Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering

Sascha Spors[a)] and Herbert Buchner
Deutsche Telekom Laboratories, Berlin University of Technology, Ernst-Reuter-Platz 7, 10587 Berlin, Germany

Rudolf Rabenstein
Multimedia Communications and Signal Processing, University Erlangen-Nuremberg, Cauerstrasse 7, 91058 Erlangen, Germany

Wolfgang Herbordt
Center of Competence Signal Processing, Rohde & Schwarz GmbH & Co. KG, 81614 Munich, Germany

(Received 11 October 2006; revised 16 April 2007; accepted 16 April 2007)

The acoustic theory for multichannel sound reproduction systems usually assumes free-field conditions for the listening environment. However, their performance in real-world listening environments may be impaired by reflections at the walls. This impairment can be reduced by suitable compensation measures. For systems with many channels, active compensation is an option, since the compensating waves can be created by the reproduction loudspeakers. Due to the time-varying nature of room acoustics, the compensation signals have to be determined by an adaptive system. The problems associated with the successful operation of multichannel adaptive systems are addressed in this contribution. First, a method for decoupling the adaptation problem is introduced. It is based on a generalized singular value decomposition and is called *eigenspace adaptive filtering*. Unfortunately, it cannot be implemented in its pure form, since the continuous adaptation of the generalized singular value decomposition matrices to the variable room acoustics is numerically very demanding. However, a combination of this mathematical technique with the physical description of wave propagation yields a realizable multichannel adaptation method with good decoupling properties. It is called *wave domain adaptive filtering* and is discussed here in the context of wave field synthesis. [DOI: 10.1121/1.2737669]

PACS number(s): 43.60.Dh, 43.60.Pt, 43.60.Tj, 43.60.Ac [EJS] Pages: 354–369

I. INTRODUCTION

The ultimate goal of sound reproduction is to create the perfect acoustic illusion. Several generations of engineers have invented various sound reproduction systems in the past decades.[1–5] The perfect acoustic illusion has never been realized by either of them. Nevertheless, sound reproduction has improved considerably in terms of quality and spatial impression. However, there are still a number of open problems. One of these, common to most reproduction systems, is the influence of the room where the reproduction takes place, the so-called *listening room*. This contribution introduces a highly efficient approach to the active compensation of the listening room for reproduction systems with a high number of reproduction channels. The problems that arise for systems with high channel numbers pose fundamental restrictions on the realizability of active listening room compensation using standard algorithms. In order to distinguish these systems and their problems from conventional multichannel reproduction systems with low channel numbers (<10) these are referred to as massive multichannel sound reproduction systems in the context of this article. The following will introduce such massive multichannel reproduction systems, will illustrate the influence of the listening room on their performance, and will briefly review existing listening room compensation systems.

A. Massive multichannel sound reproduction systems

Spatial audio reproduction systems using loudspeakers can be roughly classified into so-called stereophony based and sound field reconstruction based. Stereophony-based systems operate typically with five or seven spatial channels.[6] The loudspeaker driving signals for each channel are obtained by carefully processing the source material (usually in a sound engineering studio) and then stored or transmitted as separate tracks, typically one for each loudspeaker, in a standardized way. Consequently, the loudspeaker setup at the user's site has to conform to the corresponding standard for optimal spatial reproduction. The techniques for processing the source material into tracks for each loudspeaker channel are borrowed from conventional two-channel stereo recording and mixing. They rely mainly on amplitude panning relative to the target position of the listener. An optimal spatial reproduction is only achieved in the vicinity of this target position, the so-called *sweet spot*.

[a)]Electronic mail: sascha.spors@telekom.de

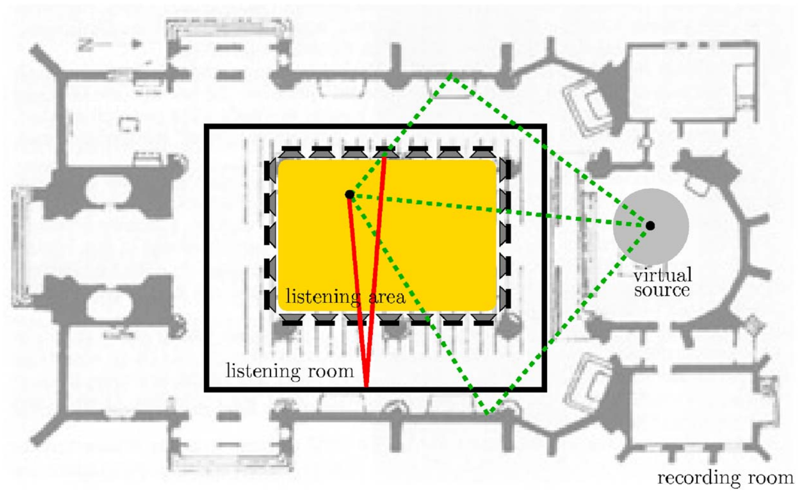

FIG. 1. (Color online) Simplified example that shows the effect of the listening room on the auralized wave field. The dashed lines from one virtual source to one exemplary listening position show the acoustic rays for the direct sound and one reflection off the sidewall of the virtual recording room. The solid line from one loudspeaker to the listening position shows a reflection of a loudspeaker wave field off the wall of the listening room.

In contrast, sound field reconstruction based systems avoid this spatial constraint. They are characterized by a high number of reproduction channels, typically some ten to several hundred. For these numbers it is neither possible nor desirable to store or transmit all the loudspeaker signals for playback. Since the loudspeaker setup often depends on the listening room architecture, each system will potentially have an individual speaker configuration. Therefore, the reproduction system cannot be built according to a certain standard for the number and placement of loudspeakers. In this situation, the driving signals for the loudspeakers have to be computed in real time by the reproduction system. The input for this processing step are the signals of the sound sources, their spatial position, and room acoustic information of the scene to be reproduced. This acoustic information is either collected during the recording of the source signals (e.g., in a concert hall) or derived from models of synthetic acoustic spaces.

Two well-known approaches for generating multichannel loudspeaker driving signals are ambisonics and wave field synthesis. Ambisonics systems represent the sound field in an enclosure by an expansion into three-dimensional basis functions. They typically use a speaker arrangement which resembles a sphere around the listening area. The contribution of the sound sources to each basis function may be recorded with special microphone arrangements.[7,8] Common realizations make use of only the contributions up to first order.[9] Wave field synthesis (WFS) is based on a physical description of wave propagation illustrated by Huygen's principle. The mathematical foundation is given by the Kirchhoff-Helmholtz integral. Certain simplifications and approximations lead to the description of a realizable multichannel sound reproduction system.[10–14]

Common to both these approaches for massive multichannel sound reproduction is that they rely on free-field acoustic wave propagation and do not consider the influence of reflections within the listening room. Since these reflections may impair the carefully designed spatial sound field, their influence should be minimized by taking appropriate countermeasures. This contribution presents an approach to addressing this issue. The following sections describe the influence of the listening room, discuss compensation methods, and give an outline of this article.

B. The influence of the listening room

The influence of the listening room on the sound scene reproduced by a multichannel reproduction system will be illustrated first in an intuitive fashion. For this purpose the simple reproduction scenario illustrated by Fig. 1 is considered in the following. The projection of an acoustic scene in a church (e.g., a singer performing in the choir) into the listening room is shown as an example. For simplicity the propagation of sound waves is illustrated by acoustic rays in Fig. 1. The dashed lines in Fig. 1 from the virtual source to one exemplary listening position show the acoustic rays for the direct sound and several reflections off the sidewalls of the church. The loudspeaker system in the listening room reproduces the direct sound and the reflections in order to create the desired spatial impression. The theory behind nearly all sound reproduction methods assumes an anechoic listening room which does not exhibit any reflections of its own. The solid line in Fig. 1 from one loudspeaker in the upper row to the listening position illustrates a possible reflection of the wave field produced by this loudspeaker off the wall of the listening room. This additional reflection caused by the listening room may impair the desired spatial impression.

The influence of the listening room on the performance of sound reproduction systems is a topic of active research.[15–26] Since the acoustic properties of the listening room and the reproduction system used may vary in a wide range, no generic conclusion can be given for the perceptual influence of the listening room. However, it is generally agreed that the reflections imposed by the listening room will have influence on the perceived properties of the reproduced scene. These influences may be, e.g., degradation of directional localization performance or sound coloration.

Summarizing, a reverberant listening room will superimpose its characteristics on the desired impression of the recorded room. Listening room compensation aims at eliminating or reducing the effect of the listening room. The following will classify and briefly introduce existing listening room compensation approaches.

C. Listening room compensation systems

Three basic classes of techniques can be identified:

1. Passive listening room compensation,
2. Consideration of the influence of the room in the rendering algorithm, and
3. Active listening room compensation.

Passive listening room compensation applies acoustic insulation materials to the listening room as countermeasure against its reflections. However, it is well known that acoustic insulation gets impractical and costly above an even rather modest level of sound absorption, especially for low frequencies. Thus, in practical setups passive room compensation alone cannot provide a sufficient suppression of listening room reflections.

The second class of approaches takes the influence of the room on the auralized field into account in calculating the loudspeaker driving signals. Almost every virtual scene will include at least some acoustic reflections, and the basic idea in this class of approaches is to modify the rendering of the scene according to the additional reflections produced by the listening room. However, no solutions or algorithms that successfully demonstrate the applicability of this idea are known to the authors at the time this article was written. A very rough sketch of some ideas can be found in Ref. 27.

The third class of approaches uses concepts from active control to perform the desired compensation. For sound reproduction, synergies with the reproduction system are used in order to control the wave field. Ideally, the desired control over the undesired reflections can be applied within the entire listening area by destructive interference. A concept shared by most active room compensation approaches, due to the causal nature of listening room reflections, is to prefilter the loudspeaker driving signals using suitable compensation filters. These filters are computed by analyzing the reproduced wave field. Most active compensation systems published in the past utilize only some few (<10) loudspeakers and analysis (microphone) positions.[28–31] As a consequence they are not able to provide sufficient control over the wave field or are they able to sufficiently analyze the reproduced wave field throughout the entire listening area. As a result, the influence of the listening room will be compensated mainly at the analyzed positions with the potential occurrence of severe artifacts at other positions.[32–35] These approaches are therefore termed *multipoint compensation* approaches.

Advanced reproduction systems like wave field synthesis and higher-order ambisonics provide an improvement in terms of control over the reproduced wave field up to a certain frequency (spatial aliasing frequency). They allow one to compensate for the reflections of the listening room throughout the entire listening area. Above the spatial aliasing frequency they may be supported easily by passive listening room compensation techniques. However, they require a large number of reproduction channels. Additionally an adequate analysis of the reproduced wave field will require a large number of analysis channels as well. Various active listening room compensation systems for such massive multichannel reproduction systems have recently been introduced.[36–46] The proposed approaches can roughly be classified into three classes depending on the algorithms used to compute the compensation filters: (1) noniterative, (2) iterative, (3) and adaptive room compensation techniques.

Most of the currently proposed listening room compensation schemes fall into the first two classes. They assume that the acoustical environment is time-invariant and has been characterized at some fixed time by impulse response measurements. The task of computing the compensation filters is typically formulated as a matrix inversion problem which is solved with noniterative or iterative methods. In general, the listening room characteristics may change over time. For instance as a result of a temperature variation in the listening room the speed of sound and hence the acoustic properties will change.[47,48] This calls for an adaptive computation of the compensation filters on the basis of an analysis of the reproduced wave field. Therefore we will focus on the derivation of an adaptive algorithm within this article.

A wide variety of problems are related to the algorithms used for the adaptation of the compensation filters. The most severe problem for a scenario with many playback and analysis channels is that the adaptation of the compensation filters is subject to fundamental problems.[49] Thus it seems that massive multichannel active room compensation may solve some of the problems of the multipoint room compensation systems like position dependence and inversion problems but only at the expense of new problems related to the multichannel adaptation.

This article will derive, discuss, and evaluate a novel efficient approach to adaptive active listening room compensation for spatial audio systems that yields an extended compensated area compared to the multipoint approaches. It also summarizes and extends our earlier work in Refs. 40–43. This approach is based on a technique that we denote as *wave domain adaptive filtering* (WDAF). In the following a short overview of this article is given.

D. Overview of this article

Section II discusses the concept of active listening room compensation in more detail and introduces the required notation. Section III presents a general concept for the adaptation of the compensation filters and lists its known problems. A framework for the solution to these problems is given in Sec. IV. It is based on a transformation into the so-called *eigenspace*, which allows a decoupling into a number of independent adaption problems. Unfortunately, this transformation turns out to be data dependent, which renders it unsuitable for practical computing. An escape is shown in Sec. V by considering the physical meaning of this transformation. Based on acoustic intuition, a new inverse filtering approach based on WDAF is presented. It is not optimal in the sense of totally decoupling the problem, but it is suitable for practical application. Finally, its application to a widely used

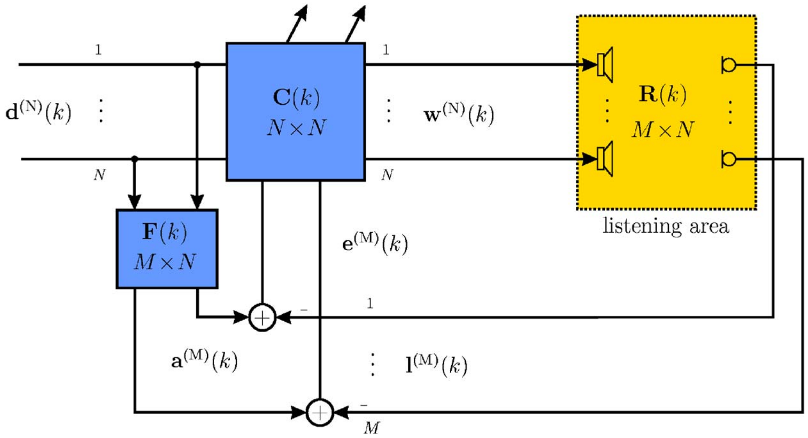

FIG. 2. (Color online) Block diagram illustrating the adaptive MIMO inverse filtering approach to room compensation.

spatial rendering technique (wave field synthesis) is shown in Sec. VI.

II. ACTIVE LISTENING ROOM COMPENSATION

The following presents the concept of active room compensation by prefiltering of the loudspeaker signals in more detail. The need for adaptation of the prefilters arises in all practical listening room situations. Moderate variations of the room temperature or the movement of a single person may change the acoustic properties of a listening room to such an extent that a previously working compensation based on fixed compensation filters is no longer valid.[48] First, the basic block diagram of an adaptive system for active room compensation system is presented and then the required notation is introduced.

A. Adaptive system for multichannel active room compensation

The block diagram of an adaptive system for multichannel active room compensation is shown in Fig. 2. The vector of input signals

$$\mathbf{d}^{(N)}(k) = [d_1(k)\ d_2(k)\ \cdots\ d_N(k)]^T \tag{1}$$

represents the N loudspeaker driving signals $d_n(k)$, $n = 1, \ldots, N$ provided by a multichannel spatial rendering system which is based on the assumption of a reflection-free listening room. Throughout this article the discrete time index is denoted by k. The matrices of impulse responses $\mathbf{C}(k)$, $\mathbf{R}(k)$, and $\mathbf{F}(k)$ represent the compensation filters, the listening room impulse responses, and the desired listening room impulses responses, respectively. In order to eliminate the effect of the listening room, free-field conditions are desired for reproduction. Hence, the matrix $\mathbf{F}(k)$ is composed from the free-field impulse responses from each loudspeaker to each microphone. This matrix is derived analytically from the solution of the wave equation for the free-field case. After the compensation filters there are N prefiltered driving signals for the loudspeakers, that are combined into a column vector $\mathbf{w}^{(N)}(k)$. A total of M microphones are used for the analysis of the resulting wave field. Their signals are combined in the vector $\mathbf{l}^{(M)}(k)$, representing spatial samples of the desired wave field on which the undesired listening room reflections are superimposed. The $N \times 1$ column vector $\mathbf{w}^{(N)}(k)$ and the $M \times 1$ vector $\mathbf{l}^{(M)}(k)$ are defined in an analogous way as given by Eq. (1). The output vector $\mathbf{a}^{(M)}(k)$ results from filtering the deriving signals $\mathbf{d}^{(N)}(k)$ with the idealized $M \times N$ matrix $\mathbf{F}(k)$ of free-field impulse responses.

The error $\mathbf{e}^{(M)}(k)$ between free-field propagation $\mathbf{a}^{(M)}(k)$ and the actual microphone signals $\mathbf{l}^{(M)}(k)$ describes the deviation of the rendered wave field from the reflection-free case. It is used to adapt the matrix of compensation filters $\mathbf{C}(k)$. After convergence of $\mathbf{C}(k)$, the response of the prefiltered reproduction system in the listening room produces the desired free-field sound field.

The matrices of impulse responses $\mathbf{R}(k)$, $\mathbf{F}(k)$, and $\mathbf{C}(k)$ describe discrete linear multiple-input/multiple-output (MIMO) systems. Assuming point source like propagation for the loudspeakers, the room $\mathbf{R}(k)$ and free-field $\mathbf{F}(k)$ matrices of impulse responses can be related to the so-called Green's function $G(\mathbf{x}|\mathbf{x}_0, \omega)$. This function characterizes the solutions of the wave equation with respect to boundary conditions. The Green's function can be regarded as the acoustic transfer function from a spatial excitation point $\mathbf{x}_0$ to a measurement point $\mathbf{x}$. Thus, its counterpart in the temporal domain $g(\mathbf{x}|\mathbf{x}_0, t) = \mathcal{F}_t^{-1}\{G(\mathbf{x}|\mathbf{x}_0, \omega)\}$ can be interpreted as the corresponding continuous-time room impulse response between the positions $\mathbf{x}$ and $\mathbf{x}_0$. This impulse response may be of infinite length, but for practical purposes it is truncated at a reasonable time (e.g., when its energy has decayed below a suitably chosen threshold) resulting in a finite impulse response (FIR). The discretized impulse response between the nth loudspeaker position and the mth analysis (microphone) position is defined as

$$r_{m,n}(k) := g(\mathbf{x}_m|\mathbf{x}_n, kT_s), \tag{2}$$

with the temporal sampling interval T_s. The $M \times N$ matrix $\mathbf{R}(k)$ captures all these impulse responses, and thus all sampling points of the Green's function $G(\mathbf{x}|\mathbf{x}_0, \omega)$ in the temporal domain. If the respective impulse responses in $\mathbf{R}(k)$ are actually measured, then these will also contain the nonideal loudspeaker and microphone characteristics (e.g., directivity and frequency response) as well as the influence of the employed hardware.[50]

Analogous definitions, as given earlier for $\mathbf{R}(k)$, apply also to the $M \times N$ matrix $\mathbf{F}(k)$ of free-field impulse responses. The matrix $\mathbf{C}(k)$ contains the sequences $c_{n,n'}(k)$ of compensation filters which have to be determined by the adaptation algorithm introduced in Sec. III. Since the respective impulse responses are finite, they will be referred to as MIMO FIR systems in the following.

B. Frequency-domain representation of signals and systems

For the temporal frequency-domain description of the signals and systems used in this article, the discrete time Fourier transform (DTFT)[51] is used. The DTFT transform and its inverse for the spatially discrete loudspeaker driving signal $d_n(k)$ are defined as follows:

$$D_n(\omega) = \sum_{k=-\infty}^{\infty} d_n(k) e^{-j\omega k T_s}, \tag{3a}$$

$$d_n(k) = \frac{T_s}{2\pi} \int_0^{2\pi/T_s} D_n(\omega) e^{j\omega k T_s} d\omega. \tag{3b}$$

Analogous definitions apply to the other signals $w_n(k)$, $l_m(k)$, $a_m(k)$, and $e_m(k)$. The vector of loudspeaker driving signals $\mathbf{d}^{(N)}(k)$ is transformed into the temporal frequency domain by transforming each element $d_n(k)$ separately using the DTFT (3a). The resulting vector is denoted by $\underline{\mathbf{d}}^{(N)}(\omega)$. Vectors and matrices of frequency-domain signals are underlined in the following. Analogous definitions as for $\mathbf{d}^{(N)}(k)$ apply to the other signals used.

The matrix of room impulse responses $\mathbf{R}(k)$ is transformed into the temporal frequency domain in the same way as $\mathbf{d}^{(N)}(k)$. The resulting matrix in the frequency domain is referred to as *room transfer matrix* and denoted by $\underline{\mathbf{R}}(\omega)$. The matrices $\mathbf{F}(k)$ and $\mathbf{C}(k)$ are transformed analogously. The transfer matrix $\underline{\mathbf{F}}(\omega)$ is referred to as *free-field transfer matrix.*

III. ADAPTATION OF THE ROOM COMPENSATION FILTERS

Room compensation can be understood as an inverse MIMO FIR filtering problem. For only few synthesis and analysis channels numerous solutions to this problem have been developed in the past.[32–35,50] However, algorithms for massive multichannel systems still remain a challenge. The following discusses the problems that arise when standard adaptive filtering algorithms are used for the computation of the matrix of compensation filters $\mathbf{C}(k)$. To this end, it is shown how to formulate the adaptation problem such that a powerful state-of-the-art least-squares algorithm can be applied. The presentation is just detailed enough to highlight the problems associated with the application of conventional adaptation approaches to active listening room compensation. It is not intended to provide all the details necessary for an implementation. Instead the discussion of the problems will lead to an alternative approach which is presented in Secs. IV and V.

A. Least-squares error adaptation of the compensation filter

The derivation of the normal equation for the adaptive pre-equalization problem introduced in Sec. II A is briefly reviewed in the following section. A detailed discussion for acoustic MIMO systems can be found, e.g., in Refs. 43 and 50. The solution of the normal equation with the filtered-x recursive least-squares algorithm (X-RLS) is additionally shown.

The adaptation of the compensation filters $\mathbf{C}(k)$ is driven by the vector of error signals $\mathbf{e}^{(M)}(k)$. Each component $e_m(k)$ is determined by the components of the signal vector $\mathbf{d}^{(N)}$ $\times(k)$ and the respective impulse response matrices as

$$e_m(k) = \sum_{n'} f_{m,n'}(k) d_{n'}(k) - \sum_{n} \sum_{n'} r_{m,n}(k) \hat{c}_{n,n'}(k) d_{n'}(k), \tag{4}$$

where the error $e_m(k)$ depends on the estimates $\hat{c}_{n,n'}(k)$ of the ideal compensation filters $c_{n,n'}(k)$. With this error signal, the following cost function is defined:

$$\xi(\hat{\mathbf{c}},k) = \sum_{\kappa=0}^{k} \lambda^{k-\kappa} \sum_{m=1}^{M} |e_m(\kappa)|^2, \tag{5}$$

where $0 < \lambda \leqslant 1$ denotes an exponential weighting factor. The cost function can be interpreted as the time-averaged energy of the error $e_m(k)$ between the desired wave field $a_m(k)$ and the actual reproduced wave field $l_m(k)$ averaged over all M analysis positions. The optimal filter coefficients in the mean-squared error (MSE) sense are found by setting the gradient of the cost function $\xi(\hat{\mathbf{c}},k)$ to zero with respect to the estimated filter coefficients $\hat{\mathbf{c}}(k)$. The normal equation is derived by expressing the error $\mathbf{e}^{(M)}(k)$ in terms of the filter coefficients, introducing the result into the cost function (5) and calculating its gradient. To express the normal equation as a linear system of equations, the estimated compensation filters are arranged in vector form,

$$\hat{\mathbf{c}}_{n,n'}(k) = [\hat{c}_{n,n'}(0), \hat{c}_{n,n'}(1), \ldots, \hat{c}_{n,n'}(N_c - 1)]^T, \tag{6a}$$

$$\hat{\mathbf{c}}_n^{(N)}(k) = [\hat{\mathbf{c}}_{n,1}^T(k) \hat{\mathbf{c}}_{n,2}^T(k) \cdots \hat{\mathbf{c}}_{n,N}^T(k)]^T, \tag{6b}$$

$$\hat{\mathbf{c}}(k) = [\hat{\mathbf{c}}_1^{(N)}(k)^T \hat{\mathbf{c}}_2^{(N)}(k)^T \cdots \hat{\mathbf{c}}_N^{(N)}(k)^T]^T, \tag{6c}$$

where N_c denotes the number of filter coefficients. Typically this number has to be chosen higher than the number of coefficients N_r of the room impulse responses $\mathbf{r}_{m,n}(k)$ since we are dealing with an inverse identification problem.[52] The resulting normal equation of the inverse filtering problem is then given as

$$\hat{\mathbf{\Phi}}_{dd}(k) \hat{\mathbf{c}}(k) = \hat{\mathbf{\Phi}}_{da}(k). \tag{7}$$

The $N^2 N_c \times N_c N^2$ matrix $\hat{\mathbf{\Phi}}_{dd}(k)$ is the time and analysis position-averaged autocorrelation matrix of the filtered driving signals

$$\hat{\mathbf{\Phi}}_{dd}(k) = \sum_{\kappa=0}^{k} \lambda^{k-\kappa} \mathbf{D}_R(\kappa) \mathbf{D}_R^T(\kappa)$$

$$= \lambda \hat{\mathbf{\Phi}}_{dd}(k-1) + \mathbf{D}_R(k) \mathbf{D}_R^T(k), \tag{8}$$

where $\mathbf{D}_R(\kappa)$ denotes the matrix of filtered driving signals. The matrix of filtered driving signals is given as follows:

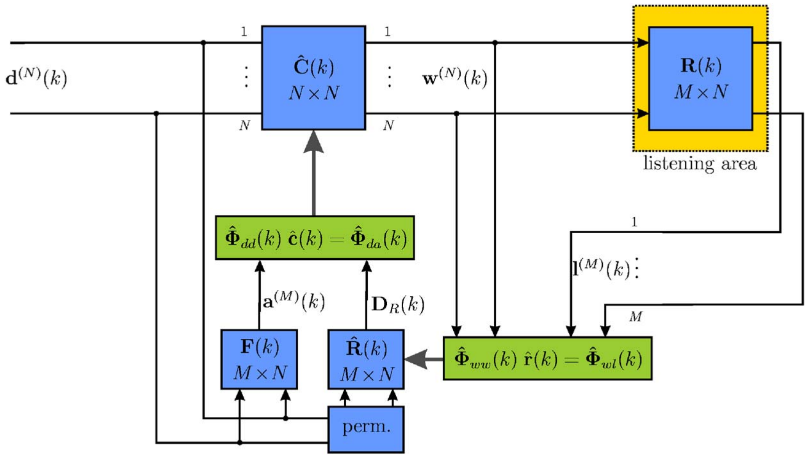

FIG. 3. (Color online) Block diagram illustrating the application of the X-RLS algorithm to active listening room compensation.

$$d_{m,n,n'}(k) = d_{n'}(k) r_{m,n}(k), \tag{9a}$$

$$\mathbf{d}_{m,n,n'}(k) = [d_{m,n,n'}(k) d_{m,n,n'}(k-1) \cdots d_{m,n,n'}(k - N_c + 1)]^T, \tag{9b}$$

$$\mathbf{d}^{(N)}_{m,n}(k) = [\mathbf{d}^T_{m,n,1}(k)\mathbf{d}^T_{m,n,2}(k) \cdots \mathbf{d}^T_{m,n,N}(k)]^T, \tag{9c}$$

$$\mathbf{d}^{(N,N)}_{M}(k) = [\mathbf{d}^{(N)}_{m,1}(k)^T \mathbf{d}^{(N)}_{m,2}(k)^T \cdots \mathbf{d}^{(N)}_{m,N}(k)^T]^T, \tag{9d}$$

$$\mathbf{D}_R(k) = [\mathbf{d}^{(N,N)}_1(k)\mathbf{d}^{(N,N)}_2(k) \cdots \mathbf{d}^{(N,N)}_M(k)]. \tag{9e}$$

Note that $\mathbf{D}_R(k)$ is composed from all N^2 combinations resulting from filtering the driving signals $d_{n'}(k)$ with all inputs of the MIMO room responses $r_{m,n}(k)$ for all possible combination of n and n'. This results from rearranging Eq. (4) in order to isolate the filter coefficients $\hat{\mathbf{c}}(k)$.

The $N^2 N_c \times 1$ vector $\hat{\mathbf{\Phi}}_{da}(k)$ can be interpreted as the time and analysis position-averaged cross-correlation vector between the filtered driving signals and the desired signals which is defined as

$$\hat{\mathbf{\Phi}}_{da}(k) = \sum_{\kappa=0}^{k} \lambda^{k-\kappa} \mathbf{D}_R(\kappa)\mathbf{a}^{(M)}(\kappa) = \lambda \hat{\mathbf{\Phi}}_{da}(k-1) + \mathbf{D}_R(k)\mathbf{a}^{(M)}(k). \tag{10}$$

The optimal pre-equalization filter with respect to the cost function (5) is given by solving the normal equation (7) with respect to the filter coefficients $\hat{\mathbf{c}}(k)$. The normal equation is typically not solved in a direct manner but by recursive updates of the filter coefficients. A recursive update equation for the filter coefficients can be derived from the normal equation (7), the recursive definitions of the correlation matrices given by Eqs. (8) and (10), and the definition of the error signal $\mathbf{e}^{(M)}(k)$ as follows:

$$\hat{\mathbf{c}}(k) = \hat{\mathbf{c}}(k-1) - \hat{\mathbf{\Phi}}^{-1}_{dd}(k)\mathbf{D}_R(k)\mathbf{e}^{(M)}(k). \tag{11}$$

Equation (11) together with the recursive definition of the correlation matrix $\hat{\mathbf{\Phi}}_{dd}(k)$ given by Eq. (8) constitutes the basis of the X-RLS algorithm. The inverse $\hat{\mathbf{\Phi}}^{-1}_{dd}(k)$ of the autocorrelation matrix is typically computed in a recursive fashion by applying the matrix inversion lemma.[53]

The X-RLS algorithm deviates from the standard RLS algorithm by using a filtered version of the driving signal for adaptation. The calculation of the filtered driving signals requires knowledge of the room response $\mathbf{R}(k)$, which is in general not known *a priori* and will be time variant. Hence, the room characteristics have to be identified additionally using a multichannel RLS algorithm. Its normal equation can be derived in the same manner as shown earlier for the compensation filters. It is given by

$$\hat{\mathbf{\Phi}}_{ww}(k)\hat{\mathbf{r}}(k) = \hat{\mathbf{\Phi}}_{wl}(k), \tag{12}$$

where $\hat{\mathbf{\Phi}}_{ww}(k)$ denotes the autocorrelation matrix of the filtered loudspeaker driving signals $\mathbf{w}^{(N)}(k)$, $\hat{\mathbf{\Phi}}_{wl}(k)$ the cross-correlation matrix between the filtered loudspeaker driving signals $\mathbf{w}^{(N)}(k)$ and the analysis signals $\mathbf{l}^{(M)}(k)$, and $\hat{\mathbf{r}}(k)$ the estimated coefficients of the room transfer matrix. Figure 3 illustrates the X-RLS algorithm applied to the active listening room compensation scenario.

B. Fundamental problems of adaptive inverse filtering

Three fundamental problems of massive multichannel adaptive pre-equalization can be identified from the normal equation (7) and the definition (8) of the autocorrelation matrix $\hat{\mathbf{\Phi}}_{dd}(k)$. These are:

1. Nonuniqueness of the solution,
2. Ill-conditioning of the autocorrelation matrix $\hat{\mathbf{\Phi}}_{dd}(k)$, and
3. Computational complexity for massive MIMO systems.

The first problem is related to minimization of the cost function $\xi(\hat{\mathbf{c}},k)$. Minimization of the cost function $\xi(\hat{\mathbf{c}},k)$ may not provide the optimal solution in terms of identifying the inverse system to the room transfer matrix. Depending on the driving signals $\mathbf{d}^{(N)}(k)$ there may be multiple solutions for $\hat{\mathbf{c}}(k)$ that minimize $\xi(\hat{\mathbf{c}},k)$.[54] This problem is often referred to as *nonuniqueness problem*.

The second and the third fundamental problems are related to the solution of the normal equation (7). The normal equation has to be solved with respect to the coefficients $\hat{\mathbf{c}}(k)$ of the room compensation filter. However, the autocorrelation matrix $\hat{\mathbf{\Phi}}_{dd}(k)$ is typically ill-conditioned for the considered multichannel reproduction scenarios.[49] The filtered driv-

TABLE I. Complexity of adaptive listening room compensation using a multichannel filtered-x RLS (X-RLS) algorithm in comparison to the proposed approach of eigenspace adaptive filtering (EAF) in conjunction with single-channel X-RLS algorithms. For the complexity of the transformations $\underline{\mathbf{X}}^H(\omega)$ and $\underline{\mathbf{V}}(\omega)$ only their application has been considered but not their computation by the GSVD.

		X-RLS	EAF
Adaption of $\hat{\underline{\mathbf{C}}}(\omega)$	Dimension of $\hat{\mathbf{\Phi}}_{dd}(k)$	$N^2N_c \times N_cN^2$	$N_c \times N_c$
	Complexity	$O(N^4N_c^2)$	$M \cdot O(N_c^2)$
Identification of $\hat{\underline{\mathbf{R}}}(\omega)$	Dimension of $\hat{\mathbf{\Phi}}_{ww}(k)$	$NN_c \times N_cN$	$N_c \times N_c$
	Complexity	$O(N^2N_c^2)$	$M \cdot O(N_c^2)$
Transformations $\underline{\mathbf{X}}^H(\omega)$ and $\underline{\mathbf{V}}(\omega)$	Complexity		$O(NM)+O(M^2)$

ing signals $\mathbf{D}_R(k)$ will contain cross-channel (spatial) correlations due to the deterministic nature of most auralization algorithms. Also temporal correlations may be present for typical virtual source signals. Besides the ill-conditioning also the dimensionality of the autocorrelation matrix $\hat{\mathbf{\Phi}}_{dd}(k)$ poses problems. Massive multichannel systems exhibit a high number of reproduction channels and additionally the length N_c of the inverse filter has to be chosen quite long for a suitable suppression of reflections.[52] As a consequence, the derivation of the compensation filters will get computationally very demanding. Table I illustrates the complexity of the X-RLS algorithm. The complexity is given on the basis of a straightforward solution of the normal equation without any optimizations. $O(N)$ denotes "in the order of N" as measure of complexity. A more detailed complexity analysis of the X-RLS algorithm can be found, e.g., in Ref. 55.

The same problems as discussed earlier apply to the identification of the room transfer matrix using a multichannel RLS algorithm.[54]

IV. EIGENSPACE ADAPTIVE FILTERING

In this section we will derive a generic framework for pre-equalization which explicitly solves the problems of complexity and cross-channel correlations by utilizing signal and system transformations. It will be shown additionally that the other problems mentioned earlier are highly alleviated by the proposed approach. The basic idea is to perform a joint decoupling of the MIMO systems $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$ in the temporal frequency domain. This decoupling yields a decoupling of the MIMO adaptation problem and the autocorrelation matrix $\hat{\mathbf{\Phi}}_{dd}(k)$ as will also be shown.

A. Nonadaptive computation of room compensation filters

In order to gain more insight into the solution of the pre-equalization problem the nonadaptive case will be regarded first. For this purpose a frequency-domain description of the pre-equalization problem depicted in Fig. 2 is used.

The error $\underline{\mathbf{e}}^{(M)}(\omega)$ between the desired $\underline{\mathbf{a}}^{(M)}(\omega)$ and the actual $\underline{\mathbf{l}}^{(M)}(\omega)$ signal at the M analysis positions in the frequency domain can be derived by a DTFT of Eq. (4) as

$$\underline{\mathbf{e}}^{(M)}(\omega) = \underline{\mathbf{a}}^{(M)}(\omega) - \mathbf{l}^{(M)}(\omega) = \underline{\mathbf{F}}(\omega)\underline{\mathbf{d}}^{(N)}(\omega) - \underline{\mathbf{R}}(\omega)\underline{\mathbf{C}}(\omega)\underline{\mathbf{d}}^{(N)}(\omega). \tag{13}$$

Optimal pre-equalization is obtained by minimizing this error: $\underline{\mathbf{e}}^{(M)}(\omega) \to \mathbf{0}$. The least-squares solution to Eq. (13) in this sense, with respect to the compensation filters, is given by[53,56]

$$\underline{\mathbf{C}}(\omega) = \underline{\mathbf{R}}^{+}(\omega)\underline{\mathbf{F}}(\omega), \tag{14}$$

where $\underline{\mathbf{R}}^{+}(\omega)$ denotes the Moore-Penrose pseudoinverse of $\underline{\mathbf{R}}(\omega)$.

B. Generalized singular value decomposition

It will be assumed in the following that $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$ have the dimensions $M \times N$ with $N \geqslant M$. Hence, a reproduction system with equal or less analysis positions than loudspeakers is assumed. This restriction is meaningful with respect to the solution of the pre-equalization problem, since in the general case an exact reproduction at M positions can only be gained for such scenarios.[57] However, the derived results can be generalized straightforwardly to arbitrary channel numbers M and N.

The singular value decomposition (SVD) states that any matrix can be decomposed into two unitary matrices and a diagonal matrix.[53,56] The concept of the SVD can be generalized to the diagonalization of a pair of matrices. This decomposition is known as generalized singular value decomposition (GSVD).[56] The GSVD of the matrices $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$ is given as follows:

$$\underline{\mathbf{R}}(\omega) = \underline{\mathbf{X}}(\omega)\underline{\tilde{\mathbf{R}}}(\omega)\underline{\mathbf{V}}^H(\omega), \tag{15a}$$

$$\underline{\mathbf{F}}(\omega) = \underline{\mathbf{X}}(\omega)\underline{\tilde{\mathbf{F}}}(\omega)\underline{\mathbf{U}}^H(\omega). \tag{15b}$$

The matrices $\underline{\mathbf{X}}(\omega)$, $\underline{\mathbf{V}}(\omega)$, and $\underline{\mathbf{U}}(\omega)$ are unitary matrices with the dimensions $M \times M$, $N \times M$, and $N \times M$, respectively. The matrix $\underline{\mathbf{X}}(\omega)$ is the generalized singular matrix of $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$, the matrices $\underline{\mathbf{V}}(\omega)$ and $\underline{\mathbf{U}}(\omega)$ the respective right singular matrices of $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$. The matrices $\underline{\tilde{\mathbf{R}}}(\omega)$ and $\underline{\tilde{\mathbf{F}}}(\omega)$ are diagonal matrices constructed from the singular values of $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$. The diagonal matrix $\underline{\tilde{\mathbf{R}}}(\omega)$ is defined as

$$\underline{\tilde{\mathbf{R}}}(\omega) = \operatorname{diag}\{[\tilde{R}_1(\omega), \tilde{R}_2(\omega), \ldots, \tilde{R}_M(\omega)]\}, \tag{16}$$

where $\tilde{R}_1(\omega) \geqslant \tilde{R}_2(\omega) \geqslant \cdots \geqslant \tilde{R}_B(\omega) > 0$ denote the B nonzero singular values $\tilde{R}_m(\omega)$ of $\underline{\mathbf{R}}(\omega)$. Their total number B is given by the rank of the matrix $\underline{\mathbf{R}}(\omega)$ with $1 \leqslant B \leqslant M$. For $B < M$ the remaining singular values $\tilde{R}_{B+1}(\omega), \tilde{R}_{B+2}(\omega), \ldots, \tilde{R}_M(\omega)$ are zero. Similar definitions as given earlier for $\underline{\tilde{\mathbf{R}}}(\omega)$ apply to the matrix $\underline{\tilde{\mathbf{F}}}(\omega)$.

The relation given by Eq. (15a) can be inverted by exploiting the unitary property of the joint and right singular matrices. This results in

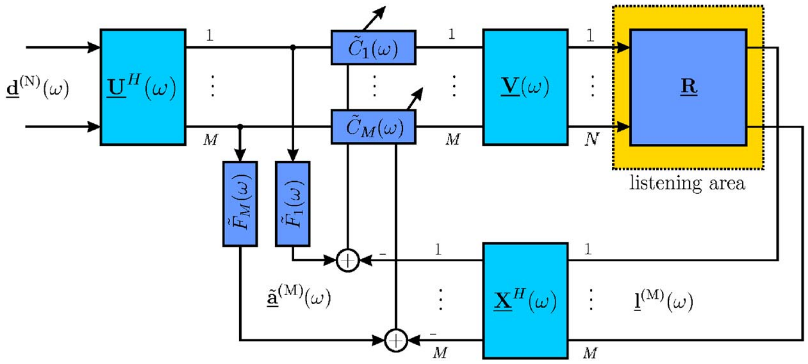

FIG. 4. (Color online) Block diagram illustrating the eigenspace adaptive inverse filtering approach to room compensation.

$$\tilde{\underline{\mathbf{R}}}(\omega) = \underline{\mathbf{X}}^H(\omega)\underline{\mathbf{R}}(\omega)\mathbf{V}(\omega). \tag{17}$$

Hence each matrix $\underline{\mathbf{R}}(\omega)$ can be transformed into a diagonal matrix $\tilde{\underline{\mathbf{R}}}(\omega)$ using the joint and right singular matrix $\underline{\mathbf{X}}(\omega)$ and $\underline{\mathbf{V}}(\omega)$. A similar relation as given by Eq. (17) can be derived straightforwardly for $\tilde{\underline{\mathbf{F}}}(\omega)$. The GSVD transforms the matrices $\mathbf{R}(\omega)$ and $\mathbf{F}(\omega)$ into their joint eigenspace using the singular matrices $\underline{\mathbf{X}}(\omega)$, $\underline{\mathbf{V}}(\omega)$, and $\underline{\mathbf{U}}(\omega)$. In general, these singular matrices depend on the matrices $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$. The GSVD is a *data-dependent transformation*. The transformation of $\underline{\mathbf{R}}(\omega)$ into its diagonal representation $\tilde{\underline{\mathbf{R}}}(\omega)$, as given by Eq. (17), can be interpreted as pre- and postfiltering the room transfer matrix $\underline{\mathbf{R}}(\omega)$ by the MIMO systems $\underline{\mathbf{V}}(\omega)$ and $\underline{\mathbf{X}}^H(\omega)$.

The SVD can be used to define the pseudoinverse $\underline{\mathbf{R}}^+(\omega)$ of the matrix $\underline{\mathbf{R}}(\omega)$,[53]

$$\underline{\mathbf{R}}^+(\omega) = \underline{\mathbf{V}}(\omega)\tilde{\underline{\mathbf{R}}}^{-1}(\omega)\underline{\mathbf{X}}^H(\omega). \tag{18}$$

Equation (15b) and Eq. (18) can be combined to derive the following result:

$$\underline{\mathbf{R}}^+(\omega)\underline{\mathbf{F}}(\omega) = \underline{\mathbf{V}}(\omega)\tilde{\underline{\mathbf{R}}}^{-1}(\omega)\tilde{\underline{\mathbf{F}}}(\omega)\mathbf{U}^H(\omega), \tag{19}$$

where it is assumed that $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$ both have full rank. Equation (19) will be utilized in the following to derive a decoupling of the room compensation filters.

C. Decoupling of the MIMO adaptation problem

The decompositions of the transfer matrices $\underline{\mathbf{F}}(\omega)$ and $\underline{\mathbf{R}}(\omega)$ are given by Eq. (15). It remains to choose a suitable decomposition of the compensation filter matrix $\underline{\mathbf{C}}(\omega)$. The nonadaptive solution for the pre-equalization filter is given by Eq. (14) in terms of the pseudoinverse $\underline{\mathbf{R}}^+(\omega)$ of the room transfer matrix. Hence, an eigenspace expansion of $\underline{\mathbf{C}}(\omega)$ is given by Eq. (19). However, the system transfer matrix $\underline{\mathbf{R}}(\omega)$ is not known in general and has to be identified additionally. An expansion of the pre-equalization filter can be given by utilizing Eq. (19) but with unknown expansion coefficients $\tilde{\underline{\mathbf{C}}}(\omega)$,

$$\underline{\mathbf{C}}(\omega) = \underline{\mathbf{V}}(\omega)\tilde{\underline{\mathbf{C}}}(\omega)\underline{\mathbf{U}}^H(\omega), \tag{20}$$

where $\tilde{\underline{\mathbf{C}}}(\omega)$ denotes a diagonal matrix, where some diagonal elements may be zero. Using Eq. (20) together with Eq. (15a) yields the transformed signal $\tilde{\underline{\mathbf{l}}}^{(M)}(\omega)$ at the analysis points

$$\tilde{\underline{\mathbf{l}}}^{(M)}(\omega) = \tilde{\underline{\mathbf{R}}}(\omega)\tilde{\underline{\mathbf{C}}}(\omega)\tilde{\underline{\mathbf{d}}}^{(M)}(\omega), \tag{21}$$

where $\tilde{\underline{\mathbf{l}}}^{(M)}(\omega) = \underline{\mathbf{X}}^H(\omega)\underline{\mathbf{l}}^{(M)}(\omega)$ and $\tilde{\underline{\mathbf{d}}}^{(M)}(\omega) = \underline{\mathbf{U}}^H(\omega)\underline{\mathbf{d}}^{(N)}(\omega)$. Decomposition of the desired system response according to Eq. (15b) yields the desired signal in the transformed domain as

$$\tilde{\underline{\mathbf{a}}}^{(M)}(\omega) = \tilde{\underline{\mathbf{F}}}(\omega)\tilde{\underline{\mathbf{d}}}^{(M)}(\omega), \tag{22}$$

where $\tilde{\underline{\mathbf{a}}}^{(M)}(\omega) = \underline{\mathbf{X}}^H(\omega)\underline{\mathbf{a}}^{(M)}(\omega)$. Equation (21) together with Eq. (22) allows one to express the error $\underline{\mathbf{e}}^{(M)}(\omega)$ in the transformed domain

$$\tilde{\underline{\mathbf{e}}}^{(M)}(\omega) = \tilde{\underline{\mathbf{a}}}^{(M)}(\omega) - \tilde{\underline{\mathbf{l}}}^{(M)}(\omega) = \tilde{\underline{\mathbf{F}}}(\omega)\tilde{\underline{\mathbf{d}}}^{(M)}(\omega) - \tilde{\underline{\mathbf{R}}}(\omega)\tilde{\underline{\mathbf{C}}}(\omega)\tilde{\underline{\mathbf{d}}}^{(M)}(\omega), \tag{23}$$

where $\tilde{\underline{\mathbf{e}}}^{(M)}(\omega)$ denotes the error signal for all M components in the transformed domain. Since $\tilde{\underline{\mathbf{R}}}(\omega)$, $\tilde{\underline{\mathbf{C}}}(\omega)$, and $\tilde{\underline{\mathbf{F}}}(\omega)$ are diagonal matrices, the mth component of the error signal $\tilde{E}_m(\omega)$ in the transformed domain is given by the following relation:

$$\tilde{E}_m(\omega) = \tilde{F}_m(\omega)\tilde{D}_m(\omega) - \tilde{R}_m(\omega)\tilde{C}_m(\omega)\tilde{D}_m(\omega), \tag{24}$$

where $\tilde{R}_m(\omega)$, $\tilde{C}_m(\omega)$, and $\tilde{F}_m(\omega)$ denote the mth component of the main diagonal of $\tilde{\underline{\mathbf{R}}}(\omega)$, $\tilde{\underline{\mathbf{C}}}(\omega)$, and $\tilde{\underline{\mathbf{F}}}(\omega)$, respectively. The error $\tilde{E}_m(\omega)$ depends only on the mth component of the respective signals and systems in the transformed domain. Thus, Eq. (24) states that the MIMO adaptive inverse filtering problem can be decomposed into M single-input/single-output (SISO) adaptive inverse filtering problems using the GSVD. The computation of the pre-equalization filters can be performed independently for each of the M transformed components. The transformation of the systems and signals is performed by transforming them into the joint eigenspace of $\underline{\mathbf{R}}(\omega)$ and $\underline{\mathbf{F}}(\omega)$ using the GSVD. Therefore this approach will be referred to as *eigenspace inverse adaptive filtering*. Please note that the transformation is not dependent on the driving signals. Figure 4 illustrates the eigenspace inverse adaptive filtering approach.

D. Eigenspace adaptive filtering

In the following, the normal equation of the multichannel adaptive pre-equalization problem presented in Sec. III A will be specialized to the decoupled MIMO system. Due to the decoupling, the cost function $\xi(\hat{\mathbf{c}},k)$ given by Eq. (5) can be minimized independently for each component $m = 1, \ldots, M$. The normal equation in the transformed domain is then given as

$$\hat{\tilde{\mathbf{\Phi}}}_{dd,m}(k)\hat{\tilde{\mathbf{c}}}_m(k) = \hat{\tilde{\mathbf{\Phi}}}_{da,m}(k), \tag{25}$$

where $\hat{\tilde{\mathbf{\Phi}}}_{dd,m}(k)$ denotes the time-averaged autocorrelation matrix of the mth component of the transformed filtered loudspeaker driving signal, $\hat{\tilde{\mathbf{\Phi}}}_{da,m}(k)$ the corresponding cross-correlation matrix between the filtered loudspeaker driving signal and the desired signal, and $\hat{\tilde{\mathbf{c}}}_m(k)$ the filter coefficients. The autocorrelation matrix $\hat{\tilde{\mathbf{\Phi}}}_{dd,m}(k)$ has the dimensions $N_c \times N_c$. Due to this reduction in dimensionality, the solution of the M equations given by Eq. (25) is much more efficient than for the adaptation using the original (not transformed) signals. Table I summarizes the complexity reduction of the eigenspace approach to MIMO inverse adaptive filtering in contrast to the X-RLS algorithm.

Equation (25) corresponds to the well-known single channel normal equation.[53] The cross-channel correlations present in $\hat{\mathbf{\Phi}}_{dd}(k)$ have been removed in the transformed domain by the spatial decoupling of the MIMO systems. Thus, the nonuniqueness and ill-conditioning problem discussed in Sec. III B are highly alleviated. There may still be time-domain correlations present in the filtered input signals which cause problems when solving the normal equation (25). However, there are numerous approaches known in the literature on single-channel adaptive (inverse) filtering to overcome these problems.[53]

In practice, however, a number of problems emerge from the GSVD-based transformations used to decouple the MIMO systems. These will be discussed in the following.

E. Fundamental problems of eigenspace adaptive filtering

The major problems of eigenspace adaptive filtering are that

1. The singular matrices $\underline{\mathbf{X}}(\omega)$, $\underline{\mathbf{V}}(\omega)$, and $\underline{\mathbf{U}}(\omega)$ depend on the room $\underline{\mathbf{R}}(\omega)$ and free-field transfer matrix $\underline{\mathbf{F}}(\omega)$, and
2. Their computation using the GSVD is quite complex.

The contents of the room transfer matrix will be time variant in general due to changes in the acoustic conditions caused by, e.g., temperature changes or persons entering the listening room. This requires that the room transfer matrix $\underline{\mathbf{R}}(\omega)$ has to be identified additionally and that the singular matrices have to be updated each time the room transfer matrix changes. Both tasks are computationally very demanding for massive multichannel systems. In general no optimizations can be performed in practice without placing restrictions on the structure of the transfer matrices $\underline{\mathbf{F}}(\omega)$ and $\underline{\mathbf{R}}(\omega)$. Section V will introduce the concept of wave-domain adaptive filtering in order to overcome these problems.

V. WAVE-DOMAIN ADAPTIVE FILTERING

In Sec. IV an eigenspace approach to adaptive inverse filtering was proposed. Its main feature was the decoupling of the MIMO adaptive inverse filtering problem into a series of single channel adaptive inverse filtering problems. This way most of the fundamental problems (see Sec. III B) of adapting the room compensation filters for massive multichannel systems were solved. However, the computation of suitable transformations using the GSVD may become very complex. In order to overcome this problem the concept of wave-domain adaptive filtering is introduced.

A. Generic concept of wave-domain adaptive filtering

Wave-domain adaptive filtering is based on two basic ideas:

1. Explicit consideration of the characteristics of the propagation medium for the derivation of suitable transformations, and
2. Approximation of the concept of perfect decoupling of the MIMO adaptation problem.

The listening room transfer matrix describes the sound transmission from the loudspeakers to the analysis positions with respect to the characteristics of the propagation medium and the boundary conditions imposed by the listening room. Hence it has to fulfill the wave equation and the homogeneous boundary conditions imposed by the room. This knowledge can be used to construct efficient transformations. Since these transformations inherently have to account for the wave nature of sound in order to perform well, this approach will be referred to as *wave domain adaptive (inverse) filtering* (WDAF) and the transformed domain as *wave domain* in the following.

The second idea is to approximate the perfect decoupling of the MIMO adaptation problem in favor of generic transformations which are to some degree independent of the listening room characteristics. In order to keep the complexity low, these generic transformations need not to be strictly diagonalizing, but they should still represent the MIMO system with as few paths as possible.

The combination of both ideas allows one to derive fixed transformations that provide nearly the same favorable properties as the optimal GSVD-based transformations used for eigenspace adaptive inverse filtering with the benefit of computational efficiency.

Based on the approach of eigenspace adaptive filtering and the above-mentioned considerations a generic block diagram of the WDAF approach can be developed. Figure 5 displays this generic block diagram. The signal and system transformations are performed by three generic transformations. Their structure is not limited to the MIMO FIR systems derived from the GSVD. Transformation $\mathcal{T}_1$ transforms the driving signals into the wave domain, $\mathcal{T}_2$ inversely transforms the filtered loudspeaker driving signals from the wave domain, and $\mathcal{T}_3$ transforms the signals at the analysis points into the wave domain. As previously for the eigenspace domain, the signals and transfer functions in the wave domain are denoted by a tilde over the respective variable, since suitable transforms will be based on the idea of a transformation into the eigenspace of the respective systems. The adaptation is performed entirely in the wave domain.

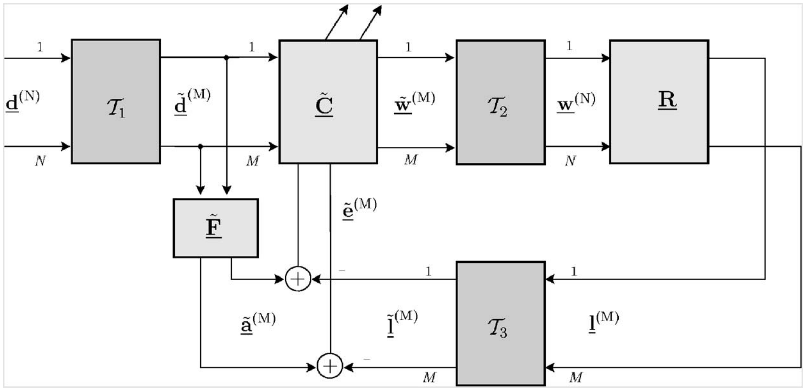

FIG. 5. Block diagram illustrating the wave domain adaptive inverse filtering approach to active room compensation.

Note that the generic block diagram depicted by Fig. 5 also includes the eigenspace adaptive inverse filtering approach. In this special case the transformations are given as the following MIMO FIR systems (see Sec. IV): $\mathcal{T}_1 = \underline{\mathbf{U}}^H(\omega)$, $\mathcal{T}_2 = \underline{\mathbf{V}}(\omega)$, and $\mathcal{T}_3 = \underline{\mathbf{X}}^H(\omega)$.

The following section introduces wave field expansions which are potential candidates as wave domain transformations. Since they are based upon the fundamental free-field solutions of the wave equation they will only decouple the free-field transfer matrix.

B. Wave field representations

A first choice to characterize an acoustic wave field is given by the acoustic pressure $P(\mathbf{x},\omega)$ captured at all positions $\mathbf{x}$. For a wide variety of applications it is convenient to represent acoustic wave fields with respect to an orthogonal basis. This basis is typically constructed from the fundamental solutions of the wave equation considering the particular problem. The general solution of the wave equation is then given as a weighted superposition of all elementary solutions. Of special interest in the following are decompositions which are based on the fundamental free-field solutions of the wave equation. These depend on the underlying coordinate system and its dimensionality. The basis functions connected to spherical, cylindrical, and Cartesian coordinates are known as spherical harmonics, cylindrical harmonics, and plane waves. An arbitrary wave field can be represented by the expansion coefficients with respect to these basis functions. If two-dimensional wave fields are considered then the polar and Cartesian coordinate systems are common. Here the basis functions are circular harmonics and (two-dimensional) plane waves. Please note that according to the term "spherical harmonics" used for the elementary solutions of the wave equation in spherical coordinates,[58–60] the solutions in cylindrical coordinates have been termed as *cylindrical harmonics* and the solutions in polar coordinates as *circular harmonics* within this article. Typical sound reproduction systems aim at the reproduction in a plane only. The analysis of the reproduced wave field is then typically performed in the reproduction plane. We will therefore limit ourselves to two-dimensional wave field representations for the following discussion of the circular harmonics and plane wave decomposition. However, since the reproduction will take place in a three-dimensional environment several artifacts of two-dimensional sound reproduction and analysis have to be considered.[61] A generalization of the proposed decompositions to three-dimensional representations can be derived straight-forwardly.[58]

1. Circular harmonics expansion

The elementary solutions of the wave equation using cylindrical coordinates are given in Refs. 58 and 62. The elementary solutions using polar coordinates can be derived from these by discarding the components which depend on the z coordinate. The circular harmonics decomposition of an arbitrary wave field is then given by

$$P_P(\mathbf{x}_p,\omega) = \sum_{\nu=-\infty}^{\infty} \left(\breve{P}^{(1)}(\nu,\omega) H_\nu^{(1)}\left(\left| \frac{\omega}{c} \right| r \right) e^{j\nu\alpha} + \breve{P}^{(2)}(\nu,\omega) H_\nu^{(2)}\left(\left| \frac{\omega}{c} \right| r \right) e^{j\nu\alpha} \right), \quad (26)$$

where $\mathbf{x}_p = [\alpha \; r]^T$ denotes the position vector in polar coordinates, $H_\nu^{(1),(2)}(\cdot)$ the νth order Hankel function of first/second kind, and ν the angular frequency. The polar coordinates are defined as $x = r\cos\alpha$ and $y = r\sin\alpha$ within this article. Quantities and functions whose arguments are given in polar coordinates are denoted by a P in their index. The infinite sum over ν can be interpreted as Fourier series with respect to the angle α. The coefficients $\breve{P}^{(1),(2)}(\nu,\omega)$ are referred to as circular harmonics expansion coefficients in the following and will be denoted by a breve over the respective variable. The Hankel function $H_\nu^{(1)}(|\omega/c|r)$ belongs to an incoming (converging) and $H_\nu^{(2)}(|\omega/c|r)$ to an outgoing (diverging) cylindrical wave.[58] Thus, the expansion coefficient $\breve{P}^{(1)}(\nu,\omega)$ describes the incoming wave field, whereas $\breve{P}^{(2)}(\nu,\omega)$ describes the outgoing wave field. According to Eq. (26) the total wave field is given as a superposition of incoming and outgoing contributions.

In order to get more insight into the circular harmonics expansion a closer look at the basis functions is taken. Each spatial variable has its own basis function. The angular coordinate α has an exponential function as basis. Figure 6 illustrates the angular basis functions for different angular frequencies. For the sake of illustration the plots only show the absolute value of the real part. It can be seen clearly that the angular basis functions exhibit a spatial selectivity in the angular coordinate. They can be interpreted as directivity

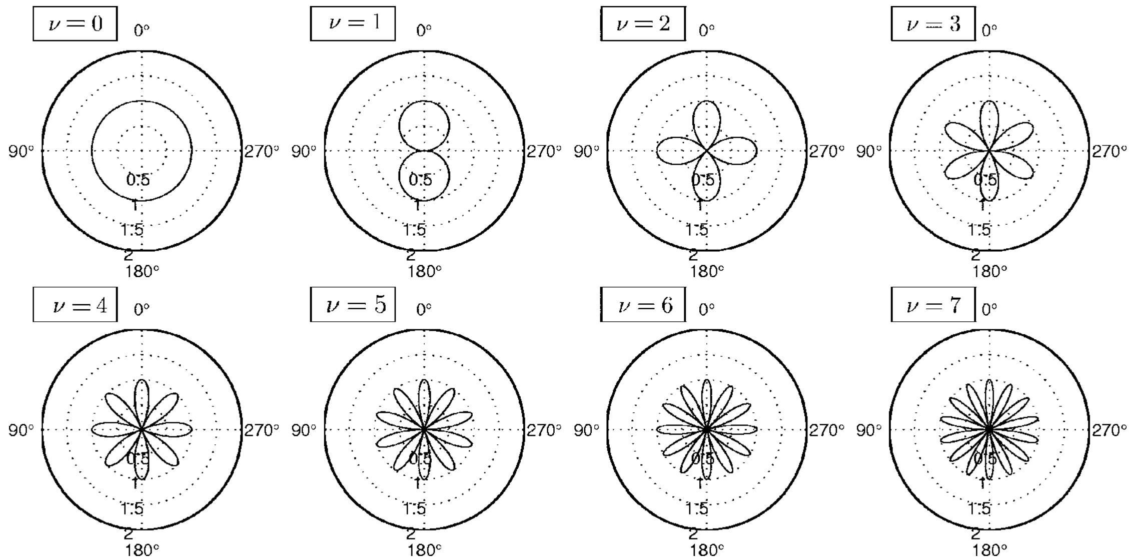

FIG. 6. Illustration of the angular basis functions $e^{j\nu\alpha}$ of the circular harmonics. The plots show the absolute value of the real part $(|\Re\{e^{j\nu\alpha}\}|)$ for different angular frequencies ν.

patterns. Hankel functions are the basis functions of the radial coordinate r. See Ref. 63 for a detailed discussion of their properties.

The circular harmonics expansion coefficients for a particular problem can be derived by introducing the expansion (26) into the underlying wave equation. However, it is also possible to measure an arbitrary wave field in a bounded region and derive the expansion coefficients from these measurements. Due to the underlying geometry circular microphone arrays are well suited for this task.[64]

2. Plane wave expansion

Arbitrary solutions of the wave equation can be expressed alternatively as superposition of plane waves traveling into all possible directions. The plane wave expansion coefficients $\bar{P}^{(1)}(\theta,\omega)$ and $\bar{P}^{(2)}(\theta,\omega)$ describe the spectrum of incoming/outgoing plane waves with incidence angle θ. They can be derived by a plane wave decomposition.[43,64] The plane wave expansion coefficients exhibit a direct link to the expansion coefficients in circular harmonics

$$\bar{P}^{(1)}(\theta,\omega) = \frac{4\pi}{k} \sum_{\nu=-\infty}^{\infty} j^{\nu} \breve{P}^{(1)}(\nu,\omega) e^{j\nu\theta}, \tag{27a}$$

$$\bar{P}^{(2)}(\theta,\omega) = \frac{4\pi}{k} \sum_{\nu=-\infty}^{\infty} j^{\nu} \breve{P}^{(2)}(\nu,\omega) e^{j\nu\theta}. \tag{27b}$$

Equation (27) states that the plane wave decomposition of a wave field is, up to the factor j^{ν}, given by the Fourier series of the expansion coefficients in terms of circular harmonics.

C. Circular harmonics as wave domain transformation

The optimal choice for the transformed signal representation depends on the acoustic characteristics of the listening room and the desired free-field response. Two different wave field representations have been introduced in Sec. IV: the decomposition into plane waves and the decomposition into circular harmonics. Both are based upon orthogonal basis functions derived from the free-field wave equation and will therefore provide a decoupling of the free-field transfer matrix $\underline{\mathbf{F}}(\omega)$ but not necessarily of the room transfer matrix $\underline{\mathbf{R}}(\omega)$. Either representation may be nevertheless a suitable choice for the wave domain representation as long as they provide an approximate decoupling of the room transfer matrix. In order to qualitatively evaluate the two proposed decompositions a closer look is taken at the room response with respect to the basis functions. In the ideal case the reproduction of a basis function will result only in contributions belonging to the same basis function after analysis. Consider now typical listening rooms with a rectangular shape as a first approximation and their response to either plane waves or circular harmonics.

The response to a plane wave of a certain angle consists in general of a mixture of plane waves with all kinds of different angles. This means that the desired effect of decoupling is not achievable with plane waves in more or less rectangular rooms.

On the other hand, the response to a circular wave depends on the wavelength. Since high frequencies can be damped effectively by passive methods, active room compensation can be restricted to low frequencies. For the corresponding long wavelengths, scattering at the corners of the listening room does not play a major role. Therefore each circular harmonics component emitted by a well-designed reproduction system will lead to reflections which can be described mainly by the same component. Consequently the desired decoupling effect will be more prominent than for plane waves. This intuitive consideration suggests that circular harmonics seem to be a more suitable choice for the wave domain representation than plane waves. This assumption holds also for reverberant rooms as long as the energy of other components than the emitted one is considerably attenuated.

Please note that the representation of a field in a transformed domain requires adequate sampling of the measured field. Sampling theorems for circular microphone arrays can be found, e.g., in Refs. 36, 43, 65, and 66.

VI. APPLICATION TO WAVE FIELD SYNTHESIS

Section V introduced wave domain adaptive filtering. Now its application to spatial reproduction techniques is shown using wave field synthesis as an example. First, this reproduction technique is discussed shortly. Then the application of wave domain adaptive filtering to wave field synthesis is shown by specifying the transformations $\mathcal{T}_1$ through $\mathcal{T}_3$ in Fig. 5. Finally some selected results of performance evaluations are presented.

A. Wave field synthesis

Wave field synthesis (WFS) is a massive multichannel sound reproduction technique which overcomes the "sweet spot" limitation well known from stereophonic surround sound methods. WFS techniques are based on Huygen's principle and are formulated in terms of the Kirchhoff-Helmholtz integral.[62] These foundations have been initially developed by the Technical University of Delft[10–12] and were later extended within the European project CARROUSO.[67] Detailed descriptions of WFS can be found in Refs. 13, 14, 43, and 68–72.

B. Application of wave domain adaptive filtering to wave field synthesis

The generic concept of WDAF is now specialized to the reproduction utilizing a WFS system and the circular harmonics decomposition as wave domain transformation. The transformations $\mathcal{T}_1$ through $\mathcal{T}_3$ are based on the decomposition into circular harmonics of the respective wave fields. For active room compensation only the incoming parts (see Sec. V B 1) of the respective wave fields are of interest. For the particular scenario considered these transformations are specialized as follows.

(1) **Transformation** $\mathcal{T}_1$: This transformation transforms the virtual source signal $S(\omega)$ using a spatial model of the virtual source into the loudspeaker driving signals $\breve{\underline{\mathbf{d}}}^{(M)}$ in the wave domain. Suitable models for the virtual source characteristics are line/point sources or plane waves.

(2) **Transformation** $\mathcal{T}_2$: This transformation generates the loudspeaker driving signals from the filtered driving signals $\breve{\underline{\mathbf{w}}}^{(M)}$. Equation (26) together with a suitable loudspeaker selection criterion[73] can be used for this purpose.

(3) **Transformation** $\mathcal{T}_3$: This transformation calculates the circular harmonics decomposition coefficients of the wave field within the listening area from the microphone array measurements.[64]

The free-field transfer matrix $\breve{\underline{\mathbf{F}}}$ models the free-field propagation in terms of circular harmonics from the loudspeakers to the microphone array. In the ideal case this matrix would only model the propagation delay. Besides this, the transfer matrix $\breve{\underline{\mathbf{F}}}$ should also include an additional delay to ensure the computation of causal room compensation filters. If desired, certain artifacts of WFS systems like, e.g., amplitude errors (see Ref. 61) can be considered in the construction of the matrix $\breve{\underline{\mathbf{F}}}$.

It was proposed by the authors in previous publications[40–42,74–76] to perform a Fourier transformation with respect to the angular variable of the plane wave decomposition to derive the signals in the wave domain. This Fourier transformation of the plane wave decomposed signals is up to the factor j^{ν} (and a frequency correction) equivalent to calculating the circular harmonics expansion coefficients, since the plane wave decomposition in terms of circular harmonics is given by the Fourier series (27).

C. Performance evaluation

After describing the application of wave domain adaptive filtering for wave field synthesis, its performance for decoupling the adaptation will now be evaluated. The evaluation of the circular harmonics decomposition for listening room compensation with wave domain adaptive filtering has been described in detail in Ref. 77. The following quote some selected results.

1. Performance measures

The performance of the proposed transformations and of the resulting active listening room compensation system is evaluated by the ability to compact the room characteristics into less coefficients than using the microphone signals directly. This ability is described by two measures: (1) the energy of the elements of the room transfer matrix and (2) the energy compaction performance. The first measure is defined by calculating the energy of each spatial transmission path for the room transfer matrix in its different representations. For its representation in the pressure domain (pressure microphones) the energy of the elements of the room transfer matrix is defined as

$$E(m,n) = \frac{1}{2\pi}\int_{-\infty}^{\infty} |R_{m,n}(\omega)|^2 d\omega. \tag{28}$$

Similar definitions apply to the room transfer matrix in its plane wave and circular harmonics representation yielding the energy representations $\bar{E}(\theta,\theta_0)$ and $\breve{E}(\nu,\nu_0)$. The second measure, the energy compaction, measures the ability of a transform to compact the energy of the room transfer matrix to as few coefficients as possible. For one particular room transfer matrix this measure is defined by calculating the ratio between the energy of the first i dominant elements and the total energy of all elements. For this purpose the energies $E(m,n)$ are sorted in descending order yielding the sorted elements $E_{\mathrm{sort}}(\eta)$. Then the ratio between the energy of the first i sorted elements and the total energy of all elements is calculated as follows:

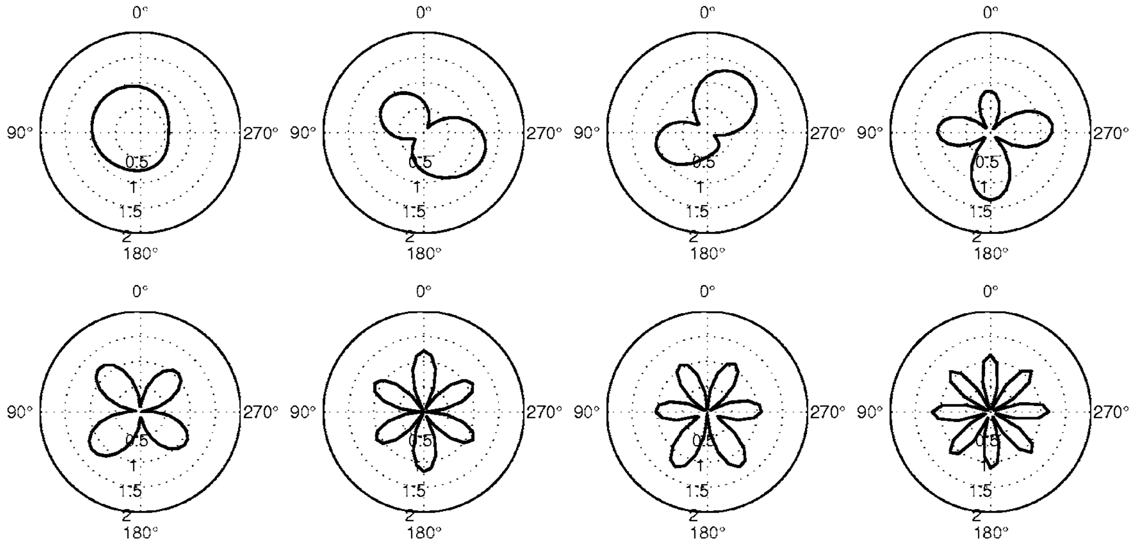

FIG. 7. Absolute value of the first eight right singular vectors (f=80 Hz) for the simulated circular WFS/wave field analysis system sorted by their descending singular values (top left to bottom right). The singular vectors for a plane wave reflection factor R_{pw}=0.8 at the walls of the simulated room are shown.

$$\mathrm{EC}(i) = \frac{\sum_{\eta=0}^{i} E_{\mathrm{sort}}(\eta)}{\sum_{\eta=0}^{MN-1} E_{\mathrm{sort}}(\eta)}, \tag{29}$$

where $0 \leqslant \mathrm{EC}(i) \leqslant 1$. Equivalent definitions apply for the transformed representations. The larger portion of the total energy that is captured by the first i elements, the better the performance in terms of energy compaction. The energy of the elements $E(m,n)$, $\bar{E}(\theta,\theta_0)$, and $\breve{E}(\nu,\nu_0)$ illustrate the distribution of the energy in the different transformed domains, while $\mathrm{EC}(i)$, $\overline{\mathrm{EC}}(i)$, and $\breve{\mathrm{EC}}(i)$ illustrate the ability of a particular transformation to compact the energy in few coefficients.

2. Evaluation results

The following illustrates the performance of active listening room compensation for WFS using the proposed WDAF approach for simulated acoustical environments. The main benefits of using simulated versus real acoustical environments are that the parameters of the simulated environment can be changed easily in order to simulate different scenarios and that practical aspects such as noise, transducer mismatch, and misplacement can be excluded for a first proof of the WDAF concept.

For an accurate simulation of the acoustic environment, a numerical simulation of wave propagation based on the functional transformation method was used.[78] The particular implementation used (wave2d) simulates the wave equation in two dimensions.[79] This method requires no spatial discretization and thus allows one to place the virtual microphones and speakers at their exact spatial positions. This is not possible when using methods which perform a spatial discretization, such as, e.g., the finite element method.[80]

The geometry of the simulated acoustical environment is a simplification of a real environment at the multimedia laboratory of the Multimedia Communications and Signal Processing Group at the University of Erlangen-Nuremberg.[81] Its geometrical and acoustical details as well as the loudspeaker and microphone setups are given in Refs. 43 and 77. The simulated room has the dimensions $5.90 \times 5.80 \times 3.10$ m ($w \times l \times h$).

A WFS system with 48 loudspeakers placed equidistant on a circle with a radius of R_{LS}=1.50 m and a wave field analysis system consisting of 48 angular sampling positions at a radius of R_{mic}=0.75 m was simulated. The simulations were performed for a reverberant environment with a plane wave reflection factor of R_{pw}=0.8 at the walls. Due to the spatial aliasing frequency of the WFS system and the microphone array all signals were low-pass filtered to a bandwidth of 650 Hz to calculate the results shown.

An indication for the suitability of circular harmonics is given by calculating the right singular vectors of the room transfer matrix $\boldsymbol{R}(\omega)$ (for the pressure microphones) using the GSVD. Figure 7 shows the absolute value of the first eight right singular vectors for the reverberant case R_{pw} =0.8. The singular vectors have been sorted by their descending singular values and thus by their energy. It can be seen that they look very similar to the circular harmonics illustrated in Fig. 6. This result and the considerations in Sec. V C give justification for using the circular harmonics representation as wave domain transformation.

Figure 8 shows the energy of the elements of the room transfer matrix for the pressure microphones and its circular harmonics representation. Figure 8(a) shows $E(m,n)$ for the pressure microphones. The direct path from the loudspeakers to the microphones can be seen clearly along the main diagonal but also the reflections of the loudspeaker wave fields at the walls of the listening room. Figure 8(b) shows the energy of the room transfer matrix represented in circular harmonics. As desired, the main diagonal elements represent a major portion of the energy. The off-diagonal elements are a result of the reverberant enclosure. Please note the different scales used for Figs. 8(a) and 8(b).

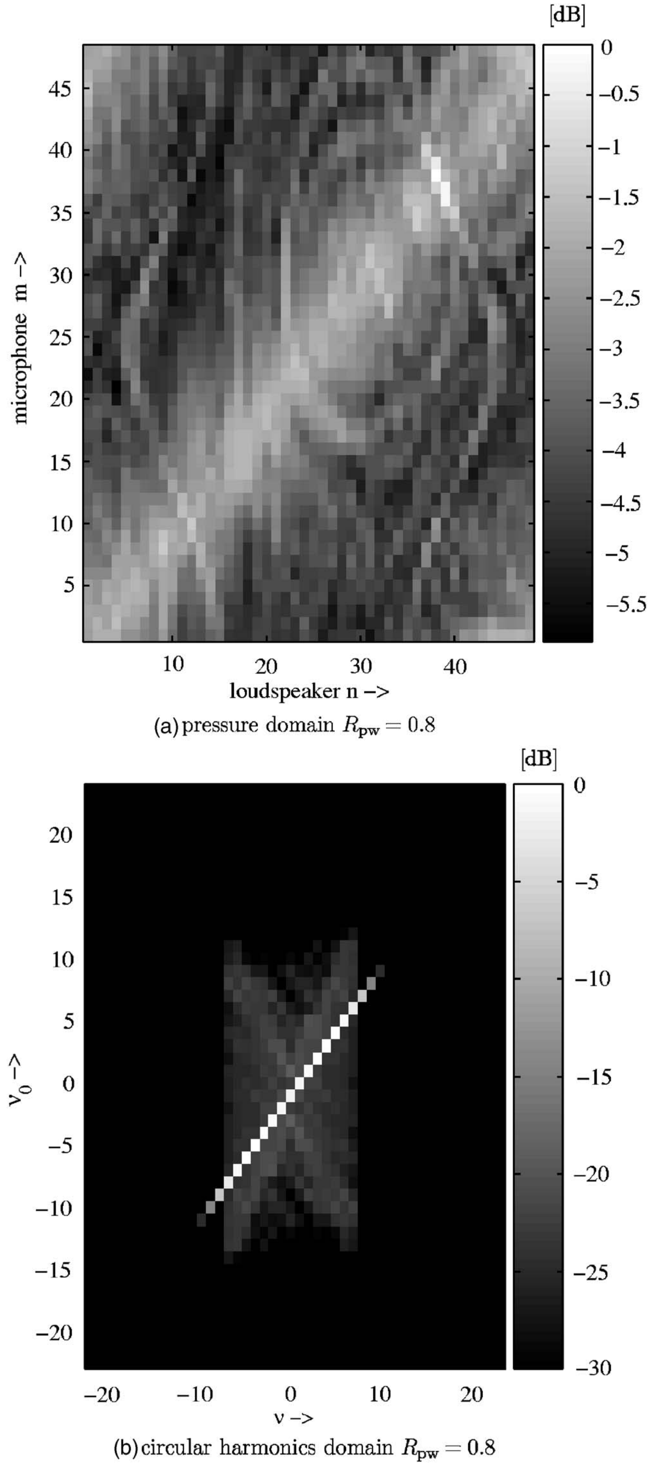

FIG. 8. Energy of the room transfer matrix of the signals captured by the pressure microphones $E(m,n)$ and in the circular harmonics domain $\breve{E}(\nu,\nu_0)$.

Not only the performance of the circular harmonics in terms of energy compaction toward the main diagonal is of interest, but also its performance to represent the MIMO system by as few coefficients as possible. This can be measured by the energy compaction of the different representations. Results are shown in Fig. 9. Figure 9 illustrates the energy compaction according to Eq. (29) for the room transfer matrix in the pressure domain $\mathrm{EC}(i)$, in the plane wave decomposed domain $\overline{\mathrm{EC}}(i)$, and in the circular harmonics domain

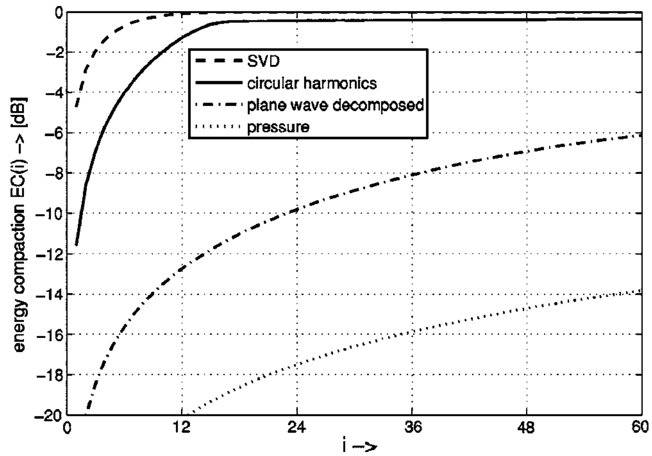

FIG. 9. Energy compaction performance $\mathrm{EC}(i)$ for the different representations of the room transfer matrix. The index i denotes the number of sorted elements which capture most of the total energy.

$\breve{\mathrm{EC}}(i)$. The energy compaction of the GSVD is shown for reference. It can be seen clearly that the circular harmonics representation of the room transfer matrix compacts the energy much better than its pressure and plane wave representation. It can also be seen that the GSVD provides the optimal transformation in this sense.

More simulation results for other values of the reflection factor and also for a rectangular loudspeaker array geometry have been reported in Ref. 77. Furthermore, measurements have been conducted for the circular wave field synthesis/analysis system described earlier. These measurements confirm the simulation results.

The presented results indicate that the circular harmonics decomposition is not the optimal transformation for the reverberant case. However, the results also show that this transformation provides a quite reasonable approximation to the optimal GSVD transformation, especially when compared to the plane wave decomposition and pressure microphone representations.

Further results of the proposed algorithm with respect to the adaptation performance and the resulting wave field inside the listening area after compensation can be found in Refs. 41 and 48. They reveal that the proposed algorithm provides fast and stable adaptation even for nonstationary virtual scenes and results in a reproduced wave field without listening room reflections throughout the entire listening area.

VII. CONCLUSIONS

This contribution has derived a method for active listening room compensation for massive multichannel sound reproduction systems. It has been shown that a major challenge is the adaptation of the matrix of compensation filters when straightforwardly applying multichannel adaptation algorithms (like the X-RLS algorithm) to massive multichannel reproduction scenarios. In addition to the computation of the compensation filters also the matrix of room impulse responses has to be estimated. The numerical expense for such a straightforward implementation would be immense: For reproduction systems with even only a few ten loudspeakers, the computation of many thousands of single adaptation problems would be required for each time step. Furthermore, the convergence of the adaptation will be slow due to a high correlation between the channels.

Suitable mathematical techniques can help to reduce the numerical expense and to improve convergence through decoupling the adaptation problem. More specifically, the matrix of compensation filters, the matrix of room impulse responses, and the free field propagation matrix can be jointly diagonalized through a GSVD. This method has been called eigenspace adaptive filtering because the diagonalization can be seen as a transformation into the eigenspace of the problem. Although eigenspace adaptive filtering provides optimal decoupling, there is a major drawback for the practical application: The GSVD matrices for the transformation into the eigenspace are data dependent, since they depend on the time-varying acoustic room response characteristics. Calculating these matrices anew in each time step would create another numerical problem.

At this point, the situation is similar as for coding of audio or image data. The optimal energy compaction and decorrelation could be achieved by the Karhunen-Loéve transform, but the correct transformation matrix is data dependent and hence not known. Thus practical coders are realized with a different transformation which does not depend on the input data. For a wide variety of problems in audio and video coding, the discrete cosine transformation has proven to approximate the properties of the Karhunen-Loève transform reasonably well. A similar route has been taken here. The search for an approximation of the transformation into the eigenspace by GSVD has been motivated by physical intuition. Inspection of the eigenvectors of the GSVD from simulations and measurements has shown that they resemble the circular harmonics decomposition of free field acoustics. This notion led us to the idea that an approach in which the optimal transformations into the eigenspace is replaced by a transformation into the space of circular harmonics, the so-called *wave domain*, may give results that are only slightly suboptimal.

The implementation of the transformation into the wave domain depends on the specific spatial reproduction/analysis technique. Details have been given for reproduction by wave field synthesis. It turned out that the decoupling properties of the transformation into the wave domain are still quite good and close to the theoretical optimum of the GSVD. In summary, it has been shown that active listening room compensation for massive multichannel sound reproduction is a tractable problem. For spatial reproduction with wave field synthesis, wave domain adaptive filtering provides a solution which is both numerically feasible and almost optimal. It remains to extend these investigations to other spatial reproduction techniques.

[1]E. Torick, "Highlights in the history of multichannel sound," J. Audio Eng. Soc. **46**, 27–31 (1998).

[2]G. Steinke, "Surround sound—The new phase. An overview," in 100th Audio Engineering, Society Convention, Copenhagen, Denmark, May 1996.

[3]D. Griesinger, "Surround: The current technological situation," in 108th Audio Engineering Society Convention, Paris, France, February 2000.

[4]S. Kiriakakis, P. Tsakalides, and T. Holman, "Surrounded by sound," IEEE Signal Process. Mag. **16**, 55–66 (1999).

[5]M. Poletti, "A unified theory of horizontal holographic sound systems," J. Audio Eng. Soc. **48**, 1155–1182 (2000).

[6]F. Rumsey, *Spatial Audio* (Focal Press, 2001).

[7]J. Meyer, "Spherical microphone arrays for 3D sound reproduction," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, edited by Y. Huang and J. Benesty (Kluwer Academic, Dordrecht, 2004).

[8]S. Moreau and J. Daniel, "3D sound field recording with higher order ambisonics—Objective measurements and validation of a 4th order spherical microphone," in 120th Audio Engineering Society Convention, Paris, France, 2006.

[9]M. Gerzon, "Ambisonics in multichannel broadcasting and video," J. Acoust. Soc. Am. **33**, 859–871 (1985).

[10]A. Berkhout, "A holographic approach to acoustic control," J. Audio Eng. Soc. **36**, 977–995 (1988).

[11]D. de Vries, E. Start, and V. Valstar, "The wave field synthesis concept applied to sound reinforcement: Restrictions and solutions," in 96th Audio Engineering Society Convention, Amsterdam, The Netherlands, February 1994.

[12]J. Sonke, D. de Vries, and J. Labeeuw, "Variable acoustics by wave field synthesis: A closer look at amplitude effects," in 104th Audio Engineering Society Convention, Amsterdam, The Netherlands, May 1998.

[13]S. Spors, H. Teutsch, A. Kuntz, and R. Rabenstein, "Sound field synthesis," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, edited by Y. Huang and J. Benesty (Kluwer Academic, Dordrecht, 2004).

[14]R. Rabenstein and S. Spors, "Wave field synthesis techniques for spatial sound reproduction," in *Topics in Acoustic Echo and Noise Control*, edited by E. Haensler and G. Schmidt (Springer, New York, 2006), Chap. 13, pp. 517–545.

[15]D. Griesinger, "Multichannel sound systems and their interaction with the room," in 16th International Conference on Audio, Acoustics, and Small Places, Audio Engineering Society, 1998, pp. 159–173.

[16]M. Dewhirst, S. Zielinski, P. Jackson, and F. Rumsey, "Objective assessment of spatial localization attributes of surround-sound reproduction systems," in 118th Audio Engineering Society Convention, Barcelona, Spain, May 2005.

[17]B. Klehs and T. Sporer, "Wave field synthesis in the real world. 1. In the living room," in 114th Audio Engineering Society Convention, Amsterdam, The Netherlands, March 2003.

[18]E. Völker, "To nearfield monitoring of multichannel reproduction—Is the acoustics of the living room sufficient?," in Proceedings of the Tonmeistertagung, Hannover, Germany, 1998.

[19]E. Völker, W. Teuber, and A. Bob, "5.1 in the living room—On acoustics of multichannel reproduction," in Proceedings of the Tonmeistertagung, Hannover, Germany, 2002.

[20]E.-J. Voelker, "Home cinema surround sound—Acoustics and neighbourhood," in 100th Audio Engineering Society Convention, Copenhagen, Denmark, May 1996.

[21]T. Sporer and B. Klehs, "Wave field synthesis in the real world. 2. In the movie theatre," in 116th Audio Engineering Society Convention, Berlin, Germany, May 2005.

[22]P. Gauthier and A. Berry, "Sound-field reproduction in-room using optimal control techniques: Simulations in the frequency domain," J. Acoust. Soc. Am. **117**, 662–678 (2005).

[23]S. Bech, "Timbral aspects of reproduced sound in small rooms. I," J. Acoust. Soc. Am. **97**, 1717–1726 (1995).

[24]S. Bech, "Timbral aspects of reproduced sound in small rooms. II," J. Acoust. Soc. Am. **99**, 3539–3549 (1996).

[25]S. Bech, "Spatial aspects of reproduced sound in small rooms," J. Acoust. Soc. Am. **103**, 434–445 (1998).

[26]T. Caulkins and O. Warusfel, "Characterization of the reverberant sound field emitted by a wave field synthesis driven loudspeaker array," in 120th Audio Engineering Society Convention, Paris, France, May 2006.

[27]E. Corteel and R. Nicol, "Listening room compensation for wave field synthesis. What can be done?," in 23rd Audio Engineering Society Convention, Copenhagen, Denmark, May 2003.

[28]L. Fielder, "Practical limits for room equalization," in 111th Audio Engineering Society Convention, New York, September 2001.

[29]L. Fielder, "Analysis of traditional and reverberation-reducing methods of room equalization," J. Audio Eng. Soc. **51**, 3–26 (2003).

[30]P. Hatziantoniou and J. Mourjopoulos, "Errors in real-time room acoustics dereverberation," J. Audio Eng. Soc. **52**, 883–899 (2004).

[31]J. Mourjopoulos, "Digital equalization of room acoustics," J. Audio Eng.

Soc. **42**, 884–900 (1994).
[32]F. Talantzis and D. Ward, "Multi-channel equalization in an acoustic reverberant environment: Established of robustness measures," in Institute of Acoustics Spring Conference, Salford, UK, March 2002.
[33]F. Talantzis and D. Ward, "Robustness of multichannel equalization in an acoustic reverberant environment," J. Acoust. Soc. Am. **114**, 833–841 (2003).
[34]S. Bharitkar, P. Hilmes, and C. Kyriakakis, "Sensitivity of multichannel room equalization to listener position," in IEEE International Conference on Multimedia and Expo (ICME), Baltimore, MD, 2003.
[35]P. Nelson, F. Orduna-Bustamante, and H. Hamada, "Inverse filter design and equalization zones in multichannel sound reproduction," IEEE Trans. Speech Audio Process. **3**, 185–192 (1995).
[36]T. Betlehem and T. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," J. Acoust. Soc. Am. **117**, 2100–2111 (2005).
[37]T. Betlehem and T. Abhayapala, "A modal approach to soundfield reproduction in reverberant rooms," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Philadelphia, PA, 2005, Vol. **III**, pp. 289–292.
[38]A. Sontacchi and R. Hoeldrich, "Enhanced 3D sound field synthesis and reproduction system by compensating interfering reflections," in Conference on Digital Audio Effects (DAFX-00), Verona, Italy, 2000.
[39]M. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," J. Audio Eng. Soc. **53**, 1004–1025 (2005).
[40]S. Spors, H. Buchner, and R. Rabenstein, "A novel approach to active listening room compensation for wave field synthesis using wave-domain adaptive filtering," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), (2004).
[41]S. Spors, H. Buchner, and R. Rabenstein, "Adaptive listening room compensation for spatial audio systems," in European Signal Processing Conference (EUSIPCO) (2004).
[42]S. Spors, H. Buchner, and R. Rabenstein, "Efficient active listening room compensation for wave field synthesis," in 116th Audio Engineering Society Convention, Berlin, Germany, 2004.
[43]S. Spors, "Active listening room compensation for spatial sound reproduction systems," Ph.D. thesis, University of Erlangen-Nuremberg, 2006.
[44]J. Lopez, A. Gonzalez, and L. Fuster, "Room compensation in wave field synthesis by means of multichannel inversion," in IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, 2005.
[45]R. von Zon, E. Corteel, D. de Vries, and O. Warusfel, "Multi-actuator panel (map) loudspeakers: How to compensate for their mutual reflections," in 116th Audio Engineering Society Convention, Berlin, Germany, 2004.
[46]P. Gauthier and A. Berry, "Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction: Theory," J. Acoust. Soc. Am. **119**, 2721–2737 (2006).
[47]M. Omura, M. Yada, H. Saruwatari, S. Kajita, K. Takeda, and F. Itakura, "Compensating of room acoustic transfer functions affected by change of room temperature," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Phoenix, AZ, 1999.
[48]S. Petrausch, S. Spors, and R. Rabenstein, "Simulation and visualization of room compensation for wave field synthesis with the functional transformation method," in 119th Audio Engineering Society Convention, New York, 2005.
[49]M. Sondhi, D. Morgan, and J. Hall, "Stereophonic acoustic echo cancellation—An overview of the fundamental problem," IEEE Signal Process. Lett. **2**, 148–151 (1995).
[50]J. Garas, *Adaptive 3D Sound Systems* (Kluwer Academic, Dordrecht, 2000).
[51]A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1999).
[52]M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," IEEE Trans. Acoust., Speech, Signal Process. **36**, 145–152 (1988).
[53]S. Haykin, *Adaptive Filter Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1996).
[54]J. Benesty, D. Morgan, and M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," IEEE Trans. Speech Audio Process. **6**, 156–165 (1998).
[55]M. Bouchard and S. Quednau, "Multichannel recursive-least-squares algorithms and fast-transversal-filter algorithms for active noise control and sound reproduction systems," IEEE Trans. Speech Audio Process. **8**, 606–618 (2000).
[56]G. Golub and C. Loan, *Matrix Computations* (The Johns Hopkins University Press, Baltimore, MD, 89).
[57]P. Flikkma, "An algebraic theory of 3D sound synthesis with loudspeakers," in 22nd International Conference on Virtual, Synthetic and Entertainment Audio, Audio Engineering Society, 2002.
[58]E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, New York, 1999).
[59]A. Pierce, *Acoustics. An Introduction to its Physical Principles and Applications* (Acoustical Society of America, New York, 1991).
[60]D. Blackstock, *Fundamentals of Physical Acoustics* (Wiley, New York, 2000).
[61]S. Spors, M. Renk, and R. Rabenstein, "Limiting effects of active room compensation using wave field synthesis," in 118th Audio Engineering Society Convention, Barcelona, Spain, 2005.
[62]P. Morse and H. Feshbach, *Methods of Theoretical Physics. Part I* (McGraw-Hill, New York, 1953).
[63]M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
[64]E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis," in 110th Audio Engineering Society Convention, Amsterdam, The Netherlands, May 2001.
[65]H. Teutsch, "Wavefield decomposition using microphone arrays and its application to acoustic scene analysis," Ph.D. thesis, University of Erlangen-Nuremberg (2005), http://www.lnt.de/lms/publications (last viewed on 16 April 2007).
[66]T. Ajdler, L. Sbaiz, and M. Vetterli, "Plenacoustic function on the circle with application to HRTF interpolation," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Philadephia, PA (2005).
[67]S. Brix, T. Sporer, and J. Plogsties, "CARROUSO—An European approach to 3D-audio," in 110th Audio Engineering Society Convention, 2001.
[68]E. Start, "Direct sound enhancement by wave field synthesis," Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 1997.
[69]E. Verheijen, "Sound reproduction by wave field synthesis," Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 1997.
[70]P. Vogel, "Application of wave field synthesis in room acoustics," Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 1993.
[71]W. de Bruijn, "Application of wave field synthesis in videoconferencing," Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2004.
[72]E. Hulsebos, "Auralization using wave field synthesis," Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2004.
[73]S. Spors, "An analytic secondary source selection criteria for wave field synthesis," in 33rd German Annual Conference on Acoustics (DAGA), Stuttgart, Germany, March 2007.
[74]H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering: Acoustic echo cancellation for full-duplex systems based on wave-field synthesis," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2004.
[75]H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering for acoustic human-machine interfaces based on wavefield analysis and synthesis," in European Signal Processing Conference (EUSIPCO), 2004.
[76]H. Buchner, S. Spors, and W. Kellermann, "Full-duplex systems for sound field recording and auralization based on wave field synthesis," in 116th Audio Engineering Society Convention, Berlin, Germany, 2004.
[77]S. Spors and R. Rabenstein, "Evaluation of the circular harmonics decomposition for WDAF-based active listening room compensation," in 28th AES Conference: The Future of Audio Technology—Surround and Beyond, Audio Engineering Society, Pitea, Sweden, June/July 2006, pp. 134–149.
[78]S. Petrausch and R. Rabenstein, "Highly efficient simulation and visualization of acoustic wave fields with the functional transformation method," in *Simulation and Visualization*, (Otto von Guericke Universität, Magdeburg, 2005), pp. 279–290.
[79]"Wave2D: Wave field simulation in 2D," http://www.lnt.de/~stepe/demo/Wave2D (last viewed on 16 April 2007).
[80]J. Reddy, *Introduction to the Finite Element Method* (McGraw Hill, New York, 2005).
[81]"Multimedia communications and signal processing at the University of Erlangen-Nuremberg," http://www.lnt.de/LMS (last viewed on 16 April 2007).