

VARIABLE CONTROL OF THE PRE-RESPONSE ERROR IN MIXED PHASE AUDIO PRECOMPENSATION

Lars-Johan Brännmark and Anders Ahlén

Uppsala University
Department of Engineering Sciences, Signals and Systems
PO Box 534, SE-751 21 Uppsala, Sweden

ABSTRACT

We introduce a method for controlling the spatial robustness of a mixed phase loudspeaker equalizer design. The emphasis is on time domain behaviour and the pre-response error that is always present in mixed phase design. Based on measurements from a small spatial region, a mixed phase compensator is regularized to be valid also over a large spatial region. The regularization can be applied gradually to match any size of listener region, while fulfilling a set of constraints for the pre-response error. The compensation thus avoids the unacceptably high pre-response error levels that generally occur outside the measurement region. The proposed compensator design, which is validated on measurements in both small and large spatial regions, is shown to produce excellent results.

Index Terms— Loudspeaker equalization, Acoustic signal processing, Robustness, All-pass filter

1. INTRODUCTION

Spatial robustness is a critical factor in the design of digital audio compensation filters. A particularly challenging aspect of this is the *pre-response errors*, or pre-rings, that occur in mixed phase filter design when the noncausal part of the filter does not exactly match the Room Transfer Function (RTF) at all points in the listening space [1, 2]. Recently, a method for cautious mixed phase equalization was introduced which admits control of the pre-rings [3]. The performance of that method depends on the existence of a near-common excess phase part among the measured RTFs. In a set of p measured RTFs, the near-common part is detectable as dense clusters of non-minimum phase zeros, each cluster containing one zero from each RTF. As the intended listening region increases in size however, such near-common zeros are less likely to exist among the RTFs, making a mixed phase design less feasible. An obvious solution would be to use a pure minimum phase design, which completely avoids pre-response errors, but this may in some cases lead to worse time domain performance.

In this paper we introduce an extension of our previous approach, featuring a variable mixed phase compensation. In order to increase the region of validity of a small-region design, a regularization of the noncausal part of the filter is introduced, along with a causal all-pass filter which controls the influence of the noncausal part. This approach admits a partial inversion of nonminimum phase zeros which are near-common only for a very limited spatial region. The efficiency of the method is verified by measurements conducted in cubic shells of increasing size.

First author is also with Dirac Research AB, Hanselligatan 6, SE-754 50 Uppsala, Sweden

Remarks on the notation: Filters and RTFs are represented by polynomials and rational functions in the backward time-shift operator q^{-1} , ($q^{-1}s(k) = s(k-1)$), where q^{-1} corresponds to z^{-1} in the frequency domain. A polynomial (polynomial matrix) is denoted by italic capital (bold capital) letters as $P(q^{-1}) = p_0 + p_1q^{-1} + \dots + p_{n_P}q^{-n_P}$ ($\mathbf{P}(q^{-1}) = \mathbf{P}_0 + \mathbf{P}_1q^{-1} + \dots + \mathbf{P}_{n_P}q^{-n_P}$), whereas rational matrices, indicated by bold calligraphic letters, are represented on common denominator form as $\mathcal{G}(q^{-1}) = \mathbf{Q}(q^{-1})/P(q^{-1})$. Scalar rational functions are denoted by normal calligraphic letters as $\mathcal{G}(q^{-1})$. For any polynomial (polynomial matrix) we define the conjugate as $P_*(q) = P(q) = p_0 + p_1q + \dots + p_{n_P}q^{n_P}$ ($\mathbf{P}_*(q) = \mathbf{P}^T(q) = \mathbf{P}_0^T + \mathbf{P}_1^Tq + \dots + \mathbf{P}_{n_P}^Tq^{n_P}$).

2. SIMO FEEDFORWARD COMPENSATION WITH PRE-RESPONSE ERROR CONSTRAINTS

As a framework for deriving spatially robust filters we shall use that of SIMO feedforward control [4]. Let the multi-point error signal of a compensated SIMO acoustical system be described by

$$\mathbf{y}(k) = \mathcal{D}w(k) - \mathcal{H}\mathcal{R}w(k) \quad (1)$$

where $\mathcal{H}(q^{-1})$ and $\mathcal{D}(q^{-1})$ are the RTFs and the desired responses, respectively, between the loudspeaker and $p \gg 1$ measurement positions, represented on matrix fraction description form as

$$\mathcal{H} = B/A = \begin{bmatrix} B_1 \\ \vdots \\ B_p \end{bmatrix} \frac{1}{A} \quad ; \quad \mathcal{D} = D/E = \begin{bmatrix} D_1 \\ \vdots \\ D_p \end{bmatrix} \frac{1}{E}. \quad (2)$$

A and E are stable and monic polynomials while \mathcal{R} is a stable and possibly noncausal compensator operating on the source signal $w(k)$. The compensator \mathcal{R} is to be designed so as to minimize the spatial MSE

$$J = E \{ \|\mathbf{y}(k)\|_2^2 \} = E \left\{ \text{tr} \left(\mathbf{y}(k)\mathbf{y}^T(k) \right) \right\}. \quad (3)$$

In other words we are required to keep the measured signals in the p measurement points as close as possible to the desired signals as defined by $\mathcal{D}(q^{-1})w(k)$. Further, assume that the individual acoustic propagation delay $q^{-\Delta_i}$ associated with each RTF $\mathcal{H}_i = B_i/A$ is removed prior to our filter design. An ideal target response is then set to be a (large) delay of d samples at all points $i = 1, \dots, p$, $D_i = q^{-d}$. Since acoustical RTFs frequently contain nonminimum phase zeros, a mixed phase compensator is required to fully invert the RTFs. However, since this can only be accomplished by the use of a noncausal filter, pre-rings are likely to occur in the filter, causing pre-response errors in the equalized system. In [3] a mixed phase compensation design was proposed with the ability to fully

control the pre-ringing effects. It was shown that a noncausal multi-point MSE optimal feedforward compensator, under the constraint of zero pre-ringing error, can be expressed as

$$\mathcal{R}(q^{-1}, q) = q^{-d} \mathcal{F}_*(q) \mathcal{R}_1(q^{-1}) = q^{-d} \mathcal{F}_*(q) \frac{Q(q^{-1})A(q^{-1})}{\beta(q^{-1})E(q^{-1})} \quad (4)$$

where $\mathcal{F}_*(q) = B_*^c/\beta_*^c$ is a noncausal all-pass filter constructed from the common factor, B_*^c , of all RTFs ($B_i = B_*^c B_i^n$), provided that such a common part exists, and β_*^c is the minimum phase equivalent of B_*^c . β is the RMS spatial average defined via

$$\beta_* \beta = \mathbf{B}_* \mathbf{B} = \sum_{i=1}^p B_{i*} B_i = B_*^c B_*^c \sum_{i=1}^p B_i^n B_i^n. \quad (5)$$

Note that β is minimum phase, and in the frequency domain $|\beta(e^{-j\omega})/A(e^{-j\omega})|$ is the square root of the spatially averaged power spectrum. If the target responses are ideal, i.e., $\mathcal{D}_i = q^{-d}$, then the polynomial Q in (4) is a scalar constant. While the compensator (4) provides a perfect mixed phase compensation for exactly common nonminimum phase factors, such factors are unlikely to occur in reality unless the p RTFs are acquired from measurements very close in space. In general this is however not the case and such densely spaced measurements are equivalent to a single-point design which is known to be spatially non-robust [5, 6]. The idea of performing mixed phase SIMO compensation is to obtain a best possible time and frequency domain equalization over a prespecified spatial volume. In order to do this without introducing pre-rings, we are interested in extracting the common (or near-common) nonminimum phase dynamics among all measured RTFs. In [3] a method for that purpose was proposed. The method is based on zero clustering, where sufficiently close nonminimum phase zeros are represented by a *nominal zero* which is compensated for by a (truncated) noncausal all-pass filter $q^{-d} \mathcal{F}_*$ according to (4) with a pole located in the position of the nominal zero. The pre-response effects of such a compensation is discussed next.

3. CONSTRAINTS ON THE PRE-RESPONSE ERROR

To quantify the effects of cancelling a near-common excess phase factor in the RTFs with a (noncausally ringing) excess phase pole $z_0 = r_0 e^{j\omega_0}$; $r_0 > 1$, the residual impulse response¹ in the different measurement positions has to be derived. It is found to be, see [3],

$$h^{e^q}(k) = \left[\delta(k) + C r_0^k \cos(-\omega_0 k + \Phi) u(-k) \right] * r_2(k) * h_2(k) \quad (6)$$

where $*$ denotes convolution, $\delta(k)$ is the Kronecker delta function, $u(k)$ is the unit step function, and

$$\Phi = \arctan \left(\frac{2\Re(\epsilon)|z_0|^2}{|z_0 + \epsilon|^2 - |z_0|^2} - \Re(z_0) \right); \quad C = \frac{|z_0 + \epsilon|^2 - |z_0|^2}{|z_0|^2 \cos \Phi} \quad (7)$$

for some small pole-zero mismatch error ϵ . The pre-ringing effects are governed by the second term within the brackets of (6) and are constrained to obey the following N inequalities:

$$20 \log_{10}(C r_0^{-\kappa_n}) < L_n \quad ; \quad n = 1, \dots, N \quad (8)$$

where L_n is the maximum allowed pre-response error level at the time instant of κ_n samples prior to the desired delay d . All clusters fulfilling the pre-ringing constraints (8) are compensated for according to Section 2. When the nominal zeros representing the near-common zero clusters are determined, the residual impulse response

¹With residual response is meant the response remaining in a certain measurement position after cancelling the nominal zero with a pole.

of each pole-zero cancellation (6) can be controlled by requiring the amplitudes C in (7) to comply with the constraints (8).

The nonminimum phase zero cluster sizes depend heavily on the size of the measurement region: The smaller the measurement region the more compact the clusters, and vice versa. This is illustrated in Fig. 1 from which we note that the clusters become larger as the measurement region size increases. It is therefore unclear whether

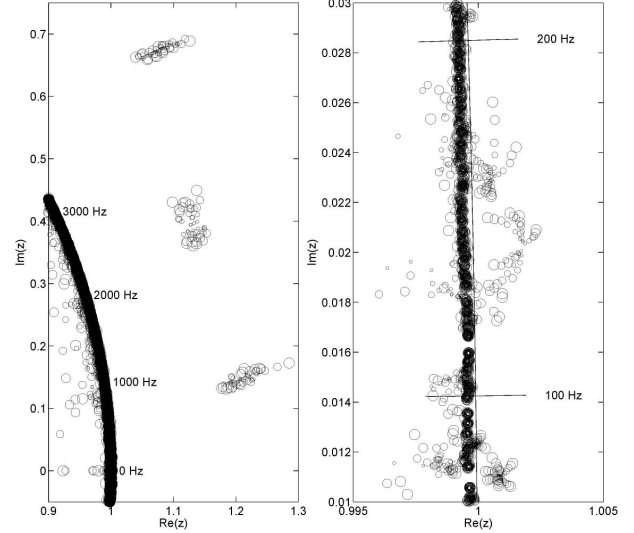


Fig. 1. Cluster performance. Left: Clusters present up to 4000Hz. Right: Clusters present along the frequency axis from 20 to 200 Hz. Small rings refer to an 8 point cubic shell of $10 \times 10 \times 10$ cm whereas the largest rings correspond to an 8 point cubic shell of $60 \times 60 \times 60$ cm. One increase in ring size corresponds to an increase by 10 cm in all directions.

the clusters will fulfill the pre-ringing constraints (8) as the clusters (and measurement region) become larger. This situation is illustrated in Fig. 2 from which we note that the clusters well outside the unit circle and around 1000–4000 Hz pass the constraints (8) also for the large region design, whereas the clusters at low frequencies close to the unit circle do not. The reason that the clusters at low frequencies do not pass the selection procedure is that the corresponding nominal zeros, when cancelled by a pole, cause a too large residual error, not compatible with (8). It is clear that the low frequency clusters would be included in the small designs but not in the large. But if we are aiming for a design which should be valid also in a larger volume, then we have to leave these clusters out of the design, at the expense of worse performance in the small volume. There is however a possible remedy to this, which will be discussed next.

4. A PRE-RESPONSE FOCUS CONTROL

Control of the pre-response artefact is essential to obtain the best possible performance over the whole pre-specified listening volume. If acceptable pre-response performance is to be obtained in the large region without completely sacrificing the time domain properties of the small-region design, then we would need to partly compensate for the zero clusters at low frequencies, even though they were discarded by the clustering process. This is accomplished by introducing a somewhat regularized version of the noncausal all-pass filter \mathcal{F}_* , say \mathcal{F}_*^r , and a companion causal all-pass filter \mathcal{G} in addition to this. In contrast to the small-region design where the pole of \mathcal{F}_* was placed on the resulting cluster nominal zero, we will here move the pole slightly, radially outward from the cluster. This will cause

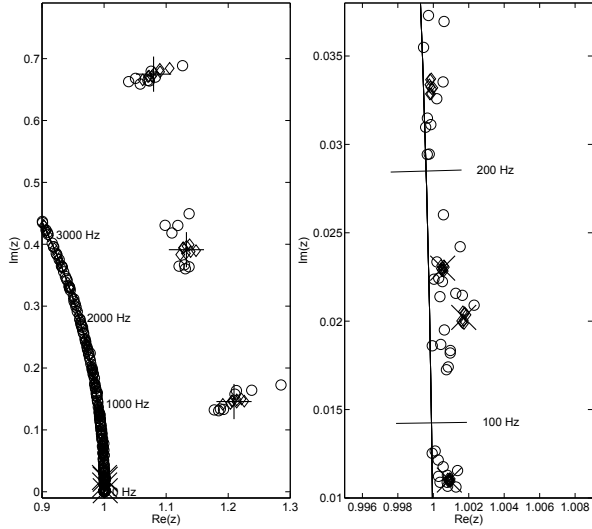


Fig. 2. Zero clusters. Left: Clusters at 0–4000 Hz. Right: Clusters close to the unit circle at 75–250 Hz. Diamonds (\diamond) refer to a measurement volume of $10 \times 10 \times 10$ cm (small design) whereas the rings correspond to a measurement volume of $60 \times 60 \times 60$ cm (large design). (+) indicate poles that pass the pre-ringing conditions (8) in the small and large regions. (\times) indicate poles passing the small-region design only.

a compression in time of the residual pre-rings but also an undesirable amplitude magnification. This magnification is however suppressed by properly adjusting the companion causal all-pass filter \mathcal{G} . In essence, \mathcal{G} neutralizes the effect that \mathcal{F}_*^r has on a broader frequency scale. By a proper selection of the pole/zero locations of \mathcal{F}_*^r , and of the companion causal all-pass filter \mathcal{G} , we can continuously control the pre-rings that are caused by partially inverting the excess phase zero clusters that were discarded in the large region design. Let $re^{j\omega}$ denote the location of a pole in the small region design, and introduce a regularization factor $\alpha > 1$ such that the regularized pole in \mathcal{F}_*^r is located at $r^\alpha e^{j\omega}$. Furthermore, denote the location of the corresponding zero in \mathcal{G} by $r^\gamma e^{j\omega}$, where $\gamma > \alpha$. The poles in \mathcal{F}_*^r and the zeros in \mathcal{G} are thus related by a factor $r^{\gamma-\alpha}$. When $\alpha = \gamma$, i.e., the reciprocal pole-zero pairs of \mathcal{F}_*^r and \mathcal{G} coincide, the compensator is identical to the cautious large-region design. As the pole-zero locations of \mathcal{G} move away from the unit circle, i.e., γ increases, the phase response correction is increased, and in the limit it is nearly identical to that of the small-region design. By optimizing over α and γ the proper selection is obtained. The procedure is further illustrated in Fig. 3. With a design as in Fig. 3 it is possible to fulfill all the pre-ringing constraints (8). Thus, based on measurements from the small cubic shell, we can regularize the design to be robust in any specified larger volume. We call this Focus Control, since it admits a method for shifting the focus of the time domain inversion between spatial regions of different sizes.

5. EQUALIZATION PERFORMANCE

For evaluation of the different mixed phase designs, measurements were collected in a geometry of six concentric cubic shells of increasing size, with eight measurements in each shell. The smallest distance between the cube corners was 10 cm, and the largest was 60 cm. For graphic evaluation we shall use the low-frequency average energy step response of $N = 8$ points, $S(k)$, and the impulse

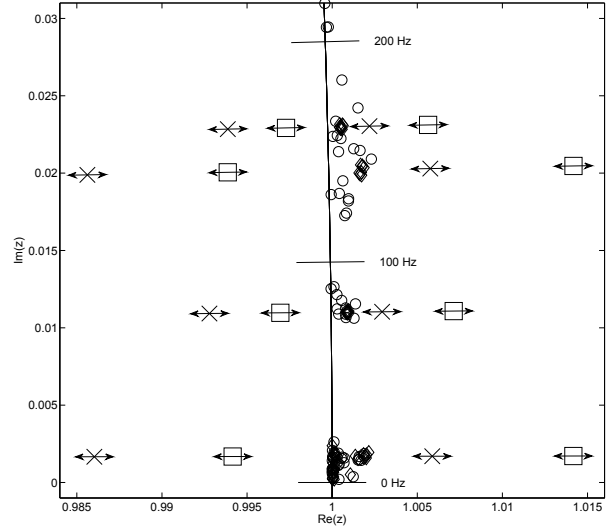


Fig. 3. Zero clusters at 0–200 Hz and phase correction filter. Diamonds refer to a measurement volume of $10 \times 10 \times 10$ cm (small design) whereas the rings correspond to a measurement volume of $60 \times 60 \times 60$ cm (large design). Crosses (\times) and boxes (\square) indicate poles and zeros of the regularized noncausal all-pass filter \mathcal{F}_*^r , and of the variable causal all-pass filter \mathcal{G} . Arrows indicate that the pole-zero patterns of the causal and noncausal all-pass filters are adjustable, to regulate the amount of phase correction.

response maximum level envelope $L(k)$, defined by

$$S(k) = \frac{1}{N} \sum_{n=1}^N \sum_{l=0}^k \frac{h_{Ln}^2(l)}{\sum_{m=0}^{M-1} h_{Ln}^2(m)} \quad (9)$$

$$L(k) = 20 \log_{10}(\max_n |h_n(k)|) ; n \in \{1, \dots, N\} \quad (10)$$

where $h_n(k)$ is the impulse response of the system at position n and $h_{Ln}(k)$ are low-frequency versions of $h_n(k)$. We shall compare the original response with the equalized responses associated with three different filters, named A, B and C, designed as follows.

- A) A strict small-region ($10 \times 10 \times 10$ cm) design, with pre-response error constrained to be at most -60 dB at -5 msec, and -80 dB at -40 msec.
- B) A mixed small/large-region ($10 \times 10 \times 10$ cm and $60 \times 60 \times 60$ cm) design with same pre-response constraint as Filter A. The nonminimum phase zero clusters from the small region are partially inverted using our new approach.
- C) A strict large-region ($60 \times 60 \times 60$ cm) design with same pre-response constraint as Filter A.

The different designs are also cross-validated on both the 60 cm and 10 cm regions. Furthermore, we will investigate impulse response maximum levels for the 10 cm, 60 cm, and mixed 10/60 cm region designs, respectively. Also here cross-validation is performed. The different performances are depicted in Fig. 4, and Fig. 5. The top diagram of Fig. 4 shows the low-frequency average energy step response for compensator designs ranging from 10 cm (A-case) to 60 cm (C-case) in addition to the unequalized response for all regions. We note that the A-design is superior in energy build-up. All the other designs are about equal except for the 20 cm design which

is slightly better. For further illustration it thus suffices to consider the A and C-designs, respectively.

In the middle diagram of Fig. 4 the A and C-designs are compared to the new regularized mixed phase B-design and the non-equalized original response. All designs are evaluated on the small (10 cm) region. We note that the A-design is, of course, superior but the new B-design is significantly better than the C-design, which is slightly better than the original response. Next, we evaluate the different designs on the large (60 cm) region.

From the bottom diagram of Fig. 4 we observe that the C-design significantly improves upon the original response. Moreover, the A-design is sometimes better and sometimes worse than the C-design. Interestingly the new B-design is always better than the C-design and mostly better than the A-design. However, as is evident from the uppermost curve in the bottom diagram of Fig. 5, the A-design does not fulfill the pre-response error constraint. Also, note that the B-design fulfills the constraints in both the small and large regions.

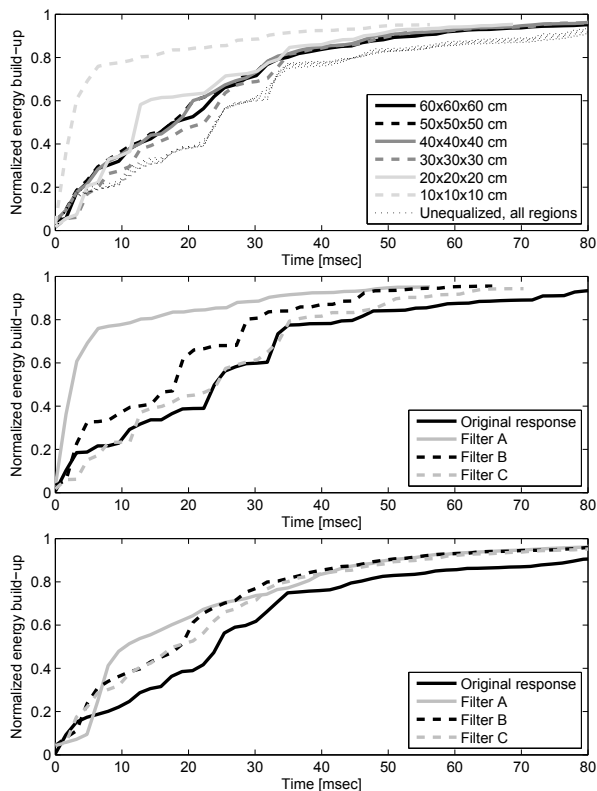


Fig. 4. Average energy step responses $S(k)$ in the low frequency band, $0 < f < 320$ Hz. Top: Performance for 10 cm–60 cm designs. Middle: Performance of filters A–C in the small region (10×10×10 cm). Bottom: Performance of filters A–C in the large region (60×60×60 cm).

6. CONCLUSIONS

We have shown that the mixed phase design proposed in this paper is superior to both small and large region designs when cross-validated over the large and small regions, respectively. With the new method, a small region design can be tuned to be valid for any spatial region by the use of a regularized noncausal all-pass filter together with a companion causal all-pass filter. By a proper selection of these all-pass filters we can control the focus of the inversion on any spatial region while fulfilling specific pre-response error constraints.

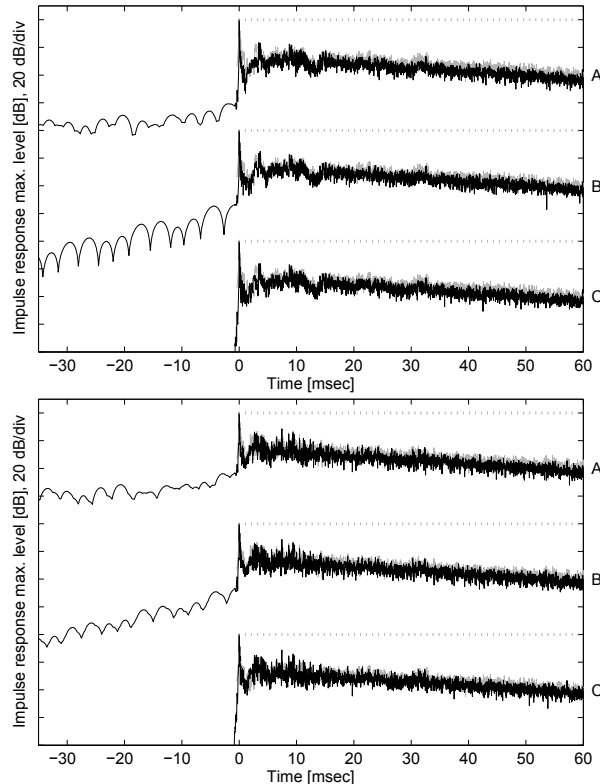


Fig. 5. Impulse response maximum level envelopes $L(k)$ for the original (grey lines) and equalized (black lines) system, using filters A–C. Top: Performance of filters A–C in the small region (10×10×10 cm). Bottom: Performance of filters A–C in the large region (60×60×60 cm).

7. REFERENCES

- [1] M. Karjalainen, T. Paatero, J. Mourjopoulos, and P. Hatziantoniou, “About room response equalization and dereverberation,” in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE, 2005, pp. 183–186.
- [2] L. D. Fielder, “Analysis of traditional and reverberation-reducing methods of room equalization,” *J. Audio Eng. Soc.*, vol. 51, no. 1/2, pp. 3–26, Jan/Feb 2004.
- [3] L.-J. Brännmark and A. Ahlén, “Spatially robust audio compensation based on SIMO feedforward control,” *IEEE Transactions on Signal Processing*, vol. 57, no. 5, May 2009.
- [4] M. Sternad and A. Ahlén, “LQ controller design and self-tuning control,” in *Polynomial Methods in Optimal Control and Filtering*, K. Hunt, Ed., pp. 56–92. Peter Peregrinus, London, 1993.
- [5] B. D. Radlovic, R. C. Williamson, and R. A. Kennedy, “Equalization in an acoustic reverberant environment: robustness results,” *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 3, pp. 311–319, May 2000.
- [6] P. Hatziantoniou and J. Mourjopoulos, “Errors in real-time room acoustics dereverberation,” *J. Audio Eng. Soc.*, vol. 52, no. 9, pp. 883–899, September 2004.