Decorrelation in Feedback Delay Networks

Sebastian J. Schlecht, *Senior Member, IEEE*, Jon Fagerström, and Vesa Välimäki, *Fellow, IEEE*

***Abstract*—The feedback delay network (FDN) is a popular filter structure to generate artificial spatial reverberation. A common requirement for multichannel late reverberation is that the output signals are well decorrelated, as too high a correlation can lead to poor reproduction of source image and uncontrolled coloration. This article presents the analysis of multichannel correlation induced by FDNs. It is shown that the correlation depends primarily on the feedforward paths, while the long reverberation tail produced by the recursive path does not contribute to the inter-channel correlation. The impact of the feedback matrix type, size, and delays on the inter-channel correlation is demonstrated. The results show that small FDNs with a few feedback channels tend to have a high inter-channel correlation, and that the use of a filter feedback matrix significantly improves the decorrelation, often leading to the lowest inter-channel correlation among the tested cases. The learnings of this work support the practical design of multichannel artificial reverberators for immersive audio applications.**

***Index Terms*—Correlation, digital filters, feedback circuits, spatial audio, spatial filters, reverberation.**

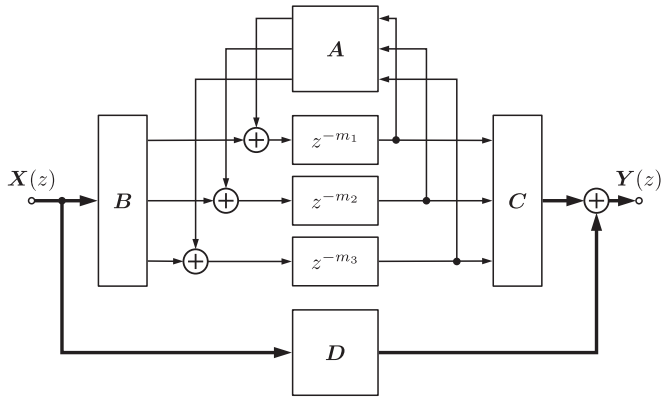


Fig. 1. MIMO FDN with three delay lines, i.e., $N = 3$, and a 3-by-3 feedback matrix $\boldsymbol{A}$. Thick lines indicate multiple channels, while thin lines indicate individual channels.

I. Introduction

INTER-CHANNEL correlation of multichannel signals is a central metric in spatial audio capture and reproduction. Correlation is also tightly connected to the room acoustic process, where late reverberation tends towards a diffuse field condition in which the positions and directions of sound waves are uncorrelated. The feedback delay network (FDN) is an efficient filter structure, which is commonly used in audio and music technology to generate multichannel late reverberation [1], [2], [3], see Fig. 1. However, relatively little is known about the inter-channel correlation of FDNs. A naive notion is that a long reverberation tail leads to low correlation, i.e. good decorrelation. This does not necessarily hold for the FDN or any other recursive delay-based filter topology. Further, this article presents a framework to compute and design the FDN inter-channel correlation.

Audio decorrelation influences the perceived stereo image and source width [4]. In spatial multichannel reproduction, inter-channel correlation leads to position-dependent coloration due to varying delays and interfering summation [5], [6]. In binaural reverberation, matching the interaural cross-correlation coefficient improves the perceived spatial extent [7]. Various spatial reverberation techniques were developed to generate spatially diffuse reverberation [8], anisotropic reverberation [9], [10], Ambisonics-based rendering [11], [12], coupled rooms [13], [14], or reverberation with prescribed correlation [15]. Low multichannel correlation can also be generated by recombining multiple uncorrelated channels by an orthogonal mixing matrix [16], [17]. Correlation can be reduced by additional decorrelation filtering, such as with a white noise sequence [4], subband techniques [18], an allpass filter [19], [20], or a velvet-noise sequence [21], [22], [23]. Occasionally, it is assumed that the FDN outputs are fully uncorrelated for multichannel playback [10], [24], [25]. There, the output is mixed with an orthogonal matrix to maintain the level of decorrelation. Alternatively, the output channels are mixed to achieve the desired correlation, such as in binaural reproduction [26], [27].

Decorrelation is often thought of as a multi-output property as we desire uncorrelated signals for spatial reproduction. Nonetheless, input correlation might be a beneficial property for multichannel inputs which are not entirely uncorrelated. For instance, a multi-microphone recording of a sound scene might exhibit correlation and lead to coloration when the channels are summed together without sufficient decorrelation [28], [29]. Similarly, digital multichannel sound productions can also contain considerable correlation between the channels due to production techniques such as panning and Ambisonics encoding. Thus, in this work, the FDN is formulated as a multi-input, multi-output (MIMO) system. As an application, we study also the single-input multi-output (SIMO) system, which is the most common configuration for spatial reverberation effects.

Manuscript received 31 August 2022; revised 16 May 2023; accepted 20 June 2023. Date of publication 14 September 2023; date of current version 20 October 2023. This work was supported by the Nordic Sound and Music Computing Network—NordicSMC, NordForsk under Grant 86892. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Romain Serizel. *(Corresponding author: Sebastian J. Schlecht.)*

Sebastian J. Schlecht is with Acoustics Lab, Department of Information and Communications Engineering, Aalto University, FI-02150 Espoo, Finland, and also with Media Lab, Department of Art and Media, Aalto University, FI-02150 Espoo, Finland (e-mail: sebastian.schlecht@aalto.fi).

Jon Fagerström and Vesa Välimäki are with Acoustics Lab, Department of Information and Communications Engineering, Aalto University, FI-02150 Espoo, Finland (e-mail: jon.fagerstrom@aalto.fi; vesa.valimaki@aalto.fi).

Digital Object Identifier 10.1109/TASLP.2023.3313440

Altogether, the inter-channel decorrelation of FDNs is a widely applicable topic [30], [31].

The main contribution of this work is a mathematical characterization of the inter-channel correlation of a MIMO FDN. We propose a feedforward-recursive decomposition of the FDN transfer function and show that only the feedforward paths contribute to the decorrelation. Different parameter sets, such as different feedback matrix types, are evaluated in terms of the decorrelation in the resulting FDN.

The rest of this article is structured as follows. Section II recapitulates the FDN and inter-channel correlation. Section III derives a novel characterization of the correlation in MIMO FDNs. In Section IV, the proposed method is applied to various FDN parameter sets and the SIMO case is presented. All figures have been made reproducible by including them in the Matlab FDN Toolbox, which is freely available [32].

II. Background

In the following, we present the necessary background on MIMO FDNs and decorrelation filters.

A. MIMO Feedback Delay Network

The MIMO FDN is given in the discrete-time domain by the difference equation in delay state-space form [2], see Fig. 1,

$$\begin{aligned} \boldsymbol{y}(n) &= \boldsymbol{C}\boldsymbol{s}(n) + \boldsymbol{D}\,\boldsymbol{x}(n), \\ \boldsymbol{s}(n+\boldsymbol{m}) &= \boldsymbol{A}\,\boldsymbol{s}(n) + \boldsymbol{B}\,\boldsymbol{x}(n), \end{aligned} \tag{1}$$

where $\boldsymbol{x}(n)$ and $\boldsymbol{y}(n)$ are the $N_{\text{in}} \times 1$ input and $N_{\text{out}} \times 1$ output vectors at time sample n, respectively. The FDN dimension N is the number of delay lines. The FDN consists of the $N \times N$ *feedback matrix* $\boldsymbol{A}$, the $N \times N_{\text{in}}$ *input gain* matrix $\boldsymbol{B}$, the $N_{\text{out}} \times N$ *output gain* matrix $\boldsymbol{C}$, and the $N_{\text{out}} \times N_{\text{in}}$ *direct gain* matrix $\boldsymbol{D}$. The lengths of the N delay lines in samples are given by the vector $\boldsymbol{m} = [m_1, \ldots, m_N]$. The $N \times 1$ vector $\boldsymbol{s}(n)$ denotes the delay-line outputs at time n. The vector argument notation $\boldsymbol{s}(n + \boldsymbol{m})$ abbreviates the vector $[s_1(n + m_1), \ldots, s_N(n + m_N)]$. We refer to an FDN where the number of delay lines is equal to the input and output channels as *full MIMO*, i.e., $N_{\text{in}} = N_{\text{out}} = N$. A single-input, single-output (SISO) FDN has $N_{\text{in}} = N_{\text{out}} = 1$.

The $N_{\text{out}} \times N_{\text{in}}$ transfer function matrix of an FDN in the z-domain [2] corresponding to (1) is

$$\boldsymbol{H}(z) = \boldsymbol{C}(\boldsymbol{D}_{\boldsymbol{m}}(z^{-1}) - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D}, \tag{2}$$

where $\boldsymbol{D}_{\boldsymbol{m}}(z) = \text{diag}[z^{-m_1}, z^{-m_2}, \ldots, z^{-m_N}]$ is the diagonal $N \times N$ delay matrix [1]. The system order is given by the sum of all delay units, i.e., $\mathfrak{N} = \sum_{i=1}^{N} m_i$ [2]. For commonly used delays $\boldsymbol{m}$, the system order is much larger than the FDN size, i.e., $\mathfrak{N} \gg N$, and, thus, the FDN is usually a sparse digital filter.

The transfer function matrix (2) can be stated as a rational polynomial [2], [33], i.e.,

$$\boldsymbol{H}(z) = \frac{\boldsymbol{Q}_{\boldsymbol{m},\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}}(z)}{p_{\boldsymbol{m},\boldsymbol{A}}(z)}, \tag{3}$$

where the denominator is a scalar-valued polynomial

$$p_{\boldsymbol{m},\boldsymbol{A}}(z) = \det(\boldsymbol{P}(z)), \tag{4}$$

where $\det$ denotes the determinant and the loop transfer function is

$$\boldsymbol{P}(z) = \boldsymbol{D}_{\boldsymbol{m}}(z^{-1}) - \boldsymbol{A}. \tag{5}$$

The numerator is a matrix-valued expression with

$$\boldsymbol{Q}_{\boldsymbol{m},\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}}(z) = \boldsymbol{D}\det(\boldsymbol{P}(z)) + \boldsymbol{C}\,\text{adj}(\boldsymbol{P}(z))\,\boldsymbol{B}, \tag{6}$$

where $\text{adj}(\boldsymbol{X})$ denotes the adjugate of any square matrix $\boldsymbol{X}$ [34]. For brevity, we occasionally omit the parameters and write $\boldsymbol{Q}(z)$ and $p(z)$.

The rational polynomial form of the transfer function (3) can be readily derived from (2) using the general identity $\boldsymbol{X}^{-1} = \text{adj}(\boldsymbol{X})/\det(\boldsymbol{X})$ [35]. Importantly, $\boldsymbol{Q}(z)$ and $p(z)$ are both polynomials of z^{-1} and not rational functions. Also, $\boldsymbol{Q}(z)$ is of size $N_{\text{out}} \times N_{\text{in}}$, and $p(z)$ is scalar-valued. In Appendix A, we present practical algorithms to compute the adjugate of a polynomial matrix.

In the following, FDN matrices are allowed to consist of filters, i.e., $\boldsymbol{A}(z)$, $\boldsymbol{B}(z)$, and $\boldsymbol{C}(z)$. In particular, the filter feedback matrix $\boldsymbol{A}(z)$ has been explored in [36].

B. Inter-Channel Correlation and Decorrelation Filtering

The inter-channel correlation between signals x_1 and x_2 is defined via the normalized cross-correlation function

$$\phi_{x_1x_2}(t) = \frac{\sum_n x_1(n)x_2(n+t)}{\|x_1\|_2\|x_2\|_2}, \tag{7}$$

where t denotes the time lag and $\|x\|_2^2 = \sum_n x(n)^2$ denotes the total energy of signal $x(n)$. For any linear and time-invariant filter $h(n)$, the cross-correlation between the input $x(n)$ and output $y(n) = x(n) * h(n)$ is [37]

$$\phi_{yx}(t) = (h * \phi_{xx})(t), \tag{8}$$

where "$*$" denotes the convolution operator.

A monophonic signal $x(n)$ can be expanded into a stereophonic signal $y_1(n)$ and $y_2(n)$ by convolving $x(n)$ with two decorrelation filters $h_1(n)$ and $h_2(n)$, i.e., $y_1(n) = x(n) * h_1(n)$ and $y_2(n) = x(n) * h_2(n)$. For a white noise signal $x(n)$, the cross-correlation is then equal to the filter cross-correlation, i.e.,

$$\phi_{y_1y_2}(k) = \phi_{h_1h_2}(k). \tag{9}$$

Thus, the cross-correlation filters determine the resulting cross-correlation; therefore, for the remaining work, we focus on the filter interrelation.

A low correlation of signals can help prevent coloration when similar signals are mixed with varying delays, which commonly occurs in multichannel sound reproduction [6]. Similarly, multiple input channels can be correlated—for example, as a microphone array produces them—and uncontrolled mixing can lead to coloration [28]. Additional decorrelation filters can reduce inter-channel correlation and prevent such unwanted coloration. To account for the worst-case time lag, we say that two signals are uncorrelated if

$$\overline{\phi}_{x_ix_j} = \max_t \left|\phi_{x_ix_j}(t)\right| \tag{10}$$

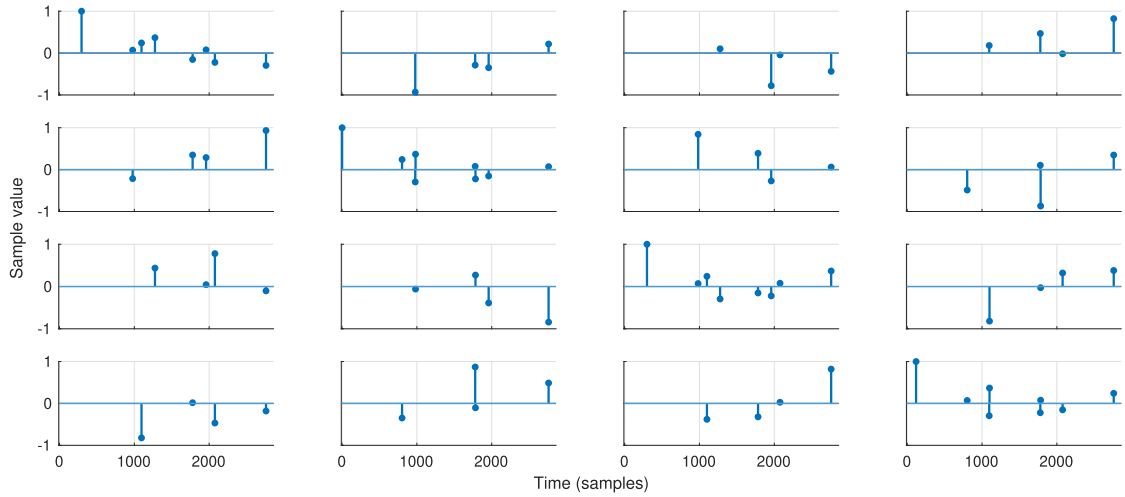


Fig. 2. Time-domain filter coefficients of the adjugate $\mathrm{adj}(\boldsymbol{P}(z))$ of a MIMO FDN with four delay lines ($N = 4$) and with a random orthogonal feedback matrix $\boldsymbol{A}$. The delays are $\boldsymbol{m} = [977, 683, 981, 801]$ samples. The entire adjugate $\mathrm{adj}(\boldsymbol{P}(z))$ is displayed without truncation. Only the non-zero values are drawn with stems for better readability.

is close to zero. Reducing the maximum correlation in (10) implicitly spreads the correlation across different time lags. In the following, we study the correlation induced by MIMO FDN filters. In particular, we study the mutual correlation between the input and output channels from a single FDN structure.

III. Feedfoward Feedback Decomposition

In the following, we derive the mathematical framework for correlation induced by FDNs.

A. Decomposition of the FDN Transfer Function

First, we decompose the FDN transfer function and isolate the relevant parts for the correlation analysis. The fundamental properties can be observed by inspecting the $N_{\text{out}} \times N_{\text{in}}$ transfer function matrix (3), i.e.,

$$\boldsymbol{H}(z) = \boldsymbol{D} + \frac{\boldsymbol{C}\,\mathrm{adj}(\boldsymbol{P}(z))\,\boldsymbol{B}}{p_{\boldsymbol{m},\boldsymbol{A}}(z)}. \tag{11}$$

The direct path $\boldsymbol{D}$ in (11) can be neglected as it only possibly increases correlation, e.g., if $\boldsymbol{D}$ is not orthogonal. Later in the design stage, $\boldsymbol{D}$ can be used to shape the dry-wet ratio and adjust the correlation to the desired amount.

As the denominator $1/p_{\boldsymbol{m},\boldsymbol{A}}(z)$ in (11) is common to all signal paths, it therefore does not impact correlation between different input-output paths. In other words, $1/p_{\boldsymbol{m},\boldsymbol{A}}(z)$ can be applied to the input signals upfront without affecting the inter-channel correlation. Thus, the recursive part $1/p_{\boldsymbol{m},\boldsymbol{A}}(z)$ simply reverberates the input.

Therefore, in can be seen from (11) that the mutual correlation of input-output paths is fully described by $\boldsymbol{C}\mathrm{adj}(\boldsymbol{P}(z))\boldsymbol{B}$. The simplest configuration is to feed each input channel into a separate delay line and also each output channel is taken from separate delay line, i.e., the input and output gains $\boldsymbol{B}$ and $\boldsymbol{C}$ are merely identity matrices. In Section IV, we also discuss cases in which $\boldsymbol{B}$ and $\boldsymbol{C}$ are delay filters, which can greatly improve the decorrelation.

Thus, the raw inter-channel correlation is given by $\mathrm{adj}(\boldsymbol{P}(z))$, which we call the MIMO feedforward paths. The feedback matrix $\boldsymbol{A}$ and delay matrix $\boldsymbol{D}_{\boldsymbol{m}}(z)$ are the main parameters of the feedforward paths. In Section IV, we show that the feedback matrix is the central governing factor to the inter-channel correlation of FDNs. The following discussion is primarily dedicated to investigating the properties of the feedforward path $\mathrm{adj}(\boldsymbol{P}(z))$.

B. Feedforward Path

In the following, we investigate the correlation of the filters in $\mathrm{adj}(\boldsymbol{P}(z))$. Fig. 2 shows an example $\mathrm{adj}(\boldsymbol{P}(z))$ for $N = 4$. The adjugate matrix $\mathrm{adj}(\boldsymbol{P}(z))$ can be expressed by co-factors, i.e.,

$$\mathrm{adj}(\boldsymbol{P}(z))_{ij} = (-1)^{i+j}\det(M_{ji}(z)), \tag{12}$$

where the $(N-1)\times(N-1)$ submatrix $M_{ij}(z)$ results from deleting row i and column j of $\boldsymbol{P}(z)$. A determinant of the form $\det(\boldsymbol{D}_{\boldsymbol{m}}(z^{-1}) - \boldsymbol{A})$ is given by [33]

$$\det(\boldsymbol{D}_{\boldsymbol{m}}(z^{-1}) - \boldsymbol{A}) = \sum_{k=0}^{\mathfrak{N}} c_k\, z^k, \text{ with}$$
$$c_k = \begin{cases} \sum_{I\in I_k} (-1)^{N-|I|} \det \boldsymbol{A}(I^c), & \text{for } I_k \neq \emptyset \\ 0, & \text{otherwise,} \end{cases} \tag{13}$$

where $I_k = \{I \subset \langle N\rangle | \sum_{i\in I} m_i = k\}$ is the set of all indices combinations whose delay lengths sum is equal to k. The empty set is denoted by $\emptyset$ and $\boldsymbol{A}(I^c)$ denotes the submatrix with column and row indices in the set I^c, where I^c is the set complementary of I. Note that $\det \boldsymbol{A}(\emptyset) = 1$. For specific choices of $\boldsymbol{m}$, $|I_k| = 1$ for $0 \leq k \leq \mathfrak{N}$, and therefore each c_k has a single summand in (13). As a consequence, the determinant (13) has at most 2^N non-zero coefficients and has a polynomial order of $\mathfrak{N}$.

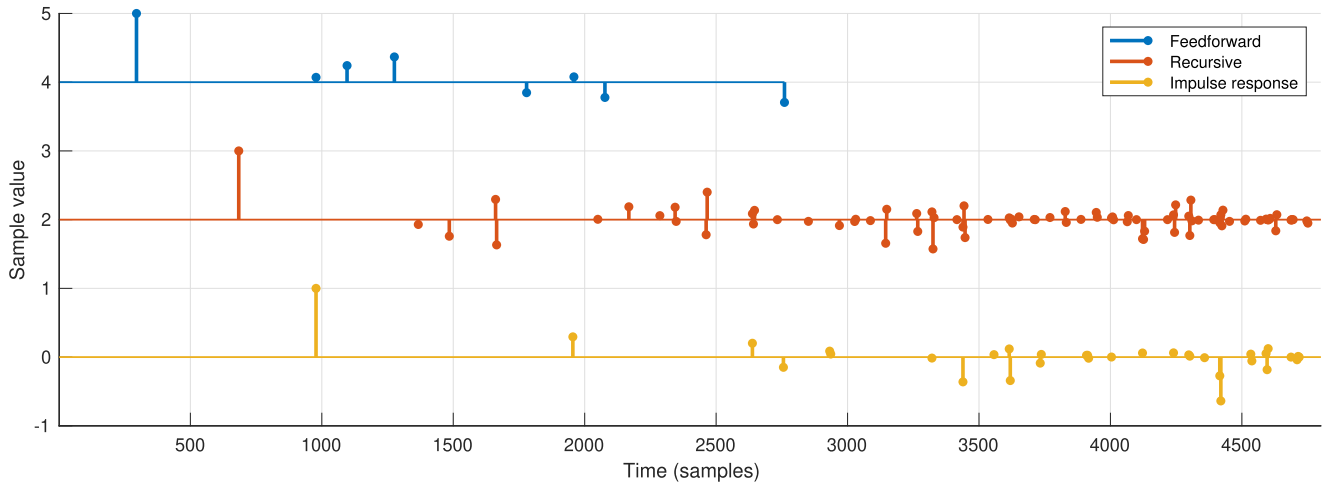


Fig. 3. A MIMO FDN decomposed impulse response with four delays, i.e., $N = 4$ and a random orthogonal feedback matrix $\boldsymbol{A}$. The delays are $\boldsymbol{m} = [977, 683, 981, 801]$ samples. The undriven impulse response is only dependent on the recursive part of the FDN and common to all possible input-output combinations and, as such, does not contribute to the correlation between channels. The feedforward response ($= \mathrm{adj}(\boldsymbol{P}(z))$) convolved with the recursive response results in the standard MIMO impulse response. Only the non-zero values are drawn with stems.

C. Co-Factor

We apply now (13) to the co-factors $M_{ij}(z)$ in (12). The co-factor notation (12) reveals that the filters in $\mathrm{adj}(\boldsymbol{P}(z))$ are strongly related. Permutation of rows and columns only alters the sign of the determinant such that we can bring the submatrix $M_{ij}(z)$ into a standard form $\hat{M}_{ij}(z) = \boldsymbol{\sigma}_\mathrm{l} M_{ij}(z) \boldsymbol{\sigma}_\mathrm{r}$ with permutation matrices $\boldsymbol{\sigma}_\mathrm{l}$ and $\boldsymbol{\sigma}_\mathrm{r}$ such that

$$\hat{M}_{ij}(z) = \begin{cases} \mathrm{diag}\big([0, z^{\boldsymbol{m}(i,j)}]\big) - \boldsymbol{\sigma}_\mathrm{l} \boldsymbol{A}(i,j) \boldsymbol{\sigma}_\mathrm{r} & \text{for } i \neq j \\ \mathrm{diag}\big(z^{\boldsymbol{m}(i)}\big) - \boldsymbol{\sigma}_\mathrm{l} \boldsymbol{A}(i,i) \boldsymbol{\sigma}_\mathrm{r} & \text{otherwise,} \end{cases} \tag{14}$$

where $\boldsymbol{A}(i,j)$ results from deleting row i and column j of $\boldsymbol{A}$ and $\boldsymbol{m}(i,j)$ is $\boldsymbol{m}$ without the ith element, and $\boldsymbol{m}(i,j)$ is $\boldsymbol{m}$ without the ith and jth elements. Thus, $\det(\hat{M}_{ij}(z)) = \pm\det(M_{ij}(z))$. For later analysis, it is important to note that the main diagonal of $\mathrm{adj}(\boldsymbol{P}(z))$, i.e., $\hat{M}_{ii}(z)$ has the form (13). In particular, $\hat{M}_{ii}(z)$ contains a coefficient value of 1 due to $\det(\boldsymbol{A}(\emptyset)) = 1$.

The principal minor representation (13) gives an explicit form for $\det(\hat{M}_{ij}(z))$. It is directly visible that the filter order of $\det(M_{ij}(z))$ is $\sum_{k \neq i,j}^{N} m_k$. Thus, the number of non-zero elements is at most 2^{N-1} for $i = j$ and 2^{N-2} otherwise. Hence, for small N, the feedforward filters tend to be sparse, while they are denser for higher N.

Computing the polynomial adjugate matrix with the co-factor formulation in (12) is expensive and intractable for large N, e.g., $N \geq 32$. Instead, the computation can be performed in the frequency domain, where the adjugate is computed for each frequency bin separately. Appendix A gives more details on this procedure.

D. Example of the Feedforward Paths

We illustrate the adjugate matrix by giving a small-scale example. Fig. 2 depicts the time-domain filter coefficients of the feedforward paths $\mathrm{adj}(\boldsymbol{P}(z))$ for an FDN with four delays, i.e., $N = 4$ and delays between 300 and 1000 samples. Each matrix entry is a sparse finite impulse response (FIR) filter, with $2^{N-1} = 8$ non-zero pulses (for the diagonal elements) and $2^{N-2} = 4$ non-zero pulses (for the non-diagonal elements). For small N, the feedforward filters tend to be sparse.

Fig. 3 depicts the decomposition (11) of the feedforward and recursive portion of the FDN impulse response for the same MIMO FDN as shown in Fig. 2. The feedforward paths are the paths from the first input to the first output channel, the recursive response is the time-domain form of $1/p_{\boldsymbol{m},\boldsymbol{A}}(z)$, and the convolution of both sequences results in the impulse response between the respective input and output channels. We observe that the recursive response is denser, i.e., with more non-zero elements than the final impulse response. The specific phase of the feedforward paths leads to cancellation.

It is important to note the temporal extent of the individual components. The adjugate $\mathrm{adj}(\boldsymbol{P}(z))$ consists of FIR filters with orders less than $\mathfrak{N}$. In contrast, $1/p_{\boldsymbol{m},\boldsymbol{A}}(z)$ is an infinite impulse response (IIR) filter with an endless temporal response (in theory), which is largely responsible for the long reverberation tail created by the FDN. However, due to the discussion above, $1/p_{\boldsymbol{m},\boldsymbol{A}}(z)$ does not contribute to the mutual decorrelation of the channels. The fact that the reverberation tail generated by $1/p_{\boldsymbol{m},\boldsymbol{A}}(z)$ does not influence the inter-channel correlation of $\boldsymbol{y}$ might come as a surprise. This contrasts the correlation between the input and output signals, $\boldsymbol{x}$ and $\boldsymbol{y}$, which is strongly impacted by a long reverberation tail.

IV. Correlation of MIMO FDNs

The adjugate $\mathrm{adj}(\boldsymbol{P}(z))$ is only a comparably short filter, and therefore it is not guaranteed to decorrelate the channels strongly. In the following, we study the inter-channel correlation of the FDN, particularly the feedforward paths. The Matlab-code

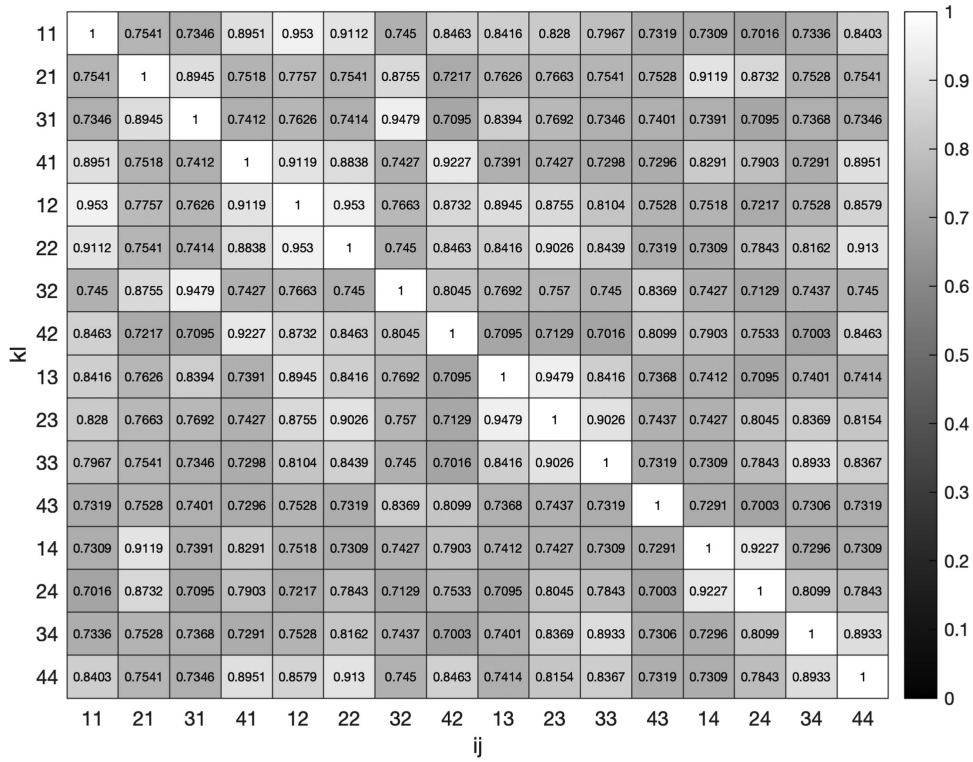


Fig. 4. Inter-channel max correlation matrix Φ_{ijkl} of the FDN shown in Fig. 2. The white cell color indicates the maximum correlation for each pair of transfer function elements, whereas a darker color indicates a lower correlation, i.e., a better decorrelation.

for repeating the numerical results of this article is available in GitHub[1].

A. Inter-Channel Correlation

An appropriate correlation definition depends on the application and is affected by the desired psychoacoustic effect. The correlation depends on the time lag, which can be integrated in various ways and up to a maximum time lag. Here, we employ a conservative metric using the maximum absolute correlation overall time lags and evaluating all input and output pairs. This correlation metric serves as an upper bound for other correlation metrics such as zero-lag correlation or average correlation.

The inter-channel correlation matrix is

$$\Phi_{ijkl} = \max_t |\phi_{ijkl}(t)|\,, \tag{15}$$

where ϕ_{ijkl} is the cross-correlation between the two filters $\mathrm{adj}(\boldsymbol{P}(z))_{ij}$ and $\mathrm{adj}(\boldsymbol{P}(z))_{kl}$. The diagonal contains the autocorrelations and is normalized to 1, whereas the remaining values are between 0 and 1.

Fig. 4 shows the 16-by-16 inter-channel correlation matrix for the same FDN as in Fig. 2, which has the size $N = 4$. The range of the correlation values is between 0.70 and 0.95. For further summarized statistics, the median is applied to the non-diagonal elements of the inter-channel correlation matrix, i.e.

$$\overline{\Phi} = \mathrm{median}\{\Phi_{ijkl} \mid ij \neq kl\}. \tag{16}$$

[1]https://github.com/Ion3rik/fdnDecorrelation

For smaller matrix sizes N, it is advisable to average multiple instances to retrieve a stable estimate of the median correlation $\overline{\Phi}$, for example, by varying the feedback matrix $\boldsymbol{A}$.

B. Influence of Feedback Delay

We note that the choice of the feedback delay-line lengths has little influence on the cross-correlation except for special constellations. From (13) and (12), we can see that the delays $\boldsymbol{m}$ affect the temporal spacing of the filter coefficients, but not their values. This is particularly true for delays $\boldsymbol{m}$, where $|I_k| = 1$ for $0 \leq k \leq \mathfrak{N}$ and therefore each c_k has a single summand in (13). On the other hand, specific choices of delays $\boldsymbol{m}$ can deviate considerably. Here, we establish a general trend for randomly chosen delay lengths.

C. Decorrelating With the Feedback Matrix

The choice of feedback matrix $\boldsymbol{A}$ is the main parameter affecting the decorrelation of FDNs, as established in Section III. Here, the inter-channel correlation of MIMO FDNs of three different sizes $N = \{4, 8, 16\}$ is analyzed using six different feedback matrices. The Random Orthogonal, Hadamard, Householder, and Circulant matrix types are traditional frequency-independent gain matrices [32], whereas Velvet Scattering and Dense scattering are filter feedback matrices introduced recently [36]. The matrix types are defined in Appendix B. The computational costs of the Velvet scattering and Dense Scattering are equal, having $K = 3$ stages, with N^K taps in each filter.

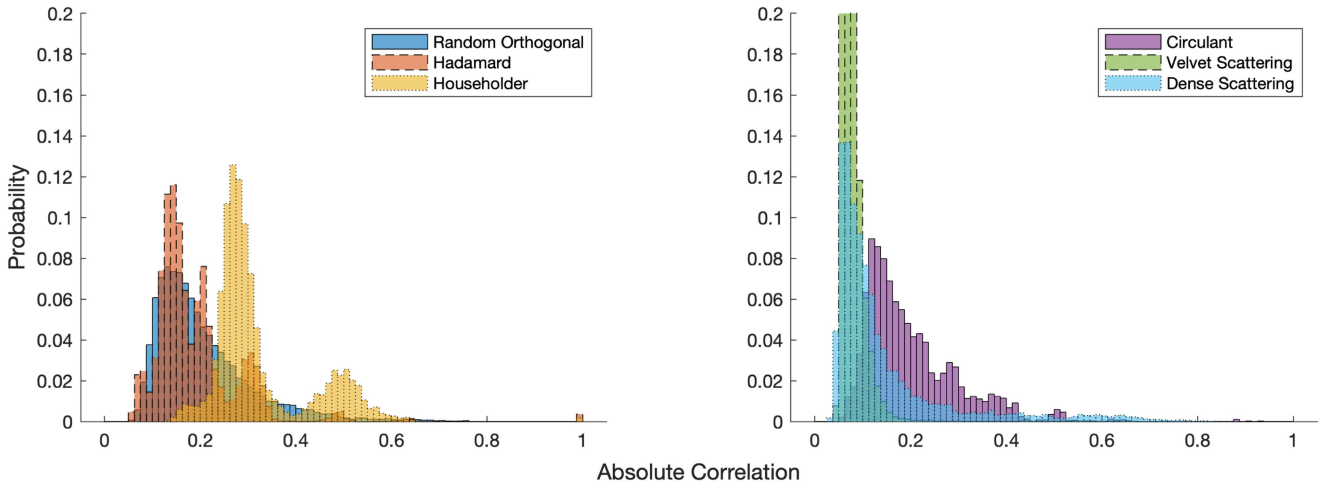


Fig. 5. Histogram of the inter-channel max correlation Φ_{ijkl} for FDN size $N = 16$ and six different feedback matrix types.

TABLE I
MEDIAN CORRELATION METRIC $\overline{\Phi}$ WITH THE IQR FOR DIFFERENT FEEDBACK MATRIX TYPES AND SIZES OF THE FDN

Feedback Matrix Type	N = 4		8		16	
	Med	IQR	Med	IQR	Med	IQR
Random Ortho.	0.712	0.209	0.368	0.178	0.178	0.112
Hadamard	0.474	0.146	0.301	0.173	0.163	0.082
Householder	0.500	0.031	0.263	0.250	0.287	0.084
Circulant	0.500	0.352	0.305	0.156	0.173	0.115
Velvet Scattering	0.129	0.029	0.063	0.031	**0.074**	0.024
Dense Scattering	**0.125**	0.031	**0.056**	0.054	0.097	0.089

The values are averages of ten different instances. Bold font indicates the lowest correlation value of each tested matrix size (columns). The values for $N = 16$ correspond to the histograms in Fig. 5.

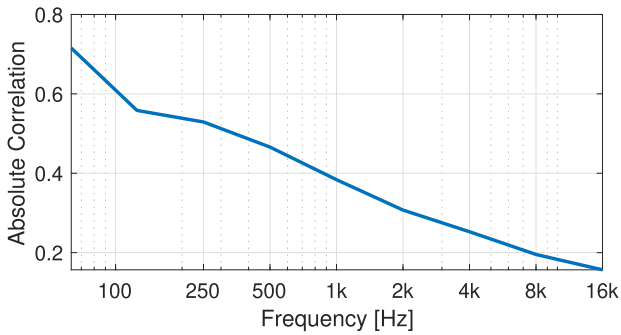


Fig. 6. Frequency-dependent median correlation metric $\overline{\Phi}$ computed at octave bands, between $63-16000$ Hz, for a MIMO FDN with the velvet scattering matrix. The values are averages of 10 different instances.

The delays are chosen uniformly randomly between 300 and 10000 samples.

Fig. 5 shows the histogram of inter-channel max correlation Φ_{ijkl} values over all input and output pairs of each tested configuration. For each of the ten instances, the feedback matrices of size $N = 16$ are randomized. The correlation values for standard feedback matrices, excluding the Householder, have a similar distribution, where the Random Orthogonal is distributed more widely than the Hadamard and Circulant matrices. The Householder matrix has the highest correlation values, which can be related to the particular structure of this matrix type. For larger N, the Householder matrix becomes increasingly closer to the identity matrix. Thus, there is a strong self-coupling of the delays and less cross-mixing. We interpret the bimodal distribution as a result of the two coupling types of the Householder matrix. Both filter feedback matrices show reduced inter-channel correlation, where the correlation between most channel pairs is less than 0.2.

Table I shows each tested configuration's estimated median correlation $\overline{\Phi}$ values with the inter-quartile range (IQR) across different matrix sizes. All numbers in the tables are generally averaged over ten random instances by varying the random seeds of the delay and matrix generation algorithms. The median maximum correlation, in general, decreases with the decreasing matrix size N in Table I. One exception is the Householder matrix which decreases for $N = 8$, but remains for $N = 16$. Due to the bimodal distribution observed in Fig. 5, the applicability of the median statistics is limited. Across all matrix sizes, the filter feedback matrices have the lowest correlation. Thus, a filter feedback matrix contributes to the impulse response density, as demonstrated previously [36] and provides excellent decorrelation performance.

D. Frequency Dependency Correlation

The correlation of a linear filter is typically frequency-dependent, where the correlation between two filters tends to decrease with frequency [22]. For a brief analysis, we compute the median correlation metric $\overline{\Phi}$ at octave bands, for the MIMO FDN of size $N = 4$, with the velvet scattering matrix.

Fig. 6 shows that the median correlation decreases with increasing frequency within $0.14-0.72$. The feedforward paths are comprised of relatively short filters (less than 2000 samples), as seen in Fig. 8(a). The short filters limit the inter-channel decorrelation performance at low frequencies.

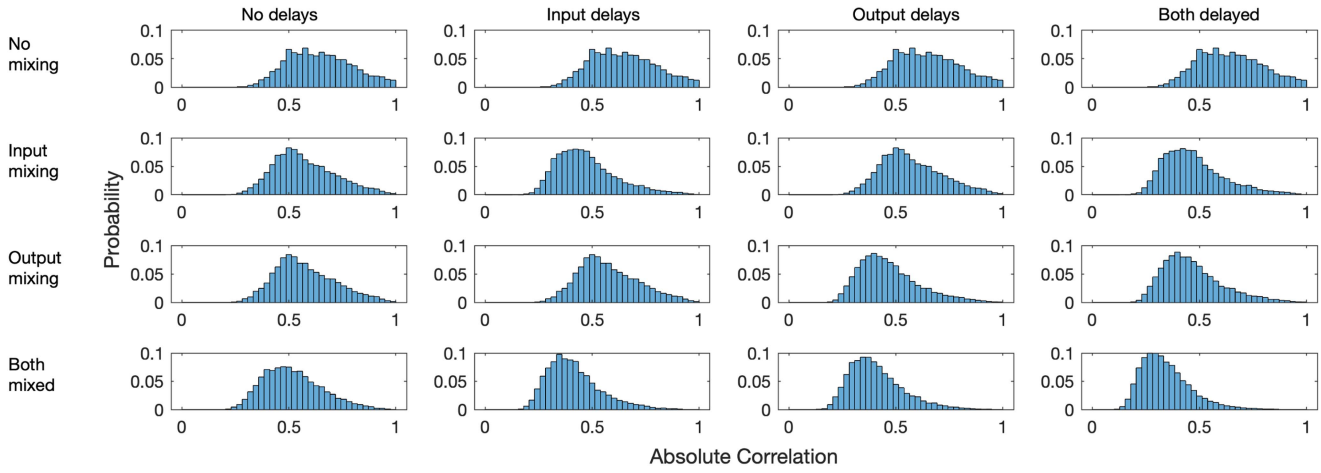


Fig. 7. Histogram of the inter-channel max correlation for different mixing configurations on the correlation metric $\overline{\Phi}$, when the FDN size is $N = 4$.

TABLE II
INFLUENCE OF DIFFERENT MIXING CONFIGURATIONS ON THE CORRELATION METRIC $\overline{\Phi}$, WHEN THE FDN SIZE IS $N = 4$

Mixing Configuration	Delay Configuration							
	No delays		Input delays		Output delays		Both delayed	
	Med	IQR	Med	IQR	Med	IQR	Med	IQR
No mixing	0.634	0.223	0.634	0.223	0.634	0.223	0.634	0.223
Input mixing	0.551	0.196	0.451	0.180	0.551	0.196	0.450	0.180
Output mixing	0.555	0.192	0.555	0.192	0.443	0.178	0.443	0.176
Both mixed	0.503	0.188	0.393	0.155	0.387	0.160	0.323	0.144

The values are the median values, and the corresponding IQR values derived from the histograms in Fig. 7.

E. Decorrelating Input and Output Filters

In this section, we replace $\boldsymbol{B}$ and $\boldsymbol{C}$ with filter matrices $\boldsymbol{B}(z)$ and $\boldsymbol{C}(z)$ to learn about their decorrelating capability. The input and output filters can be further decomposed into delay and mixing stages so that $\boldsymbol{B}(z) = \boldsymbol{D}_{\text{in}}(z)\boldsymbol{U}_{\text{in}}$ and $\boldsymbol{C}(z) = \boldsymbol{U}_{\text{out}}\boldsymbol{D}_{\text{out}}(z)$, where $\boldsymbol{D}_{\text{in}}(z)$ and $\boldsymbol{D}_{\text{out}}(z)$ are diagonal delay matrices, whereas $\boldsymbol{U}_{\text{in}}$ and $\boldsymbol{U}_{\text{out}}$ are random orthogonal mixing matrices. The correlation properties of all the combinations yielding from the decomposition are analyzed, e.g., with and without $\boldsymbol{U}_{\text{in}}$ and with and without $\boldsymbol{D}_{\text{in}}(z)$. The correlation analysis follows the same procedure as in Section IV-C. The random orthogonal matrix is used for the feedback matrix $\boldsymbol{A}$, and the input and output delays are short random delays between $[1, 100]$ samples.

Fig. 7 shows the histogram of inter-channel max correlation Φ_{ijkl} values over all input and output pairs for all different combinations of the input and output delay and mixing stages. Table II shows corresponding median correlation metric $\overline{\Phi}$. The columns contain the delay configurations and the rows contain the mixing configurations. As expected, introducing just delays (first row and last three columns in Fig. 7) has no impact on the correlation, as the delays can only affect the time lag of the maximum cross-correlation, but cannot decorrelate the channels further.

As seen in Fig. 7 and Table II, applying just the mixing matrix at either the input or output lowers the inter-channel correlation, and applying the mixing at both the input and output is the best configuration when no additional delays are applied. The most decorrelation is achieved if both mixing and delays are applied concurrently at the input and/or output as seen in the main diagonal of Fig. 7 and Table II. Thus, in the following we focus only on the configurations on the main diagonal of Table II.

TABLE III
CORRELATION $\overline{\Phi}$ FOR DIFFERENT DELAY FILTER CONFIGURATIONS AND FDN SIZES

Delay and Mixing Type	N					
	4		8		16	
	Med	IQR	Med	IQR	Med	IQR
No filters	0.634	0.233	0.381	0.180	0.184	0.110
Input filters	0.452	0.185	0.243	0.130	0.120	0.058
Output filters	0.458	0.188	0.252	0.135	0.120	0.059
Both filters	0.325	0.151	0.157	0.079	0.097	0.028

As established above, introducing delay filters at the input $\boldsymbol{B}(z)$, output $\boldsymbol{C}(z)$, or both sides of the MIMO FDN can improve the decorrelation. Median correlation metric $\overline{\Phi}$ values of the three delay filter configurations as well as for the "No filters" case without the delay filters are shown in Table III. The analysis was conducted for three different size of FDNs of sizes $N = \{4, 8, 16\}$. Note that the top left value in Tables I, II, and III, i.e., $N = 4$ with Random Orthogonal feedback matrix and no additional input and output filters, refer to the same FDN configuration and the small deviations between the values are due to randomization of the parameters.

TABLE IV
INFLUENCE OF DIFFERENT CONFIGURATIONS ON THE CORRELATION $\overline{\Phi}$ OF SIMO FDN OUTPUTS

Configuration	N					
Type	4		8		16	
	Med	IQR	Med	IQR	Med	IQR
Random Orthog.	0.598	0.205	0.468	0.214	0.280	0.145
Hadamard	0.474	0.437	0.305	0.144	0.216	0.116
Householder	0.508	0.234	0.540	0.065	0.491	0.053
Circulant	0.396	0.123	0.355	0.132	0.283	0.193
Velvet Scattering	0.350	0.019	0.251	0.025	0.152	0.024
Dense Scattering	0.400	0.050	0.307	0.038	0.236	0.064

Based on the results in Table III, having the delay filters only at the input or output side results in similar inter-channel correlation, whereas assigning both $\boldsymbol{B}(z)$ and $\boldsymbol{C}(z)$ as delay and mixing filters results in the lowest correlation for all tested FDN sizes. In general, the delay filters are effective in lowering the correlation, when the FDN size N is small, whereas with the large FDN size ($N = 16$) the correlation is already low ($\overline{\Phi} \approx 0.18$) without the filters.

F. Decorrelating SIMO FDNs

A single-input, multi-output (SIMO) FDN has practical importance, especially in the context of multichannel late reverberation. Namely, when a spatial late reverberation is applied to a mono signal, the SIMO FDN is a suitable option [9], [38]. Low correlation between the output channels is necessary to avoid coloration and to achieve a diffuse reverberation effect.

In this section, we conduct a similar analysis as for the MIMO case, utilizing the median correlation metric $\overline{\Phi}$. For the SIMO case, the mono input signal is distributed equally to all delay lines, i.e., $\boldsymbol{B} = \mathbf{1}_{N\times 1}$. Fig. 8 shows the comparison of the feedforward paths for the MIMO case $\mathrm{adj}(\boldsymbol{P}(z))$ and the SIMO case $\mathrm{adj}(\boldsymbol{P}(z))\boldsymbol{B}$ using velvet scattering matrix $\boldsymbol{A}(z)$. With this feedback matrix, it is apparent that the first peak on $\mathrm{adj}(\boldsymbol{P}(z))$ main diagonal is equal to 1, while the remaining response contains relatively low values. As a result, the correlation between the input-output pairs on the main diagonal is comparably high. A similar trend exists for larger FDN sizes N and standard feedback matrices $\boldsymbol{A}$. Consequently, all SIMO feedforward paths contain a strong first peak, see Fig. 8(b), and the inter-channel correlation between the outputs is overall higher than in the MIMO case, where only a few pairs are highly correlated.

The correlation analysis of the SIMO case across feedback matrix sizes and types is shown in Table IV. Similar to Table III, the median inter-channel correlation reduces with the matrix size N. Less expectantly, the trend of correlation is different from the MIMO case. The input channel summation in the SIMO leads to various cancellations between the feedforward paths, which are especially severe, for example, with the Householder matrix. Also, in the SIMO case, the Velvet and Dense scattering matrices do not reduce the correlation as much as in the MIMO case. Overall, the Velvet scattering matrix is still the feedback matrix with the best decorrelation performance across all of the tested FDN sizes.

V. CONCLUSION

This work presents a novel characterizing of the inter-channel decorrelation of an FDN, a property important for multichannel sound capturing and reproduction with artificial reverberation. The mathematical analysis shows that the inter-channel correlation depends only on the feedforward paths and is independent of the infinitely-long recursive reverberation tail. The inter-channel correlation in the MIMO and the SIMO FDN configurations is studied with various parameters. The feedback matrix type impacts the correlation significantly, and our analysis shows that in most cases, the filter feedback matrices, namely the Velvet and the Dense Scattering matrices [36], yield the best inter-channel decorrelation. Also, large FDN systems with many feedback delays, i.e., large N, generally have better-decorrelated output channels than small systems. Adding pre- and post-filtering using delays and mixing matrices can further improve the overall decorrelation, especially for small FDNs.

Generally speaking, long reverberation effects generated with FDNs do not necessarily result in well-decorrelated audio channels, and extra care should be taken to avoid undesirable sound coloration. The results of this study apply to the design of multichannel FDN reverberators, which are used in surround sound and immersive gaming, and virtual reality applications. For example, binaural reproduction requires a target inter-aural cross-correlation, which can be achieved by cross-mixing the output channels based on the FDN's inter-channel correlation.

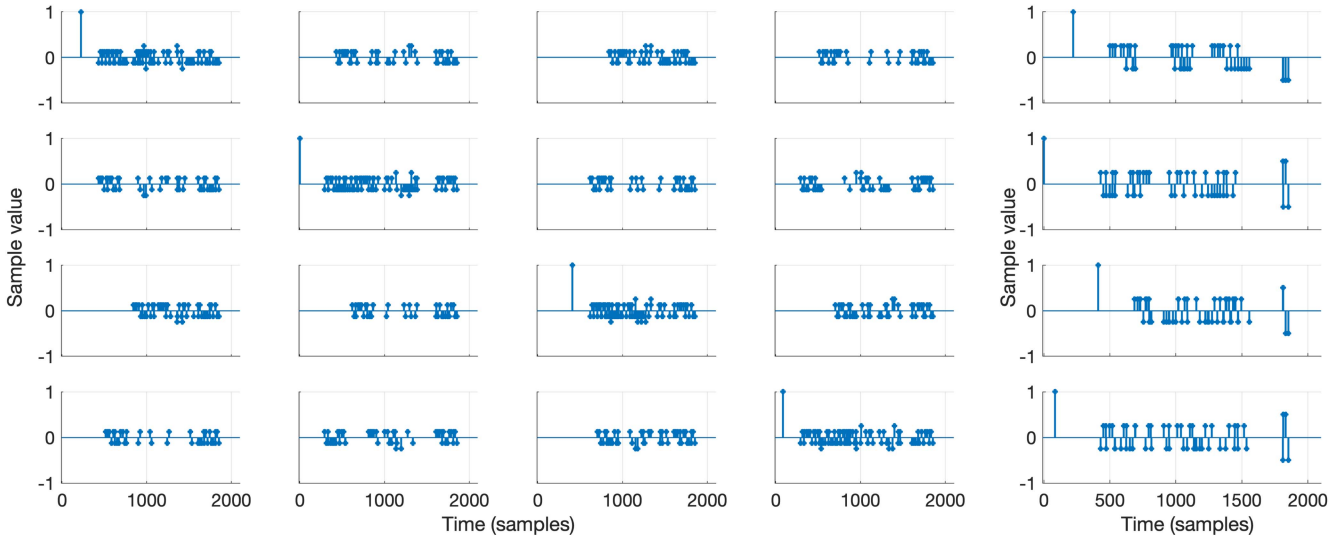
APPENDIX A
COMPUTING THE ADJUGATE

Section III demonstrates that $\mathrm{adj}(\boldsymbol{P}(z))$ plays a central role in the analysis of decorrelation of an FDN. Therefore, we review an efficient technique to compute the polynomial adjugate.

As described in [39], the determinant, the adjugate, and therefore, the inverse of a polynomial matrix can be computed in the frequency domain. Hence, for a matrix function f, e.g., $f = \det$ or $f = \mathrm{adj}$, we can apply f at uniform frequency points $e^{\imath w}$, i.e., $f(\boldsymbol{P}(e^{\imath w}))$, where $\imath$ stands for the imaginary unit. The time-domain solution can be retrieved by applying the inverse discrete Fourier transform (DFT) to each matrix entry. The DFT needs to be sufficiently zero-padded to avoid circular convolution artefacts. For example, the sum of the filter lengths of all matrix entries is sufficient. A fast implementation can be achieved by applying the fast Fourier transform, or FFT algorithm to appropriately zero-padded sequences.

APPENDIX B
FEEDBACK MATRIX TYPES

Various feedback matrix types used in FDNs are very briefly reviewed. The Random Orthogonal matrix is sampled uniformly



(a) MIMO (b) SIMO

Fig. 8. Feedforward paths of MIMO and SIMO case for velvet scattering matrix. The FDN ($N = 4$) delays are $\boldsymbol{m} = [677, 455, 865, 539]$. SIMO feedforward paths result from the MIMO case by summing each row.

from the orthogonal group $\mathrm{O}(N)$. The Hadamard matrix [40] is defined recursively:

$$\boldsymbol{H}_1 = 1,$$
$$\boldsymbol{H}_{2k} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{H}_k & \boldsymbol{H}_k \\ \boldsymbol{H}_k & -\boldsymbol{H}_k \end{bmatrix}, \text{for } k = 2, 3, \ldots \tag{17}$$

The Hadamard matrix is randomized by permuting the rows and columns of the matrix.

The Householder matrix [41] is $\boldsymbol{I} - 2\vec{v}\vec{v}^\top$ for unit vector $\vec{v}$. The Circulant matrix [2] is of the form

$$\begin{bmatrix} v_1 & v_N & \ldots & v_2 \\ v_2 & v_1 & \ldots & v_3 \\ \vdots & \vdots & \ddots & \vdots \\ v_N & v_{N-1} & \ldots & v_1 \end{bmatrix}, \tag{18}$$

where the DFT of the vector $\vec{v}$ is unimodular, i.e., $|\mathrm{DFT}(\vec{v})| \equiv 1$. FIR filter feedback matrices can be factorized as follows:

$$\boldsymbol{A}(z) = \boldsymbol{D}_{\boldsymbol{m}_K}(z)\boldsymbol{U}_K \cdots \boldsymbol{D}_{\boldsymbol{m}_1}(z)\boldsymbol{U}_1\boldsymbol{D}_{\boldsymbol{m}_0}(z), \tag{19}$$

where $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_K$ are scalar $N \times N$ unitary matrices and $\boldsymbol{m}_0, \boldsymbol{m}_1, \ldots, \boldsymbol{m}_K$ are vectors of N delays. For the Velvet Scattering matrix [36], the $\boldsymbol{U}_i$ are Hadamard, and the delays $\boldsymbol{m}_i$ are chosen such that the filter feedback matrix is sparse. In contrast, the Dense Scattering matrices $\boldsymbol{U}_i$ are random orthogonal matrices with short $\boldsymbol{m}_i$ resulting in a dense filter feedback matrix [36].

REFERENCES

[1] J. M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. Audio Eng. Soc. Conv.*, 1991, pp. 1–12.

[2] D. Rocchesso and J. O. Smith, "Circulant and elliptic feedback delay networks for artificial reverberation," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 1, pp. 51–63, Jan. 1997.

[3] V. Välimäki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel, "Fifty years of artificial reverberation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 5, pp. 1421–1448, Jul. 2012.

[4] 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.

[5] 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.

[6] J. Moore and A. Hill, "Dynamic diffuse signal processing for sound reinforcement and reproduction," *J. Audio Eng. Soc.*, vol. 66, no. 11, pp. 953–965, Nov. 2018.

[7] N. Xiang, U. Trivedi, and B. Xie, "Artificial enveloping reverberation for binaural auralization using reciprocal maximum-length sequences," *J. Acoust. Soc. Amer.*, vol. 145, no. 4, pp. 2691–2702, Apr. 2019.

[8] S. Oksanen, J. D. Parker, A. Politis, and V. Välimäki, "A directional diffuse reverberation model for excavated tunnels in rock," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2013, pp. 644–648.

[9] B. Alary, A. Politis, S. J. Schlecht, and V. Välimäki, "Directional feedback delay network," *J. Audio Eng. Soc.*, vol. 67, no. 10, pp. 752–762, Oct. 2019.

[10] C. Kirsch, J. Poppitz, T. Wendt, S. van de Par, and S. D. Ewert, "Spatial resolution of late reverberation in virtual acoustic environments," *Trends Hear.*, vol. 25, pp. 1–17, Dec. 2021.

[11] I. Engel, C. Henry, S. V. A. Garí, P. W. Robinson, and L. Picinali, "Perceptual implications of different Ambisonics-based methods for binaural reverberation," *J. Acoust. Soc. Amer.*, vol. 149, no. 2, pp. 895–910, Feb. 2021.

[12] I. Engel and L. Picinali, "Reverberation and its binaural reproduction: The trade-off between computational efficiency and perceived quality," in *Advances in Fundamental and Applied Research on Spatial Audio*, B. F. G. Katz and P. Majdak, Eds., chapter 8, 2022.

[13] O. Das and J. S. Abel, "Grouped feedback delay networks for modeling of coupled spaces," *J. Audio Eng. Soc.*, vol. 69, no. 7/8, pp. 486–496, Jul. 2021.

[14] T. B. Atalay, Z. S. Gul, E. De Sena, Z. Cvetkovic, and H. Hacihabiboglu, "Scattering delay network simulator of coupled volume acoustics," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 30, pp. 582–593, 2022.

[15] M. Kuster, "Multichannel room impulse response generation with coherence control," *IEEE Trans. Audio Speech Lang. Process.*, vol. 17, no. 4, pp. 597–606, May 2009.

[16] V. Välimäki and K. Prawda, "Late-reverberation synthesis using interleaved velvet-noise sequences," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 29, pp. 1149–1160, 2021.

[17] K. Prawda, S. J. Schlecht, and V. Välimäki, "Multichannel interleaved velvet noise," in *Proc. Int. Conf. Digit. Audio Effects*, 2022, pp. 208–215.

[18] M. Boueri and C. Kyriakakis, "Audio signal decorrelation based on a critical band approach," in *Proc. Audio Eng. Soc. Conv.*, 2004, pp. 1–6.
[19] E. K. Canfield-Dafilou and J. S. Abel, "A group delay-based method for signal decorrelation," in *Proc. Audio Eng. Soc. Conv.*, 2018, pp. 1–7.
[20] E. K. Canfield-Dafilou and J. S. Abel, "Group delay-based allpass filters for abstract sound synthesis and audio effects processing," in *Proc. Int. Conf. Digit. Audio Effects*, 2018, pp. 197–204.
[21] B. Alary, A. Politis, and V. Välimäki, "Velvet-noise decorrelator," in *Proc. Int. Conf. Digit. Audio Effects*, 2017, pp. 405–411.
[22] S. J. Schlecht, B. Alary, V. Välimäki, and E. A. P. Habets, "Optimized velvet-noise decorrelator," in *Proc. Int. Conf. Digit. Audio Effects*, 2018, pp. 87–94.
[23] J. Fagerström, B. Alary, S. J. Schlecht, and V. Välimäki, "Velvet-noise feedback delay network," in *Proc. Int. Conf. Digit. Audio Effects*, 2020, pp. 219–226.
[24] F. Menzer and C. Faller, "Binaural reverberation using a modified Jot reverberator with frequency-dependent interaural coherence matching," in *Proc. Audio Eng. Soc. 126th Conv.*, 2009, pp. 1–6.
[25] F. Menzer, "Binaural reverberation using two parallel feedback delay networks," in *Proc. Audio Eng. Soc. 40th Int. Conf.*, 2010, pp. 1–10.
[26] C. Borß and R. Martin, "An improved parametric model for perception-based design of virtual acoustics," in *Proc. Audio Eng. Soc. 40th Int. Conf.*, 2009, pp. 1–8.
[27] N. Agus, H. Anderson, J.-M. Chen, S. Lui, and D. Herremans, "Minimally simple binaural room modeling using a single feedback delay network," *J. Audio Eng. Soc.*, vol. 66, no. 10, pp. 791–807, Oct. 2018.
[28] H. Lee, "Multichannel 3D microphone arrays: A review," *J. Audio Eng. Soc.*, vol. 69, no. 1/2, pp. 5–26, Jan. 2021.
[29] L. McCormack, A. Politis, T. McKenzie, C. Hold, and V. Pulkki, "Object-based six-degrees-of-freedom rendering of sound scenes captured with multiple Ambisonic receivers," *J. Audio Eng. Soc.*, vol. 70, no. 5, pp. 355–372, May 2022.
[30] G. Potard and I. Burnett, "Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays," in *Proc. Int. Conf. Digit. Audio Effects*, 2004, pp. 280–284.
[31] M. Blochberger, F. Zotter, and M. Frank, "Sweet area size for the envelopment of a recursive and a non-recursive diffuseness rendering approach," in *Proc. Int. Conf. Spatial Audio*, 2019, pp. 151–157.
[32] S. J. Schlecht, "FDNTB: The feedback delay network toolbox," in *Proc. Int. Conf. Digit. Audio Effects*, 2020, pp. 211–218.
[33] S. J. Schlecht and E. A. P. Habets, "Time-varying feedbackmatrices in feedback delay networks and their application in artificial reverberation," *J. Acoust. Soc. Amer.*, vol. 138, no. 3, pp. 1389–1398, Sep. 2015.
[34] S. J. Schlecht and E. A. P. Habets, "Modal decomposition of feedback delay networks," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5340–5351, Oct. 2019.
[35] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1996.
[36] S. J. Schlecht and E. A. P. Habets, "Scattering in feedback delay networks," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 28, pp. 1915–1924, 2020.
[37] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2009.
[38] J. Stautner and M. Puckette, "Designing multi-channel reverberators," *Comput. Music J.*, vol. 6, no. 1, pp. 52–65, 1982.
[39] M. Hromčík and M. Šebekt, "New algorithm for polynomial matrix determinant based on FFT," in *Proc. Eur. Control Conf.*, 1999, pp. 4173–4177.
[40] D. Rocchesso, "Maximally diffusive yet efficient feedback delay networks for artificial reverberation," *IEEE Signal Process. Lett.*, vol. 4, no. 9, pp. 252–255, Sep. 1997.
[41] R. Väänänen, V. Välimäki, J. Huopaniemi, and M. Karjalainen, "Efficient and parametric reverberator for room acoustics modeling," in *Proc. Int. Comput. Music Conf.*, 1997, pp. 200–203.

Sebastian J. Schlecht (Senior Member, IEEE) received the Diploma in applied mathematics from the University of Trier, Trier, Germany, in 2010 and the M.Sc. degree in digital music processing from the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, U.K., in 2011. In 2017, he received the Doctoral degree in artificial spatial reverberation and reverberation enhancement systems with the International Audio Laboratories Erlangen, Erlangen, Germany. He is currently a Professor of Practice for Sound in virtual reality with the Acoustics Lab, Department of Information and Communications Engineering and Media Labs, Department of Art and Media, Aalto University, Espoo. From 2012 to 2019, Dr. Schlecht was also an External Research and Development Consultant and Lead developer of the 3D Reverb algorithm with the Fraunhofer IIS, Erlangen.

Jon Fagerström received the M.Sc. degree in electrical engineering, majoring in acoustics and audio technology, from Aalto University, Espoo, Finland, in 2020. He is currently working toward the Doctoral degree with the Acoustics Lab, Aalto University, Espoo. His research interests include sparse-noise modeling, artificial reverberation, and reverberation perception.

Vesa Välimäki (Fellow, IEEE) received the M.Sc. and D.Sc. degrees in electrical engineering from the Helsinki University of Technology (TKK), Espoo, Finland, in 1992 and 1995, respectively. He was a Postdoctoral Research Fellow with the University of Westminster, London, U.K., in 1996. From 1997 to 2001, he was a Senior Assistant (cf. Assistant Professor) with TKK. From 2001 to 2002, he was a Professor of signal processing with the Pori unit of the Tampere University of Technology, Tampere, Finland. From 2008 to 2009, he was a Visiting Scholar with the Stanford University Center for Computer Research in Music and Acoustics, Stanford, CA, USA. He is currently a Full Professor of audio signal processing and the Vice Dean of research in electrical engineering with Aalto University, Espoo. His research interests include audio and musical applications of signal processing and machine learning. Prof. Välimäki is a Fellow of the Audio Engineering Society. From 2007 to 2013, he was a Member of the Audio and Acoustic Signal Processing Technical Committee of the IEEE Signal Processing Society and is currently an Associate Member. From 2005 to 2009, he was an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS and from 2007 to 2011, an Associate Editor for the IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING. From 2015 to 2020, he was a Senior Area Editor of the ACM /IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING. In 2007, 2015, and 2019, he was the Guest Editor of special issues of the IEEE SIGNAL PROCESSING MAGAZINE and in 2010, of a special issue of the IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING.

He is the Editor-in-Chief of the *Journal of the Audio Engineering Society*.