

Multichannel Room Impulse Response Generation With Coherence Control

Martin Kuster

Abstract—A method is proposed by which an arbitrary number of room impulse responses (RIRs) is generated from one input RIR. The method works by convolving the input RIR with a multitude of filters, one for each desired output RIR. The filters are designed in such a way that the coherence between all possible pairs of output RIRs feature assignable frequency-dependent values. An application example is given where the values for the coherence are given by the diffuse field coherence functions for the microphones in a five-element microphone array. The limitations in terms of achieving the desired coherence values exactly and avoiding spectral coloration are also discussed. The proposed method is most suitable for the part of the RIR that does not contain strong discrete reflections.

Index Terms—Audio systems, coherence, diffuse sound, multi-channel audio reproduction, room acoustics.

I. INTRODUCTION

THE reproduction of spatial audio through multichannel loudspeaker systems has been an area of improvement and widespread application not only but especially in the past two decades. One drawback is that the extensive catalogue of existing mono or stereo recordings can currently not be played back in a way that makes effective use of the enhanced spatial reproduction capability of these systems. With a large number of channels, another disadvantage is the extensive microphone system required for an acoustic recording and the significant data volume during transmission and storage. Consequently, a method that can generate multichannel audio from a single audio channel is very desirable. In this context, two possible approaches can be distinguished. In the first, the processing occurs on the actual recording (sound source convolved with reverberation). In the second, sound source and reverberation, the latter represented by the room impulse response (RIR), are separate and multichannel RIRs are generated from the input RIR(s). The latter approach is more limited in its practical applications but has the advantage that the typically vastly different early part and reverberant tail of an RIR can be processed separately. The method presented in the current paper follows the second approach concerning the notation and the practical application examples but can, in principle, also be used with the first approach.

The method of creating a spatial impression by producing (de-correlated) copies of an audio signal has been introduced as early as the 1950s. Lauridsen [1] proposed and Schroeder [2] investigated the use of a de-correlation filter to produce an artificial stereophonic effect from a monophonic input signal. The filtering consisted in mixing the original signal with a delayed version that had a 180° phase shift between the two channels. This can be regarded as the combination of a phase-linear comb filter and an all-pass filter and Schroeder remarked that the stereophonic effect appears to be mainly caused by the comb filtering.

Contrary to this, most modern methods aim to add the spatial information without altering the time–frequency statistics of the signal. Pulkki and Merimaa [3], [4] have introduced a method based on subband processing using the short-time Fourier transform. The authors have investigated three algorithms for the rendering of the reverberant tails of the generated RIRs. The first algorithm is not relevant here because it explicitly relies on SoundField B-format data. The second algorithm consists of maintaining the magnitudes in the short-time Fourier transform of the input RIR but randomizing the phases. It is mentioned that this can lead to aliasing in the time domain. The third algorithm is based on convolving the input RIR with dedicated de-correlation filters.

These can be all pass filters or short noise bursts constructed in the frequency domain by prescribing unit magnitude and random phase frequency response. The latter have originally been introduced by Kendall [5] and will also be used in the current paper as prototype frequency responses. Alternatively, Hawksford and Harris [6] propose exponentially decaying noise bursts equalized to a uniform magnitude frequency response with a minimum phase filter. Pulkki and Merimaa [4] have used such filters with a time constant of the exponential decay of 50 ms in their investigations. In a similar context, Faller [7] has used the reverberant tail of modeled RIRs as de-correlation filters for audio signals and he controlled the amount of correlation through the maximum value of the normalized cross-correlation in sub bands. It is also worth noting that Meesawat [8] has shown that the tails of binaural RIRs measured at different positions within the same room can be swapped without audible differences.

Except maybe for Faller's method, all of these existing methods for multichannel signal generation lead to incoherent signals or allow no control over the coherence. Building upon the filters proposed by Kendall [5], the main distinguishing feature of the method derived in the current paper is that it is possible to prescribe the coherence between all possible pairs of output signals as a function of frequency. The paper is organized as follows. In Section II, the derivation of the filters

Manuscript received July 28, 2008; revised October 23, 2008. Current version published March 18, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stephen J. Elliott.

The author is with the Laboratory of Acoustic Imaging and Sound Control, Delft University of Technology, 2600 GA Delft, The Netherlands (e-mail: kuster_martin@hotmail.com).

Digital Object Identifier 10.1109/TASL.2008.2009575

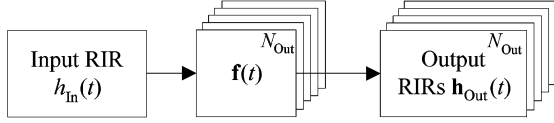


Fig. 1. Flow diagram for the generation of the N_{Out} output RIRs $h_{Out}(t)$ from the input RIR $h_{In}(t)$ by convolution with the filters $f(t)$.

possessing the prescribed coherence is presented. The specific cases of two and three output signals are solved analytically and the solution for an arbitrary number of output signals is described. The limitations due to the chosen filter design method are also investigated. In the current paper, the signals are RIRs and a practical application of the method is presented in Section III, where established analytical expressions for the diffuse field coherence functions are used to simulate the recording of multichannel RIRs with a microphone array.

II. METHODOLOGY

The proposed method generates the set of output RIRs $h_{Out}(t)$ by a convolution of the input RIR $h_{In}(t)$ with the set of FIR filters $f(t)$ as shown in Fig. 1. This can be written mathematically as

$$h_{Out}(t) = h_{In}(t) * f(t) \quad (1)$$

where $*$ denotes the convolution operator and t is the (discrete) time variable. The set $h_{Out}(t)$ contains N_{Out} output RIRs and any particular one is denoted by $h_{Out[i]}(t)$ ($i = [1, N_{Out}]$).

One requirement is that the filters $f(t)$ do neither alter the global frequency domain structure nor the global time domain structure of the RIRs. In particular, the exponential decay (corresponding to the reverberation time) in $h_{Out}(t)$ must be the same as in $h_{In}(t)$. Because the amplitudes of the filters used in the current paper are distributed evenly over its impulse response, this implies that the filter length T_F must be much smaller than the reverberation time. The current author has used values in the range $T_F \approx 30 \dots 60$ ms and together with the fact that the filter coefficients must be sampled at the same sampling frequency f_s as $h_{In}(t)$, there are thus a total of $N_F = f_s T_F$ filter coefficients per filter.

Because the coherence is defined in the frequency domain (see below), the coefficients of the filters $f(t)$ are designed from the required frequency response by using the inverse DFT method. The frequency domain analogs of $h_{In}(t)$ and $h_{Out[i]}(t)$ are the room transfer functions (RTFs) $H_{In}(\omega)$ and $H_{Out[i]}(\omega)$, respectively. Here, ω is a discrete angular frequency variable whose values are determined by f_s and T_F . Since $f(t)$ must be real-valued, the amplitudes of its frequency response at the negative frequencies are the complex conjugates of the amplitudes at the negative frequencies and the dc component is set to zero.

A. Derivation of the Filters With Prescribed Coherence Values

In the frequency domain, (1) reads

$$H_{Out}(\omega) = H_{In}(\omega)F(\omega). \quad (2)$$

With the method proposed in the current paper, any particular output RTF $H_{Out[i]}(\omega)$ is now generated as follows:

$$H_{Out[i]}(\omega) = H_{In}(\omega)F_i(\omega) \quad (3)$$

with $F_i(\omega)$ given by

$$F_i(\omega) = \left[a_{i1}(\omega)e^{j\phi_1(\omega)} + a_{i2}(\omega)e^{j\phi_2(\omega)} + \dots + a_{i(i-1)}(\omega)e^{j\phi_{i-1}(\omega)} + a_{ii}(\omega)e^{j\phi_i(\omega)} \right]. \quad (4)$$

This can be written as

$$H_{Out[i]}(\omega) = H_{In}(\omega) \sum_{n=1}^{n \leq i} a_{in}(\omega)e^{j\phi_n(\omega)}. \quad (5)$$

In the above equations, the a -coefficients are real variables and the ϕ_n 's are independent random variables of frequency with uniformly distributed values between $-\pi$ and π . Further, the terms $e^{j\phi_n(\omega)}$ represent the frequency response of the filters originally proposed by Kendall [5] and each filter $F_i(\omega)$ is a linear frequency-dependent combination of several of those.

With these definitions for $F_i(\omega)$, (2) can be rewritten in matrix form as

$$H_{Out}(\omega) = H_{In}(\omega)A(\omega)\Phi(\omega). \quad (6)$$

For notational convenience, the dependence on frequency is now omitted up to Section III but it needs to be emphasized that all variables are generally dependent on frequency. The matrix A and the vector Φ are then given by

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & 0 & \dots & 0 \\ \vdots & & & & & \vdots \\ a_{N_{Out}1} & a_{N_{Out}2} & a_{N_{Out}3} & a_{N_{Out}4} & \dots & a_{N_{Out}N_{Out}} \end{bmatrix}, \quad (7)$$

$$\Phi = \begin{bmatrix} e^{j\phi_1} \\ e^{j\phi_2} \\ \vdots \\ e^{j\phi_{N_{Out}}} \end{bmatrix}. \quad (8)$$

Note that matrix A is square with N_{Out} rows and columns and that $a_{11} = 1$. The unknowns in the filters F are the remaining weighting coefficients a_{in} in matrix A . The solution for these coefficients is found by 1) specifying the desired coherence values between all pairs of output RTFs and 2) constraining the magnitude response of the filters. The reason for the triangular structure of matrix A will become evident in the following sections.

In order to keep the magnitude in $H_{Out[i]}$ the same as in H_{In} , the following magnitude normalization constraints are introduced: ($i = [1, N_{Out}]$)

$$1 = \sum_{n=1}^{n \leq i} a_{in}^2. \quad (9)$$

It will be shown further below in Section II-F that these expressions are simplified and do not normalize the magnitude exactly but they are easily rectifiable. For a given value of N_{Out} , (9)

represents a total of $N_{\text{Out}} - 1$ equations and there are further $N_{\text{Out}}(N_{\text{Out}} - 1)/2$ coherence functions between all possible pairs of output RIRs. It can be seen from the matrix \mathbf{A} that there are also $N_{\text{Out}} - 1 + N_{\text{Out}}(N_{\text{Out}} - 1)/2$ unknowns and the system should therefore in principle be solvable. For this purpose, a formal definition of the coherence function is first required.

B. Definition of the Coherence Function

Several meanings and definitions of coherence exist in science and engineering and for a good comprehensive overview, the reader is referred to Gardner [9]. In the signal processing field, the coherence is a frequency-dependent measure of the linear time-invariant relationship between two or more signals [10]. In physics, the coherence function is linked to the variation of phase of a monochromatic wave across time or space [11], [12]. Confusingly, the mutual coherence function used in optics [12] is identical to what is known as the cross-correlation function in signal processing. In an audio context, the interaural coherence is used to describe the degree of similarity in a given frequency band between the signals present at the entrance to both ears [7], [13]. This similarity relates to the spatial auditory perception.

In the current paper, the function used to express the coherence between signals $x(t)$ and $y(t)$ is the signal processing definition given by [10]

$$\gamma_{xy}^2 = \frac{|S_{xy}|^2}{S_{xx}S_{yy}} \quad (10)$$

where S_{xx} and S_{yy} are power-spectral densities, S_{xy} is the cross-spectral density, and $|\cdot|$ denotes the magnitude function. The coherence function is a measure of the linear, time-invariant relationship between the two signals as a function of frequency. Its values are bounded between zero and one. Possible methods for the spectral density estimation are correlogram and periodogram methods. Examples of the latter are the methods attributed to Daniell and to Welch [14], [15].

In the Daniell periodogram method, the spectral densities are estimated by averaging the discrete Fourier transforms (DFTs) of $x(t)$ and $y(t)$ over a number of neighboring frequency bins (the averaging bandwidth $\Delta\omega_D$). In the more common Welch periodogram method, $x(t)$ and $y(t)$ are divided into (overlapping) sections whose discrete Fourier transforms are averaged. The DFT resolution resulting from the time sections is much coarser than the DFT resolution resulting from the whole signal length and implicitly the spectral magnitudes are distributed into the much larger frequency bins of width $\Delta\omega_W$.

With either the Daniell or Welch method for spectral density estimation, the coherence can be written as

$$\gamma_{xy}^2 = \frac{|\langle XY^* \rangle|^2}{\langle |X|^2 \rangle \langle |Y|^2 \rangle} \quad (11)$$

where $\langle \cdot \rangle$ denotes the respective averaging procedure and X and Y are the DFTs of $x(t)$ and $y(t)$. Note that the (magnitude-squared) correlation coefficient function $\rho_{xy}(\tau)$ of the bandpass filtered signals $x(t)$ and $y(t)$ evaluated at time lag $\tau = 0$ approaches the coherence function in (10) as the bandwidth $\Delta\omega$ of the bandpass filters tends to zero [9]. For a finite bandwidth,

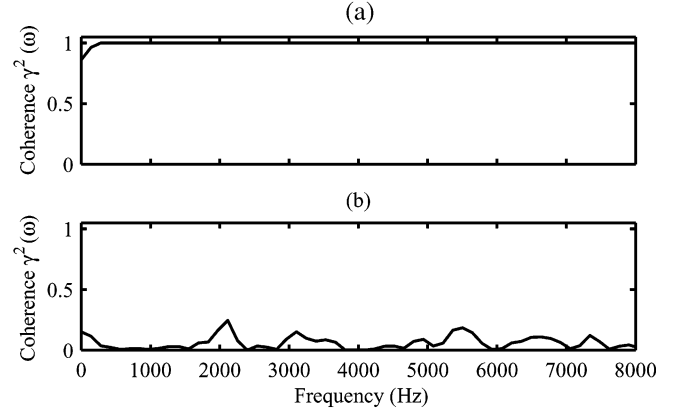


Fig. 2. Coherence between the input and output of filter $f_1(t)$ for (a) $\phi_1(\omega) = 0$ and (b) $\phi_1(\omega)$ has a uniform distribution between $-\pi$ and π . The filter has 1440 coefficient and at a sampling frequency of 48 kHz was designed with a DFT bin width of 33 Hz. The input signal consists of 115 200 samples of noise and the Welch method (Hanning window and 50% overlap) was used to estimate the coherence with a resulting DFT bin width of 141 Hz.

the coherence function can therefore also be interpreted as the correlation coefficient for that frequency band.

Since both the convolution in (1) and the reverberant sound field to be simulated are linear time-invariant systems, the coherence is by definition unity, but because of the finite spectral resolution resulting from the required averaging when calculating the spectral densities and because of the pseudo-stochastic phase response within the averaging bandwidth $\Delta\omega_D$ or $\Delta\omega_W$, this does not hold for the coherence estimate from a room transfer function (see also [16] and [17]) or the filters $\mathbf{F}(\omega)$ in (2). A pivotal feature of the filters is thus that for a zero coherence their phase response must be distributed evenly between $-\pi$ and π and the group delay must not be constant within the bandwidth $\Delta\omega_D/W$ involved in the estimation of the spectral densities. Conversely, the coherence between input and output signal of a filter whose phase variation with frequency is small or even zero within $\Delta\omega_D/W$ will be unity. This behavior is illustrated in Fig. 2, where the coherence between input and output of filter $f_1(t)$ with unit magnitude response is plotted for two different phase responses.

From a signal processing perspective, one can argue that the coherence estimate simply delivers erroneous results for these kind of signals but the key feature, and indeed the basic principle on which the method presented in the paper relies upon, is that the frequency resolution of human auditory perception is also finite. From an auditory perspective, similar arguments can thus be made if the bandwidth $\Delta\omega_D/W$ is replaced by the critical bandwidth of the auditory system. This claim is substantiated by Fig. 3 where the magnitude-squared correlation coefficient function $\rho(\tau)$ between the input and output of filter $f_1(t)$ is plotted for zero time lag τ after passing it through a Patterson filterbank [18], [19]. Note that the plots feature a logarithmic frequency axis that is more appropriate in this context.

With regards to the reader's understanding, it seems worthwhile to reinforce this major point as follows. The coherence function is used because a frequency-dependent measure of similarity is required. Calculating the correlation coefficient in sub-bands will lead to similar results as long as the bandwidth does

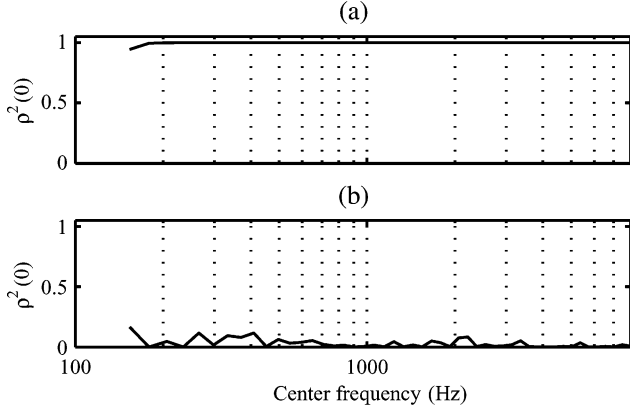


Fig. 3. Magnitude-squared correlation coefficient function at zero time lag between the input and output of filter $f_1(t)$ for (a) $\phi_1(\omega) = 0$ and (b) $\phi_1(\omega)$ has a uniform distribution between $-\pi$ and π . The filter has 1440 coefficients and at a sampling frequency of 48 kHz was designed with a DFT bin width of 33 Hz. The input signal consists of 115 200 samples of noise and the correlation coefficient was calculated for 50 bands of a Patterson auditory filterbank.

not vanish. As evidenced by the diffuse field coherence functions [16], [20], a frequency-dependent measure is necessary because the degree of similarity of the acoustic signals to be simulated is dependent on the wavelength.

C. Specific Solution for $N_{\text{Out}} = 2$

It would be possible to now proceed and assemble the $N_{\text{Out}}(N_{\text{Out}} - 1)/2$ equations and solve for the $N_{\text{Out}}(N_{\text{Out}} - 1)/2$ equations for any arbitrary value of N_{Out} , but because it is much easier to illustrate and discuss several fundamental issues for a small system, the solutions for $N_{\text{Out}} = 2$ and then $N_{\text{Out}} = 3$ are first presented before the solution process for arbitrary values of N_{Out} is outlined.

For $N_{\text{Out}} = 2$, the system defined in (2) or (6) consists of the following signals:

$$H_{\text{Out}1} = H_{\text{In}} e^{j\phi_1}, \quad (12a)$$

$$H_{\text{Out}2} = H_{\text{In}} [a_{21} e^{j\phi_1} + a_{22} e^{j\phi_2}]. \quad (12b)$$

Setting $a_{21} = 1$ and $a_{22} = 0$ results in $H_{\text{Out}1} = H_{\text{Out}2}$ and the coherence $\gamma_{H_{\text{Out}1}H_{\text{Out}2}}^2$ is then unity. The other extreme is setting $a_{21} = 0$ and $a_{22} = 1$, which is a similar case to that considered in Fig. 2(b) with a zero coherence. By setting the respective intermediate values for a_{21} and a_{22} , it should thus be possible to obtain any desired coherence value between zero and unity. Note that a_{21} and a_{22} are the minimum number of variables with which this can be achieved.

For the formal derivation of the coherence function $\gamma_{H_{\text{Out}1}H_{\text{Out}2}}^2$ that is required for the solution, it is now assumed that the spectral densities are calculated by the Daniell method and $\gamma_{H_{\text{Out}1}H_{\text{Out}2}}^2$ then follows as

$$\gamma_{H_{\text{Out}1}H_{\text{Out}2}}^2 = \frac{|\langle a_{21} + a_{22} e^{j\phi_1 - j\phi_2} \rangle|^2}{\langle a_{21}^2 + a_{22}^2 + a_{21}a_{22}(e^{j\phi_1 - j\phi_2} + e^{j\phi_2 - j\phi_1}) \rangle} \quad (13)$$

where $\langle \cdot \rangle$ denotes the averaging over the frequency bandwidth $\Delta\omega_D$. Due to the stochastic nature of ϕ_1 and ϕ_2 it is not possible to calculate the value of this equation but the factor $e^{j\phi_1 - j\phi_2}$ will vanish as $\Delta\omega_D$ tends to infinity. As long as the values of both weighting coefficients a_{21} and a_{22} within $\Delta\omega_D$

are of the same order of magnitude (and because of (9) they are bound between zero and unity), the terms $a_{22}e^{j\phi_1 - j\phi_2}$ and $a_{21}a_{22}(e^{j\phi_1 - j\phi_2} + e^{j\phi_2 - j\phi_1})$ will also vanish but with a slower convergence rate. (In the complex plane the term $\langle e^{j\phi_1 - j\phi_2} \rangle$ represents the average of a number of points distributed along the circumference of the unit circle. Similarly, $\langle a_{22}e^{j\phi_1 - j\phi_2} \rangle$ represents the average of a number of points with not only angular but also radial distribution. Because of the additional variation in radius a_{22} , it will require more averaging for the term to vanish.)

The resulting simplified expression for the coherence therefore is

$$\gamma_{H_{\text{Out}1}H_{\text{Out}2}}^2 = \frac{a_{21}^2}{a_{21}^2 + a_{22}^2}. \quad (14)$$

With the sole normalization constraint $a_{21}^2 + a_{22}^2 = 1$ from (9), the solution for the weighting coefficients in terms of the coherence is

$$a_{21} = \gamma_{H_{\text{Out}1}H_{\text{Out}2}}, \quad a_{22} = \sqrt{1 - \gamma_{H_{\text{Out}1}H_{\text{Out}2}}^2}. \quad (15)$$

This solution shows that, by setting the respective values for a_{21} and a_{22} , any desired coherence value ranging between zero and unity can be achieved.

A practical example is now considered with 61 600 samples of a RIR sampled at 48 kHz in a concert hall. Fig. 4(a) shows a time section of this RIR h_{In} . Fig. 4(b) and Fig. 4(c) show the same time sections of the output RIR $h_{\text{Out}1}$ and $h_{\text{Out}2}$ generated by convolution with the filters $f_1(t)$ and $f_2(t)$ having 1440 coefficients and a target coherence of $\gamma_{H_{\text{Out}1}H_{\text{Out}2}}^2 = 0.4$. From a visual comparison, all three RIRs are different but seem to share the same statistics. The experimental coherence estimate between the two output RIRs calculated with the Welch method and a DFT resolution of 64 Hz is shown in Fig. 4(d) and illustrates that the target coherence value has been achieved, although with some fluctuations. The causes for the fluctuations are investigated in Section II-F.

D. Specific Solution for $N_{\text{Out}} = 3$

As a second, more general but still relatively simple example, the case of $N_{\text{Out}} = 3$ is considered. The system therefore consists of the following signals:

$$H_{\text{Out}1} = H_{\text{In}} e^{j\phi_1} \quad (16a)$$

$$H_{\text{Out}2} = H_{\text{In}} [a_{21} e^{j\phi_1} + a_{22} e^{j\phi_2}] \quad (16b)$$

$$H_{\text{Out}3} = H_{\text{In}} [a_{31} e^{j\phi_1} + a_{32} e^{j\phi_2} + a_{33} e^{j\phi_3}]. \quad (16c)$$

The first two equations are identical to (12). In the third equation, the coefficients a_{31} and a_{32} allow for unity coherence between $H_{\text{Out}3}$ and $H_{\text{Out}1}$ and/or unity coherence between $H_{\text{Out}3}$ and $H_{\text{Out}2}$. The coefficient a_{33} in front of the additional phase term ϕ_3 allows for zero coherence between $H_{\text{Out}3}$ and $H_{\text{Out}1}$ as well as $H_{\text{Out}2}$. As with $N_{\text{Out}} = 2$, the coefficients appearing in (16) are the minimum number of variables with which all possible combinations of pairwise coherence can be achieved. This explains also the reason for the triangular structure of the matrix \mathbf{A} in (7) because any row in the matrix needs as many coefficients as the previous row (to allow for unity coherence) plus one additional coefficient (to allow for zero coherence).

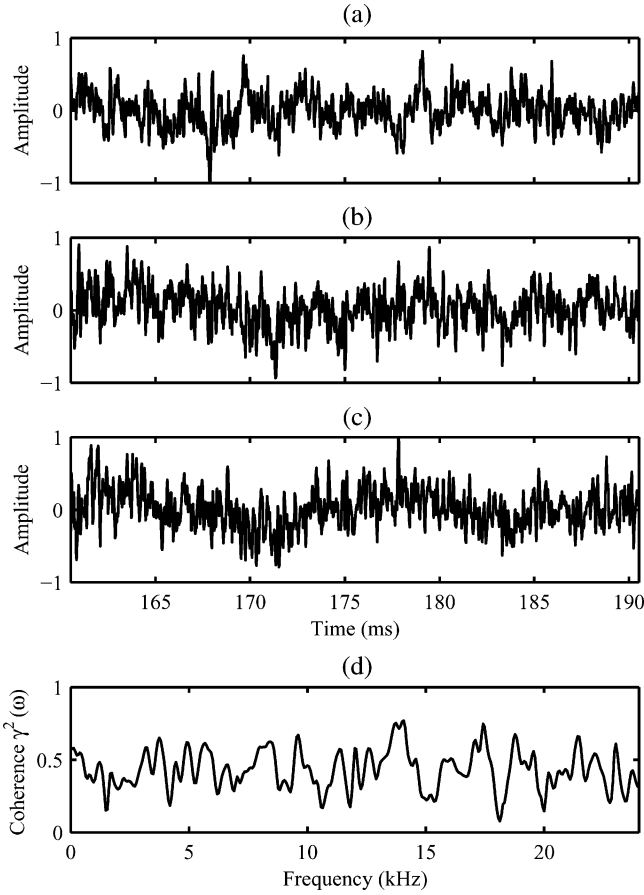


Fig. 4. Generation of RIRs with prescribed coherence on the example of a RIR measured in a concert hall. (a)–(c) Time sections of h_{In} , h_{Out1} and h_{Out2} , respectively. (d) Experimental coherence estimate $\gamma^2_{H_{Out1}H_{Out2}}(\omega)$, which fluctuates around the target value of 0.4.

Before proceeding any further with solving the system, an expression is first required for the coherence between any pair $H_{Out[i]}, H_{Out[l]}$ of output RIRs. Using analogous reasoning and assumptions that lead to (14), the coherence can be shown to be given by

$$\gamma^2_{H_{Out[i]}H_{Out[l]}} = \frac{\left| \sum_{n=1}^{N_{Out}} a_{in} a_{ln} \right|^2}{\left[\sum_{n=1}^{N_{Out}} a_{in}^2 \right] \left[\sum_{n=1}^{N_{Out}} a_{ln}^2 \right]}. \quad (17)$$

In (16), there are five unknowns, and the corresponding five equations to be solved stem from three coherence functions and two power normalization equations given by

$$\gamma^2_{H_{Out1}H_{Out2}} = \frac{a_{21}^2}{a_{21}^2 + a_{22}^2} \quad (18a)$$

$$\gamma^2_{H_{Out1}H_{Out3}} = \frac{a_{31}^2}{a_{31}^2 + a_{32}^2 + a_{33}^2} \quad (18b)$$

$$\gamma^2_{H_{Out2}H_{Out3}} = \frac{(a_{21}a_{31} + a_{22}a_{32})^2}{(a_{21}^2 + a_{22}^2)(a_{31}^2 + a_{32}^2 + a_{33}^2)} \quad (18c)$$

$$1 = a_{21}^2 + a_{22}^2 \quad (18d)$$

$$1 = a_{31}^2 + a_{32}^2 + a_{33}^2. \quad (18e)$$

Combining the first and penultimate equations and combining the second and last equation, it follows that $a_{21} = \gamma_{H_{Out1}H_{Out2}}$

and $a_{31} = \gamma_{H_{Out1}H_{Out3}}$. Inserting these two solutions into the last two equations results in $a_{22} = \sqrt{1 - \gamma_{H_{Out1}H_{Out2}}^2}$ and $a_{32} = \sqrt{1 - \gamma_{H_{Out1}H_{Out3}}^2 - a_{33}^2}$. The third equation then reads

$$\gamma_{H_{Out2}H_{Out3}} = \frac{\gamma_{H_{Out1}H_{Out2}} \gamma_{H_{Out1}H_{Out3}}}{\sqrt{1 - \gamma_{H_{Out1}H_{Out2}}^2} \sqrt{1 - \gamma_{H_{Out1}H_{Out3}}^2 - a_{33}^2}} \quad (19)$$

which can be rearranged for the unknown as

$$a_{33} = \left(1 - \frac{(\gamma_{H_{Out2}H_{Out3}} - \gamma_{H_{Out1}H_{Out2}} \gamma_{H_{Out1}H_{Out3}})^2}{1 - \gamma_{H_{Out1}H_{Out2}}^2 - \gamma_{H_{Out1}H_{Out3}}^2} \right)^{1/2}. \quad (20)$$

Note that out of the four possible solutions the one with all positive square roots has been chosen. It is also important to note that the solution found for a_{21} and a_{22} is identical to that given by (15).

It is worth emphasizing that a solution cannot be found for arbitrary values of the prescribed coherence functions. The extreme example is $\gamma_{H_{Out1}H_{Out2}}^2 = 1$, $\gamma_{H_{Out1}H_{Out3}}^2 = 1$ and $\gamma_{H_{Out2}H_{Out3}}^2 = 0$. This would imply that two signals are both fully coherent to a third signal but are fully incoherent to each other and to the current author this seems impossible to ever occur with a physical system. From a mathematical point of view, no solution can be found because (20) becomes singular. Since the purpose of the presented method is to emulate the response of a real acoustic system (an example is given in Section III), the corresponding values for the coherence functions assure that such unphysical solutions are not required.

E. General Solution Procedure for $N_{Out} > 3$

The procedure to solve for $N_{Out} > 3$ follows along analogous lines to that for $N_{Out} = 2$ and $N_{Out} = 3$. In general, the solution for $N_{Out} + 1$ recycles the solution for N_{Out} and requires the solving of another $N_{Out} + 1$ equations for the $N_{Out} + 1$ additional unknown variables. Due to the large number of terms involved in these systems, explicit expressions for the analytical solutions are not presented, but the solution strategy is always the same and given as follows:

1. Find $a_{N_{Out}1}$ from $a_{N_{Out}1}^2 = \gamma_{H_{Out1}H_{OutN_{Out}}}^2$.
2. Solve $\gamma_{H_{Out2}H_{OutN_{Out}}}^2$ for $a_{N_{Out}2}$.
3. Solve $\gamma_{H_{Out3}H_{OutN_{Out}}}^2$ for $a_{N_{Out}3}$.
- \vdots

$N_{Out} - 1$. Solve $\gamma_{H_{Out(N_{Out}-1)}H_{OutN_{Out}}}^2$ for

$$a_{N_{Out}(N_{Out}-1)}.$$

N_{Out} . Find $a_{N_{Out}N_{Out}}$ from $a_{N_{Out}N_{Out}}^2 =$

$$1 - a_{N_{Out}1}^2 - a_{N_{Out}2}^2 - \dots - a_{N_{Out}(N_{Out}-1)}^2.$$

For the practical applications concerned, the author has only searched and found solutions for $N_{Out} \leq 5$, but there is no indication that solutions for arbitrary large values of N_{Out} do not exist. Moreover, it is also conceivable and indeed expected that for large values of N_{Out} , the prescribed

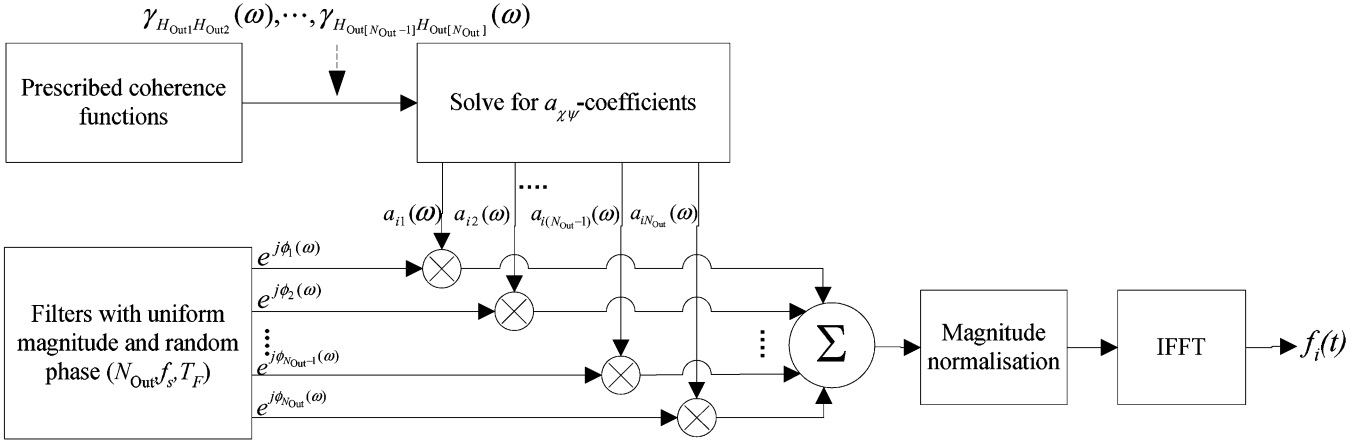


Fig. 5. Flow diagram for the generation of the filters $f_i(t)$ from the prescribed coherence functions.

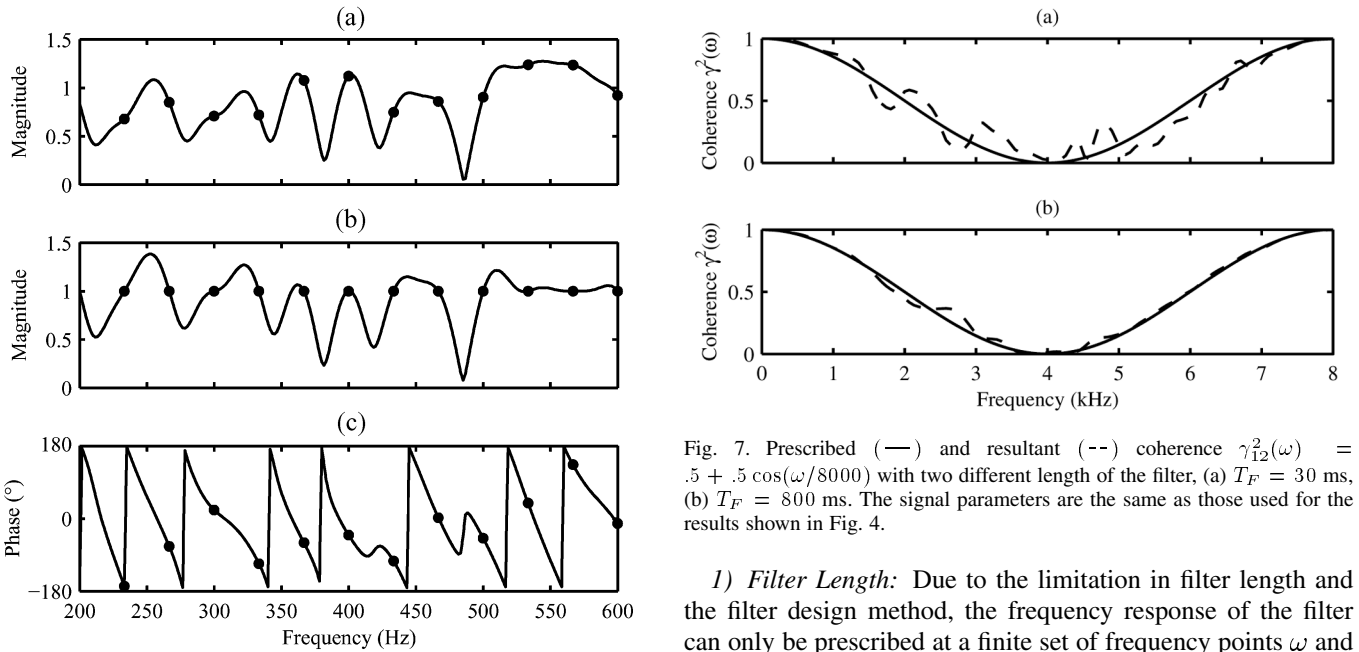


Fig. 6. Magnitude response (a) before and (b) after magnitude equalization and (c) phase response of one of the filters $f_i(t)$ at the prescribed frequency grid points (o) and with a higher frequency resolution (—).

coherence values are only significantly above zero for a limited set of RTF pairs ranging between $(H_{Out[i]}, H_{Out[i-n]})$ and $(H_{Out[i]}, H_{Out[i+n]})$, where n is a number much smaller than N_{Out} . In this case, the matrix \mathbf{A} defined in (7) is a band matrix with bandwidth $2n + 1$ and a corresponding reduced number of unknowns. For large N_{Out} problems may arise due to the “randomness” of the random number generator.

Finally, an overview of the building steps starting from the prescribed coherence function values to the resulting filters $f_i(t)$ in the time domain is given in Fig. 5. A further comment is that the method can also be extended for the case that there are two or more input RTFs available (see [21] for details).

F. Limitations

It is evident in Fig. 4(d) that the coherence estimates fluctuate around the target value and it was also remarked that the magnitude normalization constraints in (9) are not exact. Both issues are now investigated in more detail.

Fig. 7. Prescribed (—) and resultant (--) coherence $\gamma_{12}^2(\omega) = .5 + .5 \cos(\omega/8000)$ with two different length of the filter, (a) $T_F = 30$ ms, (b) $T_F = 800$ ms. The signal parameters are the same as those used for the results shown in Fig. 4.

1) *Filter Length:* Due to the limitation in filter length and the filter design method, the frequency response of the filter can only be prescribed at a finite set of frequency points ω and in-between is determined by the complex amplitudes at these points. As a consequence, the phase response of the filters is no longer random and, as Fig. 6(a) illustrates, the magnitude response does deviate from unity between the frequency points where the response is prescribed. For room acoustics applications and if these magnitude variations are not too large, this feature is actually desirable for the reason that, above the Schroeder frequency, the RTFs recorded at different locations and directions within a room also do not have their peaks and dips at the same frequencies.

The fact that the phase response in-between frequency points is no longer a random function of frequency means that the terms in (13) containing the factor $e^{j\phi_1 - j\phi_2}$ will no longer vanish and the resulting coherence will deviate from the target coherence. The larger T_F , and therefore the larger the density of frequency points with the desired complex amplitudes, the better the target coherence can be achieved. This is illustrated in Fig. 7 for two different lengths T_F of the filters. It can be seen that with the longer filter length the agreement between target and resultant coherence is improved.

2) *Complex Summation:* It has already been remarked that the magnitude normalization constraints in (9) are not exact. The

reason is that the correct calculation of the magnitude involves summing real and imaginary parts first and thereafter squaring them separately. The error introduced by the simplification in (9) is now considered.

For $N_{\text{Out}} = 2$, the correct magnitude normalization reads

$$1 = (a_{21} \cos \phi_1 + a_{22} \cos \phi_2)^2 + (a_{21} \sin \phi_1 + a_{22} \sin \phi_2)^2 \\ = a_{21}^2 + 2a_{21}a_{22} \cos(\phi_1 - \phi_2) + a_{22}^2. \quad (21)$$

The expression for the coherence remains unchanged and when solving it for a_{21} , it yields

$$a_{21} = \sqrt{\frac{a_{22}^2 \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2}{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2}}. \quad (22)$$

Inserting into the equation for the power normalization and solving for a_{22} results in

$$a_{22} = \sqrt{\frac{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2}{1 + 2\gamma_{H_{\text{Out}1} H_{\text{Out}2}} \sqrt{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2} \cos(\phi_1 - \phi_2)}}. \quad (23)$$

Finally, reinserting into the expression for a_{21} gives

$$a_{21} = \sqrt{\frac{\gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2}{1 + 2\gamma_{H_{\text{Out}1} H_{\text{Out}2}} \sqrt{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2} \cos(\phi_1 - \phi_2)}}. \quad (24)$$

The difference to the solution in (15) is the denominator, which will be unity for $\phi_1 = \pi/2$.

Alternatively, and since the coherence functions are invariant to linear scaling, it is possible to normalize the filter magnitude after having solved the system with the inaccurate expressions for the magnitude normalization. This additional normalization step simply is

$$F'_2 = \frac{F_2}{|F_2|} = \frac{a_{21} e^{j\phi_1} + a_{22} e^{j\phi_2}}{\sqrt{a_{21}^2 + 2a_{21}a_{22} \cos(\phi_1 - \phi_2) + a_{22}^2}} \quad (25)$$

and when inserting the solution of (15), it results in

$$F'_2 = \frac{\gamma_{H_{\text{Out}1} H_{\text{Out}2}} e^{j\phi_1} + \sqrt{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2} e^{j\phi_2}}{\sqrt{1 + 2\gamma_{H_{\text{Out}1} H_{\text{Out}2}} \sqrt{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2} \cos(\phi_1 - \phi_2)}} \\ = \frac{\gamma_{H_{\text{Out}1} H_{\text{Out}2}}}{\sqrt{1 + 2\gamma_{H_{\text{Out}1} H_{\text{Out}2}} \sqrt{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2} \cos(\phi_1 - \phi_2)}} e^{j\phi_1} \\ + \frac{\sqrt{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2}}{\sqrt{1 + 2\gamma_{H_{\text{Out}1} H_{\text{Out}2}} \sqrt{1 - \gamma_{H_{\text{Out}1} H_{\text{Out}2}}^2} \cos(\phi_1 - \phi_2)}} \\ \times e^{j\phi_2}. \quad (26)$$

The expressions in front of the factors $e^{j\phi_1}$ and $e^{j\phi_2}$ are exactly identical to (23) and (24) and therefore the normalization in (25) yields the same results as with the accurate magnitude normalization expression. Fig. 6(a) and (b) shows the magnitude response of the filter before and after magnitude equalization and Fig. 6(c) shows the (unaltered) phase response.

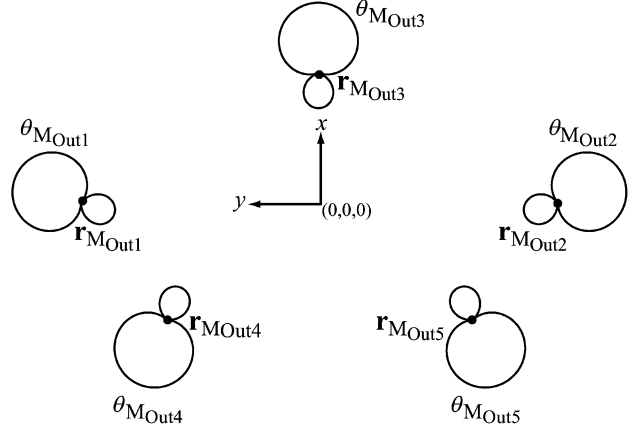


Fig. 8. Diagrammatic representation of surround sound microphone setup.

Since neither method allows control over the magnitude response of the filter between the frequency grid points, the filter length T_F is still the main limiting factor in achieving the desired coherence values as accurately as possible. A potential alternative filter design method is to directly optimize the set of FIR filter coefficients such that they yield the target coherence function values not just at single frequencies but within a given frequency range. Whether or not this approach provides better overall results is a question left open for further research.

III. APPLICATION TO VIRTUAL MICROPHONE SETUP IN REVERBERANT SOUND FIELD

One possible application of the method presented in the first part of this paper is the simulation of a multichannel RIR recording with a surround sound microphone setup. A typical setup for a five-channel reproduction system is illustrated in Fig. 8.

The position and directivity of microphone $M_{\text{Out}[i]}$ in the array are given by $\mathbf{r}_{M_{\text{Out}[i]}}$ and $(b_{w,M_{\text{Out}[i]}}, b_{xy,M_{\text{Out}[i]}})$, respectively. Measured from the x -axis, the direction is defined by the azimuth angle $\theta_{M_{\text{Out}[i]}}$. The microphone directivity and direction define the microphone directivity factor $\Gamma_{M_{\text{Out}[i]}}(\theta, \phi)$ as follows:

$$\Gamma_{M_{\text{Out}[i]}}(\theta, \phi) = b_{w,M_{\text{Out}[i]}} + b_{xy,M_{\text{Out}[i]}} \cos(\theta) \cos(\theta_{M_{\text{Out}[i]}}) \cos \phi \\ + b_{xy,M_{\text{Out}[i]}} \sin(\theta) \sin(\theta_{M_{\text{Out}[i]}}) \cos \phi \quad (27)$$

with ϕ the elevation angle from the xy -plane. A cardioid microphone $(1 + \cos(\theta - \theta_{M_{\text{Out}[i]}}))/2$ pointing in the direction $\theta_{M_{\text{Out}[i]}}$ is achieved by nominally setting $b_{w,M_{\text{Out}[i]}} = b_{xy,M_{\text{Out}[i]}} = 1/2$ but these two coefficients are also used to control the relative sensitivity of the individual microphones in the multichannel setup.

Example values for $\mathbf{r}_{M_{\text{Out}[i]}}$, $b_{w,M_{\text{Out}[i]}}$, $b_{xy,M_{\text{Out}[i]}}$, and $\theta_{M_{\text{Out}[i]}}$ for a five-channel surround sound microphone setup suggested by Williams [22], [23] are given in Table I.

A. Coherence Functions

Unfortunately, it is not possible to derive analytical expressions for the coherence between any pair of the microphones in the array without knowledge of the type of sound field in which the recording is to be simulated. But, after an initial temporal

TABLE I
POSITION $\mathbf{r}_{\text{MOut}[i]}$, DIRECTION $\theta_{\text{MOut}[i]}$ AND DIRECTIVITY
 $b_{w,\text{MOut}[i]}, b_{xy,\text{MOut}[i]}$ OF THE MICROPHONES FOR
A SPACED FIVE-CHANNEL MICROPHONE SETUP

i	$\mathbf{r}_{\text{MOut}[i]}$ (m)	$b_{w,\text{MOut}[i]}$	$b_{xy,\text{MOut}[i]}$	$\theta_{\text{MOut}[i]}$ ($^\circ$)
1	(0,0.44)	0.5	0.5	70
2	(0,-0.44)	0.5	0.5	290
3	(0.23,0)	0.5	0.5	0
4	(-0.23,0.28)	0.5	0.5	156
5	(-0.23,-0.28)	0.5	0.5	204

buildup, the sound field in many enclosed spaces can be characterized as being reverberant. With the appropriate averaging procedure, the following expression for the diffuse field coherence function between microphones $\text{MOut}[i]$ and $\text{MOut}[l]$ is then applicable [20], [24]

$$\begin{aligned} \gamma_{\text{MOut}[i]\text{MOut}[l]}^2(r, \omega) &= \frac{1}{D_i D_l} [b_{p,i} b_{p,l} \gamma_{pp}(r, \omega) + b_{v_{\perp},i} b_{v_{\perp},l} \gamma_{v_{\perp}v_{\perp}}(r, \omega) / 3 \\ &\quad + b_{v_{\parallel},i} b_{v_{\parallel},l} \gamma_{v_{\parallel}v_{\parallel}}(r, \omega) / 3]^2 \\ &\quad + \frac{1}{D_i D_l} [(b_{p,i} b_{v_{\parallel},l} + b_{v_{\parallel},i} b_{p,l}) \gamma_{pv_{\parallel}}(r, \omega) / \sqrt{3}]^2 \end{aligned} \quad (28)$$

with $r = |\mathbf{r}_{\text{MOut}[i]} - \mathbf{r}_{\text{MOut}[l]}|$, D_i given by

$$D_i = b_{p,i}^2 + \frac{b_{v_{\perp},i}^2}{3} + \frac{b_{v_{\parallel},i}^2}{3} \quad (29)$$

and an analogous expression holds for D_l . The functions $\gamma_{pp}(r, \omega)$, $\gamma_{pv_{\parallel}}(r, \omega)$, $\gamma_{v_{\perp}v_{\perp}}(r, \omega)$ and $\gamma_{v_{\parallel}v_{\parallel}}(r, \omega)$ are the diffuse field coherence functions between pressure and/or particle velocity components parallel and perpendicular to the line r . They are given by

$$\gamma_{pp}^2(r, \omega) = \left[\frac{\sin(kr)}{kr} \right]^2 \quad (30a)$$

$$\gamma_{pv_{\parallel}}^2(r, \omega) = 3 \left[\frac{\sin(kr) - (kr) \cos(kr)}{(kr)^2} \right]^2 \quad (30b)$$

$$\gamma_{v_{\perp}v_{\perp}}^2(r, \omega) = 9 \left[\frac{\sin(kr) - (kr) \cos(kr)}{(kr)^3} \right]^2 \quad (30c)$$

$$\gamma_{v_{\parallel}v_{\parallel}}^2(r, \omega) = 9 \left[\frac{(kr)^2 \sin(kr) + 2kr \cos(kr) - 2 \sin(kr)}{(kr)^3} \right]^2 \quad (30d)$$

with k the acoustic wave number corresponding to ω . Finally, the parameters $b_{p,i}$, $b_{v_{\perp},i}$, and $b_{v_{\parallel},i}$ are related to $b_{w,\text{MOut}[i]}$, $b_{xy,\text{MOut}[i]}$, and $\theta_{\text{MOut}[i]}$ as follows:

$$b_{p,i} = b_{w,\text{MOut}[i]} \quad (31a)$$

$$b_{v_{\perp},i} = b_{xy,\text{MOut}[i]} \cos(\theta_{\text{MOut}[i]} - \theta_{il}) \quad (31b)$$

$$b_{v_{\parallel},i} = b_{xy,\text{MOut}[i]} \sin(\theta_{\text{MOut}[i]} - \theta_{il}) \quad (31c)$$

(31) performs a decomposition of the microphone direction and directivity into components parallel and perpendicular to the

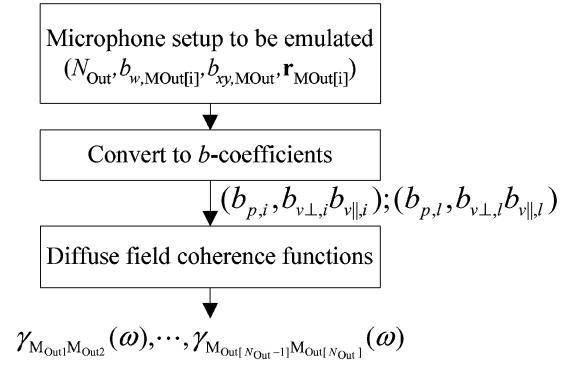


Fig. 9. Flow diagram for the calculation of the coherence function values from the parameters of the to be simulated microphone setup.

line r . The angle θ_{il} is determined by the direction of the vector $\mathbf{r}_{\text{MOut}[i]} - \mathbf{r}_{\text{MOut}[l]}$ and found from

$$\sin \theta_{il} = - \frac{(\mathbf{r}_{\text{MOut}[i]} - \mathbf{r}_{\text{MOut}[l]}) \cdot \hat{x}}{|\mathbf{r}_{\text{MOut}[i]} - \mathbf{r}_{\text{MOut}[l]}|} \quad (32)$$

where \hat{x} is the unit vector on the x -axis of Fig. 8.

An overview of the steps required to obtain the coherence function values from the parameters of the to be emulated multichannel microphone setup is given in Fig. 9. The sequence of the equations is such that (32) is used to calculate θ_{il} , which is inserted into (31). Thereafter, the results from (31) are used to calculate the values of the coherence functions using (28).

B. Example

As an illustrative example, the configuration of five omnidirectional microphones situated at the corners of a regular pentagon with side length 0.05 m is considered. The microphones are numbered consecutively along the outline of the pentagon. This example is particularly simple because 1) there are only two different inter-microphone distances of 0.05 m and 0.08 m, and 2) for every microphone $b_{w,\text{MOut}[i]} = 1$ and $b_{xy,\text{MOut}[i]} = 0$. The only nonzero term in (28) is thus the first term. The solid curves in Fig. 10 show the theoretical diffuse field coherence functions between all possible microphone combinations. Note that because there are only two different inter-microphone distances there also only two different types of curves in the figure.

The values of the theoretical coherence functions have been used to obtain the a -coefficients needed to design the filters $\mathbf{f}(t)$ as described in Section II and outlined in Fig. 5. These filters are then convolved with the reverberant tail of a single RIR measured in a concert hall to yield the total of five output RIRs. The signal parameters are the same as those used for the results shown in Fig. 4. The dashed curves in Fig. 10 show the resulting coherence estimate between all possible combinations of the output RIRs. With the exception of some fluctuations, it is evident that the target coherence functions have been achieved quite well.

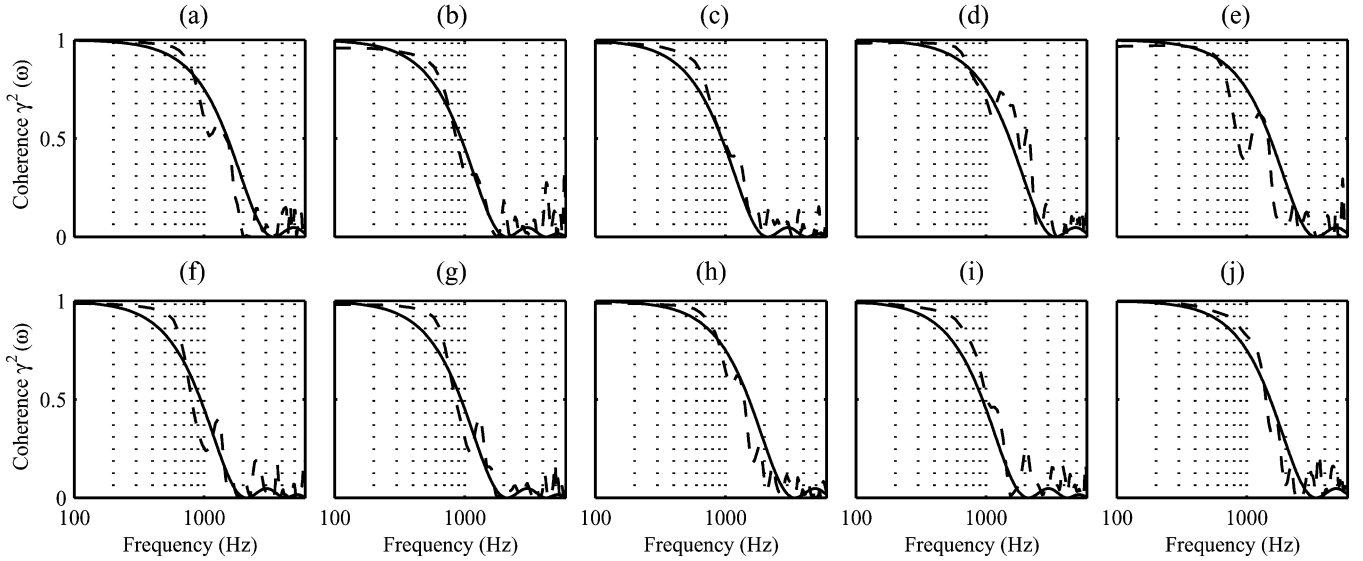


Fig. 10. Theoretical diffuse field coherence (—) for five omnidirectional microphones situated at the corners of a regular pentagon with side length 0.05 m and estimated coherence (---) between all output RIRs generated with the presented method, (a) $\gamma_{12}^2(\omega)$, (b) $\gamma_{13}^2(\omega)$, (c) $\gamma_{14}^2(\omega)$, (d) $\gamma_{15}^2(\omega)$, (e) $\gamma_{23}^2(\omega)$, (f) $\gamma_{24}^2(\omega)$, (g) $\gamma_{25}^2(\omega)$, (h) $\gamma_{34}^2(\omega)$, (i) $\gamma_{35}^2(\omega)$, (j) $\gamma_{45}^2(\omega)$.

C. Diffuse Field Power Gain

With the processing suggested up to this point, all output RIRs have the same power as the input RIR. In a further step, it is therefore necessary to adjust the relative power of the output RIRs according to the directivity and relative gain of both the output microphones to be simulated and the assumed microphone with which the input RIR has been recorded. Correcting for the latter is only necessary if the absolute gains are relevant. Suppose the directivity factor $\Gamma_{M_{In}}(\theta, \phi)$ of the input microphone is defined analogously to (27), the gain $g_{Out[i]}$ to be applied to output RIR $h_{Out[i]}$ can then be calculated from an integration of both directivity factors over the solid angle 4π

$$g_{Out[i]}^2 = \frac{\int_0^\pi \int_0^{2\pi} \Gamma_{M_{Out[i]}}^2(\theta, \phi) \cos \phi d\theta d\phi}{\int_0^\pi \int_0^{2\pi} \Gamma_{M_{In}}^2(\theta, \phi) \cos \phi d\theta d\phi}. \quad (33)$$

This equation is valid under the assumption that sound is incident from all directions with equal magnitude.

IV. CONCLUSION

A method has been presented to generate multichannel room impulse responses (RIRs) from a given input RIR by convolution with a set of filters. The filters have theoretically unit magnitude and random phase response and their coefficients are found from the user-prescribed coherence function values and magnitude normalization constraints. It was found that there are limitations in terms of achieving both unit magnitude and the desired coherence values and they are caused by the required short filter length and, potentially, the filter design method.

In the second part of the paper, it was shown how the method can be used with known diffuse field coherence functions to emulate the recording of RIRs with a surround sound microphone setup. The illustrative example consisted of five omnidirectional microphones situated at the corners of a regular pentagon and

it was shown that the prescribed coherence values between all possible pairs of output RIRs have been achieved quite closely.

A limitation of the presented method is that it is not suitable for the generation of the direct sound and early reflections in the RIRs. For this reason, the current author has developed a separate method for the generation of this part of the RIR. A companion paper presenting this method and several objective and subjective performance evaluations for both methods is currently under review.

ACKNOWLEDGMENT

The author would like to thank M. van Walstijn for general discussion on the subject matter. The author would also like to thank the two anonymous reviewers for valuable comments that lead to various improvements in the paper.

REFERENCES

- [1] H. Lauridsen, "Experiments concerning different kinds of room-acoustics recording," (in Danish) *Ingeniøren*, vol. 47, pp. 906–910, 1954.
- [2] M. R. Schroeder, "An artificial stereophonic effect obtained from a single audio signal," *J. Audio Eng. Soc.*, vol. 6, no. 2, pp. 74–79, Apr. 1958.
- [3] J. Merimaa and V. Pulkki, "Spatial impulse response rendering I: Analysis and synthesis," *J. Audio Eng. Soc.*, vol. 53, no. 12, pp. 1115–1127, Dec. 2005.
- [4] V. Pulkki and J. Merimaa, "Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests," *J. Audio Eng. Soc.*, vol. 54, no. 1/2, pp. 3–20, Jan. 2006.
- [5] G. S. Kendall, "The decorrelation of audio signals and its impact on spatial imagery," *Comput. Music J.*, vol. 19, no. 4, pp. 71–87, 1995.
- [6] M. O. J. Hawksford and N. Harris, "Diffuse signal processing and acoustic source characterization for applications in synthetic loudspeaker arrays," in *Proc. 112th Conv. Audio Eng. Soc.*, Munich, Germany, May 2002, paper 5612.
- [7] C. Faller, "Parametric multichannel audio coding: Synthesis of coherence cues," *IEEE Trans. Speech Audio Process.*, vol. 14, no. 1, pp. 299–310, Jan. 2006.
- [8] K. Meesawat, "A study of the reverberation tail in binaural room impulse responses," Ph.D. dissertation, Aalborg Univ., Aalborg, Denmark, Dec. 2004.
- [9] W. A. Gardner, "A unifying view of coherence in signal processing," *Signal Process.*, vol. 29, no. 2, pp. 113–140, Nov. 1992.

- [10] J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis*, 1st ed. New York: Wiley, 1980.
- [11] M. Born and E. Wolf, *Principles of Optics*. New York: Pergamon, 1959.
- [12] E. Hecht, *Optics*, 2nd ed. Wokingham, U.K.: Addison-Wesley, 1987.
- [13] C. Faller and J. Merimaa, "Source localization in complex listening situations: Selection of binaural cues based on interaural coherence," *J. Acoust. Soc. Amer.*, vol. 116, no. 5, pp. 3075–3089, Nov. 2004.
- [14] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*. London, U.K.: Prentice-Hall, 1975, ch. 11.
- [15] S. L. Marple, *Digital Spectral Analysis*. London, U.K.: Prentice-Hall, 1987.
- [16] F. Jacobsen and T. Roisin, "The coherence of reverberant sound fields," *J. Acoust. Soc. Amer.*, vol. 108, no. 1, pp. 204–210, Jul. 2000.
- [17] F. Jacobsen and T. G. Nielsen, "Spatial correlation and coherence in a reverberant sound field," *J. Sound Vib.*, vol. 118, no. 1, pp. 175–180, Oct. 1987.
- [18] R. D. Patterson, K. Robinson, J. Holdsworth, D. McKeown, C. Zhang, and M. H. Allerhand, "Complex sounds and auditory images," in *Auditory Physiology and Perception*, 1st ed. Oxford, U.K.: Pergamon, 1992, pp. 429–446.
- [19] M. Stanley, "An efficient implementation of the Patterson–Holdsworth auditory filter bank," Apple Computer, Inc., Tech. Rep. 35, 1993.
- [20] M. Kuster, "Spatial correlation and coherence in reverberant acoustic fields: Extension to microphones with arbitrary first-order directivity," *J. Acoust. Soc. Amer.*, vol. 123, no. 1, pp. 154–162, Jan. 2008.
- [21] M. Kuster, "Inverse methods in room acoustics with under-determined data and applications to virtual acoustics," Ph.D. dissertation, Queen's Univ. Belfast, Belfast, U.K., Nov. 2007.
- [22] M. Williams and G. L. Dû, "Microphone array analysis for multi-channel sound recording," in *Proc. 107th Conv. Audio Eng. Soc.*, New York, Sep. 1999, paper 4997.
- [23] F. Rumsey, *Spatial Audio*, 1st ed. Oxford, U.K.: Focal, 2001.
- [24] M. Kuster and M. van Walstijn, "Spatial coherence between microphones with arbitrary first-order directivity in reverberant acoustic fields," in *Proc. 19th Int. Congr. Acoust.*, Madrid, Spain, Sep. 2007, paper RBA-15–009.



Martin Kuster received the B.Eng. degree in electroacoustics from the University of Salford, Salford, U.K., in 2001, the M.Sc. degree in applied physics from Delft University of Technology, Delft, The Netherlands, in 2003, and the Ph.D. degree in acoustics from Queen's University Belfast, Belfast, U.K., in 2007.

He was with the Laboratory of Acoustic Imaging and Sound Control, Delft University of Technology, Delft, The Netherlands. He is currently working in the NVH Department of an automotive OEM parts supplier in Germany. Previously, he had been researching the use of acoustic imaging and inverse methods in room acoustics and the applicability of bending wave transducer technology in the context of wave field synthesis. He also participated in a collaborative project with a digital entertainment company for the improvement of surround sound technology.

Dr. Kuster is a member of the Acoustical Society of America, the Audio Engineering Society, and the German Acoustical Society.