IMPROVED LOUDSPEAKER–ROOM EQUALIZATION USING MULTIPLE LOUDSPEAKERS AND MIMO FEEDFORWARD CONTROL

Lars-Johan Brännmark, Adrian Bahne and Anders Ahlén

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

ABSTRACT

In this paper, a new multichannel approach to robust loudspeaker–room equalization is presented. Traditionally, the equalization (or room correction) problem has been treated mostly by single-channel methods, with loudspeaker signals being prefiltered individually by separate scalar filters. Single-channel methods can generally improve the average spectral flatness of the acoustic transfer functions in a listening region, but the variability of the transfer functions within the region cannot be affected.

Most modern audio reproduction systems, however, contain two or more loudspeakers, and in this paper we aim at improving the equalization performance by using all available loudspeakers jointly. To this end we propose a general MIMO formulation of the problem, which is a multichannel generalization of an earlier single-channel approach by the authors. The new approach is found to reduce the average reproduction error and the spatial variability of the acoustic transfer functions. Moreover, pre-ringing artifacts are avoided, and the reproduction error below 1000 Hz is significantly reduced with an amount that scales with the number of loudspeakers used.

Index Terms— Acoustic signal processing, audio systems, equalizers, MIMO, polynomials.

1. INTRODUCTION

Equalization of the loudspeaker–room response by means of digital filters has been an agile research area for decades. So far, mainly single-channel methods have been proposed and implemented, see e.g., [1]. However, in most audio systems, two or more loudspeakers are in general available, and an interesting question is whether (and how) room correction systems can be amended by using all loudspeakers in a joint design [2]. Several different approaches for controlling room modes and equalizing the room transfer functions based on multi-channel methods have been proposed [2–6].

In summary, the proposed methods can be grouped into three categories. The first includes methods based on *physical insight* about room acoustics and room modes, and it is well known that loudspeaker placement and the use of several subwoofers is critical to reduce the effect of room modes [4]. Another principle often used is the *source-sink method* [2, 5], where symmetrically positioned subwoofers are optimized so as to reduce room modes, by means of delay-, gain- and phase adjustments to the different channels. The evaluation is mostly performed by means of finite-difference time-domain (FDTD) simulations. A third important method is *modal equalization* [6], in which modal resonances and their decay times are controlled by digital prefilters. For a summary of different room correction procedures, see e.g., [1, 3].

In this paper, we propose a general multiple-input multiple-output (MIMO) formulation of the equalization problem, as an extension to an earlier single-channel approach by the authors [7, 8]. The aim is to improve the reproduction quality of an audio system by using all available loudspeakers jointly, employing not only channel inversion but also sound field superposition to reach the desired target response of each loudspeaker. Our approach makes use of a polynomial-based control systems framework, and offers a joint solution to the problems of loudspeaker–room equalization and bass management. The reproduction error and spatial variations of the acoustic transfer functions are substantially reduced and the performance scales with the number of loudspeakers used, with the most prominent effect being attained at low frequencies. Moreover, the suggested solution is robust to modeling errors and pre-ringing artifacts are avoided.

Remarks on the notation: Filters and acoustic transfer functions 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} or $e^{-\jmath\omega}$ in the frequency domain. A polynomial (polynomial matrix) is denoted by italic capital (bold italic capital) letters as $P(q^{-1}) = \mathrm{p}_0 + \mathrm{p}_1 q^{-1} + \cdots + \mathrm{p}_{n_P} q^{-n_P}$ ($\boldsymbol{P}(q^{-1}) = \mathbf{P}_0 + \mathbf{P}_1 q^{-1} + \cdots + \mathbf{P}_{n_{\boldsymbol{P}}} q^{-n_{\boldsymbol{P}}}$). Rational matrices are represented by right matrix fraction descriptions (right MFDs), and are indicated by bold calligraphic letters as $\boldsymbol{\mathcal{G}}(q^{-1}) = \boldsymbol{Q}(q^{-1})\boldsymbol{P}^{-1}(q^{-1})$. Scalar rational functions are denoted by normal calligraphic letters as $\mathcal{G}(q^{-1})$. For any polynomial (or polynomial matrix), its *conjugate* is defined as $P_*(q) = P(q) = \mathrm{p}_0 + \mathrm{p}_1 q + \cdots + \mathrm{p}_{n_P} q^{n_P}$ (or $\boldsymbol{P}_*(q) = \boldsymbol{P}^T(q) = \mathbf{P}_0^T + \mathbf{P}_1^T q + \cdots + \mathbf{P}_{n_{\boldsymbol{P}}}^T q^{n_{\boldsymbol{P}}}$), and its *reciprocal* is defined as $\overline{P}(q^{-1}) = q^{-n_P} P_*(q)$ (or $\overline{\boldsymbol{P}}(q^{-1}) = q^{-n_P}\boldsymbol{P}_*(q)$). The arguments q^{-1}, q, z^{-1}, z etc. are often omitted if there is no risk of misunderstanding. A matrix having p rows and l columns is said to be of dimension $p|l$. The notation $\mathrm{diag}(\boldsymbol{v})$, for a column vector $\boldsymbol{v}$, represents a diagonal matrix with the elements of $\boldsymbol{v}$ along the diagonal.

2. PROBLEM STATEMENT

We consider an acoustic environment with l loudspeakers positioned around a bounded three-dimensional listening region $\boldsymbol{\Omega} \subset \mathbb{R}^3$ in a room. A target impulse response is defined for one of the loudspeakers, here called the *primary loudspeaker*, to be attained in p spatial positions witin $\boldsymbol{\Omega}$. In order to improve the mean square error performance at the measurement positions, $l-1$ *support loudspeakers* are introduced. By the use of l phase compensation filters, applied individually to each loudspeaker, in conjunction with a stable and causal MIMO compensator, significant improvements relative to the

The first and second authors are also affiliated with Dirac Research AB, Hansellisgatan 6, SE-754 50 Uppsala, Sweden

single-channel case can be attained. Before taking on the MIMO set-up we shall first consider the single-channel case.

2.1. Background: The single-loudspeaker case

In [7] a mixed phase single-loudspeaker compensation was considered in a single-input multiple-output (SIMO) setting, using a general polynomial matrix framework. A further robustification of the SIMO problem was considered in [8], where a probabilistic error model was introduced for handling errors and unmeasured spatial variations in the acoustic transfer functions. It was shown in [7] that in order to avoid pre-ringing artifacts in the compensated system, the scalar mixed phase prefilter is required to contain an allpass link as a factor. We shall here briefly recall this result. Let the error signal of a SIMO system be expressed as

$$\boldsymbol{y}(k) = \frac{\boldsymbol{D}(q^{-1})}{E(q^{-1})} w(k) - \frac{\boldsymbol{B}(q^{-1})}{A(q^{-1})} \mathcal{R}(q^{-1}, q) w(k) \tag{1}$$

where

$$\begin{aligned} \boldsymbol{D}(q^{-1}) &= \left[D_1(q^{-1}) \cdots D_p(q^{-1}) \right]^T \\ \boldsymbol{B}(q^{-1}) &= \left[B_1(q^{-1}) \cdots B_p(q^{-1}) \right]^T . \end{aligned} \tag{2}$$

Here $w(k)$ is a scalar stationary white noise sequence having zero mean and covariance $\mathrm{E}\{w^2(k)\} = \psi$, $\boldsymbol{D}/E$ and $\boldsymbol{B}/A$ are the target and actual room transfer functions (RTFs), E and A being stable (i.e., minimum phase) polynomials whereas $\mathcal{R}$ is a (possibly noncausal) scalar feedforward compensator. The objective of $\mathcal{R}$ is to minimize the sum of powers of the p error signal components in $\boldsymbol{y}(k)$, i.e., to minimize $J = \mathrm{E}\{\mathrm{tr}\left[\boldsymbol{y}(k)\boldsymbol{y}^T(k)\right]\}$. According to Lemma 1 of [7], a mixed phase compensator $\mathcal{R}$ that does not generate pre-ringing errors must have the special structure

$$\mathcal{R}(q^{-1}, q) = q^{-d} \frac{\overline{F}_*(q)}{F_*(q)} \mathcal{R}_1(q^{-1}) = q^{-d} \mathcal{F}_*(q) \mathcal{R}_1(q^{-1}) \tag{3}$$

where $\overline{F}$ is such that the zeros of $\overline{F}(z^{-1})$ are the common excess phase zeros of $B_1(z^{-1}), \ldots, B_p(z^{-1})$. Since the zeros of $F(z^{-1})$ are reciprocal (with respect to the unit circle) to those of $\overline{F}(z^{-1})$ we have that $\mathcal{F}(q^{-1}) = \overline{F}(q^{-1})/F(q^{-1})$ is a causal allpass filter and its conjugate $\mathcal{F}_*(q)$, which appears in (3), is thus a noncausal allpass filter. In (3), $\mathcal{R}_1(q^{-1})$ is a stable and causal filter and d constitutes the so-called modeling delay in $\boldsymbol{D}(q^{-1})$, or equivalently, the "smoothing lag" of the compensator. The MSE-optimal mixed phase compensator (3) for the SIMO system (1)–(2) is hence given by a conjugated allpass link in series with a causal filter. The function of the allpass filter is to remove any group delay distortion that is common to all positions in the spatial region of interest. The above constitutes a necessary and sufficient structural constraint for a mixed phase compensator to avoid pre-ringing artifacts.

2.2. Multichannel extension: Introducing support loudspeakers

While [7, 8] considered robust mixed phase audio compensation by means of a single loudspeaker, we here consider a more general setting, where support loudspeakers are introduced in order to help attaining the ideal target RTF defined for the primary loudspeaker. We thereby obtain a multiple-input multiple-output (MIMO) system for which a mixed phase compensator should be designed. The derivation of this compensator is presented next.

The acoustic signal propagation between the l loudspeakers and p positions in $\boldsymbol{\Omega}$ is modeled by a $p|l$-dimensional rational matrix $\boldsymbol{\mathcal{H}}(q^{-1})$, which can be decomposed as

$$\boldsymbol{\mathcal{H}}(q^{-1}) = \boldsymbol{\mathcal{H}}_0(q^{-1}) + \Delta\boldsymbol{\mathcal{H}}(q^{-1}) \tag{4}$$

where $\boldsymbol{\mathcal{H}}_0(q^{-1})$ is the *nominal model*, representing those components of the transfer functions that are "spatially smooth" and therefore well captured by spatially sparse transfer function measurements. The additive part $\Delta\boldsymbol{\mathcal{H}}(q^{-1})$ is partly parameterized by random variables, and represents components that are not fully captured by the measurements. Typically, these spatially complex components consist of late room reflections and reverberation at high frequencies. Writing out the matrix fractions for $\boldsymbol{\mathcal{H}}(q^{-1})$ and $\Delta\boldsymbol{\mathcal{H}}(q^{-1})$, the decomposition (4) of $\boldsymbol{\mathcal{H}}(q^{-1})$ expands into

$$\begin{aligned} \boldsymbol{\mathcal{H}} &= \boldsymbol{B}_0\boldsymbol{A}_0^{-1} + \Delta\boldsymbol{B}\boldsymbol{B}_1\boldsymbol{A}_1^{-1} \\ &= (\boldsymbol{B}_0\boldsymbol{A}_1 + \Delta\boldsymbol{B}\boldsymbol{B}_1\boldsymbol{A}_0)(\boldsymbol{A}_0\boldsymbol{A}_1)^{-1} \\ &= (\hat{\boldsymbol{B}}_0 + \Delta\boldsymbol{B}\hat{\boldsymbol{B}}_1)(\boldsymbol{A}_0\boldsymbol{A}_1)^{-1} \triangleq \boldsymbol{B}\boldsymbol{A}^{-1} \end{aligned} \tag{5}$$

where $\hat{\boldsymbol{B}}_0 = \boldsymbol{B}_0\boldsymbol{A}_1$, $\hat{\boldsymbol{B}}_1 = \boldsymbol{B}_1\boldsymbol{A}_0$, $\boldsymbol{B} = \hat{\boldsymbol{B}}_0 + \Delta\boldsymbol{B}\hat{\boldsymbol{B}}_1$, and $\boldsymbol{A} = \boldsymbol{A}_0\boldsymbol{A}_1$. The matrices $\boldsymbol{B}_0$, $\Delta\boldsymbol{B}$ and $\boldsymbol{B}$ are of dimension $p|l$, whereas $\boldsymbol{B}_1$, $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ and $\boldsymbol{A}$ are of dimension $l|l$. The elements of $\Delta\boldsymbol{B}$ are polynomials with stochastic variables as coefficients and $\boldsymbol{B}_1\boldsymbol{A}_1^{-1}$ is a filter for shaping the spectral distribution of the stochastic uncertainty model. The denominators $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ and $\boldsymbol{A}$ are further assumed to be diagonal; from a physical modeling perspective this is no restriction, see [9]. For a general introduction to the above probabilistic modeling framework, the reader is referred to [10] and references therein. A detailed discussion of why and how this framework applies in the present acoustic context is given in [9].

Since (3) is a structural requirement for avoiding pre-ringing in the SIMO case, a sufficient condition for obtaining a solution without pre-ringing in the MIMO case is to apply l noncausal phase compensation filters $\mathcal{F}_{1*}, \ldots, \mathcal{F}_{l*}$, designed similarly as in the SIMO case, to each of the l loudspeakers, and then design a full causal and stable MIMO compensator $\boldsymbol{\mathcal{R}}_1$ such that the target RTF of the primary loudspeaker is attained with minimum error. The role of the filters $\mathcal{F}_{1*}, \ldots, \mathcal{F}_{l*}$ is to remove group delay distortions that are common and systematic throughout $\boldsymbol{\Omega}$ for each loudspeaker. The role of $\boldsymbol{\mathcal{R}}_1$ is to superimpose all the individual phase-corrected loudspeaker responses in an optimal way, such that the overall sum response becomes closer to the target RTF than if the primary loudspeaker would have been used alone.

Now consider the MIMO system

$$\boldsymbol{y}(k) = \boldsymbol{\mathcal{D}}(q^{-1})w(k) - \boldsymbol{\mathcal{H}}(q^{-1})\boldsymbol{u}_1(k) \tag{6}$$

where

$$\boldsymbol{\mathcal{D}}(q^{-1}) = \frac{\boldsymbol{D}(q^{-1})}{E(q^{-1})} = q^{-d_0} \frac{\widetilde{\boldsymbol{D}}(q^{-1})}{E(q^{-1})} \tag{7}$$

is the target RTF, of dimension $p|1$. In $\widetilde{\boldsymbol{D}}(q^{-1})$ above, at least one of the polynomial elements is assumed to have a nonzero leading coefficient; the second equality in (7) is included to emphasize that $\boldsymbol{\mathcal{D}}(q^{-1})$ contains an initial modeling delay of d_0 samples. Furthermore, $\boldsymbol{\mathcal{H}}(q^{-1})$ in (6) is given by (4)–(5) and $\boldsymbol{u}_1(k)$ is given by

$$\begin{aligned} \boldsymbol{u}_1(k) &= \boldsymbol{\mathcal{R}}(q^{-1}, q)w(k) = \widetilde{\boldsymbol{\Delta}}(q^{-1})\boldsymbol{\mathcal{F}}_*(q)\boldsymbol{u}(k) \\ &= \widetilde{\boldsymbol{\Delta}}(q^{-1})\boldsymbol{\mathcal{F}}_*(q)\boldsymbol{\mathcal{R}}_1(q^{-1})w(k) \end{aligned} \tag{8}$$

where

$$\begin{aligned} \widetilde{\boldsymbol{\Delta}}(q^{-1}) &= \mathrm{diag}\left(\left[q^{-(d_0-d_1)} \ \cdots \ q^{-(d_0-d_l)} \right]^T \right) \\ \boldsymbol{\mathcal{F}}(q^{-1}) &= \mathrm{diag}\left(\left[\frac{\overline{F}_1(q^{-1})}{F_1(q^{-1})} \ \cdots \ \frac{\overline{F}_l(q^{-1})}{F_l(q^{-1})} \right]^T \right) \\ \boldsymbol{\mathcal{R}}_1(q^{-1}) &= \left[\mathcal{R}_{11}(q^{-1}) \ \cdots \ \mathcal{R}_{1l}(q^{-1}) \right]^T . \end{aligned} \tag{9}$$

Analogously to the SIMO case, the matrix $\boldsymbol{\mathcal{F}}(q^{-1})$ is here constructed from excess phase zeros that are common among the RTFs of each loudspeaker for all measurement positions in $\boldsymbol{\Omega}$. That is, the elements $B_{1j}, \ldots, B_{pj}$ of the jth column of $\boldsymbol{B}$ are assumed to share a common excess phase factor $\overline{F}_j(q^{-1})$.

In (9), d_0 is the same as in (7) and constitutes the primary bulk delay (or smoothing lag) of the compensated system, whereas d_j, $j = 1, \ldots, l$ are individual delays that can be used to compensate for individual discrepancies in distances among the different loudspeakers. Since $\widetilde{\boldsymbol{\Delta}}(q^{-1})\boldsymbol{\mathcal{F}}_*(q)$ is fixed and known it can be regarded as a factor of an augmented system $\widetilde{\boldsymbol{\mathcal{H}}}(q^{-1})$, see Fig. 1,

$$\widetilde{\boldsymbol{\mathcal{H}}}(q^{-1}) \triangleq \boldsymbol{\mathcal{H}}(q^{-1})\widetilde{\boldsymbol{\Delta}}(q^{-1})\boldsymbol{\mathcal{F}}_*(q) = \widetilde{\boldsymbol{B}}(q^{-1})\boldsymbol{A}^{-1}(q^{-1}) \quad (10)$$

where $\widetilde{\boldsymbol{B}}(q^{-1}) = \boldsymbol{B}(q^{-1})\widetilde{\boldsymbol{\Delta}}(q^{-1})\boldsymbol{\mathcal{F}}_*(q)$ is still a polynomial matrix (i.e., not a rational matrix), due to cancellation of factors between $\boldsymbol{B}$ and $\boldsymbol{\mathcal{F}}_*$. The second equality of (10) is allowed because $\boldsymbol{A}$, $\widetilde{\boldsymbol{\Delta}}$ and $\boldsymbol{\mathcal{F}}_*$ are diagonal, see (5) and (9). In order to attain the primary loudspeaker target RTF, the objective is to find the optimal MIMO causal and stable compensator $\boldsymbol{\mathcal{R}}_1(q^{-1})$, see (8)–(9), that minimizes the criterion

$$J = \bar{\mathrm{E}}\{\mathrm{tr}\ \mathrm{E}[(\boldsymbol{V}\boldsymbol{y})(\boldsymbol{V}\boldsymbol{y})^T]\} + \mathrm{tr}\ \mathrm{E}[(\boldsymbol{W}\boldsymbol{u})(\boldsymbol{W}\boldsymbol{u})^T]\,. \quad (11)$$

Here $\bar{\mathrm{E}}$ and E denote, respectively, expectation with respect to the

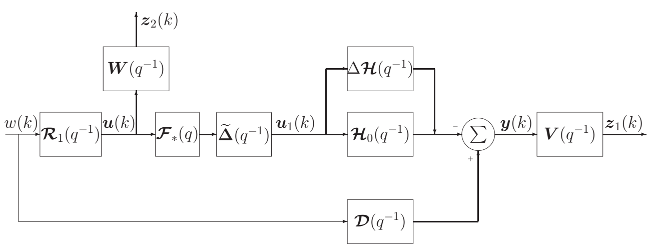

Fig. 1. Block diagram of the constrained MIMO equalizer design. The thin lines represent scalar signals, and the thick lines represent vector-valued signals of dimension l or p.

uncertain parameters in $\Delta\boldsymbol{B}$, see (5), and the driving noise $w(k)$, whereas the filters $\boldsymbol{W}$ and $\boldsymbol{V}$, of dimension $p|p$ and $l|l$, constitute *weighting matrices* for the control and error signals, respectively. We are now ready to state the main result.

Theorem 1. *Consider the system* (4)–(7)*,* (10) *and the controller structure* $\boldsymbol{\mathcal{R}}(q^{-1}, q)$ *defined in* (8)*. The optimal stable and causal compensator, which minimizes* (11)*, without residual preringings, is then given by*

$$\boldsymbol{\mathcal{R}}_1 = \boldsymbol{A}\boldsymbol{\beta}^{-1}\boldsymbol{Q}\frac{1}{E} \quad (12)$$

where $\boldsymbol{\beta}$*, of dimension* $l|l$*, is the unique (up to a unitary constant matrix) stable spectral factor of*

$$\boldsymbol{\beta}_*\boldsymbol{\beta} = \bar{\mathrm{E}}\{\widetilde{\boldsymbol{B}}_*\boldsymbol{V}_*\boldsymbol{V}\widetilde{\boldsymbol{B}} + \boldsymbol{A}_*\boldsymbol{W}_*\boldsymbol{W}\boldsymbol{A}\} \quad (13)$$

with $\widetilde{\boldsymbol{B}}$*, of dimension* $p|l$*, being as in* (10)*. The polynomial matrix* $\boldsymbol{Q}$*, together with a polynomial matrix* $\boldsymbol{L}_*$*, both of dimension* $l|1$*, constitute the unique solution to the Diophantine equation*

$$\bar{\mathrm{E}}\{\widetilde{\boldsymbol{B}}_*\}\boldsymbol{V}_*\boldsymbol{V}\boldsymbol{D} = \boldsymbol{\beta}_*\boldsymbol{Q} + q\boldsymbol{L}_*E \quad (14)$$

with generic degrees $\deg \boldsymbol{Q} = \max(\deg \boldsymbol{V} + \deg \boldsymbol{D}, \deg E - 1)$ *and* $\deg \boldsymbol{L}_* = \max(\deg \bar{\mathrm{E}}\{\widetilde{\boldsymbol{B}}_*\} + \deg \boldsymbol{V}_*, \deg \boldsymbol{\beta}_*) - 1$.

Proof. Immediate from Theorem 4 of [10] by appropriate substitutions and by setting the delay parameter m and all uncertainty models equal to zero, except $\boldsymbol{\Delta\mathcal{G}}^T(q^{-1})$. □

Using the "bracket operator" $\{\cdot\}_+$ for expressing the causal part of a rational function, the filter $\boldsymbol{\mathcal{R}}_1$ above can be written as $\boldsymbol{\mathcal{R}}_1 = \boldsymbol{A}\boldsymbol{\beta}^{-1}\{\boldsymbol{\beta}_*^{-1}\boldsymbol{\mathcal{F}}\widetilde{\boldsymbol{\Delta}}_*\hat{\boldsymbol{B}}_{0*}\boldsymbol{V}_*\boldsymbol{V}\boldsymbol{D}/E\}_+$, and the total compensator $\boldsymbol{\mathcal{R}}$ in (8) becomes

$$\boldsymbol{\mathcal{R}} = \widetilde{\boldsymbol{\Delta}}\boldsymbol{\mathcal{F}}_*\boldsymbol{A}\boldsymbol{\beta}^{-1}\left\{\boldsymbol{\beta}_*^{-1}\boldsymbol{\mathcal{F}}\widetilde{\boldsymbol{\Delta}}_*\hat{\boldsymbol{B}}_{0*}\boldsymbol{V}_*\boldsymbol{V}\boldsymbol{D}\frac{1}{E}\right\}_+. \quad (15)$$

This expression is readily verified to reduce to the scalar solution presented in [8], in the case when no support loudspeakers are used, i.e., if $l = 1$ in the present design.

3. A DESIGN EXAMPLE

The filter design proposed above is now experimentally assessed using impulse responses of a simulated multichannel system in a rectangular room.

3.1. Experimental Conditions

A 64|6 transfer function matrix $\boldsymbol{\mathcal{H}}$ describing a 6-channel audio system was generated using a 25th-order image source model. The dimensions (W×L×H) of the simulated room was 5×6×3 m and the wall reflection coefficient was 0.8. The loudspeakers were modeled as point sources located at a distance of 1.8 m around a cubic volume $\boldsymbol{\Omega}$ of 30×30×30 cm. Loudspeakers 1, 2, 3, 5 and 6, positioned in a horizontal plane in a "surround" fashion, are labeled Left (L), Center (C), Right (R), Right Surround (RS), Left Surround (LS). Loudspeaker 4, labeled Top (T), was placed close to the ceiling, straight above $\boldsymbol{\Omega}$. The receiver positions, or *control points*, were distributed on a uniform grid with 10 cm spacing, covering $\boldsymbol{\Omega}$ with totally 64 points. Element (i, j) of $\boldsymbol{\mathcal{H}}$ was obtained by evaluating the image source model for source j at control point i, for $i = 1, \ldots, 64$; $j = 1, \ldots, 6$. For details about the simulation algorithm, see [9, Appendix A]. The elements of $\boldsymbol{\mathcal{F}}_*$ in (8) were constructed by applying the framework of [7] to each of the 6 loudspeakers. The uncertainty model $\Delta\boldsymbol{\mathcal{H}}$ in (4) was constructed along the lines of [8, 10]. The penalty matrix $\boldsymbol{W}$ was designed so as to prohibit the use of support loudspeakers above 1200 Hz. As a primary source we have selected loudspeaker (L), and the target function $\boldsymbol{\mathcal{D}}(q^{-1})$ in (7) thus contains a set of delayed ideal pulses: $\mathcal{D}_i(q^{-1}) = q^{-d_0-\Delta_i}$, where Δ_i is the natural acoustic propagation delay between loudspeaker (L) and the ith control point. Based on model data obtained as above, the performance of five different compensators, employing various numbers of support loudspeakers, were assessed.

3.2. Results and discussion

From Fig. 2 it is clear that the reproduction error (defined as the difference between the target response and attained system responses) decreases significantly by the introduction of support loudspeakers. It is particularly noticeable when adding the first support loudspeaker (C), and also when the top loudspeaker (T) is added (4th curve from the top). Note that the contribution of the support loudspeakers is significant up to 400–500 Hz. Above 500 Hz the improvement is diminishing. Fig. 3 displays the magnitude frequency response of the primary loudspeaker and in Fig. 4 a single-channel compensation of the primary loudspeaker with a flat (0 dB) target is depicted. Clearly, in Fig. 4, the average response is relatively flat up to ∼1000 Hz although there is a significant variability among the measurement points. Note also that there is a significant bias in

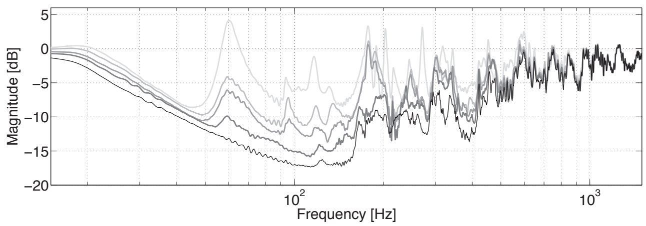

Fig. 2. The RMS error as a function of frequency, averaged over 343 points (grid spacing 5 cm) in $\mathbf{\Omega}$. The curves show, from top to bottom, the attained reproduction error for an increasing number of support loudspeakers: 1st (loudspeaker L, light grey, single-channel compensation of the primary speaker), 2nd (L+C), 3rd (L+C+R), 4th (L+C+R+T), 5th (all six loudspeakers, black curve)

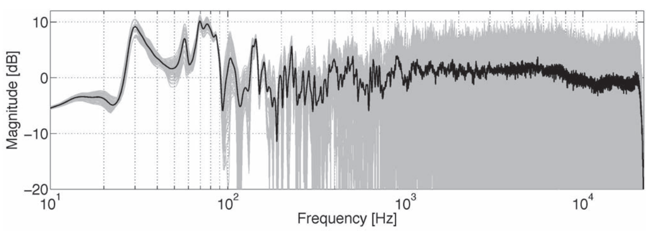

Fig. 3. Frequency responses of the uncompensated primary loudspeaker, evaluated in the 64 control points (grey lines) and their RMS average (black line).

the RMS value ($\sim$3dB) below the target in the same frequency region. By adding all the six support loudspeakers both the bias in the average RMS value and the variability is dramatically reduced, see Fig. 5. This suggests that the use of support loudspeakers can help the primary loudspeaker attaining its target response significantly.

4. CONCLUDING REMARKS

A new approach to loudspeaker–room equalizer design, inspired by concepts in multivariable robust control, has been presented. The obtained filters were shown to yield good results in a simulated acoustic environment. Investigations concerning the application to real audio systems, with associated psychoacoustic considerations, are under way. We anticipate that the use of support loudspeakers will improve the listening experience considerably in both stereo and surround settings.

5. ACKNOWLEDGMENT

The authors gratefully acknowledge the support from the Knut and Alice Wallenberg foundation (Dnr. KAW 2006.0152), which provided funding for acoustic measurement equipment.

6. 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, WASPAA’05, Proceedings*, New Paltz, NY, October 2005, pp. 183–186.

[2] J. Vanderkooy, “Multi-source room equalization: Reducing room resonances,” Presented at *AES 123rd Convention, New York*. Preprint 7262. Audio Engineering Society, October 2007.

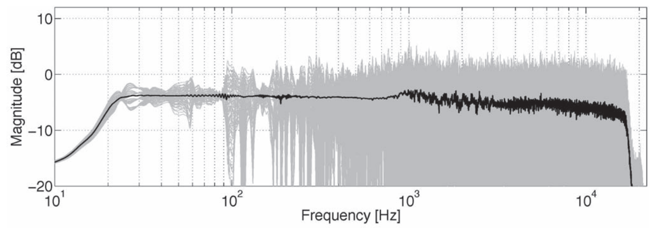

Fig. 4. Single-channel design without support loudspeakers (l=1): Frequency responses of the compensated primary loudspeaker, evaluated in the 64 control points (grey lines) and their RMS average (black line).

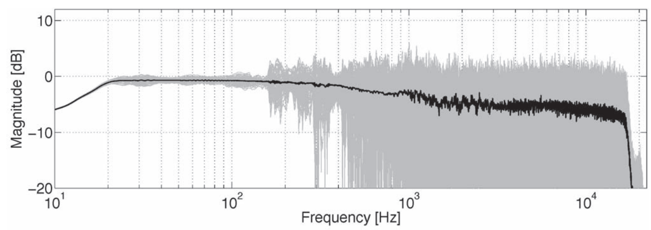

Fig. 5. Full multichannel design using all loudspeakers (l=6): Frequency responses of the compensated primary loudspeaker, evaluated in the 64 control points (grey lines) and their RMS average (black line).

[3] M. O. Hawksford and A. J. Hill, “Wide-area psychoacoustic correction for problematic room modes using non-linear bass synthesis,” Presented at *AES 129th Convention, New York*. Preprint 8313. Audio Engineering Society, November 2010.

[4] T. Welti and A. Devantier, “Low-frequency optimization using multiple subwoofers,” *J. Audio Eng. Soc*, vol. 54, no. 5, pp. 347–364, 2006.

[5] A. Celestinos and S. Birkedal Nielsen, “Time based room correction system for low frequencies using multiple loudspeakers,” Presented at the AES 32nd International Conference: DSP for Loudspeakers, September 2007.

[6] A. Mäkivirta, P. Antsalo, M. Karjalainen, and V. Välimäki, “Modal equalization of loudspeaker–room responses at low frequencies,” *J. Audio Eng. Soc*, vol. 51, no. 5, pp. 324–343, 2003.

[7] 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.

[8] L.-J. Brännmark, “Robust audio precompensation with probabilistic modeling of transfer function variability,” in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA’09, Proceedings*, New Paltz, NY, October 2009, pp. 193–196.

[9] L.-J. Brännmark, *Robust Sound Field Control for Audio Reproduction: A Polynomial Approach to Discrete-Time Acoustic Modeling and Filter Design*, Ph.D. thesis, Uppsala University, Uppsala, Sweden, 2011.

[10] K. Öhrn, A. Ahlén, and M. Sternad, “A probabilistic approach to multivariable robust filtering and open-loop control,” *IEEE Transactions on Automatic Control*, vol. 40, no. 3, pp. 405–418, March 1995.