_{\text{BP}} \mathbf{h}_{\text{opt}} = \mathbf{A} \mathbf{h}_{\text{opt}} \lambda_{\text{max}} \quad (12)$$

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

The impact of channel shortening applied to a typical RIR can be observed in Fig. 3. The parameters are as follows: The RIR has been simulated with a reverberation time of $\tau_{60} = 400$ ms using Allen and Berkley's image method [10] at a sampling frequency of 8 kHz. Its length has been truncated to $L_c = 4000$ taps. The equalizer contains $L_c = 2000$ coefficients; the maximization has been carried out for taps 149 to 548. The filter $g[n]$ has been designed as a linear-phase bandpass with $L_g = 41$ and -6 dB frequencies at 200 Hz and 3600 Hz, respectively (Matlab call `fir1(40, [0.05 0.9])`).

The first peak of the RIR occurred at tap 146 – therefore, we tried to maximize the succeeding 50 ms. Fig. 3 shows the squared impulse responses of the original and the shortened systems. As we can see from Fig. 3, the impulse response shortening is quite effective for acoustic channels. However, informal listening tests showed that the energy concentration around tap 3000 results in an audible echo.

Fig. 4 shows the magnitude frequency responses of the original RIR and its shortened version. We have to note that the frequency response of the shortened filter exposes a very distinct peak near $\omega = \pi/2$. This leads to an unacceptable amount of audible distortion.

Additional measures have to be taken to cope with both the late echo and the peaky spectral response. Both problems will be addressed in the following section.

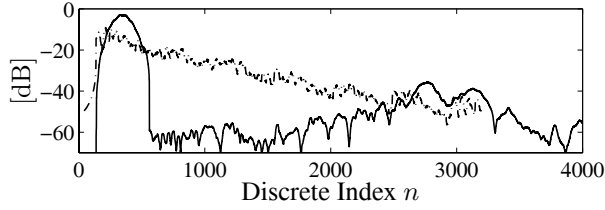


Fig. 3. Squared impulse responses of the original and the shortened systems (i.e., $c^2[n]$ and $(h * c)^2[n]$). The dash-dotted line describes the original RIR $c[n]$; the solid one is the shortened response $h[n] * c[n]$ with energy maximization between taps 149 and 548.

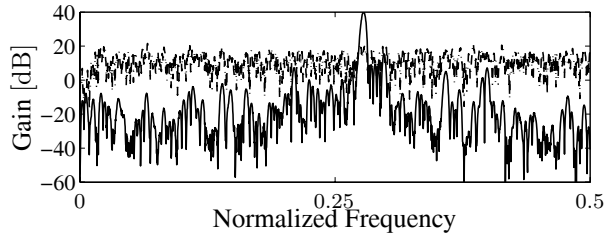


Fig. 4. Magnitude frequency responses. The dash-dotted line describes the original RIR; the solid one belongs to the shortened overall system.

4. MODIFICATIONS FOR ACOUSTIC RESPONSES

4.1. Temporal Aspects

One goal in the design of an impulse response shortening procedure is to avoid audible late echoes. Another aspect is to preserve the general shape of a natural RIR, which usually decays exponentially with time. Therefore, it would be desirable not to achieve a temporal envelope like the equalized one in Fig. 3, but one that decays more quickly than the original RIR and thus yields a shorter reverberation time.

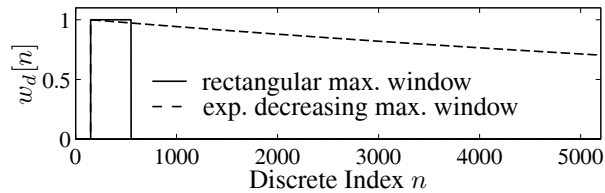


Fig. 5. Maximization windows as functions of time.

One stage for modifications is the maximization window w_d . The solid line in Fig. 5 shows the maximization window that has been used to produce Fig. 3. As an alternative we have tried an exponentially decaying window with a reverberation time that was just a little shorter than the original one. Equation (13) describes the design rule for the

novel window:

$$w_d[n] = \begin{cases} 0 & \text{for } 0 \leq n \leq n_0 - 1 \\ 10^{q(n-n_0)} & \text{for } n_0 \leq n \end{cases} \quad (13)$$

where the factor q has been chosen heuristically as $q = -3 \cdot 10^{-5}$.

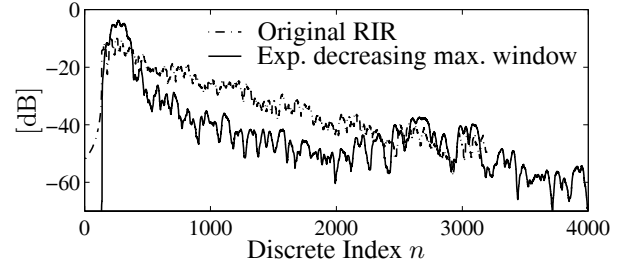


Fig. 6. Squared impulse responses of the original and the shortened systems. The solid line is produced using an exponentially decaying window.

Fig. 6 shows the power delay spectra of two equalized responses. The original RIR has already been shown in Fig. 3. The solid line results from the application of an exponentially decaying maximization window. The parameters for its design were chosen according to the paragraph above. One can see that the equalized response decreases more steeply than the RIR but not as steeply as using a rectangular window. However, we can observe a less exposed echo around tap 3000. Hence, the technique can appropriately be described as impulse response shaping rather than impulse response shortening.

4.2. Spectral Aspects

Simulation results have shown that the extent of the spectral peak shown in Fig. 4 is not crucially affected by the choice of the maximization window as discussed in Section 4.1. Therefore, we have to introduce an additional measure to overcome this problem: we propose to post-equalize the shaped response $h[n] * c[n]$ using a prediction error filter $f[n]$ that is based on a one-step linear predictor $p[n]$ with a relatively short impulse response. Fig. 7 displays the according setup.

The error signal

$$e_p[n] = s[n] * h[n] * c[n] * f[n] \quad (14)$$

is weighted by a bandpass filter $g[n]$. This measure is taken to focus the predictor's performance to the spectral area of interest. The bandpass is the same that is used with the least-squares equalizer and with the weighting in the impulse-response shortening procedure (see equation (10) for de-

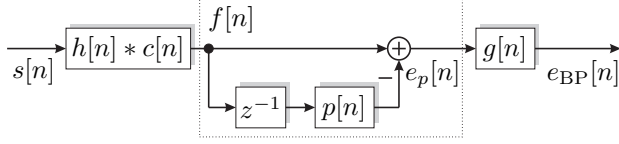


Fig. 7. Signal model of a linear predictive post-equalizer. A bandpass filter $g[n]$ is used to spectrally weight the initial error signal $e_p[n]$.

tails). The final error signal can be expressed as

$$\begin{aligned} e_{BP}[n] &= g[n] * e_p[n] \\ &= \mathbf{s}^T[n] (\mathbf{G}\mathbf{c}_{EQed} - \mathbf{G}\mathbf{c}_{EQed,-1}\mathbf{p}) \end{aligned} \quad (15)$$

with

$$\mathbf{s}[n] = [s[n], \dots, s[n - L_g - L_c - L_h - L_p + 3]]^T \quad (16)$$

$$\mathbf{g} = [g[n], \dots, g[n + L_g - 1]]^T \quad (17)$$

$$\mathbf{c}_{EQed}[n] = h[n] * c[n] \quad (18)$$

$$\mathbf{c}_{EQed} = \quad (19)$$

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

$\mathbf{c}_{EQed,-1}$ is a $(L_h + L_c + L_p - 1 \times L_p)$ -dimensional convolution matrix made of the preceding nonzero part of \mathbf{c}_{EQed} with an additional first row of zeros to take into account the delay of one sample (see Fig. 7). The convolution matrix \mathbf{G} possesses the size $(L_g + L_c + L_h + L_p - 2) \times (L_h + L_c + L_p - 1)$. 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 $E\{e_{BP}^2[n]\}$ leads to

$$\begin{aligned} \mathbf{p} &= (\mathbf{C}_{EQed,-1}^H \mathbf{G}^H E\{\mathbf{s}^*[n]\mathbf{s}^T[n]\} \mathbf{G}\mathbf{c}_{EQed,-1})^{-1} \\ &\quad \mathbf{C}_{EQed,-1}^H \mathbf{G}^H E\{\mathbf{s}^*[n]\mathbf{s}^T[n]\} \mathbf{G}\mathbf{c}_{EQed}, \end{aligned} \quad (20)$$

and finally

$$f[n] = \delta[n] - p[n - 1]. \quad (21)$$

The actual design is usually carried out under the assumption of a white and stationary excitation signal $s[n]$. In that case, the inner correlation matrices can be cancelled out. The bandpass causes a “don’t care” region outside of its cut-off frequencies. This results in a “bathtub-like” spectral shape of the signal $e_p[k]$; $e_{BP}[k]$ is spectrally flat. One further bandpass can be applied to $e_{BP}[k]$ in order to generate a bandpass-weighted signal at the loudspeaker.

5. SIMULATION RESULTS

For the following simulation results we have used the same parameters as for the generation of Fig. 6. We have set

the post-equalizer’s length to $L_p = 40$; the bandpass has already been specified before. Consequently, the post-equalizer’s overall-length amounts to 120. Fig. 8 illustrates the impact of the linear predictive equalizer on the shaped response. We can conclude that the post-equalizer weakens the shaping filter’s impact.

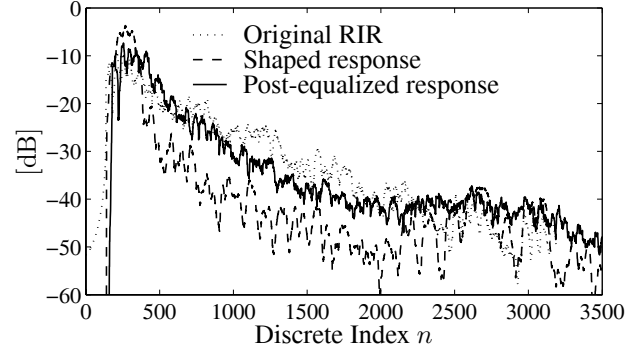


Fig. 8. Squared impulse responses of the original and two shortened systems. The dashed line describes the equalized RIR in front of the post-equalizer, the solid one is obtained at its output.

Fig. 9 contains the magnitude frequency responses for the filters whose power delay spectra were shown in Fig. 8. From this figure we see that the post equalizer could completely remove the spectral peak. However, the resulting frequency response exposes a similar amount of gaps as the original RIR.

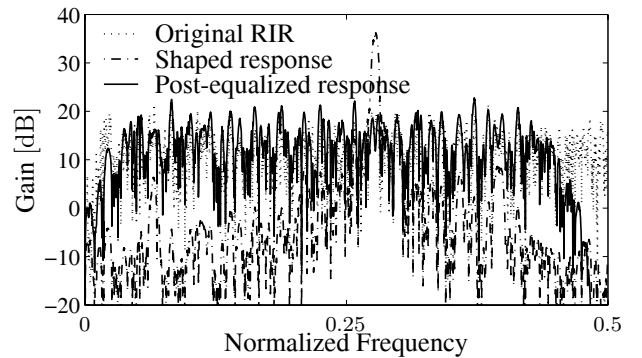


Fig. 9. Magnitude frequency responses. The dash-dotted line describes the temporally shaped spectrum in front of the post-equalizer, the solid one is obtained at its output.

Finally, we compare the novel impulse response shaping approach to the widely-used least-squares equalizer according to Section 2. The equalizer’s length was chosen to $L_h = 2000$. Fig. 10 contains previously shown functions in comparison with the equalized response using a least-squares approach. As with the impulse response shortening

procedure, fast initial decay is achieved at the cost of late echoes.

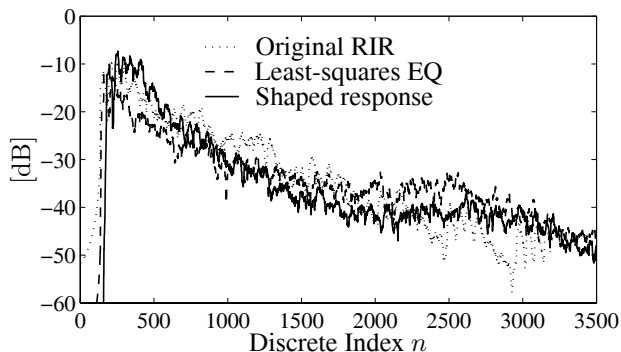


Fig. 10. Squared impulse responses of the original and two shortened systems. The dashed line describes the least-squares equalized RIR, the solid one is the shaped response.

In the spectral domain, the least-squares equalizer can compensate more small gaps than the novel shaping approach. However, the least-squares equalizer's spectral envelope exposes much wider gaps (see Fig. 11).

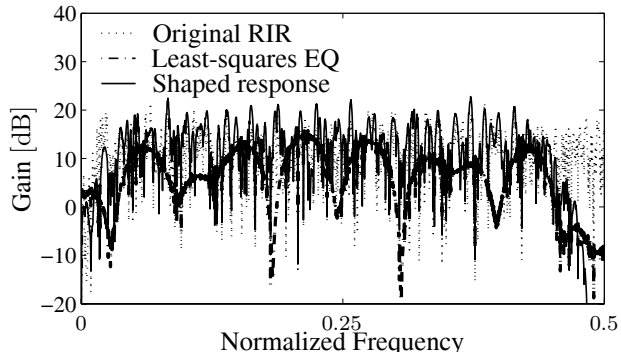


Fig. 11. Magnitude frequency responses. The solid curve was produced using a least-squares equalizer, the dashed one belongs to the novel shaping approach.

Informal listening test confirm the enhanced equalization results of the impulse response shaping technique, especially in the time domain. The clearness of speech could be enhanced; late echoes can be neglected compared to impulse response shortening and least-squares equalization. The perceived impression is that the shaping technique makes the listening room larger and less reverberant at the same time – sound samples are available for download at [11].

6. CONCLUSIONS

The investigations have shown that channel shortening techniques known from data transmission can be successfully applied to room impulse responses. However, to be suitable for audio applications, this involves some extensions to the known approaches. We proposed to use an adjusted maximization window during the filter design. The resulting equalized response is post-equalized using a low-order linear predictive filter. Third, we have introduced bandpass weighting for both equalizer stages to take into account the loudspeaker's limited transmission capabilities at very low and high frequencies. The novel equalizer shows enhanced results compared to the commonly used least-squares equalizer. An open topic remains the design of the adjusted maximization window. In the present work, the desired lower reverberation time has been chosen heuristically and the original reverberation time had to be known.

7. 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] 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.
- [9] 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.
- [10] 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.
- [11] <http://www.uni-oldenburg.de/sigproc/demos.html>