# HIGHER-ORDER LISTENING ROOM COMPENSATION WITH ADDITIVE COMPENSATION SIGNALS

*Christian Hofmann, Michael Guenther, Michael Buerger, and Walter Kellermann*

Friedrich-Alexander University Erlangen-Nürnberg (FAU),
Multimedia Communications and Signal Processing,
Cauerstr. 7, D-91058 Erlangen, Germany
{christian.hofmann,michael.guenther,michael.buerger,walter.kellermann}@FAU.de

## ABSTRACT

The performance of sound reproduction systems for spatial audio is impaired by time-variant, reverberant listening environments. To tackle this issue, the Loudspeaker-Enclosure-Microphone System (LEMS) between the loudspeakers and reference microphones in the listening environment can be identified adaptively to allow an LEMS-specific pre-processing of the loudspeaker signals. This contribution introduces a broadband implementation of a narrowband Listening Room Compensation (LRC) method with additive compensation signals, recently proposed by Talagala et al. [1], it extends the concept to higher-order compensation, and compares LRC to Listening Room Equalization (LRE) analytically. Evaluations in an image-source environment confirm the efficacy of higher-order LRC and its suitability as a complexity-reduced alternative to LRE.

***Index Terms***— listening room, compensation, equalization, dereverberation, approximate inversion

## 1. INTRODUCTION

Premium sound reproduction systems make use of a large number of loudspeakers in order to reproduce virtual acoustic scenes with high accuracy in a spatially extended listening area [2–6] or to render individual scenes in multiple listening zones within a single enclosure [7–9]. In case of real acoustical environments, the emitted target sound is reflected by the walls or reflective obstacles within the room (e.g., furniture, listeners) in an unpredictable way, which causes undesired interference in the listening zones and significantly deteriorates the reproduction performance of rendering systems [10].

Many approaches which try to maintain the rendering performance despite reverberant environments require a dedicated calibration phase where Room Impulse Responses (RIRs) within the actual acoustical listening environment are measured to allow a design of the rendering system tailored to the actual environment [11–13]. Other approaches aim at reshaping RIRs by pre-filtering in order to dereverberate only a perceptually relevant part of a known single-channel RIR [14]. Further approaches aim at crosstalk cancellation for multiple isolated points in space [15–17], some also in conjunction with the aforementioned impulse response reshaping [18]. Besides, an approach aiming at a correct synthesis of binaural cues while considering the desired acoustics of the recording room and the playback room jointly has been proposed in [19].

In practice, the RIRs of typical listening environments are time-variant due to temperature changes and possibly time-varying positions of reflecting objects and listeners. In order to cope with the time-variance of the listening environment, a set of microphones in or around the listening zone(s) can be employed to identify the RIRs between each loudspeaker and each point of interest within the room (control points) by adaptive linear filters. Based on these identified Impulse Responses (IRs), the loudspeaker signals can be pre-processed in order to reduce the impact of room reverberation. The so-called LRE approach [20] computes a Multiple-Input/Multiple-Output (MIMO) pre-equalizer in order to match the equalized IRs between loudspeakers and control points with the desired responses, which can be implemented efficiently in the wave domain [21]. Recently, the use of additive compensation signals instead of the computation of a complete pre-equalizer has been proposed [1], which additionally borrows some ideas of Wave-Domain Adaptive Filtering (WDAF) for computational efficiency. Utilizing additive compensation signals will be referred to as LRC in the following[1]. Both LRE and LRC require an adaptive estimation of the Loudspeaker-Enclosure-Microphone System (LEMS) as a prerequisite to adaptively equalize or compensate the room.

In this contribution, we investigate and extend the core idea of LRC with additive compensation signals from [1] without limiting ourselves by requiring a spatial transform. In particular, we review the state of the art and specify an actual broad-band implementation for LRC in Sec. 2. In Sec. 3, the basic concept of first-order LRC is extended to higher-order LRC by cascading multiple multiple compensation units to achieve a better compensation performance. Furthermore, differences and similarities between LRE, first-order LRE, and higher-order LRE are analyzed as well. Section 4 finally compares the efficacy of the proposed LRC methods to LRE based on the Iterative DFT-domain Inversion (IDI) algorithm [20].

## 2. STATE OF THE ART

In this section, two methods for improving the rendering performance by pre-processing loudspeaker signals according to Fig. 1 are revisited, namely LRE and LRC. All considerations use the Discrete-Time Fourier Transform (DTFT) domain, where the frequency variable $\Omega$ will be dropped convenience. Filter matrices will be typeset in uppercase boldface letters and signal vectors in lowercase boldface letters. Consider Fig. 1 with loudspeaker signals $\mathbf{x}_\mathrm{L}$, pre-processor filters $\mathbf{H}^\mathrm{PP}$, preprocessed loudspeaker signals $\mathbf{x}_\mathrm{L}^\mathrm{PP}$, the true LEMS $\mathbf{H}$, the desired (free-field) responses $\mathbf{H}_0$, and the error signals $\mathbf{e}_\mathrm{eq}$ for measuring how much the upper path in Fig. 1 deviates from the desired free-field LEMS $\mathbf{H}_0$. Although the inverses of causal systems, which are necessary for reverberation

[1]Note that in the past, the term LRC has also mistakenly been used to describe LRE systems as well.

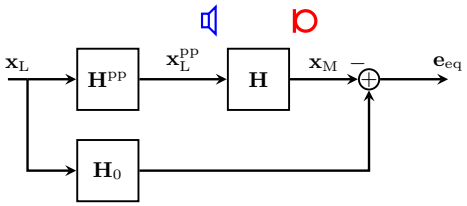
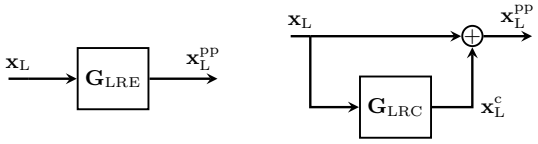



**Fig. 1**: Pre-processing loudspeaker signals to minimize the difference between the filtered responses and the desired responses $\mathbf{H}_0$.




(a) LRE pre-processing. (b) LRC pre-processing.

**Fig. 2**: Structural differences of pre-processing filters $\mathbf{H}^\mathrm{pp}$ for LRE and LRC.

compensation, require an additional delay before they can be approximated by causal discrete-time systems, we will disregard this implementation aspect for the DTFT-domain considerations.

### 2.1. Listening Room Equalization (LRE)

LRE typically minimizes a quadratic cost function derived from the error signal $\mathbf{e}_\mathrm{eq}$, employing the filter structure depicted in Fig. 2a. Such pre-equalizers ideally result in an $\mathbf{H}^\mathrm{pp} = \mathbf{G}_\mathrm{LRE}$ given by

$$\mathbf{G}_\mathrm{LRE} = \mathbf{H}^{-1}\mathbf{H}_0 \tag{1}$$

$$= \left(\mathbf{H}_0 + \mathbf{H}_\mathrm{r}\right)^{-1}\mathbf{H}_0, \tag{2}$$

where the true LEMS

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_\mathrm{r} \tag{3}$$

is represented here as superposition of the desired free-field component $\mathbf{H}_0$ and an undesired reverberation component $\mathbf{H}_\mathrm{r}$. If Eq. (1) is fulfilled, the error signal would become $\mathbf{e}_\mathrm{eq} = \mathbf{0}$. In practice, only an estimate $\hat{\mathbf{H}}$ for the true LEMS $\mathbf{H}$ is available and $\mathbf{G}_\mathrm{LRE}$ has to be estimated by an adaptive filtering algorithm, such as a filtered-X algorithm or the IDI algorithm [20–22].

### 2.2. Listening Room Compensation (LRC)

The structure of the pre-filter for LRC with additive compensation signals after [1] is depicted in Fig. 2b. As can be seen, LRC decomposes the pre-processor $\mathbf{H}^\mathrm{pp}$ into a forward path (upper path) and a compensation path producing additive compensation signals (lower path). In order to understand how this compensation method works, assume that $\mathbf{G}_\mathrm{LRC} = \mathbf{0}$ in Fig. 2b and consider the microphone signal $\mathbf{x}_\mathrm{M}$ and its reverberation component

$$\mathbf{x}_\mathrm{M,r} = \mathbf{x}_\mathrm{M} - \mathbf{H}_0\mathbf{x}_\mathrm{L} = \mathbf{H}_\mathrm{r}\mathbf{x}_\mathrm{L}, \tag{4}$$

which in Fig. 1 leads to the error signal

$$\mathbf{e}_\mathrm{eq} = \mathbf{x}_\mathrm{M,r} = \mathbf{H}_\mathrm{r}\mathbf{x}_\mathrm{L}. \tag{5}$$

Considering $\mathbf{e}_\mathrm{eq}$ as noise to be minimized by an Active Noise Cancellation (ANC) system [23] and let us assume that $\mathbf{H}$ in Fig. 1 is the secondary path, over which the compensation signals are transmitted to the error microphones. Now we assume that $\mathbf{H}$ only consists of the direct path $\mathbf{H}_0$ and ignore $\mathbf{H}_\mathrm{r}$, so that the ANC loudspeaker signal for minimizing the error would read:

$$\mathbf{x}_\mathrm{L}^\mathrm{c} = -\mathbf{H}_0^{-1}\mathbf{e}_\mathrm{eq} \tag{6}$$

$$= -\mathbf{H}_0^{-1}\mathbf{H}_\mathrm{r}\mathbf{x}_\mathrm{L}. \tag{7}$$

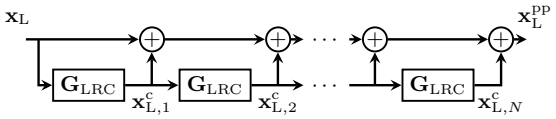



**Fig. 3**: Higher-order LRC by compensating the compensation signals repeatedly (here: order $N$).

To generate this $\mathbf{x}_\mathrm{L}^\mathrm{c}$, $\mathbf{x}_\mathrm{L}$ obviously has to be filtered with

$$\mathbf{G}_\mathrm{LRC} = -\mathbf{H}_0^{-1}\mathbf{H}_\mathrm{r} = -\mathbf{H}_0^{-1}\left(\mathbf{H} - \mathbf{H}_0\right), \tag{8}$$

which corresponds to the listening room compensation method of [1]. But as $\mathbf{H} \neq \mathbf{H}_0$, a residual error of

$$\mathbf{e}_{\mathrm{eq},1} = \mathbf{H}_\mathrm{r}\mathbf{x}_\mathrm{L} - \overbrace{\left(\mathbf{H}_0 + \mathbf{H}_\mathrm{r}\right)}^{\mathbf{H}}\overbrace{\mathbf{H}_0^{-1}\mathbf{H}_\mathrm{r}}^{-\mathbf{G}_\mathrm{LRC}}\mathbf{x}_\mathrm{L} \tag{9}$$

$$= -\mathbf{H}_\mathrm{r}\mathbf{H}_0^{-1}\mathbf{H}_\mathrm{r}\mathbf{x}_\mathrm{L} = \mathbf{H}_\mathrm{r}\mathbf{x}_\mathrm{L}^\mathrm{c} \tag{10}$$

remains. This immediately reveals the limitations of the LRC concept of [1]: the compensation signals $\mathbf{x}_\mathrm{L}^\mathrm{c}$, which should produce only the 'anti-sound' for the reverberation component $\mathbf{x}_\mathrm{M,r}$, undergo reflections themselves, such that the reverberation of the compensation signals is still present. According to [1], the aforementioned basic concept is nevertheless suited for low-reverberant environments. The impact of the systematic residual reverberation for more reverberant environments has not been investigated up to now. In [1], evaluations have been performed with complexity-reduced LEMS approximations in the modal domain for monofrequent excitations over the entire reproduction bandwidth. In this paper, will extend the work of [1] by employing actual broadband realizations of all involved systems in the form of digital Finite Impulse Response (FIR) filters and we do not consider approximate transform-domain models for the LEMS. In particular, the required system with transfer function matrices $\mathbf{H}_0^{-1}$ will be approximated in the Discrete Fourier Transform (DFT) domain by Tikhonov-regularized inverses of the form $\mathbf{H}_0^+ = \left(\mathbf{H}_0^\mathrm{H}\mathbf{H}_0 + \gamma\mathbf{I}\right)^{-1}\mathbf{H}_0^\mathrm{H}$, where a regularization parameter of $\gamma = 1$ will be assumed exemplarily in the following. For these generalizations, the basic properties of LRC with additive compensation signals and an extension to higher orders will be investigated in the following.

## 3. HIGHER-ORDER LISTENING ROOM COMPENSATION

In this section, the basic concept of LRC with additive compensation signals, extracted from [1], is extended to higher-order LRC (Sec. 3.1) and compared to LRE from an analytical point of view (Sec. 3.2). Afterwards, limitations and practical aspects of higher-order LRC are discussed in Sec. 3.3.

### 3.1. Basic Idea of Cascading Equalizers

As can be seen from Eq. (7), LRC can be interpreted as synthesizing 'anti-sound' for the reverberation at the control points, but assuming free-field conditions when playing back the compensation signals. Thus, the compensation signals $\mathbf{x}_\mathrm{L}^\mathrm{c}$ remove the reflections of the original loudspeaker signals $\mathbf{x}_\mathrm{L}$, but undergo reflections, absorptions, and attenuations themselves. That is, the original reverberation is replaced by the reverberation of the compensation signals. Simply repeating the LRC procedure for $\mathbf{x}_\mathrm{L}^\mathrm{c}$ leads to additive $2^\mathrm{nd}$-order compensation signals, which have to be compensated again, leading to additional $3^\mathrm{rd}$-order compensation signals, and so on. This cascaded structure is depicted in Fig. 3. Thereby, the $n^\mathrm{th}$-order compensation signal is $\mathbf{x}_{\mathrm{L},n}^\mathrm{c} = \mathbf{G}_\mathrm{LRC}^n\mathbf{x}_\mathrm{L}$, which are superimposed for all $n$ to yield the pre-processed loudspeaker signal

$$\mathbf{x}_\mathrm{L}^\mathrm{pp} = \sum_{n=0}^{N}\mathbf{x}_{\mathrm{L},n}^\mathrm{c} = \sum_{n=0}^{N}\mathbf{G}_\mathrm{LRC}^n\mathbf{x}_\mathrm{L}. \tag{11}$$

This obviously corresponds to a pre-filter $\mathbf{H}^{\text{pp}}$ in Fig. 1 of

$$\mathbf{H}^{\text{pp}}_{\text{LRC},N} = \sum_{n=0}^{N} \mathbf{G}^{n}_{\text{LRC}}. \tag{12}$$

Note that the complexity of this approach requires two filtering operations per order (see Eq. (7)). This corresponds to a complexity of $\mathcal{O}({N_\text{L}}^2 \cdot N)$ per frequency bin, whereas other system inversion approaches, such as IDI solving a linear system of equations, which is associated with $\mathcal{O}({N_\text{L}}^3)$ when doing so in a numerically robust way by a Cholesky decomposition [24].

### 3.2. Relation between LRE and LRC

In order to quantify the optimum performance of this higher-order LRC procedure, assume invertible transfer function matrices $\mathbf{H}$ and $\mathbf{H}_0$ and consider the limit

$$\lim_{N\to\infty} \mathbf{H}^{\text{pp}}_{N} = \lim_{N\to\infty} \sum_{n=0}^{N} \mathbf{G}^{n}_{\text{LRC}} \tag{13}$$

$$= (\mathbf{I} - \mathbf{G}_{\text{LRC}})^{-1} \tag{14}$$

$$\overset{(8)}{=} (\mathbf{I} + \mathbf{H}_0^{-1}(\mathbf{H} - \mathbf{H}_0))^{-1} \tag{15}$$

$$= \mathbf{H}^{-1}\mathbf{H}_0, \tag{16}$$

which is a geometric series of a matrix (also termed Neumann series [25, Eq. (186)]) and simplifies to Eq. (16) if all eigenvalues $\lambda_i$ of $\mathbf{G}_{\text{LRC}}$ fulfill $|\lambda_i| < 1$, $\forall i = 1, \ldots, N_\text{L}$. In this case, $\mathbf{G}_{\text{LRC}}$ is called a convergent matrix. Interestingly, Eq. (16) corresponds exactly to the LRE pre-filter $\mathbf{H}^{\text{pp}}_{\text{LRE}} = \mathbf{G}_{\text{LRE}}$ of Eq. (1), which states that LRE and higher-order LRC are asymptotically identical, given convergent matrices $\mathbf{G}_{\text{LRC}}$ and invertible $\mathbf{H}$ and $\mathbf{H}_0$. Another interpretation is that the truncation of the Neumann series to the $N^{\text{th}}$ partial sum leads to an approximate inverse. This is similar to approximating a nonlinear function by an $N^{\text{th}}$-order Taylor series expansion. Unfortunately, $|\lambda_i| < 1$ does not always hold in practice, which precludes an exact inversion by higher-order LRC. This will also be reflected by the performance measures in Sec. 4 later on. Yet, each finite-order LRC corresponds to a partial sum of Eq. (13) and will therefore be bounded, as long as $\mathbf{H}^{\text{pp}}$ is bounded, although the absolute value strongly depends on the eigenvalues of $\mathbf{G}_{\text{LRC}}$ and on the order $N$, of course. Thus, a limitation of the order of LRC may be necessary in practice.

### 3.3. Stabilization of Higher-Order LRC

As explained in Sec. 3.2, non-convergent first-order compensator matrices $\mathbf{G}_{\text{LRC}}$ limit the order of LRC. This section will introduce methods for alleviating the impact of non-convergent matrices $\mathbf{G}_{\text{LRC}}$, in order improve the performance of higher-order LRC and thereby extend the range of applicability of LRC. A first method for stabilizing the Neumann series would be to explicitly limit the eigenvalues $\lambda_i$ of the transfer function matrices $\mathbf{G}_{\text{LRC}}$. This can be achieved via the eigenvalue decomposition (EVD) $\mathbf{G}_{\text{LRC}} = \mathbf{E}\,\mathbf{diag}(\lambda_i)\mathbf{E}^{-1}$, where the matrix $\mathbf{E}$ contains the eigenvectors corresponding to $\lambda_i$ and where $\mathbf{diag}(\lambda_i)$ represents a diagonal matrix with the eigenvalues $\lambda_i$ as diagonal elements. Then, a re-synthesis of a convergent filter matrix can be done by

$$\mathbf{G}'_{\text{LRC}} = \mathbf{E}\,\mathbf{diag}(\lambda_i/|\lambda_i| \cdot \min(|\lambda_i|, 1-\epsilon))\mathbf{E}^{-1}, \tag{17}$$

where $\epsilon$ is a a small positive number. In practice, these operations can be performed in the DFT domain and an inverse DFT with subsequent temporal windowing yield the final filters. A second method for stabilizing the inversion is to introduce a soft decay of higher-order terms. This can be achieved efficiently by order-specific compensator matrices $\mathbf{G}_{\text{LRC}\,n}$ according to Eq. (8), where an order-specific Tikhonov regularization parameter $\gamma_n$ is used for the computation of an order-specific approximate inverse free field system

$$(\mathbf{H}_0)^{+}_{n} = (\mathbf{H}_0^{\text{H}}\mathbf{H}_0 + \gamma_n \mathbf{I})^{-1}\mathbf{H}_0^{\text{H}}, \tag{18}$$

as illustrated in Fig. 4. Thereby, the eigenvalues of $(\mathbf{H}_0)^{+}_{n}$ and $\mathbf{G}_{\text{LRC}\,n} = -(\mathbf{H}_0)^{+}_{n}\mathbf{H}_\text{r}$ decrease, which leads to a soft fade-out of higher-order compensation signals in practice. In the following, we will employ empirically chosen, exponentially increasing regularization parameters $\gamma_n = \gamma \cdot a^{n-1}$ with $a = 7$.

Note that in practice, the direct-path inverse $\mathbf{H}_0^{-1}$ can be assumed to not change rapidly, whereas the actual LEMS $\mathbf{H}$ may do so. Thus, in order to follow a time-variant LEMS $\mathbf{H}$, an order-specific regularization can be done offline and does not imply any computational overhead, whereas for the first method, a stabilization by the high-complexity EVD is necessary whenever a new estimate for $\mathbf{H}$ is available. Hence, the stabilization by order-specific regularization is most relevant for real-time systems with an adaptively identified $\mathbf{H}$.

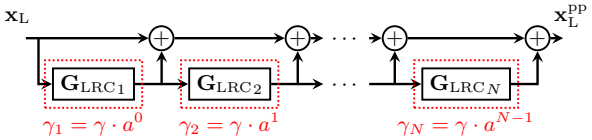



**Fig. 4**: Higher-order LRC with order-specific regularization parameters $\gamma_n$ for computation of the $n^{\text{th}}$-order compensation filter $\mathbf{G}_{\text{LRC}\,n}$.

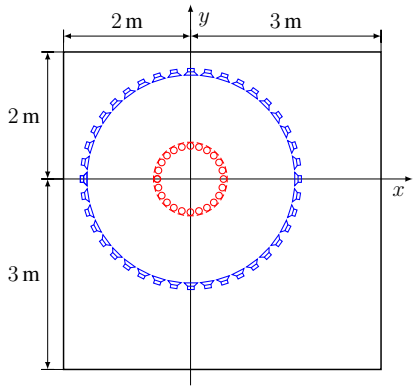



**Fig. 5**: Arrays in an image-source environment: microphones are arranged around the listening area as uniform circular array (**red**) and the loudspeakers are placed concentrically on circle (**blue**).

## 4. EXPERIMENTS

### 4.1. Experimental Setup

In this section, the setup and the performance measure for evaluating LRE and LRC are described. All evaluations in Sec. 4.2 are conducted in the image-source model [26] depicted in Fig. 5, which consists of four plane walls with reflection coefficient $r$ and fully absorptive floor and ceiling. Therein, a 30-element loudspeaker array (**blue**) with radius $R_\text{L} = 1\,\text{m}$ and an 80-element microphone array (**red**) with radius $R_\text{M} = 0.6\,\text{m}$ are placed concentrically around the origin. Note that this case with $N_\text{M} > N_\text{L}$, the reproduction error cannot be made zero at all microphones at the same time, in general, independent of the room equalization or compensation method. The listening area of a sound reproduction system for such arrays is typically the region enframed by the microphone array. Furthermore, a sampling rate of 2000 Hz is assumed, in order to avoid spatial aliasing to be able to control the sound field reasonably well. For an assessment of the capabilities of LRE and LRC themselves, without in-

terdependence to the employed system identification algorithm and without the non-uniqueness problem of multichannel system identification [27], the acoustic paths, captured by $\mathbf{H}$, are assumed to be known. Its RIRs have a length of 256 taps, the first order LRC filter is of length 335 taps, and the LRE filter is of length $5 \cdot 335$ taps for comparison to a fifth-order LRC.

Direct measures of the amount of reverberation contained in the signals are Direct-to-Reverberation Ratios (DRRs), which can be computed either from time-domain IRs or from the respective transfer function matrices. Although the measures are computed numerically in the time domain, we will define the measures in the DTFT domain employing the previously introduced transfer function matrices: the DRRs for the unequalized and for the pre-filtered (LRE or LRC) rendering system can be written as

$$\mathrm{DRR}_{\mathrm{un}} = \frac{\int_0^{2\pi} \|\mathbf{H}_0\|_{\mathrm{F}}^2 \mathrm{d}\Omega}{\int_0^{2\pi} \|\mathbf{H}_{\mathrm{r}}\|_{\mathrm{F}}^2 \mathrm{d}\Omega}, \tag{19}$$

$$\mathrm{DRR}_{\mathrm{filt}} = \frac{\int_0^{2\pi} \|\mathbf{H}_0\|_{\mathrm{F}}^2 \mathrm{d}\Omega}{\int_0^{2\pi} \|\mathbf{H}\mathbf{H}^{\mathrm{PP}} - \mathbf{H}_0\|_{\mathrm{F}}^2 \mathrm{d}\Omega}, \tag{20}$$

respectively, where $\Omega$ is the normalized angular frequency variable of the DTFT domain and parametrizes each transfer function matrix. With these intermediate definitions, a performance measure denoted as DRR gain can be defined as

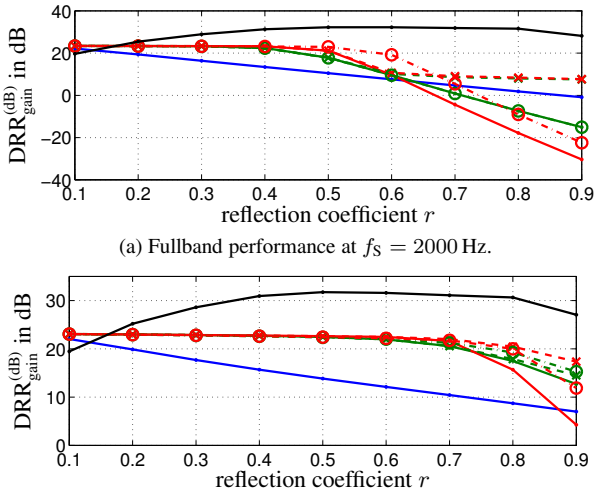
$$\begin{aligned} \mathrm{DRR}_{\mathrm{gain}}^{(\mathrm{dB})} &= 10 \cdot \log_{10} \left( \mathrm{DRR}_{\mathrm{filt}} / \mathrm{DRR}_{\mathrm{un}} \right) \mathrm{dB} \\ &= 10 \cdot \log_{10} \frac{\int_0^{2\pi} \|\mathbf{H}_{\mathrm{r}}\|_{\mathrm{F}}^2 \mathrm{d}\Omega}{\int_0^{2\pi} \|\mathbf{H}\mathbf{H}^{\mathrm{PP}} - \mathbf{H}_0\|_{\mathrm{F}}^2 \mathrm{d}\Omega} \mathrm{dB}, \end{aligned} \tag{21}$$

where a larger $\mathrm{DRR}_{\mathrm{gain}}^{(\mathrm{dB})}$ means a more significant performance increase of the rendering system due to the pre-filtering.

### 4.2. Experimental Results

In this section, the basic broadband implementation of LRC, the proposed higher-order LRC, and LRE employing equalizers determined by the efficient IDI algorithm [20] are compared in terms of DRR gain after Eq. (21). Figure 6a depicts the DRR gains as functions of the reflection coefficient $r$ of the walls of the image-source environment. Obviously, the LRE performance (black) is not reached by LRC, which only approximates LRE (recall Sec. 3.2)—unless for the case of $r = 1$. This is caused by a regularization parameter of the IDI algorithm ($\delta_b = 0.01$ according to [20]) which has been kept constant for all reflection coefficients $r$. The other extremal case of the performance is the first-order LRC (**blue**), which achieves a reverberation cancellation of more than 20 dB for the very low reflection coefficient $r = 0.1$, but monotonously drops down to 0 dB for $r = 0.9$. Higher-order LRC implementations (solid **green** and **red** curves) perform better for low-reverberation conditions and successfully maintain a 20 dB gain up to reflection factors of about 0.5. From thereon, the performances rapidly drop and even become negative for large $r$. This is caused by non-convergent transfer function matrices and therefore worsens with increasing order of the LRC. The corresponding stabilized LRC versions are represented by colored lines with cross markers (stabilization by EVD) and by circles (order-specific regularization), respectively. As can be seen, the order-specific regularization extends the range of applicability of LRC to more reflective environments (see fifth-order curves), but still exhibits the break-down. On the other hand, forcing convergent matrices by an EVD (cross markers) can ensure significant DRR gains for severe reverberation conditions as well.

Interestingly, non-convergent LRC compensator matrices appear at low frequencies first. Therefore, problems of higher-order LRE is mainly caused by very low frequencies, which are hardly audible, cannot be played back by loudspeakers, and may be removed from the loudspeaker signals anyway when protecting the loudspeakers against DC components of the signal. For this reason, the DRR gains have been computed from impulse responses filtered with a high-pass (cut-off frequency at 19 Hz) and are depicted in Fig. 6b. In terms of the practically relevant, highpass-filtered DRR gains, the LRC methods are judged significantly better. Furthermore, both stabilization methods extend the range of applicability of LRC to the whole investigated interval of reflection coefficients. In particular, already third-order LRC with the efficient order-specific regularization leads to DRR gains of more than 15 dB, even for large reflection coefficients. This is a high value compared to system distances achievable by the identification of the LEMS, which will most likely be the limiting factor for typical applications with strongly correlated loudspeaker signals. Furthermore, a 3rd-order LRC has only about 1/5 of the complexity of LRE employing the IDI algorithm.




(a) Fullband performance at $f_{\mathrm{S}} = 2000$ Hz.

(b) Frequency range above 19 Hz (high-pass filtered).

**Fig. 6**: DRR gain by pre-filtering a rendering system:

The presented results emphasize the benefit of higher-order LRC relative to first order LRC and confirm the suitability of the structurally and computationally efficient higher-order LRC to improve the rendering performance in reverberant environments.

## 5. CONCLUSION

The analyses in this paper have shown the asymptotic equivalence of LRE and the proposed higher-order LRC for the special case that the first order LRC filters' transfer function matrices are convergent. As this assumption is violated in practice very often and would lead to stability problems, countermeasures employing an eigenvalue decomposition or an order-specific regularization have been proposed. Evaluations for a fourth-order image-source model confirm the benefit of higher-order LRC and the effectiveness of the computationally efficient order-specific regularization, suggesting third-order LRC as a practical alternative to LRE. Especially, because only a single adaptive system is required to identify the LEMS—a second convergence phase for determining the equalizer after the system identification is unnecessary. After system identification, the computational complexity reduces from $\mathcal{O}({N_{\mathrm{L}}}^3)$ for filter adaptation for LRE to $\mathcal{O}({N_{\mathrm{L}}}^2 N)$ for $N^{\mathrm{th}}$-order LRC with $N_{\mathrm{L}}$ loudspeakers.

## 6. REFERENCES


[1] D. Talagala, W. Zhang, and T. Abhayapala, "Efficient multi-channel adaptive room compensation for spatial soundfield reproduction using a modal decomposition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 10, pp. 1522–1532, Oct 2014.

[2] M. A. Gerzon, "Periphony: With-height sound reproduction," *Journal of the Audio Engineering Society*, vol. 21, no. 1, pp. 2–10, 1973.

[3] M. A. Gerzon, "Ambisonics in multichannel broadcasting and video," *Journal of the Audio Engineering Society*, vol. 33, no. 11, pp. 859–871, 1985.

[4] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *Journal of the Acoustical Society of America (JASA)*, vol. 93, no. 5, pp. 2764–2778, 1993.

[5] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Audio Engineering Society Convention 124*, 2008, pp. 17–20.

[6] J. Ahrens and S. Spors, "Analytical driving functions for higher-order Ambisonics," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*. Las Vegas, United States: IEEE, April 2008, pp. 373–376.

[7] M. Poletti, "An investigation of 2-d multizone surround sound systems," in *Audio Engineering Society Convention 125*, Oct 2008.

[8] M. A. Poletti, T. D. Abhayapala, and P. Samarasinghe, "Interior and exterior sound field control using two dimensional higher-order variable-directivity sources," *Journal of the Acoustical Society of America (JASA)*, vol. 131, no. 5, pp. 3814–3823, May 2012.

[9] M. Buerger, R. Maas, H. W. Löllmann, and W. Kellermann, "Multizone sound field synthesis based on the joint optimization of the sound pressure and particle velocity vector on closed contours," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, October 2015.

[10] M. Buerger, C. Hofmann, and W. Kellermann, "The impact of loudspeaker array imperfections and reverberation on the performance of a multizone sound field synthesis system," in *3rd International Conference on Spatial Audio*, September 2015.

[11] M. A. Poletti, T. Betlehem, and T. D. Abhayapala, "Higher-order loudspeakers and active compensation for improved 2D sound field reproduction in rooms," *Journal of the Audio Engineering Society*, vol. 63, no. 1/2, pp. 31–45, 2015.

[12] T. Betlehem and C. Withers, "Sound field reproduction with energy constraint on loudspeaker weights," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 8, pp. 2388–2392, 2012.

[13] G. Lilis, D. Angelosante, and G. Giannakis, "Sound field reproduction using the lasso," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 8, pp. 1902–1912, 2010.

[14] J. Jungmann, R. Mazur, and A. Mertins, "Joint time-domain reshaping and frequency-domain equalization of room impulse responses," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6642–6646.

[15] B. S. Atal and M. R. Schroeder, "Apparent sound source translator," US Patent US3 236 949A, February 22, 1966.

[16] D. B. Ward and G. Elko, "Optimum loudspeaker spacing for robust crosstalk cancellation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 6, May 1998, pp. 3541–3544.

[17] A. Sakurai and S. Trautmann, "Cross-talk cancellation," US Patent 2005/0 254 660 A1, November, 2005.

[18] L. Krishnan, P. D. Teal, and T. Betlehem, "A robust sparse approach to acoustic impulse response shaping," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2015, pp. 738–742.

[19] J. Grosse and S. van de Par, "Perceptually accurate reproduction of recorded sound fields in a reverberant room using spatially distributed loudspeakers," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 867–880, Aug 2015.

[20] M. Schneider and W. Kellermann, "Iterative DFT-domain inverse filter determination for adaptive listening room equalization," in *International Workshop on Acoustic Signal Enhancement (IWAENC)*, Aachen, Germany, September 2012, pp. 1–4.

[21] M. Schneider and W. Kellermann, "Adaptive listening room equalization using a scalable filtering structure in the wave domain," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2012.

[22] S. Goetze, M. Kallinger, A. Mertins, and K. Kammeyer, "Multichannel listening-room compensation using a decoupled filtered-X LMS algorithm," in *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, Califonia, USA, October 2008, pp. 811–815.

[23] S. M. Kuo and D. Morgan, *Active noise control systems: algorithms and DSP implementations*. John Wiley & Sons, Inc., 1995.

[24] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Johns Hopkins University Press, 1996.

[25] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Technical University of Denmark, nov 2012, version 20121115. [Online]. Available: http://www2.imm.dtu.dk/pubdb/p.php?3274

[26] J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America (JASA)*, vol. 65, no. 4, pp. 943–950, 1979.

[27] J. Benesty, D. Morgan, and M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Transactions on Speech and Audio Processing*, vol. 6, no. 2, pp. 156–165, 1998.