Feedback Delay Networks: Echo Density and Mixing Time

Sebastian J. Schlecht and Emanuël A. P. Habets, *Senior Member, IEEE*

***Abstract*—Feedback delay networks (FDNs) are frequently used to generate artificial reverberation. This paper discusses the temporal features of impulse responses produced by FDNs, i.e., the number of echoes per time unit and its evolution over time. This so-called echo density is related to known measures of mixing time and their psychoacoustic correlates such as auditive perception of the room size. It is shown that the echo density of FDNs follows a polynomial function, whereby the polynomial coefficients can be derived from the lengths of the delays for which an explicit method is given. The mixing time of impulse responses can be predicted from the echo density, and conversely, a desired mixing time can be achieved by a derived mean delay length. A Monte Carlo simulation confirms the accuracy of the derived relation of mixing time and delay lengths.**

***Index Terms*—Artificial reverberation, Ehrhart theory, feedback delay network (FDN), mixing time.**

I. Introduction

A SOUND emitted in an enclosed space is repeatedly reflected by its boundaries and results in *reverberation*. The number and distribution of reflections over time, the so-called *echo density* is important to characterize the reverberation as it determines the perception of the sound texture [1], [2] and the perception of the room size and shape [3]. Therefore, to create a convincing artificial reverberator, it is crucial to match a desired echo density.

Two types of echo densities can be distinguished: *absolute* echo density, denoted by $\mathcal{A}$, which is the expected number of echoes reaching the receiver per time unit, and *empirical* echo density, denoted by $\mathcal{E}$, which is the probability of at least one echo reaching the receiver within a time unit.[1] Throughout the paper, the time unit for measuring echo densities is a *sampling period* $1/f_s$ where f_s denotes the sampling frequency. The empirical echo density can be determined from a measured impulse response at the receiver, where overlapping echoes cannot be distinguished and counted separately. On the contrary, because of the overlaps, it is not possible to derive the absolute echo density from a measured impulse response.

If an impulse sound is emitted in a rectangular room, the absolute echo density can be estimated as the ratio of the time derived volume of the expanding "time sphere" with radius cn/f_s to the room volume V [1, p. 110]:

$$\mathcal{A}_{\text{geometric}}(n) = \frac{4\pi c^3}{V f_s^3} n^2, \tag{1}$$

where c is the speed of sound and n is the time in samples. The echo density $\mathcal{A}_{\text{geometric}}(n)$ in a physical room is a polynomial of degree two and it is statistically similar for any room geometry [5]. This is also supported by experimental results by Defrance et al. [6]. Equal polynomial degree can be obtained by counting image-sources depending on the reflection order [7], [8]. The echo density can be retrieved also from a measured impulse response by the matching pursuit algorithm [6]. As the echo density increases over time, the impulse response evolves from sparse early reflections to a dense late reverberation, whereas the transition point has been coined *mixing time* [9].

In the context of delay based reverberators that were first introduced in [10], [11], Jot and Chaigne described in [4] the echo density of a parallel comb-filter section as typically used in the Schroeder reverberator [11]

$$\mathcal{A}_{\text{parallelComb}}(n) = \sum_{i=1}^{N} \frac{1}{m_i}, \tag{2}$$

where $m_1, \dots m_N$ are the lengths of the N delays in samples. The echo density of the parallel comb-filter section is constant over time and it was soon realized that the echo density can be increased by interconnecting the feedback of the comb-filters resulting in the *feedback delay network* (FDN). Since its introduction by Stautner and Puckette [12] and its essential development by Jot and Chaigne [4], the popularity of employing *feedback delay networks* (FDNs) to create reverberation artificially has increased steadily during the last five decades [13]–[17]. This can be accounted to the FDN's flexibility in design, computational efficiency, and independent control over energy decay, amount of diffusion and total energy in each frequency band [13]. Further work has been done to maximize the echo density of FDNs by using specific feedback matrices [18]–[20].

In this contribution, we determine the echo density of an FDN over time and by this provide insight into the temporal evolution and the mixing time of an FDN's impulse response. Opposed to counting echoes in the impulse response, the presented

Manuscript received May 24, 2016; revised October 7, 2016; accepted November 22, 2016. Date of publication December 1, 2016; date of current version January 10, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Augusto Sarti.

The authors are with the International Audio Laboratories Erlangen (A joint institution of the University of Erlangen-Nuremberg and Fraunhofer Institut für Integrierte Schaltungen), Erlangen 91058, Germany (e-mail: sebastian.schlecht@audiolabs-erlangen.de; emanuel.habets@audiolabs-erlangen.de).

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

Digital Object Identifier 10.1109/TASLP.2016.2635027

[1]The definitions of absolute and empirical echo densities are the authors' own definitions. However, similar terms have been used before like *echo density* and *time density* [4].

prediction method gives means to determine the analytic properties of the echo density. It is shown that, under certain assumptions, the echo density follows a polynomial function. In addition, a method to calculate the polynomial coefficients from the lengths of the feedback delays is provided. Further, two approximations of the echo density, which are easier to compute, are developed. The accuracy of the mixing time predicted from the analytical echo density is evaluated by a Monte Carlo simulation.

This paper is organized as follows. In Section II, various definitions of mixing time and their connection to echo density are reviewed. In Section III, the echo density of the FDN is presented alongside two approximation methods for the echo density. In Section IV, the empirical echo density is derived from the absolute echo density. In Section V, the derived absolute and empirical echo densities are applied to the estimation of the FDN mixing time and, in reverse, the derivation of delay lengths based on a specified mixing time.

II. Background

A. Mixing Time

Mixing time represents a perceptual phenomenon caused by the limited temporal resolution of the auditory system and its inability to distinguish successive reflections in an increasingly diffuse sound field [21]. Although, spatial diffuseness is important for the perceptual mixing time [22], the present study evaluates the mixing time solely using monophonic and therefore non-spatial criteria. In the following, three established mixing time measures are reviewed and subsequently related to the absolute and empirical echo densities.

Polack [23] suggested a mixing time which is heuristically based on auditory perception: Perceptually mixing is reached if more than 10 echoes arrive within 24 ms, where the time interval has been chosen according to the auditory integration interval, i.e.,

$$t_{\text{Polack}} = \arg\min_n \left[\mathcal{E}(n) \geq \frac{10}{0.024 f_s} \right], \tag{3}$$

where f_s is the sampling frequency. Defrance *et al.* [24] based the definition of the mixing time on the notion of ergodicity, i.e., understanding the reverberation as a transition from a concentrated impulse to a diffuse field. The mixing time is then defined as the time when the echoes begin to overlap in time, or in other words, when the empirical echo density becomes a linear function:

$$t_{\text{Defrance}} = \arg\min_n \left[\mathcal{E}(n) \geq \mathcal{T} \right], \tag{4}$$

where $\mathcal{T} = 0.9$ denotes the density threshold that was chosen empirically. Within the context of this work, the echo bandwidth is assumed to be full bandwidth of $f_s/2$ such that the expansion in time of an echo is a single sampling period of $1/f_s$ and an overlap occurs only if multiple echoes arrive at the same sample instance.

In ergodic rooms, the impulse response eventually converges to a Gaussian process. Abel and Huang [3] defined the mixing time as the time instant at which the levels of the impulse response are Gaussian distributed. The so-called *echo density profile* is defined in [3] as

$$\eta(n) = \frac{1}{\text{erfc}(1/\sqrt{2})} \sum_{\tau=n-\delta}^{n+\delta} w(\tau) |\{|h(\tau)| > \sigma\}|, \tag{5}$$

where $|\{\cdot\}|$ indicates the cardinality of a set, $2\delta + 1$ is the length of the window w in samples, σ is the standard deviation of the windowed response, $\text{erfc}(1/\sqrt{2}) = 0.3173$ is the expected fraction of samples lying outside a standard deviation from the mean of a Gaussian distribution. A closely related measure was introduced by the same authors [25] by relating the absolute echo density to the echo density profile:

$$\eta(n) \approx \frac{\mathcal{A}(n)}{\mathcal{A}(n) + 1}. \tag{6}$$

The mixing time is then defined as exceeding a threshold value, e.g., $\mathcal{T} = 0.9$:

$$t_{\text{Abel}} = \arg\min_n \left[\eta(n) \geq \mathcal{T} \right]. \tag{7}$$

In summary, all three definitions of mixing time can be derived from either the empirical or absolute echo density. Sections III–IV are dedicated to the estimation of the absolute and empirical echo densities of the FDN.

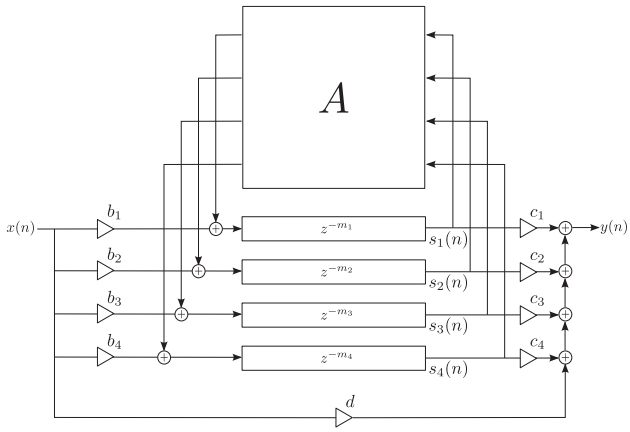

Fig. 1. Feedback delay network structure for artificial reverberation as proposed by Stautner and Puckette [12].

B. Feedback Delay Network

As depicted in Fig. 1, an FDN is built around N delays with lengths $\boldsymbol{m} = [m_1, \ldots, m_N]$ in samples. The output of the delay lines is mixed and returned back to the delay lines through the *feedback matrix* $\boldsymbol{A}$. The state-space definition of an FDN is

$$\begin{aligned} y(n) &= \sum_{i=1}^{N} c_i s_i(n) + d x(n) \\ s_i(n + m_i) &= \sum_{j=1}^{N} a_{ij} s_j(n) + b_i x(n) \end{aligned} \tag{8}$$

where $x(n)$ and $y(n)$ are the input and output values respectively and $s_i(n), 1 \leq i \leq N$, are the delay line outputs at time sample n [14]. The double-indexed a_{ij} weights form an

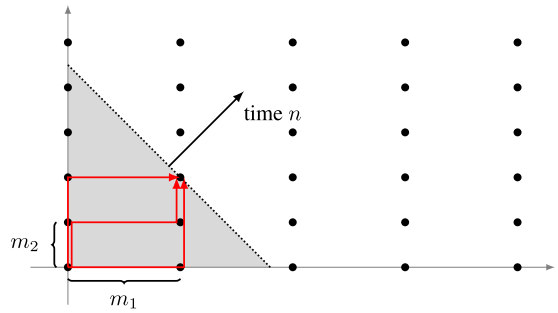

Fig. 2. Echo lattice of an FDN with $N = 2$ and $\boldsymbol{m} = [m_1, m_2]$. The dots indicate the coordinates of the echoes in terms of the delay lengths. The red arrows indicate three paths $\boldsymbol{p} = [2, 2, 1], [2, 1, 2]$ and $[1, 2, 2]$ with the same echo time $n_{\boldsymbol{p}}$. The dotted line indicates the expanding time polygon, whereas the gray shaded triangle indicates the time simplex.

$N \times N$ feedback matrix $\mathbf{A}$. In a similar manner, we define $\mathbf{b} = [b_1\ b_2\ \ldots b_N]^\top$ and $\mathbf{c} = [c_1\ c_2\ \ldots c_N]^\top$, where $(\cdot)^\top$ is the transpose operation.

III. Absolute Echo Density

A. Absolute Echo Density of the FDN

An impulse starting for instance in the first delay line, is distributed through the feedback matrix to other delay lines and after it has traversed the successive delay line, the impulse is send to other delay lines again. The path the impulse goes through can be described by an ordered list of delay line numbers. Let us define the sequence $\boldsymbol{p} \in \{1, 2, \ldots, N\}^l$ as a *path* of length l and p_i the i^{th} element of the sequence. The *echo time* $n_{\boldsymbol{p}}$ of $\boldsymbol{p}$ gives the time at which the echo appears at the output of the FDN: $n_{\boldsymbol{p}} = \sum_{i=1}^{l} m_{p_i}$ (see Fig. 2). The impulse response of the FDN, denoted by h, can therefore be expressed as

$$h(n) = \sum_{\boldsymbol{p} \in \mathcal{P}, n_{\boldsymbol{p}} = n} a_{\boldsymbol{p}}, \tag{9}$$

where $\mathcal{P}$ is the set of all paths $\mathcal{P} = \bigcup_{l \geq 1} \{1, 2, \ldots, N\}^l$ and $a_{\boldsymbol{p}}$ is the level of the echo path $\boldsymbol{p}$, i.e.,

$$a_{\boldsymbol{p}} = b_{p_1} \left(\prod_{k=1}^{l-1} a_{p_{k+1}\, p_k} \right) c_{p_l} \tag{10}$$

resulting from the input and output gains $\boldsymbol{b}$ and $\boldsymbol{c}$, and the gains in the feedback matrix $\boldsymbol{A}$. The echo time $n_{\boldsymbol{p}}$ is not dependent on the particular order of the path $\boldsymbol{p}$ such that we collect all paths with the same number of occurrences of the i^{th} delay line $\boldsymbol{q} \in \mathbb{N}_0^N$ in $\mathcal{P}_{\boldsymbol{q}}$:

$$\mathcal{P}_{\boldsymbol{q}} = \{\boldsymbol{p} \in \mathcal{P} \mid |\{p_k = i\}| = q_i \text{ for } 1 \leq i \leq N\}, \tag{11}$$

where $|\{\cdot\}|$ indicates the cardinality of a set. An *echo* corresponds then to $\mathcal{P}_{\boldsymbol{q}}$ or $\boldsymbol{q}$, respectively. The level of the echo $a_{\boldsymbol{q}}$ is then given by:

$$a_{\boldsymbol{q}} = \sum_{\boldsymbol{p} \in \mathcal{P}_{\boldsymbol{q}}} a_{\boldsymbol{p}}. \tag{12}$$

Assuming that all echo levels $a_{\boldsymbol{q}}$ are non-zero, the absolute echo density of an FDN is

$$\mathcal{A}_{\text{FDN}}(n) = \left| \left\{ \boldsymbol{q} \in \mathbb{N}_0^N \mid \boldsymbol{q}^\top \boldsymbol{m} = n \right\} \right| \tag{13}$$

The number of echoes per time unit in (13) is deterministic and coincide always with the expected number of echoes.

To ensure non-zero echo levels $a_{\boldsymbol{q}}$ in (12), it is not necessary to assume an entirely non-zero feedback matrix $\boldsymbol{A}$. Take for instance an FDN with $\boldsymbol{b} = \boldsymbol{c} = \mathbf{1}, d = 1$ and with $\boldsymbol{A}$ being a lower triangular matrix occupied with ones. In this case, all paths $\boldsymbol{p} \in \mathcal{P}_{\boldsymbol{q}}$ have $a_{\boldsymbol{p}} = 0$ in (10) except one path $\boldsymbol{p}_0$ for which $a_{\boldsymbol{p}_0} = 1$. The path $\boldsymbol{p}_0$ is the only path which traverses the delay lines in non-descending order. Therefore, the echo level in (12) is $a_{\boldsymbol{q}} = 1$. In fact, the impulse response h_Δ of such an FDN is identical to the absolute echo density:

$$h_\Delta(n) = \mathcal{A}_{\text{FDN}}(n) \text{ for all } n \geq 0. \tag{14}$$

In [26], the authors have shown that a series connection of comb-filters is equivalent to an FDN with an lower triangular feedback matrix whose rows and columns are simultaneously permuted. Hence, the series connection of allpass comb-filters, as for example used by the Schroeder reverberator [27], has the same echo density than the FDN with an entirely non-zero feedback matrix $\boldsymbol{A}$ and with the same number of delay lines.

B. Ehrhart Theory

The formulation of the absolute echo density $\mathcal{A}_{\text{FDN}}$ in (13) suggests that the echoes are placed on a *echo lattice* where the spacing of the echoes in each dimension is given by the delays $\boldsymbol{m}$. The echo time is given by the distance to the origin in the so-called Manhattan norm, i.e., the sum of the absolute cartesian coordinates. In reverse, all echoes with equal echo time n are enclosed on a skewed $N-1$-dimensional *time polygon* with vertices $(n, 0, \ldots, 0), (0, n, 0, \ldots, 0) \ldots, (0, \ldots, 0, n)$. Accordingly, all past echoes are enclosed in the N-dimensional *time simplex* with vertices $(0, 0, \ldots, 0), (n, 0, \ldots, 0), \ldots, (0, \ldots, 0, n)$. The echo lattice splits the space into equal *echo boxes* with volume $\prod_{i=1}^{N} m_i$ such that each echo box corresponds to a single echo. In a first approximation, the larger the volume of the echo boxes are the lower the absolute echo density will be. Figure 2 depicts the echo lattice and the moving time polygon for $N = 2$. Geometrically speaking, the absolute echo density $\mathcal{A}_{\text{FDN}}$ describes the number of lattice points on the time polygon as it moves away from the origin.

The study of counting lattice points in expanding polygons, the so-called Ehrhart theory, has been an active field of discrete linear programming since the formulation by its eponym Eugene Ehrhart [28], [29]. In Ehrhart theory, (13) is called the *restricted partition problem*: given an integer n, how many positive integer partitions of n into the parts $\boldsymbol{m}$ exists?

It has been shown that the number of solutions to the restricted partition problem is a quasi-polynomial [30], i.e., there are periodic coefficients $c_0(n), \ldots, c_{N-1}(n)$ such that

$$\mathcal{A}_{\text{FDN}}(n) = n^{N-1} c_{N-1}(n) + \cdots + n c_1(n) + c_0(n). \tag{15}$$

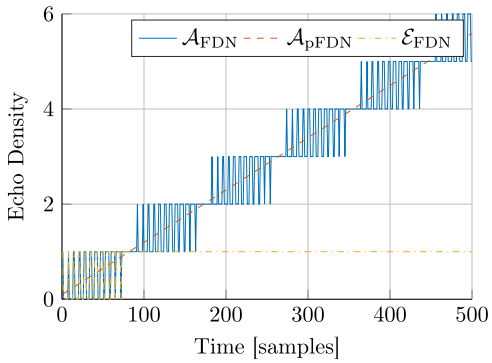

Fig. 3. Absolute, polynomial absolute, and empirical echo density for FDN with $N = 2$ and delays $\boldsymbol{m} = [7, 13]$. Absolute and empirical echo densities are identical up to the point of saturation $t_{\text{saturate}} = 73$. The polynomial part $\mathcal{A}_{\text{pFDN}}(n)$ is given by (21).

The period of the coefficients is at most the least common multiple of $\boldsymbol{m}$ [31, p. 57]. Because of the periodic nature, the periodic coefficient $c_i(n)$ can be divided into a constant part c_i and a periodic part with average zero. The polynomial part of $\mathcal{A}_{\text{FDN}}$ with constant coefficients is given by:

$$\mathcal{A}_{\text{pFDN}}(n) = n^{N-1}c_{N-1} + \cdots + nc_1 + c_0. \tag{16}$$

The polynomial part $\mathcal{A}_{\text{pFDN}}$ is particularly close to $\mathcal{A}_{\text{FDN}}$ if the delays $\boldsymbol{m}$ are pairwise relatively prime, i.e., when $\gcd(m_i, m_j) = 1$ for any i and j, where $\gcd$ yields the greatest common divisor, then all periodic coefficients of $\mathcal{A}_{\text{FDN}}$ except $c_0(n)$ are constant [30]. In other words, if the delays $\boldsymbol{m}$ are pairwise relatively prime, the polynomial part $\mathcal{A}_{\text{pFDN}}$ averages the only periodic coefficient $c_0(n)$ of $\mathcal{A}_{\text{FDN}}$. Figure 3 shows the discrete absolute echo density $\mathcal{A}_{\text{FDN}}(n)$ and the polynomial absolute echo density $\mathcal{A}_{\text{pFDN}}(n)$ for $N = 2$ and delays $\boldsymbol{m} = [7, 13]$. Whereas the discrete absolute echo density $\mathcal{A}_{\text{FDN}}(n)$ gives the exact number of echoes per samples, the polynomial part is an average number of echoes over the period of the quasi-polynomial in (15).

C. Ehrhart Polynomial for the Absolute Echo Density

The computation of the periodic coefficient functions $c_0(n), c_1(n), \ldots, c_{N-1}(n)$ required for $\mathcal{A}_{\text{FDN}}$ in (15) demands elaborate algorithms, whose description would outreach the scope of this work [32] and which are practically intractable for the high dimensions used in the context of artificial reverberation design. Fortunately, the polynomial part $\mathcal{A}_{\text{pFDN}}(n)$ has been recently stated in closed form [31, pp. 151]:

$$\mathcal{A}_{\text{pFDN}}(n) = \frac{(-1)^N}{(N-1)!} B_N^{\boldsymbol{m}}(-n) \tag{17}$$

where $B_N^{\boldsymbol{m}}(n)$ is a so-called Bernoulli-Barnes polynomial, which is defined as

$$B_N^{\boldsymbol{m}}(n) = n^{N-1}c_{N-1} + \cdots + nc_1 + c_0$$
$$c_l = \binom{N-1}{l} B_{N-1-l}(\boldsymbol{m}) \tag{18}$$

where $B_k(\boldsymbol{m})$ are the so-called Bernoulli-Barnes numbers with the definition:

$$B_k(\boldsymbol{m}) = \sum_{q_1+\cdots+q_N=k} \binom{k}{q_1, \ldots, q_N} \prod_{i=1}^{N} m_i^{q_i - 1} B_{q_i}, \tag{19}$$

where B_m are the Bernoulli numbers given by [33]

$$B_m = \sum_{k=0}^{m} \sum_{v=0}^{k} (-1)^v \binom{k}{v} \frac{v^m}{k+1}. \tag{20}$$

The closed forms of $\mathcal{A}_{\text{pFDN}}$ for $N = 1, 2$ and 3 [31, p. 152] are for example:

$$N = 1: \quad \frac{1}{m_1}$$
$$N = 2: \quad \frac{n}{m_1 m_2} + \frac{1}{2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right)$$
$$N = 3: \quad \frac{n^2}{2m_1 m_2 m_3} + \frac{n}{2}\left(\frac{1}{m_1 m_2} + \frac{1}{m_1 m_3} + \frac{1}{m_2 m_3}\right)$$
$$+ \frac{1}{12}\left(\frac{3}{m_1} + \frac{3}{m_2} + \frac{3}{m_3} + \frac{m_3}{m_1 m_2} + \frac{m_2}{m_1 m_3} + \frac{m_1}{m_2 m_3}\right). \tag{21}$$

The closed form expression for $N = 2$ was employed to determine $\mathcal{A}_{\text{pFDN}}$ in Fig. 3. The number of summands in (19) grows exponentially such that the numerical implementation for high N is a challenging task. Moved this section here. We have formulated the Ehrhart $\mathcal{A}_{\text{pFDN}}$ in (17), however, the computation of the coefficients $c_0, \ldots, c_{N-1}$ remains rather complicated in the general case. Solely, the leading coefficient c_{N-1} can be easily given by [31]

$$c_{N-1} = \frac{1}{(N-1)! \prod_{i=1}^{N} m_i}. \tag{22}$$

In the following, we present two approximation methods to give a simplified estimate of the polynomial part of the absolute echo density $\mathcal{A}_{\text{pFDN}}$ that in reverse also allows simple computation of the delay lengths $\boldsymbol{m}$ from a desired mixing time.

D. Volume Approximation

We introduce an approximation of the polynomial part of the absolute echo density $\mathcal{A}_{\text{pFDN}}$ based on the leading coefficient. The following approximation can be derived from the resemblance with the echo counting in geometric room acoustics as given in (1). The echo density of the geometric room acoustic derives from the the ratio of the time derived volume of the "time sphere" $4\pi c^3 n^2 / f_s^3$ to the volume of the room $V = L_x \times L_y \times L_z$, where L_x, L_y and L_z are the dimensions of the room. Equivalently, the echo density of the FDN can be estimated from the ratio of time derived time simplex to the volume of the echo boxes. The time simplex as it was defined in Section III-B is the Ehrhart theory pendant to the time sphere in geometric room acoustics. Similarly, the volume of the echo boxes in Ehrhart theory, as well as the room volume in geometric room acoustics both indicate the volume that contains exactly one echo. The volume of the *time simplex* is

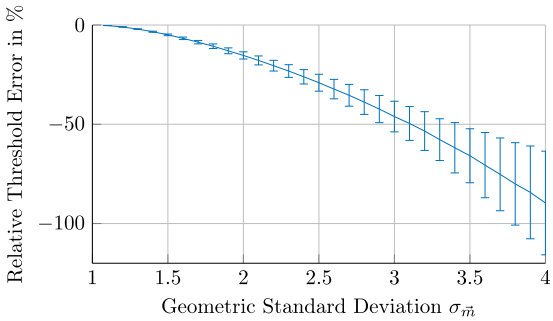

Fig. 4. Mean relative threshold error (solid) and error standard deviation (error bars) of $\left(t_{\text{equi}} - t_{\text{pFDN}}\right)/t_{\text{pFDN}}$ between the equilateral approximation $\mathcal{A}_{\text{equi}}$ and the absolute echo density $\mathcal{A}_{\text{pFDN}}$ for $N = 8$.

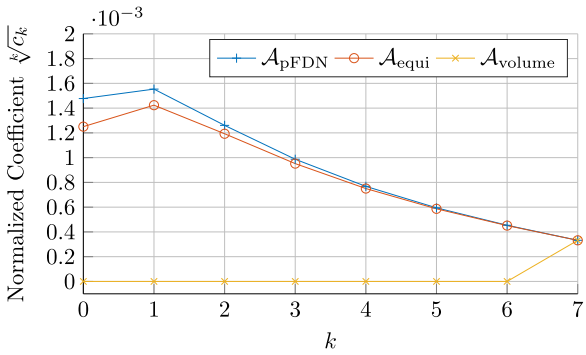

Fig. 5. Normalized coefficient of the absolute echo density and its approximations with $\boldsymbol{m} = [479, 587, 673, 773, 881, 991, 1091, 1201]$ and $N = 8$.

$\det(n\boldsymbol{I}_N)/N!$, where det denotes the determinant and $\boldsymbol{I}_N$ is the $N \times N$ identity matrix and therefore its time derivative is $\frac{n^{N-1}}{(N-1)!}$. The volume of the echo boxes is $\prod_{i=1}^{N} m_i$. The ratio of time derived volume of the time simplex and the echo boxes corresponds to the very leading coefficient c_{N-1} of $\mathcal{A}_{\text{pFDN}}$, as given in (22), such that we call this intuitive simplification of $\mathcal{A}_{\text{pFDN}}$ the *volume approximation*:

$$\mathcal{A}_{\text{volume}}(n) = n^{N-1} c_{N-1}, \tag{23}$$

with c_{N-1} given by (22).

E. Equilateral Approximation

To balance simplicity and accuracy, we introduce another approximation to $\mathcal{A}_{\text{pFDN}}$. If all delays are equal to the geometric mean $\overline{m} = \sqrt[N]{\prod_{i=1}^{N} m_i}$, then it can be shown that (17) simplifies to:

$$\begin{aligned} \mathcal{A}_{\text{equi}}(n) &= n^{N-1} c_{N-1} + \cdots + n c_1 + c_0 \\ c_k &= \frac{(-1)^{N-k}}{(N-1)!} \frac{\text{stirl}(N+1, k+1)}{\overline{m}^{k+1}} \end{aligned} \tag{24}$$

where $\text{stirl}(i, j)$ is the Stirling number of the first kind [31, p. 33]. As $\text{stirl}(N+1, N) = 1$, the leading coefficients of $\mathcal{A}_{\text{equi}}$ and $\mathcal{A}_{\text{pFDN}}$ are equal. This approximation corresponds in Ehrhart theory to an equilateral echo lattice such that the equilateral echo boxes have equal volume to the original echo boxes. The quality of the approximation $\mathcal{A}_{\text{equi}}$ is governed by the model violation, i.e., the more similar the delays are to each other the better the approximation will be. Therefore, the geometric standard deviation of the delays $\boldsymbol{m}$, i.e.,

$$\sigma_{\boldsymbol{m}} = \exp \sqrt{\frac{\sum_{i=1}^{N} (\ln \frac{m_i}{\overline{m}})^2}{N}} \tag{25}$$

quantifies the deviation from the approximation assumption. To evaluate the quality of the equilateral approximation, a Monte Carlo simulation is performed with $N = 8$ and uniform random delays $\boldsymbol{m}$ between 50 and 5000 samples. The results as shown in Fig. 4 evaluate the approximation along the geometric standard deviation $\sigma_{\boldsymbol{m}}$ by comparing a threshold value of the absolute densities:

$$\left(t_{\text{equi}} - t_{\text{pFDN}}\right)/t_{\text{pFDN}} \tag{26}$$

is the relative threshold error, where

$$\begin{aligned} t_{\text{equi}} &= \arg\min_n \left[\mathcal{A}_{\text{equi}}(n) \geq 1\right] \\ t_{\text{pFDN}} &= \arg\min_n \left[\mathcal{A}_{\text{pFDN}}(n) \geq 1\right]. \end{aligned} \tag{27}$$

It can be observed that the equilateral approximation underestimates the absolute echo density with increasing geometric standard deviations. The mean error is a systematic underestimation and can be therefore potentially compensated. However, the increasing standard deviation of the error degrades the quality of estimation for higher standard deviations of the delays $\sigma_{\boldsymbol{m}}$.

Figure 5 shows the normalized coefficients, i.e., $\sqrt[k]{c_k}$, of the polynomials of $\mathcal{A}_{\text{pFDN}}$, $\mathcal{A}_{\text{equi}}$ and $\mathcal{A}_{\text{volume}}$. By construction, all coefficients except the leading coefficient of $\mathcal{A}_{\text{volume}}$ are zero, resulting in a large approximation error, which is particularly large for higher polynomial degrees and larger number of delays, respectively.

F. Condition on Echo Saturation

The impulse response of the FDN *saturates* at a point in time if there is an echo at every successive time unit. The so called *saturation time* t_{saturate} is defined as

$$t_{\text{saturate}} = \arg\min_n \left[\forall k \geq n \mid \mathcal{A}_{\text{FDN}}(k) \geq 1\right]. \tag{28}$$

In the context of Ehrhart theory, t_{saturate} is also known as the Frobenius number of $\boldsymbol{m}$ [31]. For the Frobenius number to be finite, it is sufficient and necessary that the delays $\boldsymbol{m}$ are relatively prime positive integers [30]. The computation of the Frobenius number is a close relative to the determination of the Ehrhart polynomial. Figure 3 shows the saturation point $t_{\text{saturate}} = 73$ for $\boldsymbol{m} = [7, 13]$.

The saturation point t_{saturate} could be understood as the system's immanent notion of mixing time. However, due to its lacking relation to perceptual cues in the context of artificial reverberation, the echo saturation time is not investigated in detail besides its mere existence.

IV. Empirical Echo Density

The empirical echo density can be based on counting delay paths similar to (13) except that instead of counting multiple paths with equal n_p, every time instance n can only exhibit a single echo. Thus, the empirical echo density is

$$\mathcal{E}_{\text{FDN}}(n) = \begin{cases} 1 \text{ if } \mathcal{A}_{\text{FDN}}(n) \geq 1 \\ 0 \text{ otherwise .} \end{cases} \tag{29}$$

Figure 3 shows the different echo measures of an FDN with $N = 2$ and delays $\boldsymbol{m} = [7, 13]$. A smoothed version of the empirical echo density $\mathcal{E}_{\text{FDN}}$ is given by a convolution with a smoothing kernel w:

$$\mathcal{E}_{\text{wFDN}} = \mathcal{E}_{\text{FDN}} * w. \tag{30}$$

For the simulation experiments in this work, w is a Hamming window with 256 samples which corresponds to 5.3 ms at a sampling frequency of 48 kHz. In the following, we discuss how the smoothed empirical echo density can be derived from its absolute counter part.

A. Expected Empirical Echo Density

As the detection of overlapping echoes requires knowledge on the, difficult to obtain, non-polynomial part of $\mathcal{A}_{\text{FDN}}$, we present a statistical estimation of the empirical echo density. Given the average echo probability λ, the Poisson distribution describes the probability that k echoes occur at a certain time instance:

$$P_{k \text{ echoes}} = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{31}$$

Therefore, the probability that there is at least one echo at a given time instance is

$$P_{\text{echo}} = 1 - P_{0 \text{ echoes}} = 1 - e^{-\lambda}. \tag{32}$$

The absolute echo density $\mathcal{A}(n)$ is the expected number of echoes to occur per time unit. Accordingly, the *expected empirical echo density* is

$$\tilde{\mathcal{E}}(n) = 1 - e^{-\mathcal{A}(n)}. \tag{33}$$

For both approximations, $\mathcal{A}_{\text{equi}}(n)$ and $\mathcal{A}_{\text{volume}}(n)$, of the absolute echo density $\mathcal{A}_{\text{pFDN}}(n)$, there are no meaningful empirical echo densities connected, but only expected empirical echo densities, which can be derived directly from the absolute echo density by (33).

As the smoothed empirical echo density $\mathcal{E}_{\text{wFDN}}$ is derived from a deterministic process, the expected empirical echo density $\tilde{\mathcal{E}}_{\text{pFDN}}$ may differ considerably. This deviation is called *degeneration* of the delays. In the following, we give some examples of delay patterns, which can introduce strong degenerations.

B. Degeneration

The following three degenerated delay patterns: clustering, common prime factors, and low-order dependency are by no means a complete list, but have been found by the authors to be the most prominent.

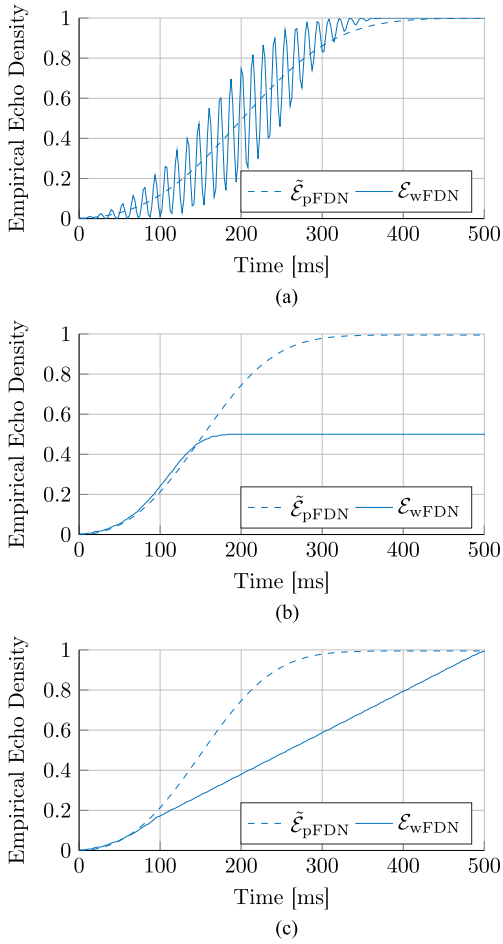

Fig. 6. Examples of degenerated delays $\boldsymbol{m}$. (a) Clustering: $m = [635; 657; 627; 1321]$. (b) Common Primes: $m = [308; 558; 908; 1022]$. (c) Low-order Dependency: $m = [307; 557; 907; 1021]$.

Clustering is the concentration of delay lengths around a certain value and its integer multiples. Because of the clustering, the echoes are emitted in concentrated batches and therefore the absolute echo density fluctuates strongly. Figure 6(a) shows an example of clustered delays $\boldsymbol{m} = [635, 657, 627, 1321]$. Although the smoothed empirical echo density follows the large trend of the expected empirical echo density, it fluctuates around the expected value. In the resulting impulse response, it is possible to perceive a fluctuation in energy over time. Even though such a fluctuation would be evaluated as a defect in concert hall acoustics, it may be appropriate when modeling an almost spherical room.[2] Heuristically, clustering can be largely avoided by choosing a sufficiently large geometric standard deviation of the delays, e.g. $\sigma_{\boldsymbol{m}} \geq 1.2$.

Common prime factors of the delays diminish the maximum empirical echo density to lower than 1 as the saturation point t_{saturate} is never reached. If all delays share a certain prime factor q, then all echo times n_p are integer multiples of q, such that on the contrary all time instances, which are not integer multiples of q are not emitting an echo. Figure 6(b) shows an example of

[2]In a spherical room, the echoes will reach the listener at regular intervals and therefore there is a fluctuation of echo density and energy.

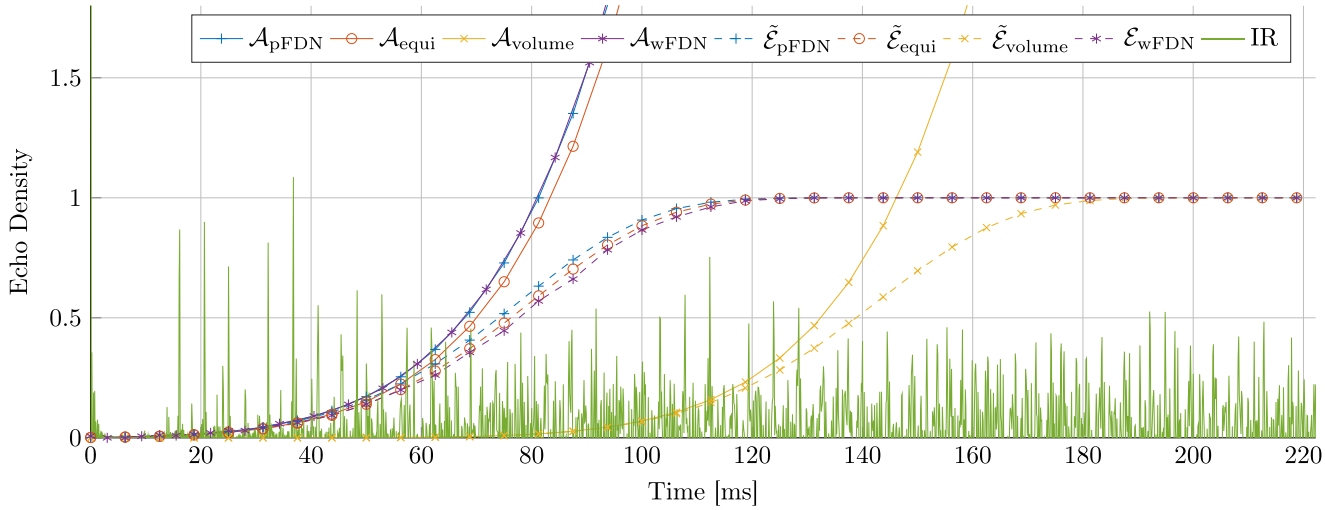

Fig. 7. Absolute echo density and the volume and equilateral approximations (solid lines) with $\boldsymbol{m} = [479, 587, 673, 773, 881, 991, 1091, 1201]$ and $N = 8$. The expected empirical echo densities are added along side the smoothed empirical echo density (dashed lines). The absolute value of the impulse response is added for reference.

delays with the common prime factor 2. The smoothed empirical echo density cannot increase over 0.5. Whereas the degeneration is most severe if all delays share a common prime, similar degeneration can also occur in the beginning if almost all delays share a common prime. Although it is sufficient to choose all delays as primes, it is important to notice that it is not necessary to do so.

Low-order dependency is the integer linear combination of delays which coincide with another integer linear combination of delays with small coefficients. For $\boldsymbol{m} = [49, 51, 100]$, every echo path involving $m_3 = 100$ can be as well represented by echo paths only involving $m_1 = 49$ and $m_2 = 51$. The resulting empirical echo density is equal to an FDN with only the delays $\boldsymbol{m} = [49, 51]$. We will call $\boldsymbol{m}$ having a *k-order dependency*, if there are $\boldsymbol{q_1}, \boldsymbol{q_2} \in \mathbb{N}^N$ with

$$\begin{aligned} \boldsymbol{q}_1^\top \boldsymbol{m} &= \boldsymbol{q}_2^\top \boldsymbol{m}, \\ \sum \boldsymbol{q_1} &= k, \end{aligned} \tag{34}$$

where $\boldsymbol{q_1}, \boldsymbol{q_2}$ are the number of occurrences of the delays in corresponding paths p_1, p_2 as employed in (13). In this sense for $N = 3$, $\boldsymbol{m} = [49, 51, 100]$ has a first-order dependency.

Figure 6 depicts the expected and smoothed empirical echo densities for $N = 4$ with delays $\boldsymbol{m} = [307, 557, 907, 1021]$. The low k-order dependencies of $\boldsymbol{m}$ are

1) $k = 3, \boldsymbol{q}_1 = [0, 2, 0, 1], \boldsymbol{q}_2 = [4, 0, 1, 0]$
2) $k = 4, \boldsymbol{q}_1 = [1, 2, 0, 1], \boldsymbol{q}_2 = [5, 0, 1, 0]$
3) $k = 4, \boldsymbol{q}_1 = [0, 3, 0, 1], \boldsymbol{q}_2 = [4, 1, 1, 0]$
4) $k = 4, \boldsymbol{q}_1 = [0, 2, 1, 1], \boldsymbol{q}_2 = [4, 0, 2, 0]$
5) $k = 4, \boldsymbol{q}_1 = [0, 2, 0, 2], \boldsymbol{q}_2 = [4, 0, 1, 1]$

It can be seen that for $N = 4$, dependencies of order up to 4 or 5 can degenerate the empirical echo density. For larger N, few low-order dependencies tend to be less severe as there are many more paths. The given example also shows that it can be rather difficult to detect dependencies among delays without actually computing all possibilities. Heuristically speaking, it is easier to form dependencies if delays span a wider range. On the contrary, first-order dependencies can be avoided entirely for $\frac{\max \boldsymbol{m}}{\min \boldsymbol{m}} < 2$. The authors suggest as a rule of thumb to strictly avoid dependencies of order up to 2, and for $N \leq 8$ to avoid even dependencies of order up to 4.

C. Example

Take for instance an FDN with $N = 8$ delays $\boldsymbol{m} = [479, 587, 673, 773, 881, 991, 1091, 1201]$ with low degeneration and an orthogonal and fully occupied feedback matrix $\boldsymbol{A}$. Figure 7 depicts the impulse response and the corresponding echo densities: polynomial part of the absolute density $\mathcal{A}_{\text{pFDN}}$, equilateral approximation $\mathcal{A}_{\text{equi}}$ and volume approximation $\mathcal{A}_{\text{volume}}$, the smoothed empirical echo density $\mathcal{E}_{\text{wFDN}}$ and the expected empirical echo densities $\tilde{\mathcal{E}}_{\text{pFDN}}$, $\tilde{\mathcal{E}}_{\text{equi}}$ and $\tilde{\mathcal{E}}_{\text{volume}}$ derived from the absolute echo densities with (33).

Both approximations, $\mathcal{A}_{\text{equi}}$ and $\mathcal{A}_{\text{volume}}$, underestimate the absolute echo density $\mathcal{A}_{\text{pFDN}}$, e.g. the deviation of the time instance for the approximation densities to reach 1 with respect to $\mathcal{A}_{\text{pFDN}}$, is smaller than 5 ms for the equilateral approximation $\mathcal{A}_{\text{equi}}$, but is over 60 ms for the volume approximation $\mathcal{A}_{\text{volume}}$. The corresponding expected echo densities deviate accordingly. The smoothed empirical echo density $\mathcal{E}_{\text{wFDN}}$ is close to both the $\tilde{\mathcal{E}}_{\text{pFDN}}$ and the equilateral expected empirical densities $\tilde{\mathcal{E}}_{\text{equi}}$, and is therefore well predicted by these measures. On the contrary, the volume approximation is a poor predictor of the smoothed empirical echo density.

V. Mixing Time

In this section, we evaluate the quality of the mixing time estimation based on the absolute and expected empirical echo density given by (17) and (33), respectively. As the mixing time of Polack (3) is technically similar to the approach of Defrance (4), only the latter measure is discussed.

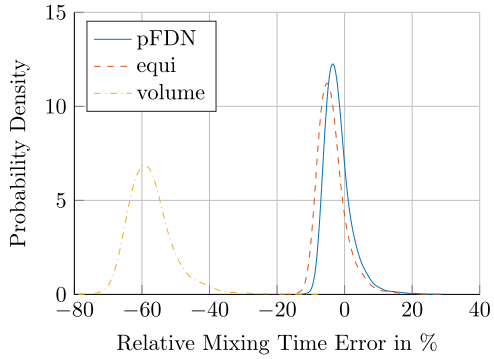

Fig. 8. Probability density function of the relative mixing time error $\epsilon_{\text{Defrance}}$ based on a Monte Carlo simulation.

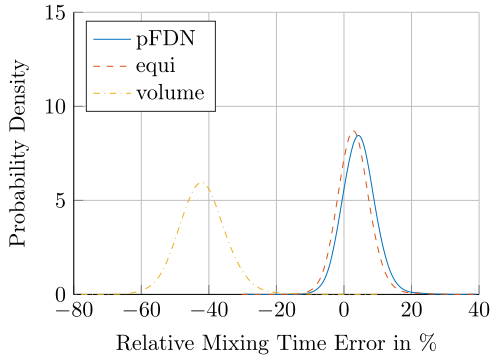

Fig. 9. Probability density function of the relative mixing time error ϵ_{Abel} based on a Monte Carlo simulation.

To evaluate the quality of estimation, we perform a Monte Carlo simulation by choosing $N = 8$ uniformly random delays $\boldsymbol{m}$ with lengths between 1000 and 3000 samples. According to heuristic rules, delays with

1) more than half of the delays sharing the same common prime,
2) first-order and second-order dependencies and
3) geometric standard deviation $\sigma_{\boldsymbol{m}}$ larger than 2 and smaller than 1.2

have been excluded.

The orthogonal feedback matrix $\boldsymbol{A}$ is chosen randomly from the set orthogonal matrices with a uniform distribution [34]: applying the QR factorization to a matrix $\boldsymbol{X} = \boldsymbol{QR}$ whose elements are independently normally distributed with zero mean and common variance and choosing the diagonal elements of the matrix $\boldsymbol{R}$ to be positive yields the orthogonal matrix $\boldsymbol{Q}$ with the desired distribution. The input, output and direct gains, $\boldsymbol{b}$, $\boldsymbol{c}$ and d, respectively are all ones. The measured mixing times t_{Defrance} and t_{Abel} are derived from the impulse response of the FDN using (8) and the definitions (4) and (7), respectively. The predicted mixing time $\tilde{t}_{\text{Defrance}}$ is based on the expected empirical echo density using (4) and the predicted mixing time $\tilde{t}_{\text{Abel}}$ is based on the absolute echo density using (6). As the echo density profile proposed by Abel has typically strong fluctuations, the mixing time prediction is based on the arithmetic mean of 20 echo density profiles derived from FDN impulse responses with identical delays, but random feedback matrices.

The relative prediction error is given by

$$\epsilon = \frac{\tilde{t} - t}{t}. \tag{35}$$

The mixing time threshold for both measures is $T = 0.9$.

A. Prediction of Mixing Time

Figure 8 shows the probability distribution of the relative mixing time error $\epsilon_{\text{Defrance}}$. The mixing time prediction derived via the empirical density from the absolute density $\mathcal{A}_{\text{pFDN}}$ and the equilateral approximation $\mathcal{A}_{\text{equi}}$ lie within a $\pm 10\%$ error, with a mean error of -1.9% and -3.7%, respectively. The $\mathcal{A}_{\text{pFDN}}$ underestimates the echo density, which is due to small degenerations. The equilateral approximation introduces further underestimation because of the geometric standard deviation. The mean error of the volume approximation $\mathcal{A}_{\text{volume}}$ is -57.3%, and for this it largely underestimates the mixing time.

Figure 9 shows the probability distribution of the relative mixing time error ϵ_{Abel}. Both $\mathcal{A}_{\text{pFDN}}$ and $\mathcal{A}_{\text{equi}}$ overestimates the mixing time with 4.4% and 2.9%, respectively. Similarly to the Defrance mixing time, the volume approximation highly underestimates the Abel mixing time with a mean of -41.8%. The overestimation of the Abel mixing time is due to the deterministic gain distribution of the FDN impulse response. The estimation made by Huang and Abel [25] in (6) assumes normally distributed echo gains. The Gaussian distribution of the echo levels in the FDN is a consequence of multiple echoes overlapping with identical echo time, such that more echoes are necessary than in the theoretical assumption.

B. Specify Mixing Time with Equilateral Approximation

Both $\mathcal{A}_{\text{pFDN}}$ and $\mathcal{A}_{\text{equi}}$ show potential to estimate the mixing time reasonably well for the usage in artificial reverberators. Due to the simplicity of the equilateral approximation, it is also possible to specify a desired mixing time and determine the mean delay length, which can serve in turn as a basis to determine the actual delays.

For a given Abel mixing time t_{Abel} and mixing threshold $\mathcal{T}$, from the specification (6) and (7)

$$\frac{\mathcal{A}_{\text{equi}}(t_{\text{Abel}})}{\mathcal{A}_{\text{equi}}(t_{\text{Abel}}) + 1} = \mathcal{T}, \tag{36}$$

the mean delay length $\overline{m}$ can be determined as the real root of the polynomial equation

$$(1 - \mathcal{T})\mathcal{A}_{\text{equi},t_{\text{Abel}}}(\overline{m}) - \mathcal{T} = 0. \tag{37}$$

Similarly, for a given Defrance mixing time t_{Defrance} and mixing threshold $\mathcal{T}$, from the specification (4)

$$\tilde{\mathcal{E}}_{\text{equi}}(t_{\text{Defrance}}) = 1 - e^{-\mathcal{A}_{\text{equi}}(t_{\text{Defrance}})} = \mathcal{T}, \tag{38}$$

the mean delay length $\overline{m}$ can be determined as the real root of the polynomial function

$$\mathcal{A}_{\text{equi},t_{\text{Defrance}}}(\overline{m}) + \ln(1 - \mathcal{T}) = 0. \tag{39}$$

Please note that in both equations (37) and (39), the mean delay lengths $\overline{m}$ is the polynomial variable and the mixing time t

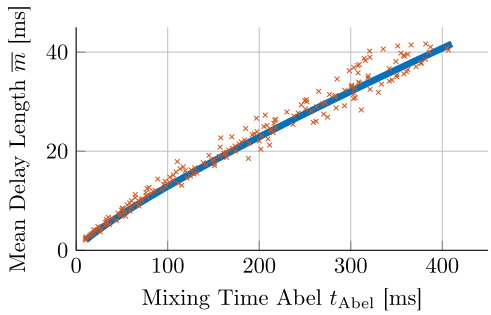

Fig. 10. Specifying Abel mixing time t_{Abel} and resulting mean delay length $\overline{m}$ (solid line) according to ((37)) for $N = 8$ and simulation results with Hadamard feedback matrix ($\times$) and delays according to (40).

is a fixed coefficient parameter of the polynomial $\mathcal{A}_{\text{equi,t}}(\overline{m})$. The non-degenerated delays $\boldsymbol{m}$ can be derived from the mean value $\overline{m}$ by following the guidelines, which have been discussed throughout the paper:

1) Choose pairwise relative prime delays $\boldsymbol{m}$.
2) Choose moderate geometric standard deviation: $1.2 \leq \sigma_{\boldsymbol{m}} \leq 3$.
3) Choose delays without low-order dependencies.

Following, these guidelines, as shown in Fig. 8 and Fig. 9, mixing time specification within $\pm 10\%$ relative error can be achieved.

For example, let us assume the FDN dimension of $N = 8$ and the mixing threshold $\mathcal{T} = 0.9$. The mean delay $\overline{m}$ is derived from the Abel mixing time t_{Abel} with ((37)), and this relation is depicted in Figure 10 by the solid line. There are many ways to derive the actual delays $\boldsymbol{m}$ from the mean delay $\overline{m}$, here we scale a normalized array:

$$\boldsymbol{m} = \left\lfloor \overline{m}\, e^{\frac{\widetilde{\boldsymbol{m}}_N}{\log 3}} \right\rfloor, \tag{40}$$

where $\lfloor \cdot \rfloor$ is the floor operation and the delay scalers are

$$\begin{aligned} \boldsymbol{m}_N &= \log([1, 2, \ldots, N]) \text{ and} \\ \widetilde{\boldsymbol{m}}_N &= \frac{\boldsymbol{m}_N - \overline{\boldsymbol{m}}_N}{\sigma_{\boldsymbol{m}_N}} \end{aligned} \tag{41}$$

are the normalized delay scalers with arithmetic mean and standard deviation. The actual delays $\boldsymbol{m}$ defined in (40) have the specified mean delay $\overline{m}$ and a geometric standard deviation $\sigma_{\boldsymbol{m}} \approx 2.8$. For instance, the actual delays derived from $\overline{m} = 20$ ms are $\boldsymbol{m} = [2.52, 7.46, 14.02, 21.98, 31.16, 41.42, 52.70, 64.91]$ ms. The choice of $\boldsymbol{m}$ in (40) is simple and practical, but by no means the only viable choice as long the described condition are satisfied.

Figure 10 shows the measured Abel mixing time t_{Abel}, computed from (5) and (7), for delays $\boldsymbol{m}$ as defined in (40) along different mean delay lengths $\overline{m}$ and an orthogonal Hadamard feedback matrix. The close match between the specified and measured Abel mixing time confirms the accuracy of the presented approach.

VI. Conclusion

An analysis of the echo density and mixing time of impulse responses as they are typically produced by FDNs was presented. It was shown that the echo density over time is a polynomial function, whereby the polynomial coefficients depend on the delays. As the computation of the exact polynomial has exponential computational complexity, two more practical approximations, equilateral and volume, were proposed. The expected empirical echo density was derived from the absolute echo density to predict the echo density of an impulse response. The echo density was applied to predict the mixing time of an FDN according to the definitions of Abel and Defrance. It was shown that exact and equilateral echo density can predict both mixing times within an $\pm 10\%$ relative error. Conversely, the equilateral echo density can be employed to compute a mean delay for a desired mixing time. Additionally, heuristics for choosing delays with pseudo-random echo distribution were discussed.

The influence of the mixing matrix were minimized by employing random matrices throughout the Monte Carlo experiments. Further, the maximally dense impulse response requires that all echo paths are non-zero. In the literature on FDNs, various matrices were proposed which do not comply with these requirements and therefore the presented theory has to be adjusted to their matrix structure.

Acknowledgment

The authors would like to thank A. Silzle and B. Neugebauer who have contributed to this work through their acoustical expertise. The authors also would like to thank M. Beck for pointing out the Bernoulli–Barnes polynomials and for writing an accessible and insightful book on Ehrhart theory. Further, the authors would like to thank the anonymous reviewers whose comments helped to considerably raise the quality of this work.

References

[1] H. Kuttruff, *Room Acoustics, 5th ed*. Boca Raton, FL, USA: CRC Press, Jun. 2009.
[2] P. Huang, J. S. Abel, H. Terasawa, and J. Berger, "Reverberation echo density psychoacoustics," in *Proc. Audio Eng. Soc. Conv.*, San Francisco, CA, USA, Oct. 2008, pp. 1–10.
[3] J. S. Abel and P. Huang, "A simple, robust measure of reverberation echo density," in *Proc. Audio Eng. Soc. Conv.*, San Francisco, CA, USA, Oct. 2006, pp. 1–10.
[4] J. M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. 94th Audio Eng. Soc. Conv.*, Paris, France, Feb. 1991, pp.1–16.
[5] J. H. Rindel, "The use of computer modeling in room acoustics," *J. Vibroengineering*, vol. 3, no. 4, pp. 41–72, 2000.
[6] G. Defrance, L. Daudet, and J.-D. Polack, "Detecting arrivals within room impulse responses using matching pursuit," in *Proc. Int. Conf. Digital Audio Effects*, Espoo, Finland, 2008, pp. 1–4.
[7] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943–950, 1979.
[8] C. Borß, "An improved parametric model for the design of virtual acoustics and its applications," Ph.D. dissertation, Institut für Kommunikationsakustik, Ruhr-Universität Bochum, Bochum, 2011.
[9] J.-D. Polack, "Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 235–244, Jan. 1993.
[10] M. R. Schroeder, "Improved quasi-stereophony and 'Colorless' artificial reverberation," *J. Acoust. Soc. Amer.*, vol. 33, no. 8, pp. 1061–1064, 1961.

[11] M. R. Schroeder, "Natural sounding artificial reverberation," *J. Audio Eng. Soc.*, vol. 10, no. 3, pp. 219–223, 1962.
[12] J. Stautner and M. Puckette, "Designing multi-channel reverberators," *Comput. Music J.*, vol. 6, no. 1, pp. 52–65, 1982.
[13] V. Välimäki, J. D. Parker, L. Savioja, J. O. Smith, III, and J. S. Abel, "Fifty years of artificial reverberation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 5, pp. 1421–1448, Jul. 2012.
[14] D. Rocchesso and J. O. Smith, III,"Circulant and elliptic feedback delay networks for artificial reverberation," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 1, pp. 51–63, Jan. 1997.
[15] M. Kahrs and K. Brandenburg, *Applications of Digital Signal Processing to Audio and Acoustics*. Berlin, Germany: Springer Science & Business Media, Mar. 1998.
[16] U. Zölzer *et al.*, *DAFX—Digital Audio Effects*, U. Zölzer, Ed. Chichester, UK: Wiley, Apr. 2002.
[17] F. Menzer and C. Faller, "Binaural reverberation using a modified Jot reverberator with frequency-dependent interaural coherence matching," in *Proc. Audio Eng. Soc. Conv.*, May 2009, pp. 1–6.
[18] J. M. Jot, "Efficient models for reverberation and distance rendering in computer music and virtual audio reality," in *Proc. Int. Comput. Music Conf.*, Thessaloniki, Greece, 1997, pp. 1–8.
[19] 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.
[20] F. Menzer and C. Faller, "Unitary matrix design for diffuse jot reverberators," in *Proc. Audio Eng. Soc. Conv.*, London, UK, May 2010, pp. 1–16.
[21] J. Blauert, *Spatial Hearing* (The Psychophysics of Human Sound Localization). Cambridge, MA, USA: MIT Press, 1997.
[22] P. Götz, K. Kowalczyk, A. Silzle, and E. A. P. Habets, "Mixing time prediction using spherical microphone arrays," *J. Acoust. Soc. Amer.*, vol. 137, no. 2, pp. L206–EL212, Feb. 2015.
[23] J.-D. Polack, "La transmission de l'energie sonore dans les salles," Ph.D. dissertation, Université du Maine, Sound energy transmission in rooms, 1988.
[24] G. Defrance, L. Daudet, and J.-D. Polack, "Using matching pursuit for estimating mixing time within room impulse responses," *Acta Acust. United Acust.*, vol. 95, no. 6, pp. 1071–1081, Nov. 2009.
[25] P. Huang and J. S. Abel, "Aspects of reverberation echo density," in *Proc. Audio Eng. Soc. Conv.*, New York, NY, USA, Oct. 2007, pp. 1–7.
[26] S. J. Schlecht and E. A. P. Habets, "Connections between parallel and serial combinations of comb filters and feedback delay networks," in *Proc. Int. Workshop Acoust. Signal Enhancement*, Aachen, Germany, 2012, pp. 1–4.
[27] M. R. Schroeder and B. F. Logan, "'Colorless' artificial reverberation," *Audio, IRE Trans*, vol. AU-9, no. 6, pp. 209–214, 1961.
[28] E. Ehrhart, "Sur un problème de géométrie diophantienne linéaire. I.," *J Für Die Reine Und Angew. Math. (Crelles Journal)*, no. 226, pp. 1–29, 1967.
[29] E. Ehrhart, "Sur un problème de géométrie diophantienne linéaire. II." *J. Für Die Reine Und Angew. Math. (Crelles Journal)*, no. 227, 1967.
[30] M. Beck, "The frobenius problem, rational polytopes, and fourier–dedekind sums,," *J. Number Theory*, vol. 96, no. 1, pp. 1–21, Sep. 2002.
[31] M. Beckнd and S. Robins, *Computing the Continuous Discretely* (Undergraduate Texts in Mathematics). New York, NY, USA: Springer, 2015.
[32] A. I. Barvinok, "Computing the Ehrhart polynomial of a convex lattice polytope," *Discrete Comput. Geom.*, vol. 12, no. 1, pp. 35–48, Sep. 2005.
[33] L. Jiu, V. H. Moll, and C. Vignat, "A symbolic approach to some identities for Bernoulli–Barnes polynomials," *Int. J. Number Theory*, pp. 1–14, Aug. 2015.
[34] G. W. Stewart, "The efficient generation of random orthogonal matrices with an application to condition estimators," *SIAM J. Numer. Anal.*, vol. 17, no. 3, pp. 403–409, 1980.

Sebastian J. Schlecht 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. He is currently working toward the Doctoral degree at the International Audio Laboratories Erlangen, Erlangen, Germany, on the topic of spatial artificial reverberation and reverberation enhancement systems.

His research interests are in the area of acoustic modeling and auditory perception of acoustics, spatial audio signal processing, analysis and synthesis of feedback systems, computer music, and augmented performance techniques.

Emanuël A. P. Habets (S'02–M'07–SM'11) received the B.Sc. degree in electrical engineering from the Hogeschool Limburg, The Netherlands, in 1999, and the M.Sc. and Ph.D. degrees in electrical engineering from the Technische Universiteit Eindhoven, Eindhoven, The Netherlands, in 2002 and 2007, respectively.

From 2007 until 2009, he was a Postdoctoral Fellow at the Technion–Israel Institute of Technology and at the Bar-Ilan University, Israel. From 2009 until 2010, he was a Research Fellow in the Communication and Signal Processing Group, Imperial College London, London, U.K. He is currently an Associate Professor at the International Audio Laboratories Erlangen [A joint institution of the Friedrich-Alexander-University Erlangen-Nürnberg and Fraunhofer Institut für Integrierte Schaltungen (IIS)], Erlangen, Germany, and the Head of the Spatial Audio Research Group, Fraunhofer IIS, Germany. His research activities center around audio and acoustic signal processing, and include spatial audio signal processing, spatial sound recording and reproduction, speech enhancement (dereverberation, noise reduction, echo reduction), and sound localization and tracking.

Dr. Habets was a member of the organization committee of the 2005 International Workshop on Acoustic Echo and Noise Control in Eindhoven, The Netherlands, and the General Co-Chair of the 2013 International Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) in New Paltz, New York, and the 2014 International Conference on Spatial Audio in Erlangen, Germany. He was a member of the IEEE Signal Processing Society Standing Committee on Industry Digital Signal Processing Technology (2013–2015), and a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING and the *EURASIP Journal on Advances in Signal Processing*. He received, as a co-author with S. Gannot and I. Cohen, the 2014 IEEE Signal Processing Letters Best Paper Award. He is a member of the IEEE Signal Processing Society Technical Committee on Audio and Acoustic Signal Processing (2011–2016), Vice-Chair of the EURASIP Special Area Team on Acoustic, Sound and Music Signal Processing, an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS, and the Editor-in-Chief of the *EURASIP Journal on Audio, Speech, and Music Processing*.