Efficient Multi-Channel Adaptive Room Compensation for Spatial Soundfield Reproduction Using a Modal Decomposition

Dumidu S. Talagala, *Member, IEEE*, Wen Zhang, *Member, IEEE*, and Thushara D. Abhayapala, *Senior Member, IEEE*

***Abstract*—Mitigating the effects of reverberation is a significant challenge for real-world spatial soundfield reproduction, but the necessity of a large number of reproduction channels increases the complexity and presents several challenges to existing listening room compensation techniques. In this paper, we present an adaptive room compensation method to overcome the effects of reverberation within a region, using a model description of the reverberant soundfield. We propose the reverberant channel estimation and compensation be carried out in a single step using completely decoupled adaptive filters; thus, reducing the complexity of the overall process. We compare the soundfield reproduction performance with existing adaptive and nonadaptive room compensation methods through several simulation examples. The performance of the proposed method is comparable to existing techniques, and achieves a normalized wideband region reproduction error of 1% at a signal-to-noise ratio of 50 dB, within a 1 m radius region of interest using 60 loudspeakers and 55 microphones at frequencies below 1 kHz. Robust behavior of the room compensator is demonstrated down to direct-to-reverberant-path power ratios of -5 dB. Overall, the results suggest that the proposed method can diagonalize the room compensation system, leading to a more robust and parallel implementation for spatial soundfield reproduction.**

***Index Terms*—Adaptive channel estimation, decoupled adaptive filtering, modal decomposition, multi-channel audio, reverberation, room compensation, room equalization, soundfield reproduction.**

I. Introduction

REPRODUCTION of a desired soundfield within a region of interest is the ultimate objective of all soundfield reproduction systems. However, environmental noise and reverberation caused by the listening room will result in a less than ideal reproduction of the desired soundfield. In applications such as teleconferencing or gaming, smaller rooms and the materials used in their construction may result in significantly higher levels of reverberation; thus, requiring compensation by the system designer. However, the lack of knowledge of the reverberant channel, its time-varying nature, and the large number of loudspeakers required by many soundfield reproduction systems, complicates the application of classical active control techniques for reverberation control. This paper introduces an adaptive reverberant channel estimation technique that uses a modal description of the reverberant soundfield to simplify the listening room compensation problem within a region of interest.

Spatial soundfield reproduction within a desired region has been an active field of research for many years. Reproduction techniques can be broadly classified into two types; those based on Ambisonics [2], [3] or spatial harmonics [4]–[7] and others based on Wave Field Synthesis (WFS) [8], [9]. Each method requires a number loudspeakers to be placed at discrete locations outside a region of interest. Typically, this number is proportional to the size of the region of interest and the maximum operating frequency; hence, a reasonably large reproduction region will require a large number of loudspeakers. Although existing soundfield reproduction methods achieve good reproduction performance under free-field conditions, reverberation often results in the drastic degradation of performance.

The techniques employed to reduce the effects of reverberation on the reproduced soundfield can be broadly grouped into three categories [10]–[12]; passive techniques that minimize reflections, compensation schemes based on models of the reverberant room or channel behavior, and adaptive room compensation methods. The passive approaches can be further categorized into two subtypes. Although the simplest of these, the use of acoustic insulation materials, can produce a modest reduction in reverberation, it is often outweighed by the associated costs and impracticality in many real-world application scenarios (e.g., soundfield reproduction in an office or home environment). The more complex passive approaches may use fixed or variable directivity higher order loudspeakers in order to minimize the acoustic energy directed towards the walls of a room [13], [14], or exploit the reverberant sound sources to reproduce a desired soundfield [15]. In general, this requires some control of the listening room geometry and sound reproduction apparatus, which may not always be possible. In contrast, equalizing the effects of reverberation, i.e., inverting the

Manuscript received March 26, 2013; revised February 20, 2014; accepted July 04, 2014. Date of publication July 14, 2014; date of current version July 23, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Patrick A. Naylor.

D. S. Talagala was with the Applied Signal Processing Group, Research School of Engineering, College of Engineering and Computer Science, Australian National University, Canberra ACT 0200, Australia. He is now with the Centre for Vision, Speech, and Signal Processing, University of Surrey, Guildford, Surrey GU2 7XH, U.K. (e-mail: dumidu.talagala@anu.edu.au; d.talagala@surrey.ac.uk).

W. Zhang and T. D. Abhayapala are with the Applied Signal Processing Group, Research School of Engineering, College of Engineering and Computer Science, Australian National University, Canberra ACT 0200, Australia (e-mail: wen.zhang@anu.edu.au; thushara.abhayapala@anu.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASLP.2014.2339195

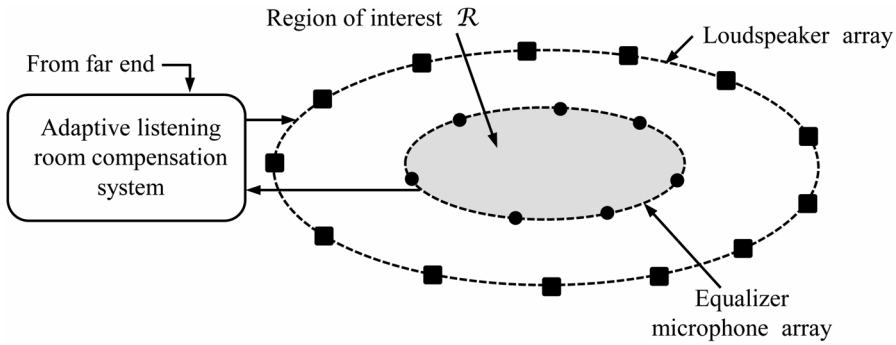


Fig. 1. Loudspeaker and microphone array configuration of the proposed listening room compensation system.

reverberant channel response at a set of control points prior to driving the loudspeakers, was shown to be theoretically capable of good performance [16], [17] over a region of interest. However, inversion of the reverberant channel requires an accurate description of the reverberant room, and imperfections in the modelling or measurement process may lead to a reduction of the room compensation performance. This is further exacerbated by non-static room conditions, rapid variations in the reverberant channel at different locations [18], and the sensitivity to the loudspeaker-microphone placement and orientation [19]. The recent work by Brännmark *et al.* [20], [21] has extended the basic concept of equalizing the effects of reverberation to a polynomial based Multiple Input Multiple Output (MIMO) channel equalization problem by applying a probabilistic model of the channel variability. Although the approach is shown to be quite robust, a prior measurement of the room transfer functions at a set of control points is still necessitated.

In this context, adaptive techniques are better suited for the general problem of mitigating the effects of reverberation within a spatial region, but require a large number of loudspeakers for a region of an appreciable size [16], [22]. Increasing the number of loudspeakers results in significantly higher computational complexity [23], and the high correlation between these reproduction channels can lead to the ill-conditioning of matrices used during the adaptive channel estimation. Thus, the convergence behavior of conventional time and frequency domain adaptive filters is adversely affected by the large number of loudspeakers used for soundfield reproduction. Eigenspace Adaptive Filtering (EAF) [24], [25] was proposed to overcome these limitations by decoupling the loudspeaker signals from MIMO system that represents the reverberant room. Ideally, EAF requires data dependent transformations, but it was shown that a wave-domain transformation [25] could be used as a practical alternative to these transformations. Initially proposed for multi-channel acoustic echo cancellation, Wave Domain Adaptive Filtering (WDAF) [26], [27] has since been used for adaptive listening room compensation in WFS systems [25], [28]–[32]. The work by Schneider *et al.* [31] has further reduced the computational complexity of the basic EAF/WDAF adaptation process by considering the coupling between soundfield modes to be minimal, i.e., coupling is limited to just one or two adjacent soundfield modes. An alternate approach known as Adaptive WFS (AWFS) was proposed by Gauthier *et al.* [33]–[35], where the effects of reverberation could be compensated using an optimal control approach. This was achieved using a weighted cost function, and is equivalent to the independent control of individual radiation modes. Together, the work by Schneider *et al.* and Gauthier *et al.* provides some insights into the underlying structure of the reverberant soundfield in the mode-domain. We use this inspiration to model the reverberant soundfield, and propose an adaptive listening room compensation method for spatial soundfield reproduction within a region by extending our previous work on acoustic echo cancellation in [1]. The estimation of the reverberant channel and the compensation signal calculation is proposed as a single adaptive operation on individual soundfield modes, and its performance and robustness are evaluated.

The remainder of this paper is organized as follows. In Section II we describe the structure of the listening room compensation system and the mode-domain representation of the desired and measured soundfields. Section III describes our model of the reverberant soundfield; a collection of linear transformations of the desired soundfield modes. We then describe the adaptive channel estimation process used to identify these transformations and derive the required loudspeaker compensation signals. The simulation setup and the various measures of the soundfield reproduction performance are described next in Section IV. Section V compares the performance of the proposed method with existing adaptive and non-adaptive listening room compensation techniques using simulations of a reverberant room, and is followed by the concluding remarks in Section VI.

II. Structure of the Spatial Soundfield Reproduction System

Consider the problem of recreating a desired soundfield within a region of interest $\mathcal{R}$ of a two-dimensional reverberant room. In this paper, we propose the use of a feedback control system that consists of two concentric circular arrays of loudspeakers and microphones. Fig. 1 illustrates the array configuration while the block diagram in Fig. 2 indicates the signal flow within the system. The outer array consist of P loudspeakers and the inner array consists of Q equalizer microphones. Together these arrays reproduce the desired soundfield in $\mathcal{R}$ using an iterative feedback control approach. A similar configuration can be used in teleconferencing or virtual reality gaming applications, where the virtual source positioning is performed by the loudspeaker and equalizer microphone arrays. In this section, we describe the signal and channel model used characterize the spatial soundfield in $\mathcal{R}$.

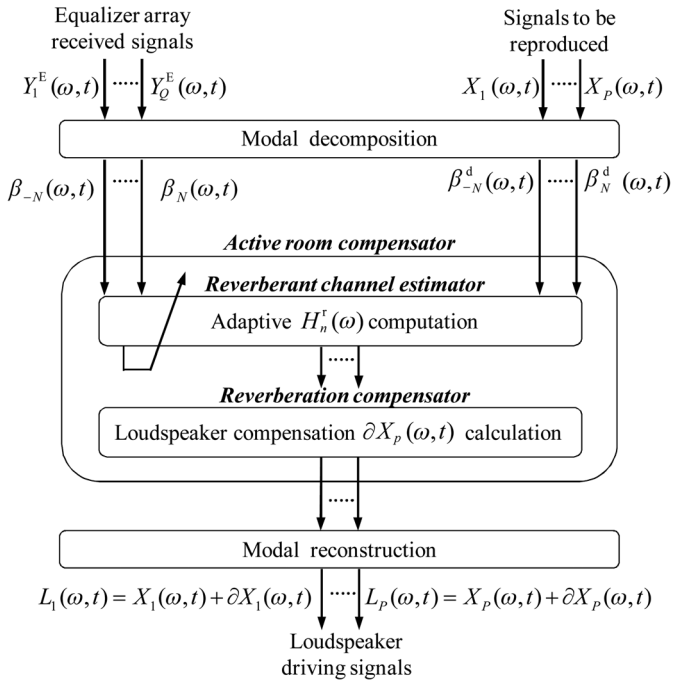


Fig. 2. Signal flow block diagram of the proposed listening room compensation system.

A. The Signal and Channel Model

Consider the scenario where $x_p(t)$ $(p = 1 \ldots P)$ are P loudspeaker driving signals, that have been preconditioned to recreate a particular soundfield in $\mathcal{R}$. The time domain signals $x_p(t)$ can be described as a collection of time-varying signals of different frequencies using a Short Time Fourier Transform (STFT). The Fourier coefficients of the pth loudspeaker signal at a frequency ω can therefore be expressed as

$$X_p(\omega,t) = \int_{-\infty}^{\infty} x_p(t')w(t'-t)e^{-j\omega t'}\mathrm{d}t', \tag{1}$$

where $w(t')$ represents a square window function of an appropriate length for the intended frequency bin size.

The interaction between the reverberant room and each loudspeaker-microphone pair represents a convolution in the time domain, which becomes a multiplication operation in the frequency domain representation in (1). Thus, the received signal at the qth equalizer microphone $(q = 1 \ldots Q)$ at a frequency ω becomes

$$Y_q^{\mathrm{E}}(\omega,t) = \sum_{p=1}^{P} H_{pq}(\omega)X_p(\omega,t) + V_q^{\mathrm{E}}(\omega,t), \tag{2}$$

where $H_{pq}(\cdot)$ represents the time-invariant transfer function (time-invariant during short time periods) of the reverberant channel between the (p,q)th loudspeaker-microphone pair and $V_q^{\mathrm{E}}(\omega,t)$ is the ambient noise. The received signals at each equalizer microphone can be expressed in the more convenient matrix form

$$\mathbf{Y}^{\mathbf{E}}(\omega,t) = \mathbf{H}(\omega)\mathbf{X}(\omega,t) + \mathbf{V}^{\mathbf{E}}(\omega,t), \tag{3}$$

where

$$\mathbf{Y}^{\mathbf{E}}(\omega,t) = [\,Y_1^{\mathrm{E}}(\omega,t) \quad Y_2^{\mathrm{E}}(\omega,t) \quad \cdots \quad Y_Q^{\mathrm{E}}(\omega,t)\,]^T,$$
$$\mathbf{H}(\omega) = \begin{bmatrix} H_{11}(\omega) & H_{21}(\omega) & \cdots & H_{P1}(\omega) \\ H_{12}(\omega) & H_{22}(\omega) & \cdots & H_{P2}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ H_{1Q}(\omega) & H_{2Q}(\omega) & \cdots & H_{PQ}(\omega) \end{bmatrix},$$
$$\mathbf{X}(\omega,t) = [\,X_1(\omega,t) \quad X_2(\omega,t) \quad \cdots \quad X_P(\omega,t)\,]^T.$$

B. Modal Representation of a Soundfield

Sound pressure at any point $\mathbf{x}$ within a source-free region can be expressed using the interior solution to the wave equation [36]. In the plane of $\mathcal{R}$, the sound pressure at a location $\mathbf{x} \equiv (x, \phi_x)$ is given by the summation

$$Y(\mathbf{x};\omega,t) = \sum_{n=-\infty}^{\infty} \beta_n(\omega,t)J_n(kx)e^{in\phi_x}, \tag{4}$$

where the Bessel function $J_n(\cdot)$ and exponential term $e^{(\cdot)}$ represent the orthogonal basis function. $k = \omega/c$ is the wave number where the speed of sound in the medium is given by c. The implication is that the soundfield within a region of radius x is characterized by the soundfield coefficients $\beta_n(\omega,t)$, and is a property we use to solve the room compensation problem.

Consider a region of interest $\mathcal{R}$ within the circular equalizer array of radius r_e, where the desired and measured soundfields at any point within that region can be expressed in terms of its soundfield coefficients recorded at the equalizer array [16], [36]. By applying an appropriate truncation length N to the number of active basis functions for a specified error bound [16], [22], the desired and measured soundfields at the equalizer array can be expressed as

$$Y^{\mathrm{d}}(\mathbf{x}_{\mathbf{e}};\omega,t) = \sum_{n=-N}^{N} \beta_n^{\mathrm{d}}(\omega,t)J_n(kr_e)e^{in\phi_x} \tag{5}$$

and

$$Y(\mathbf{x}_{\mathbf{e}};\omega,t) = \sum_{n=-N}^{N} \beta_n(\omega,t)J_n(kr_e)e^{in\phi_x}, \tag{6}$$

respectively. $\beta_n^{\mathrm{d}}(\omega,t)$ and $\beta_n(\omega,t)$ now represent the desired and measured soundfield coefficients at the equalizer array enclosing the region $\mathcal{R}$. Hence, by comparing (5) and (6) it can be seen that the desired soundfield within $\mathcal{R}$ can be reproduced by satisfying the condition

$$\beta_n(\omega,t) = \beta_n^{\mathrm{d}}(\omega,t). \tag{7}$$

For the room compensation problem in this work, the desired soundfield given by $\beta_n^{\mathrm{d}}(\omega,t)$ represents the free-field or direct-path propagation behavior. Thus, $\beta_n^{\mathrm{d}}(\omega,t)$ can be computed to recreate a virtual source at a desired location outside $\mathcal{R}$.

The measured soundfield coefficients at the equalizer microphone array can be obtained from the analysis equation

$$\beta_n(\omega,t) = \frac{1}{2\pi J_n(kr_e)} \int_0^{2\pi} Y^{\mathrm{E}}(\mathbf{x}_{\mathbf{e}};\omega,t)e^{-in\phi_x}\mathrm{d}\phi_x, \tag{8}$$

where $Y^{\mathrm{E}}(\mathbf{x_e};\omega,t)$ are the soundfield measurements along the array. It should be noted that the zero crossings of the Bessel functions can affect the accuracy of the estimated soundfield coefficients, especially at higher frequencies. Although we disregard this problem in the context of this paper by selecting r_e or restricting the range of k such that $J_n(kr_e) \neq 0$, more robust techniques such as the use of multiple or rigid microphone arrays [16] can be used to overcome this problem in a practical scenario.

Since the Q microphones are evenly spaced in the azimuth of the equalizer array (i.e., $\mathrm{d}\phi_x = 2\pi/Q$), (8) can be approximated by a Discrete Fourier Transform (DFT) of (3) [16]. Thus, the measured soundfield coefficients $\beta_n(\omega,t)$ can be expressed in the matrix form

$$\begin{aligned}\boldsymbol{\beta}(\omega,t) &= \mathbf{T_{CH}}\mathbf{Y^E}(\omega,t)\\ &= \mathbf{T_{CH}}\mathbf{H}(\omega)\mathbf{X}(\omega,t) + \mathbf{T_{CH}}\mathbf{V^E}(\omega,t),\end{aligned} \tag{9}$$

where

$$\boldsymbol{\beta}(\omega,t) = \begin{bmatrix}\beta_{-N}(\omega,t) & \cdots & \beta_N(\omega,t)\end{bmatrix}^T,$$

$$\mathbf{T_{CH}} = \frac{1}{Q}\mathbf{J^{-1}}\begin{bmatrix} e^{jN\phi_1} & e^{jN\phi_2} & \cdots & e^{jN\phi_Q}\\ \vdots & \vdots & \ddots & \vdots\\ e^{-j0\phi_1} & e^{-j0\phi_2} & \cdots & e^{-j0\phi_Q}\\ \vdots & \vdots & \ddots & \vdots\\ e^{-jN\phi_1} & e^{-jN\phi_2} & \cdots & e^{-jN\phi_Q}\end{bmatrix}$$

represents the transformation into the circular harmonic mode-domain and

$$\mathbf{J^{-1}} = \mathrm{diag}\begin{bmatrix}J_{-N}(kr_e) & \cdots & J_N(kr_e)\end{bmatrix}^{-1}.$$

Desired soundfield coefficients can be expressed similarly as

$$\boldsymbol{\beta}^{\mathbf{d}}(\omega,t) = \mathbf{T_{CH}}\mathbf{H^d}(\omega)\mathbf{X}(\omega,t), \tag{10}$$

where

$$\boldsymbol{\beta}^{\mathbf{d}}(\omega,t) = \begin{bmatrix}\beta^{\mathrm{d}}_{-N}(\omega,t) & \cdots & \beta^{\mathrm{d}}_{N}(\omega,t)\end{bmatrix}^T$$

and $\mathbf{H^d}(\cdot)$ is the direct-path room transfer function coefficients of the loudspeaker-microphone pairs.

III. Room Compensation and Channel Estimation

In the previous section we described how the desired and measured soundfields in $\mathcal{R}$ can be characterized using a collection of soundfield coefficients measured at an equalizer microphone array. This section describes a model of the reverberant channel, the computation of the compensation signals required at the loudspeakers, and the channel estimation processes in the mode-domain.

A. Reverberant Channel Model

Comparing (9) and (10), the measured soundfield coefficients at the equalizer array can be described as a collection of modes given by

$$\boldsymbol{\beta}(\omega,t) \triangleq \boldsymbol{\beta}^{\mathrm{d}}(\omega,t) + \boldsymbol{\beta}^{\mathrm{r}}(\omega,t) + \boldsymbol{\beta}^{\mathrm{v}}(\omega,t). \tag{11}$$

Each mode represents a different effect measured at the equalizer array, where the desired and reverberant soundfield coefficients are $\boldsymbol{\beta}^{\mathrm{d}}(\omega,t)$ and $\boldsymbol{\beta}^{\mathrm{r}}(\omega,t)$ respectively. The effects of any other sources (independent of the loudspeaker signals) as well as any ambient noise effects are collectively described by the soundfield coefficients $\boldsymbol{\beta}^{\mathrm{v}}(\omega,t) = \mathbf{T_{CH}}\mathbf{V^E}(\omega,t)$.

Consider the reverberant soundfield as a linear transformation of the desired soundfield. $\boldsymbol{\beta}^{\mathrm{r}}(\omega,t)$ can then be expressed as

$$\boldsymbol{\beta}^{\mathbf{r}}(\omega,t) = \mathbf{H^r}(\omega)\boldsymbol{\beta}^{\mathbf{d}}(\omega,t), \tag{12}$$

where $\mathbf{H^r}(\omega) = \mathrm{diag}[H^{\mathrm{r}}_{-N}(\omega)\ldots H^{\mathrm{r}}_{N}(\omega)]$ is a transformation matrix that describes the effect of reverberation. Since the individual soundfield modes are coefficients of the orthogonal basis functions in (4), the different modes do not interact with each other in a two-dimensional soundfield described in the previous section [36], and produces the diagonal structure of $\mathbf{H^r}(\omega)$. This can be visualized intuitively as follows. Consider each independent soundfield mode as a separate source. The cumulative effect of reverberation on each mode (caused by multiple scaled and delayed images of the itself) can then be represented by the corresponding diagonal element of $\mathbf{H^r}(\omega)$. If we assume that this effect differs for each mode, the effect of reverberation on the reverberant soundfield coefficients is described by the linear transformation in (12).

B. Loudspeaker Compensation Signals

Consider the active listening room compensation system shown in Fig. 2, where reverberation is controlled using an active noise control approach at the loudspeakers. The required compensation signals at the loudspeaker can now be described using the reverberation transformation matrix $\mathbf{H^r}(\omega)$ in (12). If we assume that an estimate of $\mathbf{H^r}(\omega)$ is available, the reverberation compensation signals at the loudspeaker $\boldsymbol{\delta}\mathbf{X}(\omega,t)$ can be derived as follows.

First, the channel effects of $\mathbf{H}(\omega)$ are expressed as the sum of the direct-path effects $\mathbf{H^d}(\omega)$ and the reverberant path effects $\mathbf{H^r}(\omega)$. Next, in order to derive these compensation signals, we ignore the effects of other sound sources, i.e., $\boldsymbol{\beta}^{\mathbf{v}}(\omega,t) = 0$ (we will show that these signals can still be used to compensate the effects of reverberation when $\boldsymbol{\beta}^{\mathbf{v}}(\omega,t) \neq 0$ in Section V). Thus, (9) can be reformulated as

$$\begin{aligned}\boldsymbol{\beta}(\omega,t) &= \mathbf{T_{CH}}\left[\mathbf{H^d}(\omega) + \mathbf{H^r}(\omega)\right]\mathbf{X}(\omega,t)\\ &= \boldsymbol{\beta}^{\mathbf{d}}(\omega,t) + \boldsymbol{\beta}^{\mathbf{r}}(\omega,t).\end{aligned} \tag{13}$$

Substituting the results in (10) and (12) leads to

$$\boldsymbol{\beta}(\omega,t) = [\mathbf{I} + \mathbf{H^r}(\omega)]\mathbf{T_{CH}}\mathbf{H^d}(\omega)\mathbf{X}(\omega,t), \tag{14}$$

where $\mathbf{I}$ represents the identity matrix. The compensation signals at the loudspeaker $\boldsymbol{\delta}\mathbf{X}(\omega,t)$ are then introduced into (14), such that $\boldsymbol{\beta}(\omega,t) = \boldsymbol{\beta}^{\mathbf{d}}(\omega,t)$. Thus,

$$\boldsymbol{\beta}^{\mathbf{d}}(\omega,t) = [\mathbf{I} + \mathbf{H^r}(\omega)]\mathbf{T_{CH}}\mathbf{H^d}(\omega)[\mathbf{X}(\omega,t) + \boldsymbol{\delta}\mathbf{X}(\omega,t)], \tag{15}$$

which when simplified further using (10) becomes

$$\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)=\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)+\mathbf{H}^{\mathbf{r}}(\omega)\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)\boldsymbol{\delta}\mathbf{X}(\omega,t)+\mathbf{H}^{\mathbf{r}}(\omega)\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)+\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)\boldsymbol{\delta}\mathbf{X}(\omega,t). \tag{16}$$

The second term in the right hand side of (16) now represents the reverberated loudspeaker compensation signals measured at the equalizer microphone array. We neglect the effects of this term in the subsequent steps (i.e., assume that $\mathbf{H}^{\mathbf{r}}(\omega)\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)\boldsymbol{\delta}\mathbf{X}(\omega,t)\rightarrow\mathbf{0}$), under the assumption that its effects can be mitigated by the adaptive control process[1]. Equation (16) can therefore be simplified as

$$\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)\boldsymbol{\delta}\mathbf{X}(\omega,t)=-\mathbf{H}^{\mathbf{r}}(\omega)\boldsymbol{\beta}^{\mathbf{d}}(\omega,t). \tag{17}$$

The reverberation compensation signals at the loudspeaker are then given by

$$\boldsymbol{\delta}\mathbf{X}(\omega,t)=-\left[\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)\right]^{\dagger}\mathbf{H}^{\mathbf{r}}(\omega)\boldsymbol{\beta}^{\mathbf{d}}(\omega,t), \tag{18}$$

where $[\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)]^{\dagger}$ is the Moore-Penrose pseudoinverse of $\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)$. Thus, calculating the loudspeaker compensation signals becomes a matrix multiplication problem related to the estimates of the reverberation transformation matrix $\mathbf{H}^{\mathbf{r}}(\omega)$ and the desired soundfield coefficients $\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)$.

C. Reverberant Channel Estimation

From (18), it is seen that the effective compensation of reverberation requires the knowledge of the diagonal elements of $\mathbf{H}^{\mathbf{r}}(\omega)$. If we assume that $\widehat{\mathbf{H}}^{\mathbf{r}}(\omega)$ represents an estimate of $\mathbf{H}^{\mathbf{r}}(\omega)$, the channel estimation problem can be described as a classical adaptive filtering problem.

Including the effect of the loudspeaker compensation signals in (17), the measured soundfield coefficients in (11) can be expressed as

$$\boldsymbol{\beta}(\omega,t)=\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)+\boldsymbol{\beta}^{\mathbf{r}}(\omega,t)-\widehat{\mathbf{H}}^{\mathbf{r}}(\omega)\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)+\boldsymbol{\beta}^{\mathbf{v}}(\omega,t). \tag{19}$$

The error between the measured and desired soundfield coefficients is given by

$$\boldsymbol{\beta}^{\mathrm{e}}(\omega,t)=\left[\mathbf{H}^{\mathbf{r}}(\omega)-\widehat{\mathbf{H}}^{\mathbf{r}}(\omega)\right]\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)+\boldsymbol{\beta}^{\mathbf{v}}(\omega,t), \tag{20}$$

where

$$\boldsymbol{\beta}^{\mathrm{e}}(\omega,t)=\boldsymbol{\beta}(\omega,t)-\boldsymbol{\beta}^{\mathbf{d}}(\omega,t).$$

Equation (20) can be characterized as an adaptive filtering problem [37], where minimizing the squared error of (20) leads to an estimate of the reverberant channel coefficients in $\widehat{\mathbf{H}}^{\mathbf{r}}(\omega)$. Estimated in an iterative fashion, the coefficients at the time step t_m are given by the adaptation equation

$$\widehat{\mathbf{H}}^{\mathbf{r}}(\omega,t_m)^{H}=\widehat{\mathbf{H}}^{\mathbf{r}}(\omega,t_{m-1})^{H}+\Phi\boldsymbol{\beta}^{\mathbf{d}}(\omega,t_m)\boldsymbol{\beta}^{\mathbf{e}}(\omega,t_m)^{H}. \tag{21}$$

The adaptation gain Φ is determined by the adaptation technique, and can be a constant for the Least Mean Squares (LMS) approach or a variable quantity $E\{\boldsymbol{\beta}^{\mathbf{d}}(\omega,t_m)\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)^{H}\}^{-1}$ for the Recursive Least Squares (RLS) approach.

Since $\widehat{\mathbf{H}}^{\mathbf{r}}(\omega)$ is a diagonal matrix, this knowledge can be used to further reduce the computational complexity of calculating Φ. The diagonal structure of $\widehat{\mathbf{H}}^{\mathbf{r}}(\omega)$ implies that the transformation coefficients $H_n^{\mathrm{r}}(\omega)$ are independent of each other; thus, calculating the individual coefficients becomes a single-tap adaptive filtering problem. Hence, the adaptation equation for a diagonal element of $\widehat{\mathbf{H}}^{\mathbf{r}}(\omega)$ is given by

$$\widehat{H}_n^{\mathrm{r}}(\omega,t_m)^{H}=\widehat{H}_n^{\mathrm{r}}(\omega,t_{m-1})^{H}+\phi_n\beta_n^{\mathrm{d}}(\omega,t_m)\beta_n^{\mathrm{e}}(\omega,t_m)^{H}, \tag{22}$$

where ϕ_n is the adaptation gain of the nth mode. For a single-tap Normalized Least Mean Squares (N-LMS) adaptive filter [37], ϕ_n can be expressed as

$$\phi_n=\frac{\lambda}{a+\beta_n^{\mathrm{d}}(\omega,t_m)\beta_n^{\mathrm{d}}(\omega,t_m)^{H}},$$

where $0<\lambda<2$ is an adaptation constant and a is a positive constant. Although the evaluations in this paper use the N-LMS algorithm to calculate the filter gain due to its low complexity and fast convergence, any adaptive technique may be used to obtain the values of ϕ_n.

IV. Simulation Setup

A. Simulation Parameters

In this paper, we consider the reproduction of a soundfield in a two-dimensional 5 m $\times$ 6.4 m reverberant room with a wall absorption coefficients of 0.36, that corresponds to a carpeted concrete environment. The floor and ceiling are assumed to be non-reflective, and the reverberation due to the four other walls are simulated using the image-source method [38] for an image depth of 5 (i.e., 60 image sources). The 1 m radius circular region of interest $\mathcal{R}$ is centred at (2.4 m, 3.8 m), within a 2 m circular array of 60 loudspeakers simulated by vertical line sources. The region $\mathcal{R}$ is enclosed by an equalizer array of 55 microphones at a radial distance of 1 m from the centre of $\mathcal{R}$. The number of loudspeakers and microphones are selected to satisfy the mode truncation length $N=\lceil ekr_e\rceil+1$ [22] at a maximum frequency of 1 kHz. Fig. 3 illustrates the array configuration in the reverberant room.

In order to evaluate the performance of the room compensation process, we consider the reproduction of two types of narrowband sources at an angular frequency ω (wideband sources are simulated by integrating narrowband sources across frequency); a plane wave source with the desired soundfield coefficients

$$\beta_n^{\mathrm{d}}(\omega,t)=(-i)^n e^{-i(n\phi_y+\omega t)}, \tag{23}$$

[1]It should be noted that this may not always be possible, and as shown in Section V, the algorithm may diverge for low direct-to-reverberant-path power ratios. Although the loudspeaker compensation signals in (18) can be updated as $\boldsymbol{\delta}\mathbf{X}(\omega,t)=-[\mathbf{H}^{\mathbf{r}}(\omega)\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)+\mathbf{T}_{\mathbf{CH}}\mathbf{H}^{\mathbf{d}}(\omega)]^{\dagger}\mathbf{H}^{\mathbf{r}}(\omega)\boldsymbol{\beta}^{\mathbf{d}}(\omega,t)$ to account for this condition and ensure that the algorithm converges, it will introduce additional computational complexity due to the computation of a pseudoinverse in each iteration.

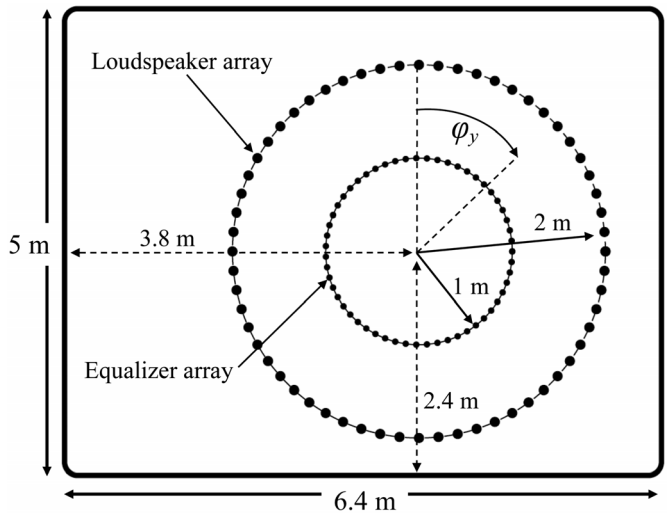


Fig. 3. Loudspeaker and microphone configuration in the reverberant room.

where $\phi_y = \pi/3$, and a monopole source given by

$$\beta_n^{\mathrm{d}}(\omega, t) = \mathcal{H}_n^{(2)}\left(\frac{\omega y}{c}\right) e^{-i(n\phi_y + \omega t)}, \tag{24}$$

where $(y, \phi_y) \equiv (1.7\ \mathrm{m}, \pi/3)$, $c = 343$ m/s is the speed of sound in air and $\mathcal{H}_n^{(2)}(\cdot)$ is the nth order Hankel function of the second kind [36]. Two scenarios of each source is simulated; a narrowband scenario that reproduces a 1 kHz waveform is used to evaluate the room compensation performance at the design frequency, and a broadband case that reproduces a signal of 1 kHz bandwidth (i.e., frequencies in the 100 Hz to 1 kHz range) is used to simulate a real-world soundfield reproduction scenario. We consider both the single and multiple source scenarios, where the reproduction of multiple virtual sound sources is achieved through the superposition of the relevant $\beta_n^{\mathrm{d}}(\omega, t)$ for each desired source location.

The source power is normalized to 0 dB at the centre of the array, and white Gaussian noise is introduced in order to maintain a specific Signal-to-Noise Ratios (SNRs) with respect to this location. We also investigate the impact of varying the direct-to-reverberant-path power ratios (DRRs) at the centre of the region of interest to simulate different room conditions. This is achieved by varying the wall absorption coefficients as required at a fixed SNR of 50 dB. The behavior of the equalization error at the equalizer microphone array is then used to investigate the robustness of the room compensation process for the two narrowband sound source reproduction scenarios described above.

The performance of the proposed technique is compared with the Multi-Point Equalization method [39]; a non-adaptive room compensation approach which requires knowledge of the reverberant channel, and Filtered-X Recursive Least Squares (FxRLS) [23], [25]; an adaptive algorithm which requires an estimate of the reverberant channel. The Normalized Least Mean Squares (N-LMS) is used as the adaptation algorithm of the proposed technique, where $\lambda = 0.1$ and a is chosen as the standard deviation of the noise.

B. Performance Measures

The performance of the proposed technique can be measured in terms of the equalization error at the equalizer microphone array, i.e., $\boldsymbol{\beta}^{\mathrm{e}}(\omega, t)$ defined previously in (20), and the normalized reproduction error within the region. The narrowband and wideband definitions of each measure can be obtained from the soundfield coefficients as follows.

1) Equalization Error at the Equalizer Array:

Narrowband :

$$\mathcal{E}_{\mathrm{NB}}^{\mathrm{E}}(\omega, t) = 10\log_{10}\left|\boldsymbol{\beta}^{\mathrm{e}}(\omega, t)^H \boldsymbol{\beta}^{\mathrm{e}}(\omega, t)\right| \tag{25}$$

Wideband :

$$\mathcal{E}_{\mathrm{WB}}^{\mathrm{E}}(t) = 10\log_{10}\left|\int_{\omega_1}^{\omega_2} \boldsymbol{\beta}^{\mathrm{e}}(\omega, t)^H \boldsymbol{\beta}^{\mathrm{e}}(\omega, t)\mathrm{d}\omega\right| \tag{26}$$

where $\{\omega_1, \omega_2\}$ denotes the range of equalized frequencies.

2) Normalized Region Reproduction Error: Reproduction error in the region $\mathcal{R}$ can be defined using the cumulative difference of (5) and (6) over each location in $\mathcal{R}$. Hence, the normalized region reproduction error $\mathcal{T}$ is given by

Narrowband :

$$\mathcal{T}_{\mathrm{NB}}(\omega, t) = 10\log_{10}\frac{\mathcal{N}(\omega, t)}{\mathcal{D}(\omega, t)} \tag{27}$$

Wideband :

$$\mathcal{T}_{\mathrm{WB}}(t) = 10\log_{10}\frac{\int_{\omega_1}^{\omega_2}\mathcal{N}(\omega, t)\mathrm{d}\omega}{\int_{\omega_1}^{\omega_2}\mathcal{D}(\omega, t)\mathrm{d}\omega} \tag{28}$$

where

$$\mathcal{N}(\omega, t) = \int_{\mathcal{R}} \left|Y(\mathbf{x}; \omega, t) - Y^{\mathrm{d}}(\mathbf{x}; \omega, t)\right|^2 \mathrm{d}a(\mathbf{x})$$

$$\mathcal{D}(\omega, t) = \int_{\mathcal{R}} \left|Y^{\mathrm{d}}(\mathbf{x}; \omega, t)\right|^2 \mathrm{d}a(\mathbf{x}),$$

and $\mathrm{d}a(\mathbf{x}) = x\ dx\ d\phi_x$ is the differential area element of $\mathbf{x}$.

V. Discussion

A. Narrowband Performance

Figs. 4 and 5 illustrate the reproduction of a 1 kHz monopole source 1.7 m from the centre of the region $\mathcal{R}$ from the direction $\pi/3$, and a complex soundfield generated by three 1 kHz plane wave sources incident from the directions 0, $\pi/3$ and $3\pi/4$ (with the amplitudes 1, 1.5 and 0.5, respectively), at a 50 dB SNR at the centre of the equalizer microphone array. The desired, reverberant and compensated soundfields are shown in subfigures (a), (b) and (c) respectively. The location of the equalizer microphone array is denoted by the dotted circle. A good reproduction of the desired soundfield is observed within $\mathcal{R}$, in comparison with the reverberant soundfield in Figs. 4(b) and 5(b).

The equalization error at the equalizer microphone array is shown in Fig. 6, averaged over 10 trial runs after 500 adaptation steps. It is seen that the error begins to converge to the noise floor created by the external uncorrelated noise. The normalized region reproduction error within $\mathcal{R}$, seen in Fig. 7, follows the equalization error curve in Fig. 6. A minimum SNR of

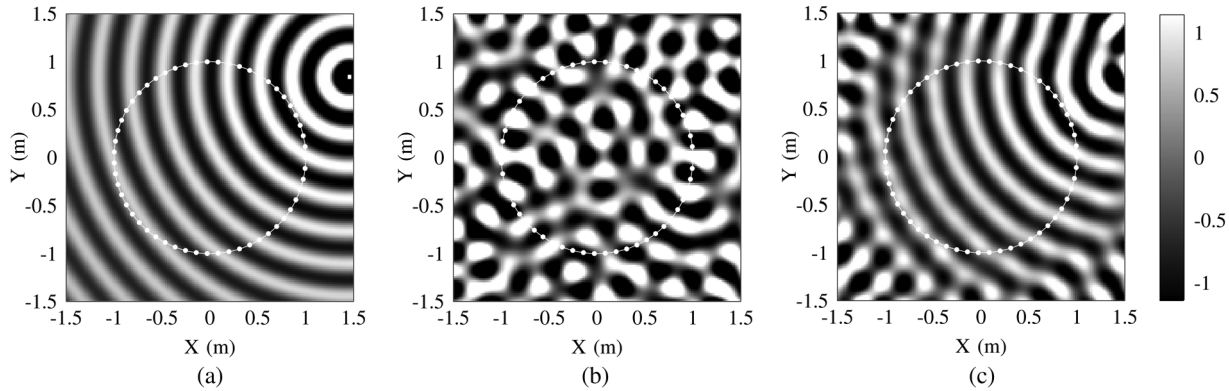


Fig. 4. Reproduction of a 1 kHz monopole source at (1.7 m, $\pi/3$) within a region of 1 m radius at 50 dB SNR. The dotted circle indicates the equalization array which encloses the region of interest $\mathcal{R}$.

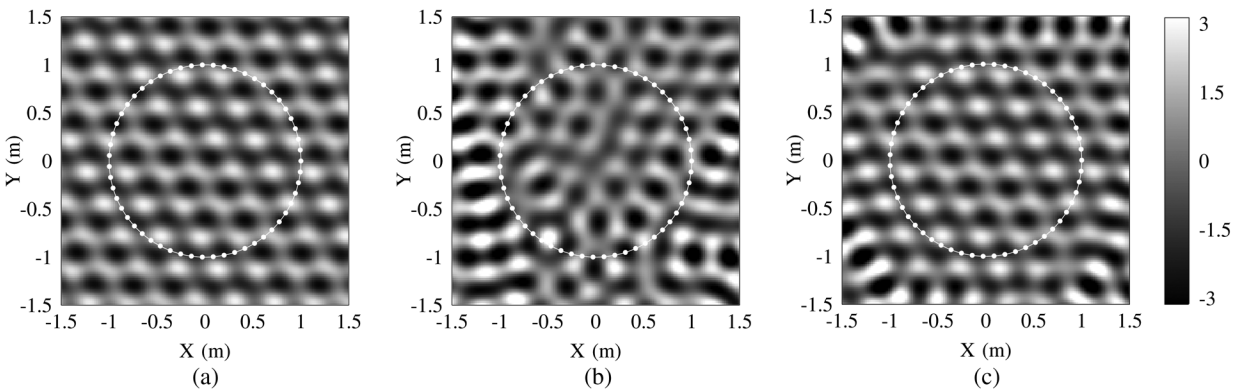


Fig. 5. Reproduction of a complex soundfield generated by three 1 kHz plane wave sources (incident from azimuths 0, $\pi/3$ and $3\pi/4$) within a region of 1 m radius at 50 dB SNR. The dotted circle indicates the equalization array which encloses the region of interest $\mathcal{R}$.

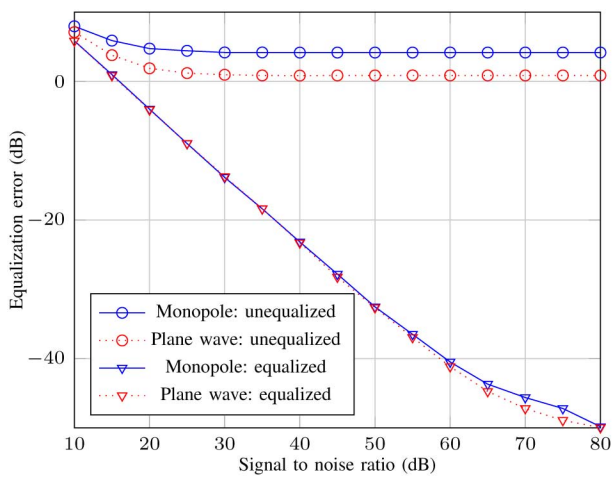


Fig. 6. Equalization error of 1 kHz plane wave (dotted line) and monopole sources (solid line) at the equalization array vs. SNR at the centre of the region of interest. Circular and triangular markers indicate the unequalized and equalized equalization error after 500 adaptation steps, averaged over 10 trial runs.

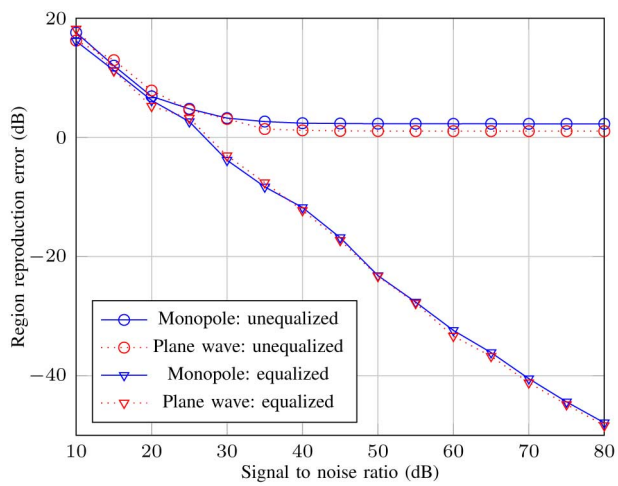


Fig. 7. Normalized region reproduction error of equalized and unequalized 1 kHz plane wave (dotted line) and monopole (solid line) sources within the 1 m radius region of interest. Circular and triangular markers indicate unequalized and equalized normalized region reproduction error, averaged over 10 trial runs.

approximately 20 dB is required for the adaptive room compensator to converge, while a SNR between 40 - 50 dB is necessary to maintain a normalized region reproduction error below 1%. Fig. 8 illustrates the equalization error at the equalizer microphone array with respect to the DRR at the centre of the region of interest at a fixed SNR of 50 dB. A similar behavior is exhibited for both sources, and as expected, the equalization error marginally increases with the decreasing DRRs. This is mainly due to the slower speed of convergence of the adaptation process. However, the room compensator no longer converges at DRRs below -5 dB, as indicated by the rapidly increasing equalization error.

These results indicate that the proposed algorithm provides good narrowband soundfield reproduction for the simulated reverberant rooms. They further suggest that the normalized region reproduction error can be maintained below 10% for a

TABLE I
COMPUTATIONAL COMPLEXITY OF ADAPTIVE ALGORITHMS

	FxRLS	EAF	Proposed
Compensation Filter Adaptation	$\mathcal{O}\left(P^4 N_c^2\right)$	$Q \cdot \mathcal{O}\left(N_c^2\right)$	$\mathcal{O}\left(\sum_{f=0}^{N_c/2} N_f\right) \approx \mathcal{O}\left(QN_c\right)$
Reverberant Channel Identification	$\mathcal{O}\left(P^2 N_c^2\right)$	$Q \cdot \mathcal{O}\left(N_c^2\right)$	
Orthogonal Transformations	-	$\mathcal{O}\left(PQ\right) + \mathcal{O}\left(Q^2\right)$	-

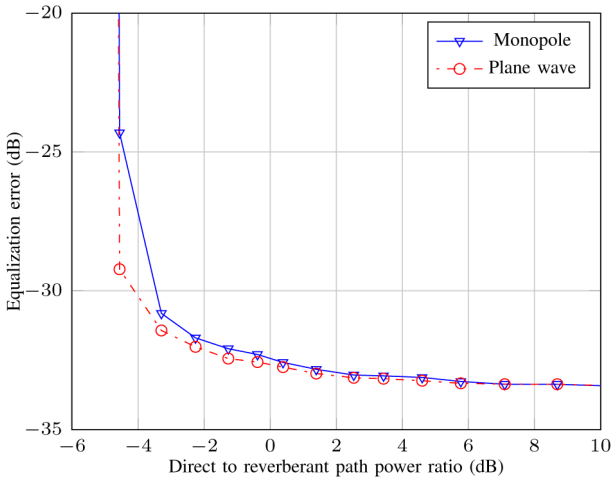


Fig. 8. Equalization error of 1 kHz plane wave (dot dash line, circular markers) and monopole sources (solid line, triangular markers) at the equalization array vs. the DRR at the centre of the region of interest. The illustrated equalization error after 500 adaptation steps is the average over 10 trial runs.

moderate SNR of 30 dB, and that the adaptation process is robust to different room conditions, down to -5 dB DRRs. Similar or better performance can be expected at frequencies below 1 kHz.

B. Wideband Performance

Fig. 9 illustrates the wideband performance of the proposed method using sources of 1 kHz bandwidth at different SNRs. The monopole and plane wave sources are positioned at the same locations as in the previous narrowband scenario. The performance is compared with the FxRLS and Multi-Point equalization methods, both of which are assumed to have perfect knowledge of the reverberant channel.

The behavior of the wideband normalized region reproduction error closely follows the performance of the FxRLS method for both source types as seen in Fig. 9. The performance is comparable to the multi-point equalization method for the plane wave source, although it diverges by up to 10 dB at certain SNRs in the case of the monopole source. Overall, the wideband performance is consistent with the narrowband behavior, achieving a wideband reproduction error within $\mathcal{R}$ of less than 10% for SNRs above 35 dB. The performance of the proposed method approaches that of the non-adaptive multi-point technique and appears to be similar to the more complex adaptive algorithm, FxRLS. The performance of the proposed method may be adversely affected at some frequencies due to the zeros in the Bessel function in (9). However, as stated previously, a dual equalizer array arrangement could be implemented [16], and may improve the soundfield reproduction performance by avoiding the amplification of errors at modes with low SNR.

C. Computational Complexity and Convergence Behavior

The reduction of computational complexity and good convergence behavior are desirable properties that are difficult to achieve in massive multi-channel soundfield reproduction systems. Since the computational complexity is related to the number of unknown coefficients to be calculated by the adaptive filters, it can be expressed as a function of the number of unknown elements in these matrices. Thus, the computational complexity of the different adaptive algorithms can be summarized using the big-$\mathcal{O}$ notation as shown in Table I. The computational complexity of FxRLS and EAF are obtained from the derivations in [23] and [25]. It should be noted that the computational complexity of the EAF approximation using WDAF has since been improved upon by Schneider *et al.* [31], and varies with the number of off-diagonal coupled modes considered. The complexity of each algorithm has been categorized under three main operations; operations required to compute the loudspeaker compensation signals, operations required to identify the reverberant channel and operations required to calculate relevant data dependent orthogonal transformations. Since the proposed algorithm does not consist of separate reverberant channel estimation and compensation filter adaptation operations, the computational complexity is listed as part of the channel identification operation.

The computational complexity of the proposed algorithm can be derived as follows. In order for the complexity of the different algorithms to be comparable, we consider a similar number of time-domain filter taps for each method. Hence, let N_c represent the number of filter coefficients used to model the reverberant channel, and P, Q be the number of loudspeakers and microphones required to reproduce the soundfield up to a design frequency F_0. The number of unknown channel coefficients at a frequency F corresponds to the number of diagonal elements of $\mathbf{H}^{\mathbf{r}}(2\pi F)$, and is given by $N_F = \lceil e 2\pi F r_e / c \rceil + 1$ [16]. Thus, for wideband channel estimation at the frequency bins fF_s/N_c, for $f = 0 \ldots N_c/2$ at a sampling frequency F_s, the total number of unknown coefficients is given by

$$\sum_{f=0}^{N_c/2} N_f = \sum_{f=0}^{N_c/2} \left\lceil er_e \left(\frac{2\pi f F_s}{c N_c} \right) \right\rceil + 1. \tag{29}$$

The number of equalizer microphones required can be derived similarly, and is given by

$$Q = \left\lceil er_e \left(\frac{2\pi F_0}{c} \right) \right\rceil + 1. \tag{30}$$

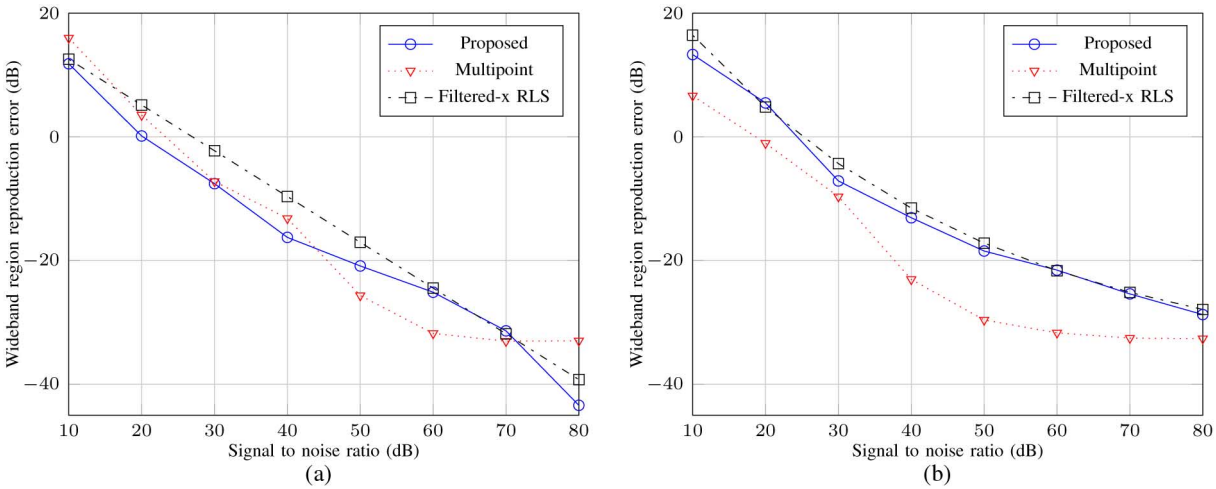


Fig. 9. Normalized region reproduction error of (a) plane wave and (b) monopole sources of 1 kHz bandwidth within the 1 m radius region of interest. The proposed technique (solid line) is compared with the multi-point equalization (dotted line) and Filtered-x RLS (dot dash line) techniques after 500 adaptation steps, averaged over 10 trial runs.

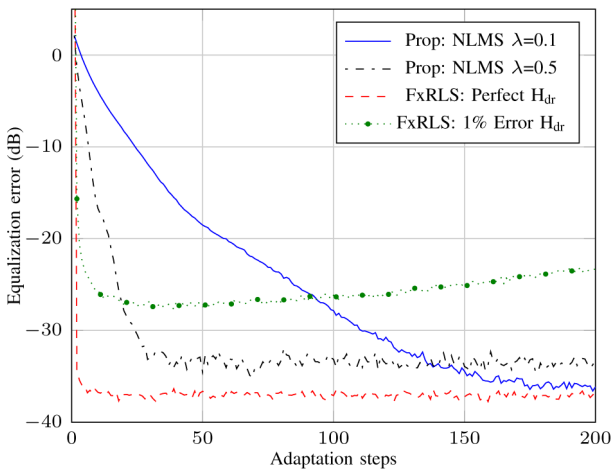


Fig. 10. Filter adaptation performance of the proposed and FxRLS techniques at the equalization array averaged over 10 trials for a 500 Hz monopole source at 50 dB SNR. The proposed technique is evaluated for an adaptation constant of 0.1 and 0.5. FxRLS is evaluated for a forgetting factor of 0.95 with perfect knowledge of the forward channel, as well as a forward channel estimate with a 1% error.

If $F_0 = F_s/2$, substituting (30) in (29), the computational complexity can be obtained as

$$\sum_{f=0}^{N_c/2} N_f \approx \frac{1}{4}\{QN_c + Q + N_c + 2\} \approx \mathcal{O}(QN_c), \qquad (31)$$

and implies a linear increase of the computational complexity with respect to the reverberation time.

The adaptive filter convergence behavior of the proposed technique using N-LMS and the performance of the FxRLS approach is illustrated in Fig. 10, for a monopole source at 1 kHz and a DRR of 10 dB. The FxRLS algorithm can provide very fast convergence with perfect knowledge of the room response, but imperfect channel information can affect the numerical stability of the algorithm as illustrated. Thus, the use of FxRLS for large multi-channel systems requires longer memory, which results in slower adaptation to changing channel conditions. The proposed method using N-LMS performs similar to the FxRLS algorithm with perfect knowledge of the room responses, albeit at a reduced convergence speed. Faster convergence can be achieved with a larger convergence factor, but this is achieved at the expense of increased region reproduction error and reduced numerical stability. Decreasing DRRs will marginally increase the convergence time of the filters, but the overall impact is negligible for mid to high DRRs.

VI. Conclusion

In this paper, we propose a multi-channel adaptive room compensation technique for reproduction of a desired soundfield within a region of a two-dimensional reverberant room. We describe how the measured soundfield can be expressed as a collection of modes. The reverberation effects are then modelled as transformations of the individual desired soundfield modes, which provides a convenient mechanism to diagonalize the reverberant channel estimation process and reduce the computational complexity. Next, we derive the required reverberation compensation signals at the loudspeakers, and describe how an adaptive algorithm such as N-LMS can be used to estimate the diagonal transformation matrix that describes the effects of reverberation. This enabled the combination of the channel estimation and compensation filter adaptation processes into a single simple adaptive filtering operation. Simulation results suggest that the proposed technique can produce good region reproduction performance for narrowband and wideband sources at moderate SNRs. In addition, it is also observed that the proposed room compensator is robust to changing room conditions, down to DRRs of -5 dB in the region of interest. A reduction in the computational complexity can also be achieved for a similar reproduction accuracy in comparison with the existing adaptive and non-adaptive listening room compensation techniques.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions that helped to improve the clarity and quality of this manuscript.

REFERENCES

[1] D. S. Talagala, W. Zhang, and T. D. Abhayapala, "Active acoustic echo cancellation in spatial soundfield reproduction," in *Proc. IEEE Int. Conf. Acoust,, Speech, Signal Process. (ICASSP '13)*, Vancouver, BC, Canada, May 2013, pp. 620–624.

[2] M. A. Gerzon, "Periphony: With-height sound reproduction," *J. Audio Eng. Soc.*, vol. 21, pp. 2–10, Feb. 1973.

[3] J. Daniel, S. Moreau, and R. Nicol, "Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging," in *Proc. Audio Eng. Society 114th Conv.*, Amsterdam, The Netherlands, Mar. 2003.

[4] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 6, pp. 697–707, Sep. 2001.

[5] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, Nov. 2005.

[6] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 1, pp. 107–116, Jan. 2009.

[7] A. Gupta and T. D. Abhayapala, "Three-dimensional sound field reproduction using multiple circular loudspeaker arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 5, pp. 1149–1159, Jul. 2011.

[8] A. J. Berkhout, "A holographic approach to acoustic control," *J. Audio Eng. Soc.*, vol. 36, no. 12, pp. 977–995, Dec. 1988.

[9] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Amer.*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[10] 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, Feb. 2003.

[11] S. M. Kuo and D. R. Morgan, "Active noise control: A tutorial review," *Proc. IEEE*, vol. 87, no. 6, pp. 943–973, Jun. 1999.

[12] S. J. Elliott, *Signal Processing for Active Control*. New York, NY, USA: Academic, 2000.

[13] M. Poletti, F. M. Fazi, and P. A. Nelson, "Sound-field reproduction systems using fixed-directivity loudspeakers," *J. Acoust. Soc. Amer.*, vol. 127, no. 6, pp. 3590–3601, Jun. 2010.

[14] M. A. Poletti and T. D. Abhayapala, "Interior and exterior sound field control using general two-dimensional first-order sources," *J. Acoust. Soc. Amer.*, vol. 129, no. 1, pp. 234–244, Jan. 2011.

[15] T. Betlehem, C. Anderson, and M. A. Poletti, "A directional loudspeaker array for surround sound in reverberant rooms," in *Proc. Int. Conf. Acoust. (ICA '10)*, Sydney, Australia, Aug. 2010, pp. 1–6.

[16] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Amer.*, vol. 117, no. 4, pp. 2100–2111, Apr. 2005.

[17] M. Kolundzija, C. Faller, and M. Vetterli, "Reproducing sound fields using MIMO acoustic channel inversion," *J. Audio Eng. Soc.*, vol. 59, no. 10, pp. 721–734, Oct. 2011.

[18] J. Mourjopoulos, "On the variation and invertibility of room impulse response functions," *J. Sound Vibr.*, vol. 102, no. 2, pp. 217–228, Sep. 1985.

[19] B. D. Radlovic, R. C. Williamson, and R. A. Kennedy, "Equalization in an acoustic reverberant environment: Robustness results," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 3, pp. 311–319, May 2000.

[20] L.-J. Brännmark, "Robust audio precompensation with probabilistic modeling of transfer function variability," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA '09)*, New Paltz, NY, USA, Oct. 2009, pp. 193–196.

[21] L.-J. Brännmark, A. Bahne, and A. Ahlen, "Compensation of loudspeaker-room responses in a robust MIMO control framework," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 6, pp. 1201–1216, Jun. 2013.

[22] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2542–2556, Jun. 2007.

[23] M. Bouchard and S. Quednau, "Multichannel recursive-least-square algorithms and fast-transversal-filter algorithms for active noise control and sound reproduction systems," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 606–618, Sep. 2000.

[24] S. Spors, H. Buchner, and R. Rabenstein, "Eigenspace adaptive filtering for efficient pre-equalization of acoustic MIMO systems," in *Proc. Eur. Signal Process. Conf. (EUSIPCO '06)*, Florence, Italy, Sep. 2006.

[25] S. Spors, H. Buchner, R. Rabenstein, and W. Herbordt, "Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering," *J. Acoust. Soc. Amer.*, vol. 122, no. 1, pp. 354–369, Jul. 2007.

[26] H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering: Acoustic echo cancellation for full-duplex systems based on wave-field synthesis," in *Proc. IEEE Int. Conf. Acoust,, Speech, Signal Process. (ICASSP '04)*, May 2004, vol. 4, pp. 117–120.

[27] H. Buchner and S. Spors, "A general derivation of wave-domain adaptive filtering and application to acoustic echo cancellation," in *Proc. 42nd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Oct. 2008, pp. 816–823.

[28] S. Spors and H. Buchner, "An approach to massive multichannel broadband feedforward active noise control using wave-domain adaptive filtering," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA '07)*, New Paltz, NY, USA, Oct. 2007, pp. 171–174.

[29] M. Schneider and W. Kellermann, "Adaptive listening room equalization using a scalable filtering structure in the wave domain," in *Proc. IEEE Int. Conf. Acoust,, Speech, Signal Process. (ICASSP '12)*, Kyoto, Japan, Mar. 2012, pp. 13–16.

[30] M. Schneider and W. Kellermann, "A direct derivation of transforms for wave-domain adaptive filtering based on circular harmonics," in *Proc. Eur. Signal Process. Conf. (EUSIPCO '12)*, Bucharest, Romania, Aug. 2012, pp. 1034–1038.

[31] M. Schneider and W. Kellermann, "A wave-domain model for acoustic MIMO systems with reduced complexity," in *Proc. Joint Workshop Hands-Free Speech Commun. Microphone Arrays (HSCMA '11)*, Edinburgh, U.K., Jun. 2011, pp. 133–138.

[32] M. Schneider and W. Kellermann, "Iterative DFT-domain inverse filter determination for adaptive listening room equalization," in *Proc. Int. Workshop Acoust. Signal Enhance. (IWAENC '12)*, Aachen, Germany, Sep. 2012, pp. 1–4.

[33] P.-A. Gauthier and A. Berry, "Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction: Theory," *J. Acoust. Soc. Amer.*, vol. 119, no. 5, pp. 2721–2737, May 2006.

[34] P.-A. Gauthier and A. Berry, "Adaptive wave field synthesis for sound field reproduction: Theory, experiments, and future perspectives," *J. Audio Eng. Soc.*, vol. 55, no. 12, pp. 1107–1124, Dec. 2007.

[35] P.-A. Gauthier and A. Berry, "Adaptive wave field synthesis for active sound field reproduction: Experimental results," *J. Acoust. Soc. Amer.*, vol. 123, no. 4, pp. 1991–2002, Apr. 2008.

[36] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. New York, NY, USA: Academic, 1999.

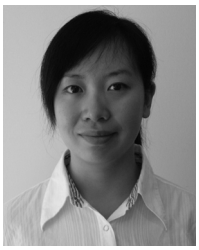
[37] S. O. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.

[38] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.

[39] S. J. Elliott and P. A. Nelson, "Multiple-point equalization in a room using adaptive digital filters," *J. Audio Eng. Soc.*, vol. 37, no. 11, pp. 899–907, Nov. 1989.

Dumidu S. Talagala (S'11–M'14) received the B.Sc. Eng (Hons) in electronic and telecommunication engineering from the University of Moratuwa, Sri Lanka, in 2007. From 2007 to 2009, he was an Engineer at Dialog Axiata PLC, Sri Lanka. He completed his Ph.D. degree within the Applied Signal Processing Group, College of Engineering and Computer Science, at the Australian National University, Canberra, in 2013.

He is currently a research fellow in the Centre for Vision, Speech and Signal Processing at the University of Surrey, United Kingdom. His research interests are in the areas of sound source localization, spatial soundfield reproduction, active noise control, array signal processing and convex optimization.



Wen Zhang (S'06–M'09) received the B.E. degree in telecommunication engineering from Xidian University, Xi'an, China, in 2003 and the M.E. degree in electrical engineering (with first class honors) and the Ph.D. degree from the Australian National University, Canberra, in 2005 and 2010, respectively.

From 2010 to 2012, she was an OCE Postdoctoral Fellow at CSIRO Process Science and Engineering, Sydney, Australia. She is currently a research fellow in the Research School of Engineering at the Australian National University. Her primary research interests are in the field of spatial sound-field recording and reconstruction, source separation and localization, and active noise cancellation. She is currently an editorial board member for *Science Journal of Circuits, Systems and Signal Processing*.

Thushara D. Abhayapala (M'00–SM'08) received the B.E. degree (with honors) in interdisciplinary systems engineering and the Ph.D. degree in telecommunications engineering, from the Australian National University (ANU), Canberra, in 1994 and 1999, respectively. From 1995 to 1997, he was a Research Engineer with the Arthur C. Clarke Center for Modern Technologies, Sri Lanka. Since December 1999, he has been a faculty member at ANU.

He was the Leader of the Wireless Signal Processing (WSP) Program at the National ICT Australia (NICTA) from November 2005 to June 2007. Currently, he is the Director of the Research School of Engineering at the Australian National University.

His research interests are in the areas of spatial audio and acoustic signal processing, space-time signal processing for wireless communication systems, and array signal processing. Professor Abhayapala has supervised 28 research students and coauthored over 180 peer-reviewed papers. He is an Associate Editor for the *EURASIP Journal on Wireless Communications and Networking*. Professor Abhayapala is a Member of the Audio and Acoustic Signal Processing Technical Committee (2011-2013) of the IEEE Signal Processing Society.