FREQUENCY-DEPENDENT DIRECTIONAL FEEDBACK DELAY NETWORK

Benoit Alary[*]

Aalto University
Dept. of Signal Processing and Acoustics
Espoo, Finland

Archontis Politis

Tampere University
Faculty of Information Technology and
Communication Sciences
Tampere, Finland

ABSTRACT

A recent publication introduced the Directional Feedback Delay Network, a parametric artificial reverberation algorithm capable of producing direction-dependent energy decay. This method extends the capabilities of Feedback Delay Networks by using multichannel delay-line groups and a spatial transform to produce direction-dependent reverberation in the Ambisonics domain. In this paper, we present a modified formulation of the Directional Feedback Delay Network method that allows both frequency- and direction-dependent reverberation. Multichannel delay-line groups are used to manipulate signals incident on a spherical grid through independent recursive signal paths, while an early reflection module gives a physically-motivated spatial distribution of input signals in the system. The overall number of delay lines is reduced from the previous formulation, and the design allows flexibility to favor a lower computational cost over accuracy.

Index Terms— Artificial reverberation, multichannel sound reproduction, spatial audio

1. INTRODUCTION

Delay-based artificial reverberators are a popular family of reverberation methods that have been used for over fifty years [1, 2, 3, 4, 5]. Today, they remain widely popular since they offer flexible parametric properties suitable for aesthetic design as well as being computationally efficient [5]. One popular method, the Feedback Delay Network (FDN), generalized the use of multiple recirculating delay lines [2, 3, 4]. For multichannel sound reproduction, networks of recirculating delay lines have been used to produce output signals that are mutually decorrelated [2, 3]. This design follows the assumption that the reverberant sound field is isotropic after an initial mixing time [6, 7]; a decaying sound field is considered isotropic once its mean energy distribution is uniform in all directions.

Recent studies have demonstrated that anisotropy in the decaying sound field is observable both in simulations [8, 9] and recorded signals [10, 11]. The perception of anisotropic decays has also been investigated through listening experiments [12]. A factor contributing to anisotropy in the decay is the lack of diffusion in a room [13, 8].

Convolution with a captured multichannel impulse response [14] and simulations [15] can both produce anisotropy to some extent. However, parametric delay-based reverberators offer key benefits in terms of computational cost, interactivity, and tuning capabilities. For these reasons, several designs for multichannel reverberators have been proposed. In [16], the recirculating matrix of an FDN was used to alter the positioning of the output and to create multiple independent reverberators capable of reproducing scenarios such as decoupled rooms. In [17, 18], methods are presented to produce weighted directional output using decorrelated signals with an isotropic decay rate.

In [19], the authors proposed a Directional Feedback Delay Matrix (DFDN), a parametric method allowing direction-dependent energy decay by using multichannel delay lines and spatial transformation matrices in the recirculation path. However, the possibility of producing frequency-dependent characteristics as well is greatly hindered by the processing of Ambisonics signals. Since sound propagation in a room is inherently a frequency-dependent phenomena, due to acoustical properties being relative to wavelengths, this is also an important feature in artificial reverberation.

In this paper, we propose improvements to the DFDN method to simplify and generalize the parametric control of both frequency and directional properties. First, the FDN and the previous DFDN formulation are introduced, then we present the new formulation using absorbent filters within independent signal paths, representing different incident directions. The system also includes a formal definition of early reflections to produce physically-informed multichannel signals and to obtain perceptually important correlated signals during the early stage of reverberation. Finally, the accuracy of the system, and the impact of reducing the spatial resolution and delay-line groups are evaluated.

[*]Benoit's work has been funded by the Academy of Finland (ICHO project, Aalto University project No. 13296390).

2. BACKGROUND

An FDN consists of a set of recirculating delay lines interconnected through a feedback matrix $\mathbf{A}$, which defines the gains for each connection [4, 3]. The matrix is constrained to be unitary to create a lossless prototype, while the energy decay is controlled using attenuation gains or absorbent filters in the recirculation path. The decay is commonly characterized using a set of frequency-dependent target reverberation times $T_{60}(\omega)$, corresponding to the time required to reach 60 dB of attenuation in the decay. The output sample $y(n)$ of the recursive system, for an input $x(n)$ is formulated as

$$y(n) = \sum_{i=1}^{N} c_i g_i s_i(n) + dx(n), \tag{1}$$

$$s_i(n+m_i) = \sum_{j=1}^{N} A_{ij} g_j s_j(n) + b_i x(n), \tag{2}$$

where $s_i(n)$ is the summed output of the ith delay line of length m_i, weighted by input and output parametric gains b_i and c_i, respectively, d is the gain coefficient for the direct path, and g_i is the attenuation gain of the i-th delay line. Upon recirculation, the state values $s_j(n)$ are scaled by the corresponding value from the unitary matrix A_{ij}.

The DFDN [19] is an extension of the FDN that uses multichannel delay lines, called delay-line groups, to allow the directivity of the signal to persist through the recirculation. A set of target decay times $T_{60}(\phi, \theta)$ is used to specify the direction-dependent decay characteristics of the system. In [19], the multichannel delay-line groups are encoded in Ambisonics, and a directional weighting transform $\mathbf{T}$ is added in the recirculation path to apply direction-dependent attenuation.

Due to the spherical harmonic representation of the spatial signals in Ambisonics, a weighted sum of channels is necessary to produce an output in a given direction. As such, the spectral processing of the decay-per-direction requires either the introduction of absorbent filters in the directional weighting matrices $\mathbf{T}$ or the use of a filter bank to have a different $\mathbf{T}$ matrix for each frequency band. This has a significant impact on the complexity and computing cost of the reverberator.

3. PROPOSED METHOD

To generalize the design of a frequency-dependent DFDN without the drawback of Ambisonics processing, we propose to use a multichannel representation of delay-line groups where each channel corresponds to an incident direction on a spatial grid of Q directions. The spatial grid is defined as a uniform distribution of points on a sphere, which simplifies the modeling of a target reverberation profile. Similar to Ambisonics, we rely on the notion of spherical harmonic order to obtain a simple parameter for the grid generation. For a given order L we need at least $Q \geq (L+1)^2$ uniform points, such as spherical t-designs of degree $t = 2L$ [20]. The grid directions are collected in a matrix of unit vectors $\mathbf{U}_{\text{grid}} = [\mathbf{u}_1, ..., \mathbf{u}_Q]$, required in the spatialization of the signals in the system.

The signal flow of the reverberator is illustrated in Fig. 1. Starting with a mono signal $x(n)$, the input is split into multiple branches. First, the input is processed through an omnidirectional signal path, scaling the input to each delay-line group through a vector $\mathbf{b}_i$ of size $Q \times 1$. This scaling vector sets an initial gain for each channel, corresponding to the distribution of early energy during the scattering phase of the reverberation. The direct sound is encoded to the output using its incident direction (ϕ, θ) through

$$\mathbf{d} = \text{vbap}(\mathbf{U}_{\text{grid}}, (\phi, \theta)), \tag{3}$$

where vbap denotes an operator returning the vector-base amplitude panning (VBAP) weights [21] distributing the panned signal to a triplet of grid points enclosing the panning direction.

Finally, an early reflection module (ER) generates the first reflections of the reverberator. Initial attenuation values λ_r are calculated for a set of R reflections before being encoded to multichannel signals with $\boldsymbol{\lambda}_r = \lambda_r \, \text{vbap}(\mathbf{U}_{\text{grid}}, (\phi_r, \theta_r))$, (see Fig. 2). Individual delay times l_r are calculated for each reflection, and the output of the module is defined as

$$\mathbf{s}'_r(n + l_r) = \boldsymbol{\lambda}_r x(n) \tag{4}$$

$$\mathbf{s}'(n) = \sum_{r=1}^{R} \mathbf{s}'_r(n), \tag{5}$$

Each l_r value represents the input location of the signal to a given delay line [22], meaning the initial delay will be shortened to synchronize these reflections on the output channels (Fig. 2). These values are either randomized or physically informed through the use of a virtual acoustics algorithm [15].

Within a delay-line group, individual delay lengths are randomized as $\mathbf{m}_i \in [z_{\min}, z_{\max}]$. Shorter delays yield faster increase in echo density, whereas longer delays create higher modal density, both desirable properties.

The characterization of the decay is achieved through a set of target frequency- and direction-dependent decay parameters $T_{60}(\omega, \phi, \theta)$. These parameters are used to calculate per-sample attenuation coefficients using

$$g_{\text{dB}}(\omega, \phi, \theta) = \frac{-60}{T_{60}(\omega, \phi, \theta) f_{\text{s}}}, \tag{6}$$

$$g_{\text{lin}}(\omega, \phi, \theta) = 10^{\frac{g_{\text{dB}}(\omega, \phi, \theta)}{20}}, \tag{7}$$

which are raised to the power of the delay lengths to establish the necessary attenuation at the output of each delay line

$$\mathbf{g}_i(\omega, \phi, \theta) = (g_{\text{lin}}(\omega, \phi, \theta))^{\mathbf{m}_i}. \tag{8}$$

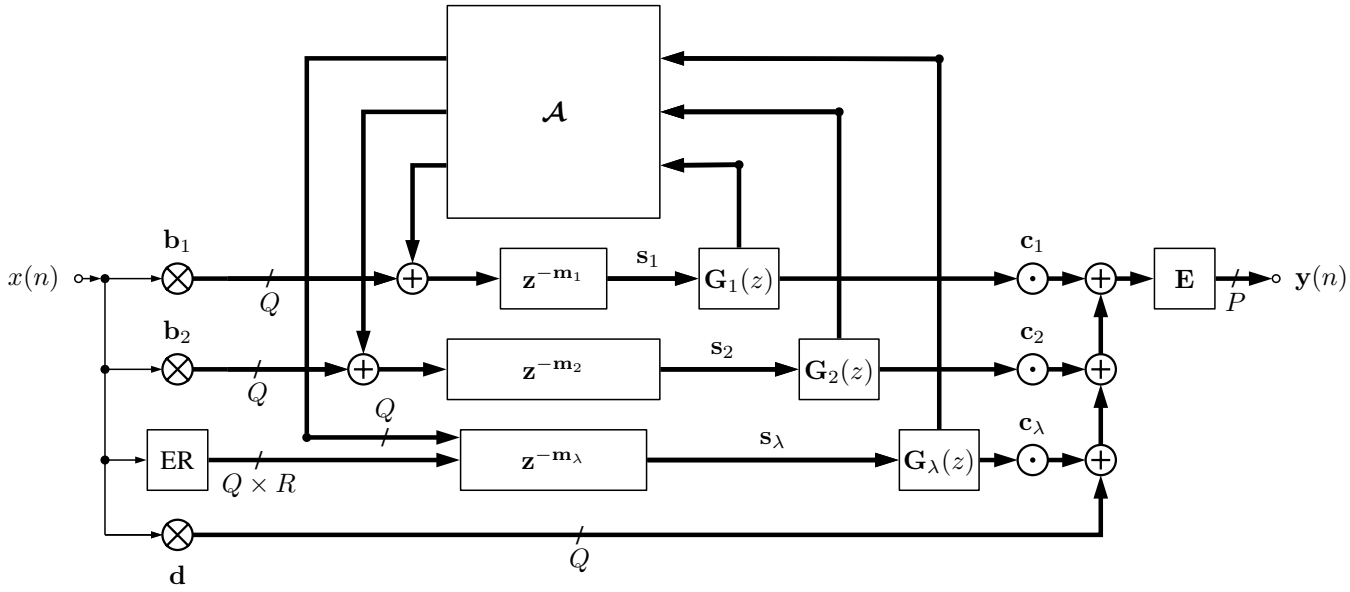

Fig. 1. *DFDN flow diagram. The thick lines represent multiple connections associated with a delay-line group containing Q delay lines.*

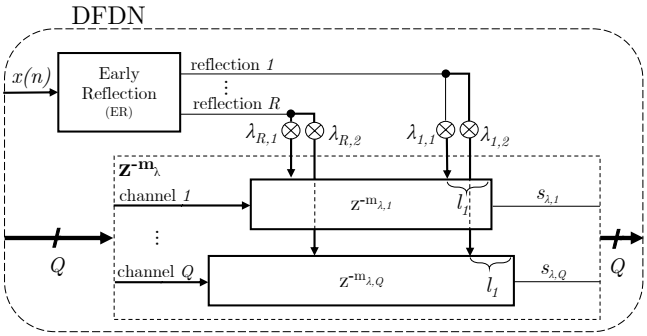

Fig. 2. *The early reflection module generates gain values $\boldsymbol{\lambda}_r$ and time delays l_r that determine the tap-in location to individual delay lines $(\mathbf{m}_i - l_r)$.*

Each channel is attenuated using a set of absorbent filters $G_i(z)$, one per frequency band [23]. As an example, a first-order low-shelf filter is created using [24]

$$G(z) = \frac{g\tan(\frac{\omega}{2}) + \sqrt{g} + [g\tan(\frac{\omega}{2}) - \sqrt{g}]z^{-1}}{\tan(\frac{\omega}{2}) + \sqrt{g} + [\tan(\frac{\omega}{2}) - \sqrt{g}]z^{-1}}, \quad (9)$$

where ω is the center frequency and g is the target attenuation corresponding to $\mathbf{g}_i(\omega, \phi, \theta)$. Following this procedure, K filters are used to form a multi-band equalizer which are cascaded through

$$\mathrm{G}(z) = \prod_{k=1}^{K} G_k(z). \quad (10)$$

The bold $\mathbf{G}_i(z)$ in Fig. 1 indicates that a delay-line group contains a set of filters per q channel, since each filter is dependent on its related delay length.

The number of delay-line groups N is chosen to maximize the echo density while keeping the computational cost of the system acceptable for a given application. The overall system will contain $N \times Q$ delay lines. A unitary matrix $\mathbf{A}$ of order N is created to control the recirculation of the signals between delay-line groups. This matrix is then expanded to a size of $NQ \times NQ$ while preventing recirculation between different channels with $\boldsymbol{\mathcal{A}} = \mathbf{A} \otimes \mathbf{I}_Q$, where $\otimes$ is the Kronecker product and $\mathbf{I}_Q$ is the identity matrix of order Q [19].

Before the signals are mixed to obtain the final output, the delay lines are scaled by a weighting vector $\mathbf{c}_i$, controlling the contribution of each channel and delay-line group. The mixed signal $\mathbf{y}'(z)$, in the z-domain, is given by

$$\mathbf{y}'(z) = \sum_{i=1}^{N} \mathbf{C}_i \mathbf{G}_i(z)\, \mathbf{s}_i(z) + \mathbf{d}\, x(z), \quad (11)$$

$$\mathbf{s}_i(z) = \begin{cases} \mathbf{s}'(z) + \mathbf{Z}_i(z)\Big[\sum_{j=1}^{N} A_{ij}\mathbf{G}_j(z)\mathbf{s}_j(z)\Big], i = \lambda \\ \mathbf{Z}_i(z)\Big[\mathbf{b}_i\, x(z) + \sum_{j=1}^{N} A_{ij}\mathbf{G}_j(z)\mathbf{s}_j(z)\Big], \text{ else} \end{cases} \quad (12)$$

where $\mathbf{C}_i = \mathrm{diag}(c_{i1}, ..., c_{iQ})$ and $\mathbf{Z}_i = \mathrm{diag}(z^{-m_{i1}}, ..., z^{-m_{iQ}})$ are diagonal matrices. In Fig. 1, $\odot$ denotes element-wise multiplication (Hadamard product).

Finally, $\mathbf{y}'$, containing the Q grid signals, is encoded to a desired playback system of P channels using a $P \times Q$ encod-

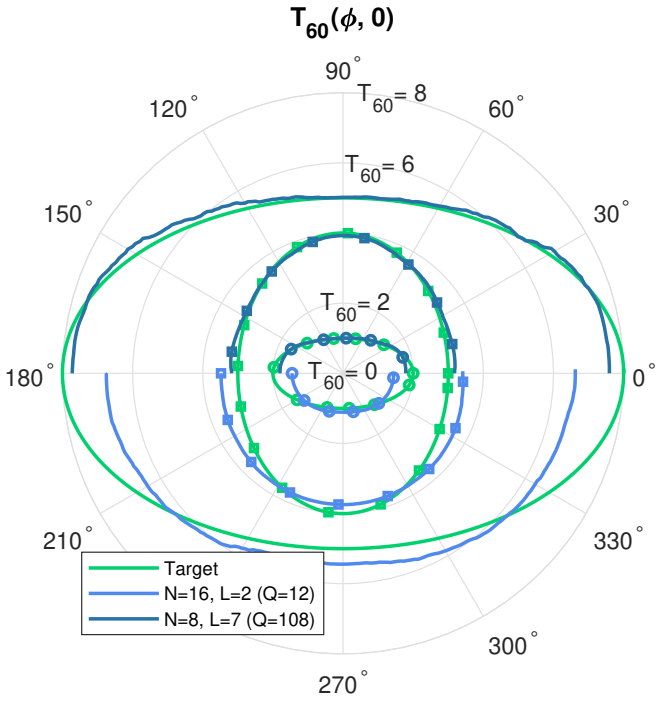

Fig. 3. *Directional T_{60} on the lateral plane at different frequency bands, centered at 125 Hz (line), 1 kHz (square), and 8 kHz (circle). Due to symmetry and for better readability, only half of the results are shown.*

ing matrix $\mathbf{E}$ as $\mathbf{y}(n) = \mathbf{E}\,\mathbf{y}'(n)$, where $\mathbf{E}$ can be

$$\mathbf{E} = \begin{cases} \mathbf{Y}(L, \mathbf{U}_{\text{grid}}), & (L+1)^2 \times Q, \text{ambisonics}, \\ \text{vbap}(\mathbf{U}_{\text{ls}}, \mathbf{U}_{\text{grid}}), & P \times Q, \text{loudspeakers}, \\ \text{hrtf}(\mathbf{U}_{\text{grid}}), & 2 \times Q, \text{binaural}. \end{cases} \tag{13}$$

$\mathbf{Y}$ contains real spherical harmonic values up to order L for the grid directions, the $\text{vbap}()$ matrix pan the grid directions to a loudspeaker set-up $\mathbf{U}_{\text{ls}}$, and the $\text{hrtf}()$ filter matrix performs binaural encoding of the grid signals.

4. EVALUATION

The number of grid points determine how accurately the system can approximate a target decay shape. Since a very large system would be unpractical, we evaluated the impact of using smaller configurations. To obtain reliable results in the frequency domain, we used octave-band absorbent filters [25, 26] and focused the evaluation on the directional properties per frequency bands. For this, the early reflections were omitted, the number of delay-line groups (N) and grid points (Q) were varied, and a single impulse was injected to the system. The output was encoded to Ambisonics using Eq. (13), and the directional T_{60} was analyzed from signals obtained through plane-wave decomposition at various angles on the lateral plane. To define the attenuation in the system, a set of

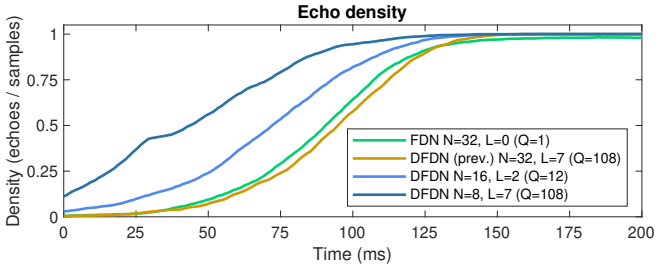

Fig. 4. *Comparison of echo density over time of different configurations, (prev.) refers to the previous formulation [19].*

three target decay shapes was created, each corresponding to a distribution of $T_{60}(\boldsymbol{\omega}, \phi, \theta)$ with $\boldsymbol{\omega} = [125\,\text{Hz}, 1\,\text{kHz}, 8\,\text{kHz}]$. In Fig. 3, we compare two configurations of the reverberator to the frequency- and direction-dependent target shape. With a grid size of $Q = 108$, the target shape was closely reproduced, whereas a grid size of $Q = 12$ demonstrates the smoothing effect of lower-order spatial resolution.

An important attribute of reverberation is the echo density build up over time. Since the multichannel signals integrate at the listener's ears, the DFDN requires less density from individual channels when compared to a monophonic reverberator, meaning that the number of delay-line groups can be reduced. To measure the integrated echo density, an DFDN without absorbent filters was created, its output channels were summed, and the echo density was calculated by counting the non-zero elements of short-time windows. As a reference, we used a single-channel configuration with 32 delay lines, which corresponds to a typical configuration of an FDN. Fig. 4 shows that an echo density higher than the reference can be achieved with a smaller number of delay-line groups. Although using a lower N did not impact the directional accuracy (Fig. 3), a very low value should be avoided to prevent an adverse effect on the modal density. In the previous formulation, [19], the channels of a delay-line group shared the same delay length, thus not contributing to a higher integrated sum, and requiring the same N value as the FDN.

5. CONCLUSION

In conclusion, we presented an artificial reverberation algorithm capable of producing precise frequency- and direction-dependent decay characteristics using multichannel signals on a spatial grid and absorbent filters. An early reflection module was proposed to combine the anisotropic late reverberation to a set of physically-informed directional reflections. We demonstrated that the number of delay-line groups can be reduced while retaining sufficient echo density, and the order of the spatial grid can be adjusted to favor a lower computing cost over accuracy.

6. REFERENCES

[1] M. R. Schroeder and B. F. Logan, "'Colorless' artificial reverberation," *J. Audio Eng. Soc.*, vol. 9, no. 3, pp. 192–197, Jul. 1961.

[2] M. A. Gerzon, "Synthetic stereo reverberation: Part I," *Studio Sound*, vol. 13, pp. 632–635, Dec. 1971.

[3] J. Stautner and M. Puckette, "Designing multi-channel reverberators," *Computer Music J.*, vol. 6, no. 1, pp. 52–65, Spring 1982.

[4] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. 90th Conv. of the Audio Eng. Soc.*, Paris, France, Feb. 1991.

[5] 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.

[6] V. Pulkki, "Spatial sound reproduction with directional audio coding," *J. Audio Eng. Soc.*, vol. 55, no. 6, pp. 503–516, Jun. 2007.

[7] A. Lindau, L. Kosanke, and S. Weinzierl, "Perceptual evaluation of model- and signal-based predictors of the mixing time in binaural room impulse responses," *J. Audio Eng. Soc*, vol. 60, no. 11, pp. 887–898, Dec. 2012.

[8] R. Badeau, "General stochastic reverberation model.," *Research Report, Télécom ParisTech*, 2019.

[9] M. Berzborn and M. Vorländer, "Investigations on the directional energy decay curves in reverberation rooms," in *Proc. of Euronoise*. Crete, Greece, May 2018.

[10] B. Alary, P. Massé, V. Välimäki, and M. Noisternig, "Assessing the anisotropic features of spatial impulse responses," in *Proc. EAA Spatial Audio Sig. Process. Symp.*, Sep. 2019, pp. 43–48.

[11] M. Berzborn, M. Nolan, E. Fernandez-Grande, and M. Vorländer, "On the directional properties of energy decay curves," in *Proc. 23rd Int. Cong. of Acoustics*. Aachen, Germany, Sep. 2019.

[12] D. Romblom, C. Guastavino, and P. Depalle, "Perceptual thresholds for non-ideal diffuse field reverberation," *J. Acoust. Soc. Am.*, vol. 140, no. 5, pp. 3908–3916, Nov. 2016.

[13] C. G. Balachandran and D. W. Robinson, "Diffusion of the decaying sound field," *Acta Acust.*, vol. 19, no. 5, pp. 245–257, Jan. 1967.

[14] M. A. Gerzon, "Recording concert hall acoustics for posterity," *J. Audio Eng. Soc.*, vol. 23, no. 7, pp. 569–571, Sep. 1975.

[15] L. Savioja and U. P. Svensson, "Overview of geometrical room acoustic modeling techniques," *J. Acoust. Soc. Am.*, vol. 138, no. 2, pp. 708–730, Aug. 2015.

[16] S. J. Schlecht and E. A. P. Habets, "Sign-agnostic matrix design for spatial artificial reverberation with Feedback Delay Networks," in *Proc. Audio Eng. Soc. Int. Conf. Spatial Reproduction—Aesthetics and Science*, Tokyo, Japan, Aug. 2018.

[17] J. Anderson and S. Costello, "Adapting artificial reverberation architectures for b-format signal processing," in *Proc. of the Ambisonics Symp.* Graz, Austria, Jun. 2009.

[18] B. Wiggins and M. Dring, "Ambifreeverb 2-development of a 3D ambisonic reverb with spatial warping and variable scattering," in *Proc. Audio Eng. Soc. Int. Conf. on Sound Field Control*, Guildford, UK, Jul. 2016.

[19] B. Alary, A. Politis, S. J. Schlecht, and Vesa Välimäki, "Directional Feedback Delay Network," *J. Audio Eng. Soc.*, vol. 67, no. 10, Oct. 2019.

[20] R. H. Hardin and N. J. A. Sloane, "McLaren's improved snub cube and other new spherical designs in three dimensions," *Discrete Comput. Geom.*, vol. 15, no. 4, pp. 429–441, Jul. 1996.

[21] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, Jun. 1997.

[22] S. Schlecht, A. Silzle, E. Habets, C. Borss, B. Neugebauer, and H. Stenzel, "Apparatus and method for generating output signals based on an audio source signal, sound reproduction system and loudspeaker signal," Patent:PCT/EP2015/075141, May 2016.

[23] J.-M. Jot, "Proportional parametric equalizers—application to digital reverberation and environmental audio processing," in *Proc. Audio Eng. Soc. Conv. 139*, Oct. 2015.

[24] V. Välimäki and J. D. Reiss, "All about audio equalization: Solutions and frontiers," *Appl. Sci.*, vol. 6, no. 5, May 2016.

[25] V. Välimäki and J. Liski, "Accurate cascade graphic equalizer," *IEEE Signal Processing Letters*, vol. 24, no. 2, pp. 5; 176–180, Feb. 2017.

[26] K. Prawda, V. Välimäki, and S. J. Schlecht, "Improved reverberation time control for Feedback Delay Networks," in *Proc. 22th Int. Conf. Digital Audio Effects (DAFx-19)*, Birmingham, UK, Sept. 2019.