Valimaki, Vesa; Prawda, Karolina

Late-Reverberation Synthesis using Interleaved Velvet-Noise Sequences

Late-Reverberation Synthesis Using Interleaved Velvet-Noise Sequences

Vesa Välimäki ⓘ, *Fellow, IEEE*, and Karolina Prawda

Abstract—This paper proposes a novel algorithm for simulating the late part of room reverberation. A well-known fact is that a room impulse response sounds similar to exponentially decaying filtered noise some time after the beginning. The algorithm proposed here employs several velvet-noise sequences in parallel and combines them so that their non-zero samples never occur at the same time. Each velvet-noise sequence is driven by the same input signal but is filtered with its own feedback filter which has the same delay-line length as the velvet-noise sequence. The resulting response is sparse and consists of filtered noise that decays approximately exponentially with a given frequency-dependent reverberation time profile. We show via a formal listening test that four interleaved branches are sufficient to produce a smooth high-quality response. The outputs of the branches connected in different combinations produce decorrelated output signals for multichannel reproduction. The proposed method is compared with a state-of-the-art delay-based reverberation method and its advantages are pointed out. The computational load of the method is 60% smaller than that of a comparable existing method, the feedback delay network. The proposed method is well suited to the synthesis of diffuse late reverberation in audio and music production.

Index Terms—Acoustics, audio systems, digital signal processing, filtering algorithms.

I. INTRODUCTION

ARTIFICIAL reverberation algorithms often produce an impulse response that is reminiscent of exponentially decaying filtered noise [1]. Moorer first suggested that the tail of a room impulse response could be described well using pseudo-random noise [2]. This paper proposes a novel noise-based reverberation algorithm, which is easy to design and efficiently filters the input signal with its response.

Room impulse responses are generally known to sound similar to exponentially decaying noise after a short time from the onset. In practice, the noise must be filtered and not white noise, and different frequency bands must decay at a different rate [1], [2]. This observation led to the idea that an artificial reverberation algorithm could be based on a random-noise generator. In fact, some reverberation algorithms, such as the well-known feedback

delay network (FDN) [3]–[5], can be converted into a pseudo-random noise generator by turning off the decay of sound. In practice, replacing all filters and attenuating coefficients in the algorithm with unity gains (i.e., no signal-processing operation) leaves only delay line and summation operations. A reverberation algorithm must also be able to efficiently convolve an arbitrary input signal—the signal to be processed—with its noisy response.

Rubak and Johansen proposed to use a finite-impulse response (FIR) filter with random coefficients as a loop filter in the feedback loop of a reverberation algorithm [6], [7]. This appears to be a computationally efficient method to generate a decaying response with the help of random noise, but, unfortunately, the decay rate of the system is also affected by the randomness, because the loop filter has a random magnitude response. Karjalainen and Järveläinen further elaborated this idea by cascading a random FIR filter with the feedback loop structure, allowing accurate control of the reverberation time (RT) using a low-order loop filter [8]. Additionally, Karjalainen and Järveläinen introduced the concept of velvet noise, a smooth-sounding sparse random noise [8]. More recently, Lee *et al.* [9] and Oksanen *et al.* [10] have investigated variants of recursive structures employing velvet noise. In addition to artificial reverberation, velvet noise has been applied recently to audio decorrelation [11], [12], time-expansion of sounds [13], [14], music synthesis [15], speech synthesis [16]–[18], and acoustic measurements [19].

A recurring problem in previous velvet-noise-based recursive algorithms is that a single sequence is filtered and attenuated over time [8]–[10]. Human hearing is sensitive to repetitions, and the produced repetitive noise sounds similar to flutter echo, a well-known problem in room acoustics. The fluttering is easiest to perceive in percussive sounds. Earlier solutions to reduce the flutter problem include time-varying randomization of the impulses, and cross-fading sequences drawn from a small collection of velvet-noise sequences [8]–[10]. However, these time-variant techniques lead to another problem: warbling, which is apparent when a stationary sound is processed using the reverberation algorithm. Suppressing both the flutter and the warbling simultaneously seems hard in a recursive velvet-noise reverberation algorithm.

This paper proposes to hide the repetition in the noisy response by constructing a velvet-noise sequence from several sequences that are combined. The positive and negative impulses must not be allowed to accumulate or cancel each other in the combination sequence so as not to destroy the advantageous properties of velvet noise. An extended velvet-noise sequence [20] can leave gaps

in the sequence at regular intervals so that it can be added to other similar sequences maintaining the prescribed pulse density. This work shows that when several repetitive sequences of different length are combined, the repetitions become inaudible. This principle is then used to devise a novel recursive reverberation algorithm. A graphic equalizer is used to control the RT of each sequence.

The rest of this paper is organized as follows. Sec. II introduces the idea of interleaving velvet-noise sequences to hide repetitive patterns and describes a listening test to verify its perceived quality. Sec. III proposes a novel reverberation algorithm that uses interleaved velvet-noise sequences as sparse FIR filters. A smearing technique to soften the onset and a segmented decay technique for smooth approximation of the exponential decay are also proposed as additions to the basic algorithm. Sec. IV presents a validation and comparison to previous methods. Sec. V concludes the paper.

II. INTERLEAVED VELVET-NOISE SEQUENCES

This section introduces an interleaving technique for velvet-noise sequences, which provides a lossless prototype for the new reverberation technique. Additionally, this section describes a listening test to evaluate the perceptual quality of the produced white noise.

A. Interleaving Extended Velvet-Noise Sequences

Velvet noise consists of 1's, -1's, and zeros. The locations $k_{\text{vn}}(m)$ of the non-zero samples in the sequence are determined as

$$k_{\text{vn}}(m) = \text{round}[mT_{\text{d}} + r_1(m)(T_{\text{d}} - 1)], \qquad (1)$$

where $m = 0, 1, 2, \ldots$ is the pulse counter, T_{d} is the grid size, and $r_1(m)$ is a value produced with a random-number generator having uniform distribution $(0,1)$ [20]. Another random number sequence $r_2(m)$ is used to select the sign of the impulse, so that the sample inserted at index $k_{\text{vn}}(m)$ is either 1 or -1. The remainder of the samples in the velvet-noise sequence are zero. This method places exactly one non-zero sample in every range of T_{d} samples. When the number of non-zero impulses is sufficiently large, or, equivalently, T_{d} is small enough, velvet noise sounds similar to white noise. In fact, experiments have established that it sounds even slightly less rough than Gaussian white noise when there are at least 2000 impulses per second, when the sample rate is 44.1 kHz [8], [20], which is also used in this work.

The interleaving technique is based on the use of extended velvet-noise sequences (EVN) introduced in [20]. In an EVN, the range where the impulse can appear is further limited, leaving specified times between the impulses always empty. The impulse locations in an EVN are determined as follows:

$$k_{\text{evn}}(m) = \text{round}[mT_{\text{d}} + \Delta r_1(m)(T_{\text{d}} - 1)], \qquad (2)$$

where $0 < \Delta < 1$ is a scale factor limiting the range where the impulse can be located [20]. For example, when $\Delta = 0.25$, the impulses can only appear in the first quarter of the grid, so that the remaining 75% of the samples are always known to be zeros.
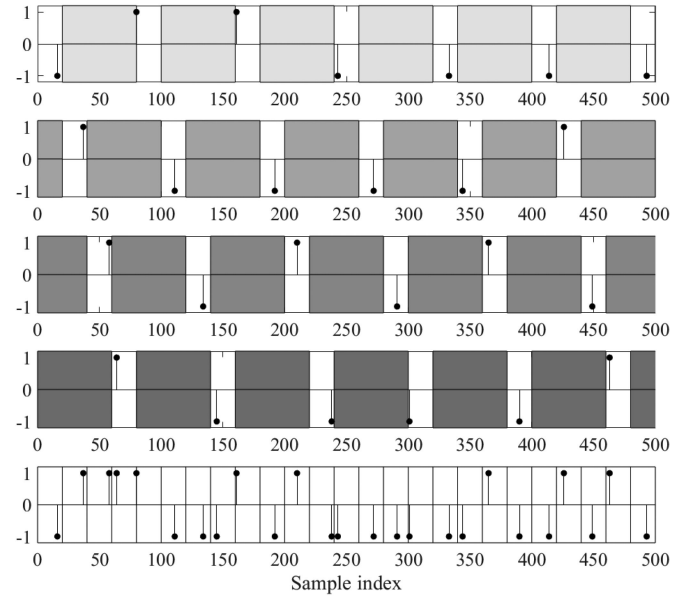

Fig. 1. Example of delaying and interleaving four extended velvet-noise sequences (the four upmost subfigures) so that impulses never collide in their sum (bottom). The shaded areas indicate the ranges in which impulses do not appear in each sequence. Only the non-zero samples are plotted.

The algorithm proposed here employs several EVNs in parallel and combines them so that their non-zero samples never occur at the same time. This is possible by delaying each sequence by a different number of sampling intervals so that the non-zero samples in each sparse sequence occur at different times. Fig. 1 shows an example of four interleaved EVNs. The grid size of each of them is four times the targeted overall number, $\hat{T}_{\text{d}} = 4T_{\text{d}}$, and $\Delta = 1/4$. Here, we have chosen $T_{\text{d}} = 20$, so each EVN has a grid size of $\hat{T}_{\text{d}} = 80$ samples, but an impulse can only appear within the first 20 samples of the grid range in each EVN. Furthermore, in Fig. 1, the second, third, and fourth sequences are delayed by $\hat{T}_{\text{d}}/4$, $2\hat{T}_{\text{d}}/4$, and $3\hat{T}_{\text{d}}/4$ samples, respectively.

The final sequence shown at the bottom of Fig. 1 is the resulting interleaved sequence in which samples originating from the four EVNs appear alternately but never collide. The grid size of the final sequence is $T_{\text{d}} = 20$.

B. Smooth Noise Generation Using Interleaved Sequences

To produce a long interleaved noise sequence, each EVN can have a finite length and they can be repeated indefinitely. Such a repetitive noise sequence is called frozen noise [21]–[23]. The EVN lengths must be selected to be different from each other and to be coprime with each other, so that there is no simple relation between any two EVN lengths, such as one is twice longer than the other. This problem is equivalent to that of choosing the delay-line lengths of comb filters in traditional delay-based reverberation algorithms, such as in an FDN [3], to be mutually incommensurate. Otherwise, a repetitive disturbance similar to flutter echo appears. A safe option is to use prime-number based sequence lengths of the form

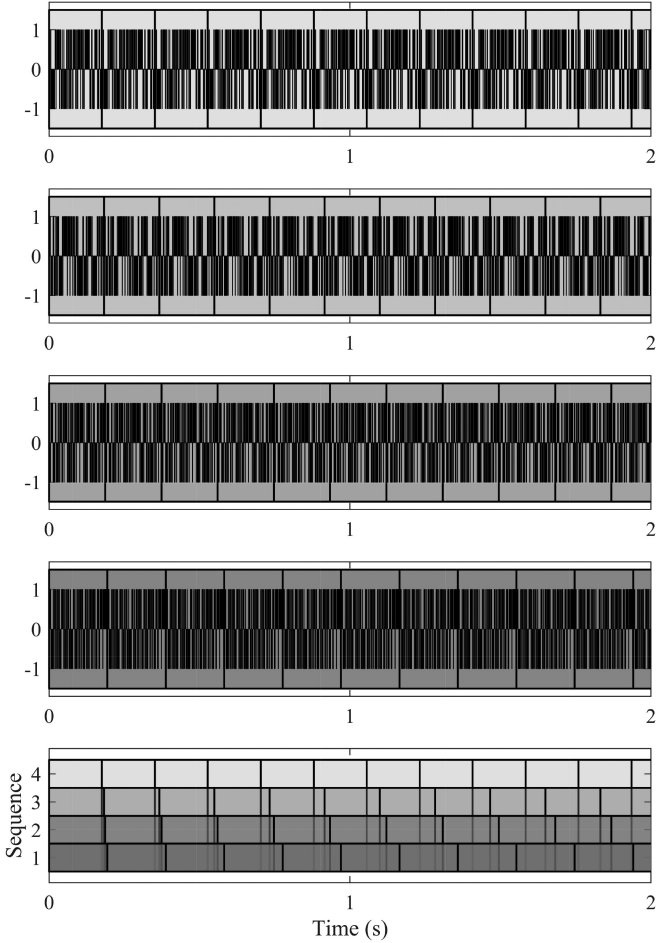$$L_i = C_i \hat{T}_{\text{d}}, \qquad (3)$$

Fig. 2.   Example of interleaving four finite-length EVN signals. In the top four panes, each EVN has a different prime-number based duration and restarts at the vertical lines. In the bottom pane, the combined boxes show that, except for the beginning, the EVN start times do not overlap.

where $i = 1, 2, \ldots, M$ is the sequence index, M is the number of sequences combined, C_i are different prime numbers, and $\hat{T}_d = M T_d$.

The density of the interleaved velvet-noise sequence must be sufficient so as not to sound rough. We know from previous studies that a density of about 2000 samples/s is sufficient [8], [20], and therefore, in this work we have chosen the grid size to be $T_d = 20$, which gives a density of 2205 samples/s at the sample rate of 44.1 kHz. Each EVN in an interleaved velvet-noise sequence will then have a grid size of $\hat{T}_d = 20 M$ samples and $\Delta = 1/M$.

The example in Fig. 2 shows four repeating EVN signals interleaved to produce a velvet-noise sequence that does not repeat for a long time. Here, the chosen primes are $C_1 = 97$, $C_2 = 101$, $C_3 = 103$, and $C_4 = 107$, so when they are multiplied by $\hat{T}_d = 80$, the sequence lengths of 176 ms, 183 ms, 187 ms, and 194 ms, respectively, are obtained. The combined velvet-noise sequence repeats after about 8.6 billion samples, or 54.4 hours, which is obtained by multiplying the four sequence lengths.

We extensively tested different numbers of interleaved sequences M and various delay-line lengths L_i. The repetitions seem hard to mask using less than four interleaved sequences. Furthermore, if any of the delay lines is very short, such as shorter than about 5000 samples or 110 ms, the repetition will easily become audible.

C. Listening Test

A Multiple Stimuli with Hidden Reference and Anchor (MUSHRA) listening test [24] was conducted to verify the perceptual qualities of interleaved repetitive velvet-noise samples. Before the start of the experiment, sufficiently many interleaved velvet-noise sequences were assumed to produce a smooth, or non-repetitive, sound and would receive a high average score.

The data set used in the listening test was created as sequential velvet-noise sounds. The stimuli consisted of one to five interleaved sequences, whose lengths were determined by the use of consecutive primes from 83 to 139 multiplied by the number of parallel lines and the grid size of the velvet-noise sequence. The length of each type of sequence was that of the shortest delay line determined by a prime number in the range 83–113. However, the sequences created with the numbers 101, 103, 107 and 109, primes that are very close to one another, were excluded from the experiment to avoid the audible repetition in the sound produced with these numbers.

The task in the listening test was to assess the smoothness of the sound stimuli compared to the reference. The possible grades ranged from 0 to 100, with 0 given to the most repetitive sound and 100 to a perfectly smooth stimulus. Additionally, text descriptions for five grade ranges were used: *Very annoying* (0–20), *Annoying* (20–40), *Slightly annoying* (40–60), *Perceptible, but not annoying* (60–80), and *Imperceptible* (80–100).

A regular, infinitely long, velvet-noise sequence having a density of 2205 samples/s was used as a reference sound that was meant to receive the maximum score. The test items were constructed using interleaved velvet-noise sequences with one delay line as the low-quality anchor and with two to five delay lines as MUSHRA conditions. These six sounds were assessed in every question. Each page in the MUSHRA test consisted of samples having the same prime number that determined the length of the shortest delay line, which, together with the anchor and the reference, were presented in random order. All sounds used in the experiment were four seconds long. The test was carried out using the web audio API-based experiment software webMUSHRA developed by International Audio Laboratories Erlangen [25].

The experiment was conducted in sound-proof listening booths at the Aalto Acoustics Lab using Sennheiser HD-650 reference headphones. In all, 26 people participated in the test. Five of the results, however, were excluded from the analysis for giving a score under 100 to the reference item more than three times. None of the 21 participants whose results were analyzed reported a hearing impairment. Their average age was 30.7 years (the standard deviation was 6.1). All the participants were either students or employees of the Aalto University Department of Signal Processing and Acoustics, and the majority of them had prior experience with MUSHRA tests.
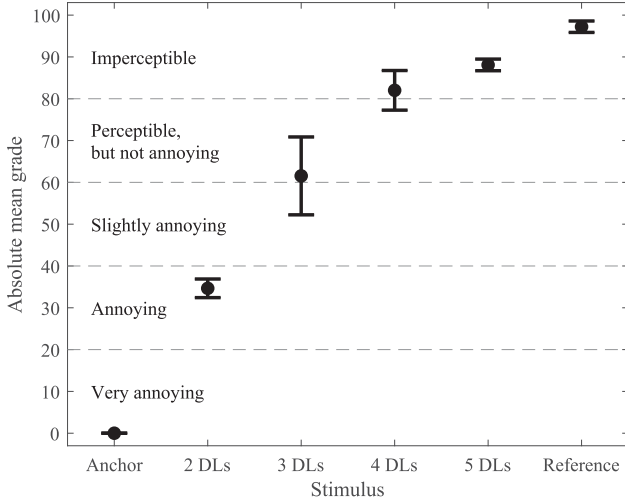
Fig. 3. Results of the MUSHRA listening test with 21 subjects. The absolute mean grades and the 95% confidence intervals are shown for the reference noise, the anchor, and the interleaved EVN signals produced with different numbers of delay lines (DLs), which is equal to the number of interleaved sequences. The horizontal dashed lines divide the grading ranges labeled with text descriptions.

The subjects were allowed to adjust the volume of the sound before starting the test. They were also familiarized with the task, its structure, and some of the test sounds in a short training session. The scores of the training session were excluded from the results. During the test, the subjects were presented with seven questions, all of which were doubled, resulting in 14 MUSHRA test pages. After completing the listening part of the task, the subjects were asked to answer questions about the strategy they used to distinguish between test items and the type of differences they heard.

The scores granted by the listeners to each sound sample were averaged based on the number of delay lines in the sequence. Fig. 3 shows the results with the 95% confidence intervals that reveal that in general the participants had no difficulties in recognizing the reference and the anchor, which received the score 0 (*Very annoying*) in all trials. They also easily distinguished between the repetitive and the smooth sounds. The samples with velvet-noise sequences composed of two delay lines were considered very repetitive and received the average score 34.7 (*Annoying*). The stimuli with three delay lines were perceived as less repetitive, receiving the average score of 61.5 (*Perceptible, but not annoying/Slightly annoying*). The sounds containing four or five delay lines, on the other hand, scored much higher and were perceived as almost equally smooth. The average scores for the stimuli with four and five delay lines were 82.0 and 88.1 (*Imperceptible*), respectively, and the confidence intervals overlapped, as shown in Fig. 3.

After the listening test, the participants were interviewed, and most of them classified the stimuli on each page into two groups: three samples were repetitive and of low quality, and the remaining three sounded smooth and were very similar to one another. The subjects described the samples that received low scores as "engine-like", "buzzy" or "flutter-like". On the other hand, the stimuli consisting of four or five delay lines were characterized as hardly distinguishable from the reference
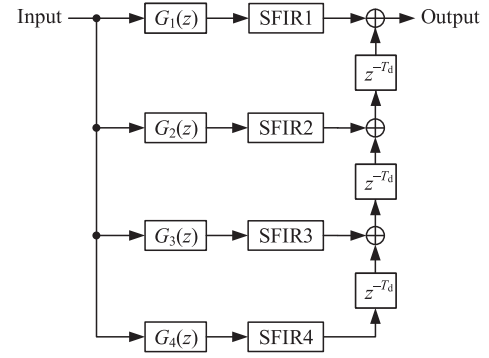

Fig. 4. Structure of the proposed reverberation algorithm based on interleaved velvet-noise filters ($M = 4$).
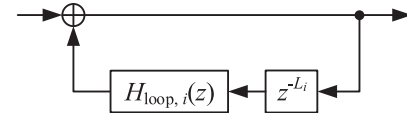

Fig. 5. Structure of transfer function $G_i(z)$ consisting of a delay line and a loop filter in a feedback loop. In practice, the delay line of L_i samples can be shared with the SFIR filter in the same branch (see Fig. 4).

and requiring more time and effort to assess. The stimuli used in the listening test are available online at [26].

III. NOVEL REVERBERATION ALGORITHM

This section introduces the new reverberation algorithm, which is based on the interleaving of velvet-noise signals.

A. Interleaved Velvet-Noise Reverberator

Fig. 4 shows the basic form of the proposed structure, called the interleaved velvet-noise sequence (IVN) reverberation algorithm, that consists of M parallel signal-processing branches. This example has $M = 4$ branches, which appears to be a sufficiently large number according to the listening test of Sec. II. Each branch includes a feedback structure $G_i(z)$ and a sparse FIR (SFIR) filter $S_i(z)$, the coefficients of which are samples taken from different EVNs. The branch impulse responses are interleaved by delaying them appropriately. The blocks marked z^{-T_d} in Fig. 4 are delay lines of T_d samples. When the delay-line length is set to $T_d = \hat{T}_d/M$, it is guaranteed that the non-zero samples produced by each SFIR filter will not occur at the same time, in the same manner as in Fig. 1.

Fig. 5 shows the structure of transfer function $G_i(z)$, which contains a feedback loop with a delay of L_i samples and a loop filter $H_{\text{loop},i}(z)$. Each SFIR filter also requires a delay line of L_i samples, but this delay memory can be shared with the comb filter: the SFIR filtering, which is equivalent to the convolution of the input signal with velvet-noise sequence, is in practice implemented with multiple output taps, which are added [11], [13], [27], [28]. The signal passing through the whole delay length of L_i samples, not processed with the multi-tap delay system, is the output signal fed to the loop filter and thereafter added to the input of feedback loop. Thus, only one delay line of L_i samples is needed for each branch.
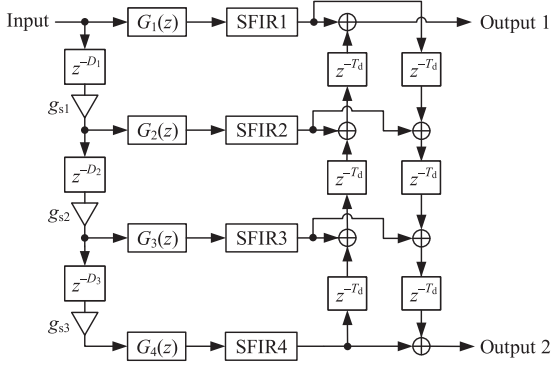
Fig. 6. Stereo IVN reverberator is obtained by interleaving the same velvet-noise sequences a second time in a different order, which is accomplished with additional delay lines of T_d samples on the right-hand side. The delays of D_1, D_2, and D_3 samples and gain factors g_{s1}, g_{s2}, and g_{s3} implement the smearing technique.

The EVN signals can be arranged in all different permutations, and not only in the order shown in Fig. 4. To obtain multiple decorrelated outputs, additional delay lines can be used to interleave the branch impulse responses. Fig. 6 shows one possible way to obtain two decorrelated outputs, which leads to a stereo effect. The impulse responses of outputs 1 and 2 consists of the same interleaved sequences but in opposite order, i.e., 1-2-3-4 vs. 4-3-2-1.

B. Controlling the Decay Rate

Each branch has its own loop filter $H_{\text{loop},i}(z)$ inside the feedback loop, as shown in Fig. 5. These filters must be designed collectively so that they all approximate the same frequency-dependent RT. One possibility is to devise a target response for each loop filter based on either a unit-sample response or a per-second response. The latter approach is chosen here, so that each target response is obtained by cascading a fraction $\nu = L_i/f_s$ of the prototype magnitude response $H_{\text{prot}}(\omega)$, when f_s is the sample rate used (e.g., 44.1 kHz):

$$H_{\text{target},i}(\omega) = |H_{\text{prot}}(\omega)|^{\nu}. \tag{4}$$

Each target response is then approximated using the same filter-design technique to obtain commensurate loop filters. Naturally, the order of each loop filter must be sufficiently large, so that they approximate well the target response. Otherwise, some components in the response will decay faster than the others, which will lead to a metallic disturbance in the sound [3].

Fig. 7 shows an example with $M = 4$, where the comb filters $G_i(z)$ are replaced with a delay loop having a constant gain factor (i.e., no filtering). The grid size for all EVNs is $\hat{T}_d = 80$ samples, which yields a sufficient velvet-noise density of 2205 non-zero samples at the sample rate of 44.1 kHz. The lengths of the four EVN sequences are again prime multiples of 80 samples: $L_1 = 97 \times 80$, $L_2 = 101 \times 80$, $L_3 = 103 \times 80$, and $L_4 = 107 \times 80$ samples. The gain factors controlling the decay of the four branches are 0.6669, 0.6558, 0.6504, and 0.6396, respectively, which correspond to a decay of $20\log(0.1) = 20$ dB per second, or a RT T60 of 3.0 s. Note that the steps
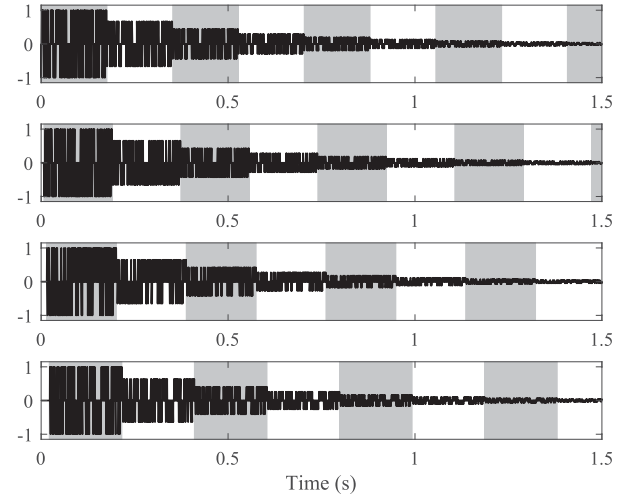

Fig. 7. Four repetitive EVN signals having lengths differing from each other and various starting times but decaying at the same rate. The decaying steps are accented with gray and white blocks.
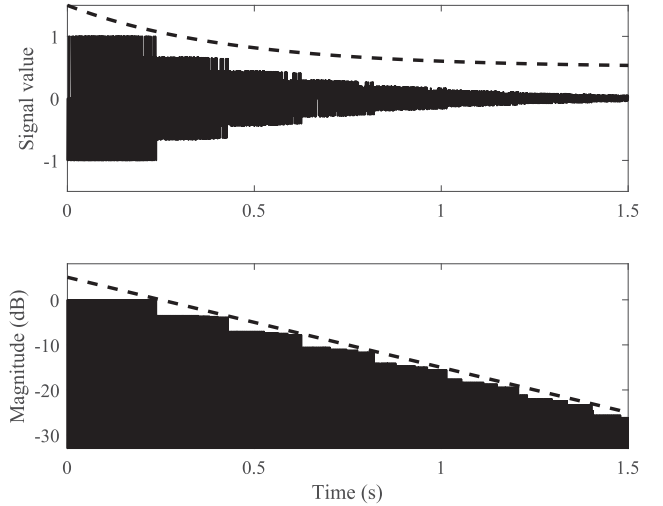

Fig. 8. (Top) The linear and (bottom) logarithmic impulse response obtained as a sum of the four sequences of Fig. 7. The dashed lines mark exponential decay with an offset of 0.5 and 5 dB in the top and bottom pane, respectively.

in the different signals of Fig. 7 appear always at different time instances, since sequences of different lengths are used.

Fig. 8 shows the impulse response obtained by adding the signals of Fig. 7. The signals in Figs. 7(b), (c), and (d) have been delayed by 20, 40, and 60 samples, respectively, with respect to the sequence in Fig. 7(a), i.e. $T_d = 20$ samples in Fig. 4. For this reason, the non-zero samples never appear at the same sample times in Fig. 8(a), but are always interleaved. Thus, the density of the sum signal is four times that of any of the individual EVN signals. Additionally, the decay of the logarithmic response in Fig. 8(b) is approximately linear (i.e., exponential on the linear scale), because the steps caused by gradual attenuation appear at different times. This impulse response is produced by the feedback structure in Fig. 4, having a total of $97 + 101 + 103 + 107 = 408$ taps in the four SFIR filters.
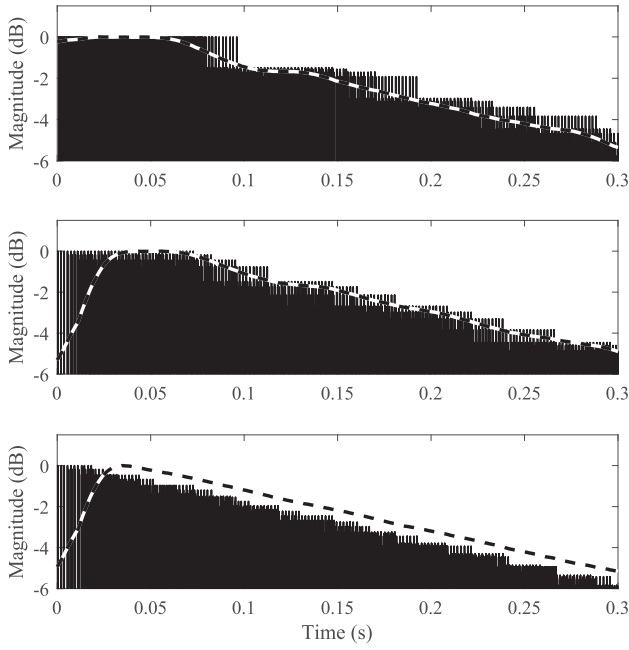
Fig. 9. (a) Synthetic impulse response and its envelope (dashed) produced with (a) all sequences starting at time zero without segmentation, (b) with smearing of the starting times, and (c) with both smearing and segmentation.
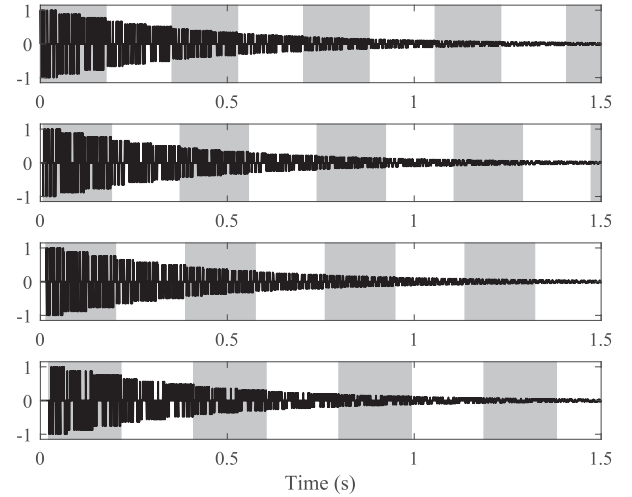

Fig. 10. Four EVN signals with segmented decay. The EVN frames accented with gray and white blocks now decay in three steps, leading to a closer approximation of an exponential decay than in Fig. 7.

C. Smearing by Delaying the Sequences

The difference in magnitude between the first and second step is the steepest in the entire signal and may be audible. To provide smoother changes and fade-in, a smearing technique, which has most impact on the first few steps, is introduced by delaying the start of all except the first EVN signals, as shown in Fig. 6, where delay lines of D_1, D_2, and D_3 samples implement this. In order to ensure that the non-zero samples will not occur at the same time, the delays must be multiples of the grid size $\hat{T}_d$.

An example of the delayed sequences is shown in Fig. 7. Each one of them is delayed by a multiple of $D_1 = 240$ samples, which equals $3\hat{T}_d$, where the integer multiplier is a free parameter set manually. Fig. 9 presents the comparison of the few first steps of the impulse response obtained with the IVN reverberator without (top pane) and with smearing (middle pane). This method is seen to introduce a fade-in technique, which requires more operations in other recursive reverberation algorithms, such as in the FDN [29]–[31]. Different starting times, however, do not change the fact that all sequences must approximate the same T60 value. Thus, the gain of each branch should be reduced proportionally to the delay in its starting time. This is realized using gain factors g_{s1}, g_{s2}, and g_{s3}, shown in Fig. 6.

D. Segmented Decay

Although the proposed method works best when synthesizing a very long reverberation tail, shorter impulse responses appear more frequently in practice. The disadvantage of the basic method in such cases is the step-like nature of the beginning of the synthetic response, as seen in Figs. 7 and 8. This problem is emphasized, when the RT is very small or the decay is very fast, leading to big differences between consecutive steps.

To achieve a smoother decay in the early part of the synthetic reverb, a segmentation method can be applied to each EVN sequence, as suggested by Alary *et al.* [11]. The decaying steps of L_i samples are divided into segments, which are attenuated in smaller steps by inserting a multiplier between two adjacent segments. This means that introducing one more segment to the EVN sequence requires just one multiplication. In the extreme case, every impulse in the EVN can have its own multiplying coefficient to obtain a completely smooth exponential decay, but then the computational efficiency is lost.

To avoid audible jumps in loudness between the steps, the level of the segments within one step should gradually decay from the initial level of this step to the initial level of the next one. The smallest number of segments producing a virtually stepless decay was empirically found to be three. This number prevents the segments from being too long and does not add excessive multiplication operations to the computational cost. Non-uniform segmentation is beneficial in terms of reducing the risk of audible periodic changes and achieving an exponential-like decay.

In the example shown in Fig. 10, the lengths of the segments were set to 25%, 35%, and 40% of the initial step length. To obtain small differences between consecutive steps in the presented case, gains for the segments were based on the difference in the magnitudes of two first steps. The first segment was left unaltered, the second was attenuated by 1/3 of the magnitude difference, and the third one by 2/3 of the magnitude difference. This way the change between the segments is always the same within one step.

Since each step is treated separately and the segmentation does not change the initial level of the EVN sequence, the overall decay rate of the sum of sequences is not affected by this operation. The segmentation of each frame into three requires only two extra multiplications per EVN branch. The bottom pane of Fig. 9 shows that adding the four EVN sequences having three
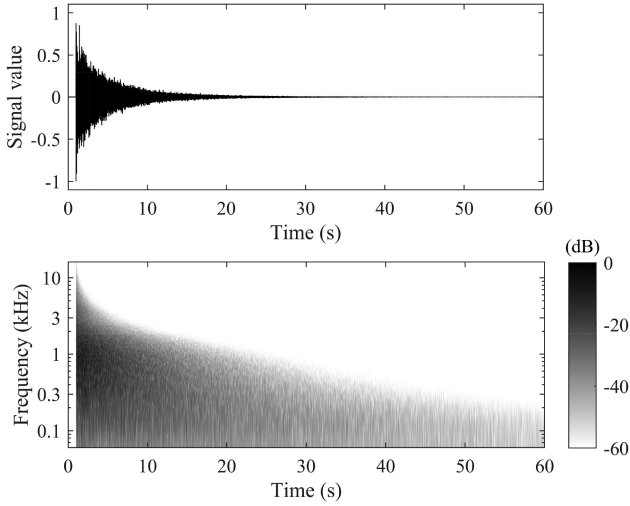
Fig. 11. (Top) Measured impulse response of the "world's longest echo" and (bottom) its spectrogram.

Fig. 12. (Top) Synthetic late reverberation and (bottom) spectrogram of the "world's longest echo" produced with the IVN reverberator, cf. Fig. 11.

decaying segments each leads to an approximately exponential decay pattern, i.e. a linear decay on the dB scale.

E. Loop-Filter Design

Historically, the first attempt to produce frequency-dependent reverberation was made by inserting a one-pole lowpass filter into a feedback structure [2], [3]. Later, controlling the decay rate in three independent frequency bands was possible by introducing biquadratic filters with adjustable crossover frequencies [32]. In [33], a 13th-order filter comprising single bandpass filters with a second-order Butterworth bandpass filter was proposed. Recently, Jot [34], Schlecht and Habets [35], and Prawda *et al.* [36] have considered using graphic equalizers to control the frequency-dependent reverberation time in FDN reverberators.

The approach adopted in this work used the cascaded graphic equalizer as proposed in [37] as an attenuation filter in order to accurately control the decay rate in ten octave frequency bands. The prototype per-second response was determined based on the reference RT values and transformed into the target responses for each of the delay lines es presented in (4). Shifting and scaling of the magnitude response by the median of gains was included, as suggested in [36]. The first-order high-shelf filter for attenuating frequencies above 16 kHz was also inserted in the loop filter.

IV. VALIDATION AND COMPARISON

This section presents results of synthesizing the reverberation tail of "the world's longest echo" and a concert hall response using the IVN reverberator. The properties and the computational cost of the proposed method and the FDN reverberator are also compared.

A. Synthesizing the Tail of the "World's Longest Echo"

To examine the ability of the proposed algorithm to reproduce the reverberation tail of a real impulse response, the extreme case of the world's longest reverberation was chosen. The sample
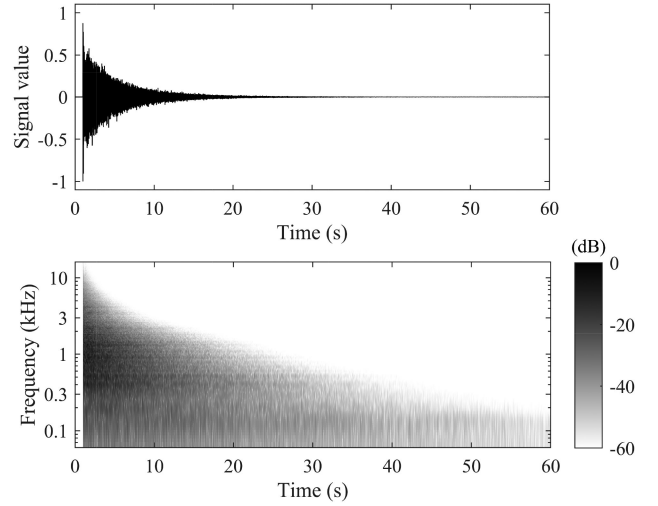
selected for analysis was recorded in tank number 1 at the Inchindown oil depository, Ross-shire, Scotland, U.K. [38] and was obtained from [39]. The average RT of the tank is 1 min 15 s [38]. Since the purpose of the experiment was to control the reverberation in as wide a frequency range as possible, the values were not taken directly from [38], where the T20 and T30 were given for seven and six octave frequency bands, respectively. Moreover, since the measurements of the world's longest impulse response were performed according to [40], the numbers provided are the result of analysis of several impulse responses, and thus may vary considerably from the sample chosen for the purpose of this work. Therefore, the reference RT values were calculated directly from the impulse response used in the experiment.

The IVN reverberator used for the experiment comprised four ($M = 4$) delay lines. This number was proven to be sufficient to obtain smooth, non-repetitive sound, as described earlier in Sec. II-C. The grid size of the IVN reverberator was set to $\hat{T}_d = 80$, and the lengths of the EVN sequences in samples were $L_1 = 97\hat{T}_d$, $L_2 = 101\hat{T}_d$, $L_2 = 103\hat{T}_d$, and $L_4 = 107\hat{T}_d$, which led to a total of 408 taps in the four SFIR filters ($97 + 101 + 103 + 107$). Because of small differences between the magnitude of the consecutive steps in the EVN sequences, there was no need to perform the segmentation mentioned in Sec. III-C.

The original impulse response of the "world's longest echo" and the corresponding spectrogram are shown in the top and bottom panes of Fig. 11, respectively. The reverberation tail synthesized with the IVN reverberator is depicted in the upper pane of Fig. 12, whereas its spectrogram is shown in the bottom pane. The impulse response synthesized with the proposed algorithm appears to be smoother than the original sample and its shape is more regular. However, the spectrograms reveal almost identical decay characteristics for both responses, with differences noticeable only in the low frequencies between 100 Hz and 300 Hz.

The impulse response with the reverberation tail created with the proposed method was also analyzed in terms of the RT values in octave bands. The obtained T60 values were compared to the reference and are depicted in Fig. 13. Minor differences between
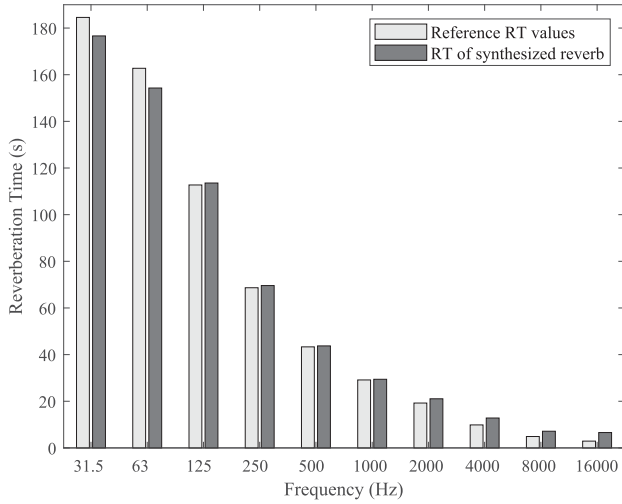
Fig. 13.   Comparison between reference values of the RT of the "world's longest echo" and the values obtained with the proposed method.

Fig. 14.   Segmented decay of SFIR coefficients improves the exponential shaping of the impulse response in the proposed IVN algorithm.

the measured and obtained values are visible. At low frequencies below 2 kHz, which are perceptually important, the difference does not exceed 5% of the target, making the dissimilarity imperceptible in this respect, according to [41]. Audio examples of both impulse responses—the original and the one with late reverberation created using the IVN reverberator—are available online at [26].

B. Synthesis of Short Reverberation

To test the performance of the algorithm on a practical case of a short impulse response, the reverberation of a concert hall in Pori, Finland, was synthesized. The T60 values in this case stretched from over 2 s in the low frequencies to less than 0.2 s in the high frequencies. The IVN reverberator's configuration was the same as that for reproducing the "world's longest echo" described in Sec. IV-A. To synthesize a short reverberation, however, the fast attenuation forced steep transitions from step to step in the interleaved sequences. Therefore, the segmentation of the steps was necessary to avoid audible artifacts in the produced reverberation. Each step was divided into three parts, as described in Sec. III-C. The effect of the segmentation, compared to the unsegmented IVN algorithm output, is shown in Fig. 14. Segmentation makes the transition between the consecutive steps more gradual, which results in a smoother sounding reverberation.

The spectrograms of the impulse responses with the original and the synthesized late reverberation are depicted in Fig. 15. The reverberation reproduced using the IVN reverberator follows the decay characteristics of the measured impulse response well up to around 8–9 kHz, where it is visibly slower. This is due to the step-like nature of the IVN reverberation, which means that the sound decay cannot be shorter than the longest EVN sequence, in this case 8560 samples, or 194 ms at the sampling rate of 44.1 kHz. Both impulse responses, together with other samples of short reverberation tails synthesized with the IVN reverberator, are available online at [26].
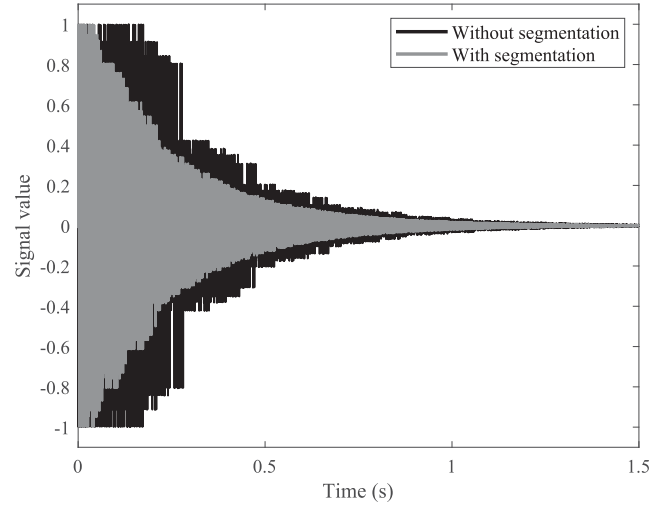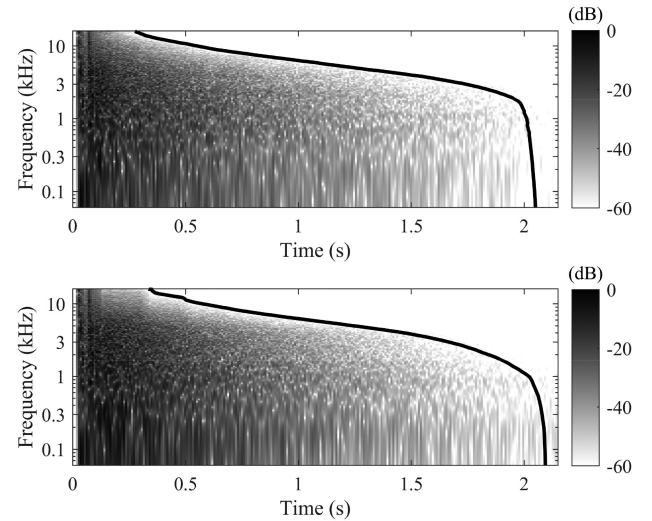

Fig. 15.   Spectrograms of the (a) original and (b) synthetic impulse response of the Pori concert hall. The RT estimation is shown for both with a black solid line.

C. Echo Density

Two aspects of the proposed method were compared to an FDN, which is considered to be a leading artificial reverberation algorithm. The first was the echo density of the reverberation tail produced with both methods. To avoid bias in the number of echoes caused by the smearing created by the attenuation filters, they were all removed. This way, only the echoes generated by the circulation of the delay lines in feedback structures were counted.

The main difference between the IVN and FDN reverberators is that, except for the beginning of the signal, the former algorithm has a fixed number of echoes, which is determined by the grid size T_d. Therefore, for the IVN reverberator used in Sec. IV-A with $T_d = 20$ and $f_s = 44.1$ kHz, the number of impulses per second is 2205. This was proven to produce perceptually smooth and random noise [20]. Thus, after the initial increase, determined by the delay after which the sequences
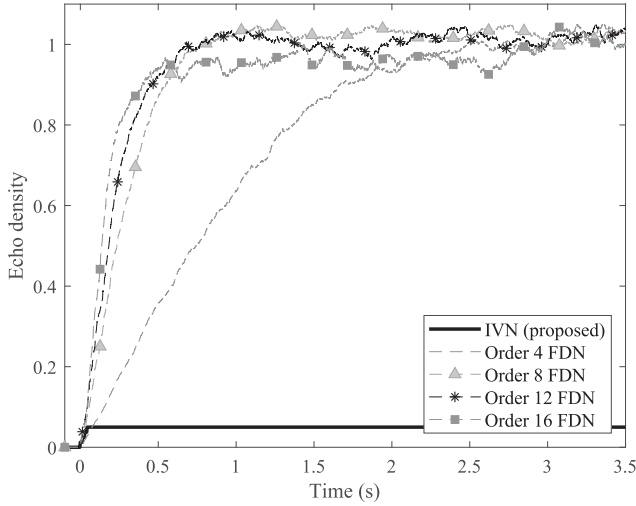
Fig. 16. Comparison of echo density between FDN reverberators of different order and the proposed IVN reverberator with four branches.
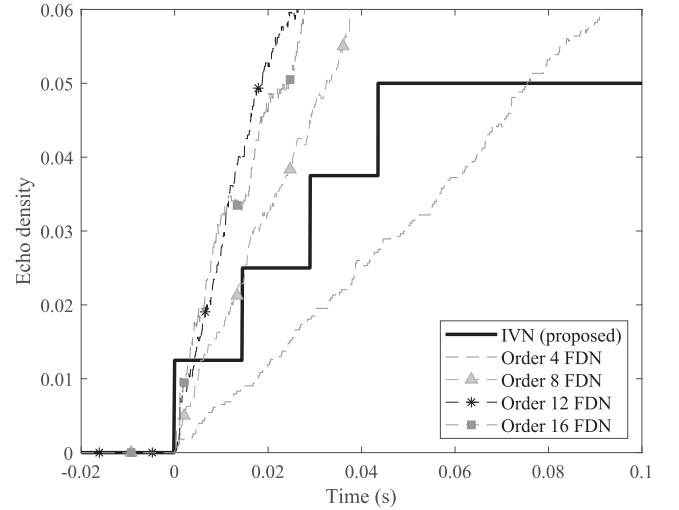

Fig. 17. Comparison of echo density growth in the beginning of the response, cf. Fig. 16.

TABLE I
NUMBER OF SFIR FILTER TAPS FOR DIFFERENT NUMBERS OF DELAY LINES IN THE IVN STRUCTURE

Delay Lines	4	6	8	10	12	14	16	18	20
Number of Taps	408	410	408	424	412	430	438	501	639

begin, the echo density in the synthesized late reverberation is 0.05 echoes per sample. The echo density is independent of the length of the signal and the number of delay lines in the IVN reverberator's structure.

On the other hand, the echo density of the FDN reverberator depends on the order of the structure and accumulates in time. Fig. 16 shows the normalized echo density for different orders of the FDN, calculated with the method proposed in [42]–[44], compared with that of the IVN reverberator. Fig. 17 shows the first 0.1 s of the echo density build-up. A very small number of impulses in the beginning of the impulse response of the FDN is usually clearly noticeable as a series of audible clicks and artifacts. As the build-up of the echo density is directly proportional to the order of the structure, a reverberator with a small number of delay lines is usually unsuitable for producing quality reverberation.

Additionally, the rise in the echo count of the FDN continues until the impulse response is saturated, i.e., there is an echo at every successive time unit [45]. The very dense impulse response contributes to production of the smooth reverberation tail, which is similar to white noise. However, this adds a considerable number of operations to the overall computational cost. At the same time, [20] proves that such a high echo density is unnecessary, since smooth noise-like sound can be obtained with a much smaller number of echoes per sample, when the impulses are appropriately distributed over time, i.e., never too densely or sparsely. This is how impulses appear in the velvet-noise sequences used in the IVN algorithm, which quickly reaches the perceptually sufficient density, as shown in Fig. 17.

D. Computational Cost

The second aspect assessed here is the computational cost. The computational efficiency of reverberation algorithms is of great importance, since they are often used in real-time applications. The cost is usually presented as the number of floating-point operations (FLOP) per processed sample, specified as a

sum of additions and multiplications required by the algorithm to produce one output sample.

The computational complexity of the proposed IVN reverberator depends on the number of EVN sequences, number of taps in the SFIR filters, and the complexity of the attenuation filter. Adding together the prime numbers C_i that specify the number of taps in the SFIR filters, the FLOPs required by the attenuation filter for every delay line, the addition per delay line from each comb filter (M), two multiplications per branch that account for segmentation ($2M$), $M-1$ multiplications that adjust the gain for the delay introduced by smearing, and the $M-1$ additions that form the output signal. The number of operations for an attenuation filter consisting of 10 second-order infinite impulse response sections and a first-order shelving filter is 53 multiplications and 41 additions, or 94 operations in total.

To account for the fact that the number of taps should be estimated using unique prime numbers, the number of FLOPs was determined when there were at least 400 taps in the SFIR filters. Of course, this number increases with the number of EVN sequences, since finding 18 or more unique primes that add up to less than 500 is impossible. The required number of taps for IVN structures with 4 to 20 branches is shown in Table I.

To calculate the total FLOP/sample for an FDN, the formula proposed in [46] was used. Since the approach of comparing the number of operations per processed sample was adopted, the multiplication by the sample rate was omitted.

The costs of both methods were determined for different numbers of delay lines, including operations required by the attenuation filter. The results of the comparison are shown in Fig. 18 and Table II. Both algorithms are almost equal in complexity
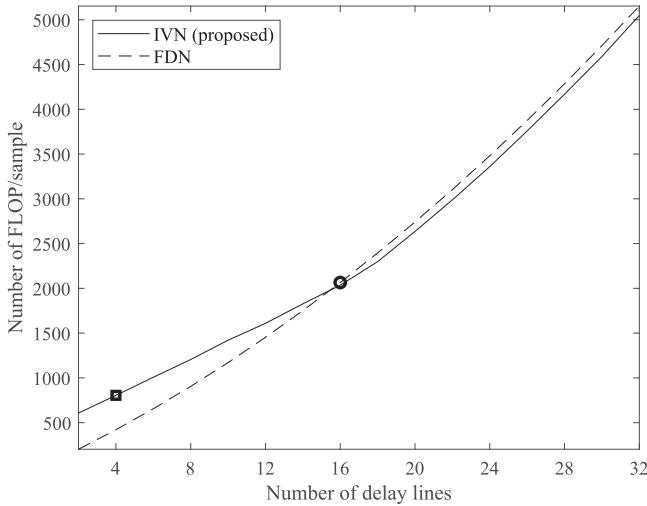
Fig. 18. Comparison between the computational cost of the proposed method (IVN) and the FDN. The conditions in which each method gives smooth reverberation tail are 4 delay lines and 802 FLOP/sample for the IVN (marked with a square), and 16 delay lines and 2065 FLOP/sample for the FDN (indicated with a circle).

TABLE II
NUMBER OF OPERATIONS PER OUTPUT SAMPLE FOR IVN
AND FDN REVERBERATORS. THE NUMBER OF DELAY LINES FOR WHICH
A SMOOTH REVERBERATION TAIL IS ENSURED IS EMPHASIZED IN
BOLD FOR EACH METHOD

Number of DLs	Number of FLOP/sample	
	IVN (proposed)	FDN
4	**802**	421
6	1002	655
8	1198	905
10	1412	1171
12	1598	1453
14	1814	1751
16	2020	**2065**
18	2281	2395
20	2617	2741

computational cost is 2065 operations per sample, which is marked with a circle in Fig. 18.

A comparison of the highlighted computational costs of 802 and 2065 operations per sample in Table II shows that the number of operations required by the IVN reverberator is just 40% of that required by the FDN. This study indicates that the proposed IVN reverberator can provide a high-quality response with a much smaller number of parallel systems than the FDN. This reduces similarly the number of loop filters needed.

V. DISCUSSION

The proposed algorithm synthesizes the late part of the reverberation well in terms of reverberation time values and echo density. As the sound examples presented online [26] prove, the synthetic impulse responses are very similar to the original ones, but there are still perceptible differences.

The evaluation of synthetic reverberation is a complex issue. On one hand, objective measures, such as reverberation time or frequency characteristics of the reverberation [50], provide quantification which, to some extent, is relevant to the acoustic quality of the produced sound [51], [52]. Such parameters, however, do not consider the whole spectrum of perceptual aspects of sound within a room, which has been studied for well over a century [53]. Recent research suggests that there are tens of attributes associated with small listening rooms only [54] and at least as many used to describe concert hall acoustics [55]–[58]. Of these, the term "reverberance," which is used as a descriptor for the perception of reverberation, is not solely dependent on the T60 values, in the same way as "diffuseness" is not only determined by echo density.

On the other hand, numerous studies have compared synthesized reverberation with measured impulse responses [27], [59]–[62]. Participants of listening tests can usually distinguish between the target and synthesized sound [27], [59], [60], [62]. This shows that the parameters regulating artificial reverberation algorithms are insufficient to impeccably imitate the complexity of a real-world sound. However, the practical use of artificial reverberation hardly ever involves direct comparison between measured and synthetic impulse responses of the same space. Thus, plausible yet slightly modified artificial reverberation can still be highly useful in music production and gaming, where it is used for artistic effect or spatial impression.

VI. CONCLUSION

This paper proposes a novel algorithm for synthesizing late reverberation by interleaving extended velvet-noise sequences. Each of the EVNs has one non-zero sample in each $MT_{\rm d}$ samples, where M is the number of interleaved sequences and $T_{\rm d}$ is the grid size, which in this work was set to 20 samples. Delaying each EVN by a different number of sampling intervals ensures that the non-zero samples never occur at the same time when several parallel sequences are combined. This is a new principle in audio processing. The results of the listening test presented in this paper show that interleaving four EVNs having a total density of 2205 samples/s is sufficient to obtain perceptually smooth, non-repetitive noise.

for 16 delay lines, with the FDN being less costly for small and the IVN for big numbers of delay lines. However, because the proposed method works well for as few as four delay lines, the number of FLOPs required for good-quality late reverberation is only 802 operations (794 without segmenting). This number is marked in Fig. 18 with a square on the corresponding curve as well as emphasized in bold font in Table II.

The smallest FDN order that is sufficient for high-quality reverberation is still a controversial question. Our recent work shows that the smallest useful order of the FDN is 16 [47], whereas Alary *et al.* point out that an order as high as 32 may be necessary to achieve sufficient echo and modal densities, depending on the algorithm implementation [48]. Fagerström *et al.* consider the reverberation produced by an FDN of order 32 as sufficiently dense and the one synthesized with a 16th-order FDN as slightly too sparse [49]. For fairness, we choose here the order 16 as the smallest order for the FDN that is useful for high-quality audio. Table II indicates that the corresponding

Each branch of the IVN reverberator includes an SFIR filter and a feedback structure containing a delay line and a loop filter. Because SFIR filtering is equivalent to convolving the input with a velvet-noise sequence that can be implemented by adding and subtracting delayed samples, only one delay line is needed for each branch of the reverberator. To control the decay rate in ten octave bands, this work proposes to use an accurate graphic equalizer as a loop filter in every branch.

The proposed IVN reverberator is best suited for producing long reverberation, and therefore the "world's longest" impulse response was used for validation. In this case, the perceptually important RT values for frequencies below 2 kHz deviated by less than 5% of the target values. For the synthesis of shorter impulse responses, this paper proposes a segmented decay technique, that helps to attenuate the velvet-noise sequences in fine steps, approaching a continuous exponential decay. An example design showed how accurately a concert hall impulse response could be reproduced.

The proposed IVN algorithm produces late reverberation with an optimal echo density that ensures smooth sound and is economical to compute. The number of FLOPs per processed sample in the proposed IVN algorithm is about 60% lower than that of the FDN algorithm synthesizing reverberation of comparable smoothness.

Future work may investigate how well the different permutations of the branches of the IVN reverberator are decorrelated. This will lead to a better understanding how to use the proposed method in multichannel setups. Future research may also aim at improving the knowledge of the perception of reverberation.

Acknowledgment

References

[1] 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.

[2] J. Moorer, "About this reverberation business," *Comput. Music J.*, vol. 3, no. 2, pp. 13–28, 1979.

[3] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. Audio Eng. Soc. 90th Conv.*, Paris, France, Feb. 1991, pp. 1–16, Paper 3030.

[4] 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.

[5] S. J. Schlecht and E. A. P. Habets, "On lossless feedback delay networks," *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1554–1564, Mar. 2017.

[6] P. Rubak and L. G. Johansen, "Artificial reverberation based on a pseudo-random impulse response, Part I," in *Proc. Audio Eng. Soc. 104th Conv.*, Amsterdam, The Netherlands, May 1998, pp. 1–13, Paper 4725.

[7] P. Rubak and L. G. Johansen, "Artificial reverberation based on a pseudo-random impulse response, Part II," in *Proc. Audio Eng. Soc. 106th Conv.*, Munich, Germany, May 1999, pp. 1–15, Paper 4900.

[8] M. Karjalainen and H. Järveläinen, "Reverberation modeling using velvet noise," in *Proc. Audio Eng. Soc. 30th Int. Conf. Intell. Audio Environ.*, Saariselkä, Finland, pp. 1–9, Oct. 2007.

[9] K. S. Lee, J. S. Abel, V. Välimäki, T. Stilson, and D. B. Berners, "The switched convolution reverberator," *J. Audio Eng. Soc.*, vol. 60, no. 4, pp. 227–236, Apr. 2012.

[10] S. Oksanen, J. 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.*, Vancouver, Canada, May 2013, pp. 644–648.

[11] B. Alary, A. Politis, and V. Välimäki, "Velvet-noise decorrelator," in *Proc. Int. Conf. Digit. Audio Effects*, Edinburgh, U.K., Sep. 2017, pp. 405–411.

[12] S. J. Schlecht, B. Alary, V. Välimäki, and E. A. P. Habets, "Optimized velvet-noise decorrelator," in *Proc. Int. Conf. Digit. Audio Effects*, Aveiro, Portugal, Sep. 2018, pp. 87–94.

[13] V. Välimäki, J. Rämö, and F. Esqueda, "Creating endless sounds," in *Proc. Int. Conf. Digit. Audio Effects*, Aveiro, Portugal, Sep. 2018, pp. 32–39.

[14] S. D'Angelo and L. Gabrielli, "Efficient signal extrapolation by granulation and convolution with velvet noise," in *Proc. Int. Conf. Digit. Audio Effects*, Aveiro, Portugal, Sep. 2018, pp. 107–112.

[15] K. J. Werner, "Generalizations of velvet noise and their use in 1-bit music," in *Proc. Int. Conf. Digit. Audio Effects*, Birmingham, U.K., pp. 1–8, Sep. 2019.

[16] H. Kawahara, K.-I. Sakakibara, M. Morise, H. Banno, T. Toda, and T. Irino, "Frequency domain variants of velvet noise and their application to speech processing and synthesis," in *Proc. Interspeech*, Hyderabad, India, Sep. 2018, pp. 2027–2031.

[17] H. Kawahara, "Application of the velvet noise and its variant for synthetic speech and singing," IPSJ SIG Tech. Rep., vol. 2018-MUS-118, no. 3, 2018, pp. 1–5.

[18] M. Morise, "Modification of velvet noise for speech waveform generation by using vocoder-based speech synthesizer," *IEICE Trans. Inf. Syst.*, vol. E 102D, no. 3, pp. 663–665, Mar. 2019.

[19] H. Kawahara, K. Sakakibara, M. Mizumachi, H. Banno, M. Morise, and T. Irino, "Frequency domain variant of velvet noise and its application to acoustic measurements," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, Lanzhou, China, Nov. 2019, pp. 1523–1532.

[20] V. Välimäki, H.-M. Lehtonen, and M. Takanen, "A perceptual study on velvet noise and its variants at different pulse densities," *IEEE Trans. Audio Speech Lang. Process.*, vol. 21, no. 7, pp. 1481–1488, Jul. 2013.

[21] N. Guttman and B. Julesz, "Lower limits of auditory periodicity analysis," *J. Acoust. Soc. Amer.*, vol. 35, no. 4, Apr. 1963, Art. no. 610.

[22] R. M. Warren and J. A. Bashford, "Perception of acoustic iterance: Pitch and infrapitch," *Percept. Psychophys.*, vol. 29, no. 4, pp. 395–402, May 1981.

[23] R. M. Warren, J. A. Bashford, J. M. Cooley, and B. S. Brubacker, "Detection of acoustic repetition for very long stochastic patterns," *Percept. Psychophys.*, vol. 63, no. 1, pp. 175–182, Jan. 2001.

[24] International Telecommunication Union, "Method for the subjective assessment of intermediate quality level of audio systems," Recommendation ITU-R BS.1534-3, Tech. Rep., Oct. 2015.

[25] M. Schoeffler, S. Bartoschek, F.-R. Stöter, M. Roess, S. W. B. Edler, and J. Herre, "WebMUSHRA—A comprehensive framework for web-based listening tests," *J. Open Res. Softw.*, vol. 6, no. 1, pp. 1–8, Feb. 2018.

[26] K. Prawda and V. Välimäki, "Companion page: Late reverberation synthesis using interleaved velvet-noise sequences," Accessed: Dec. 18, 2020. [Online]. Available: http://research.spa.aalto.fi/publications/papers/ieee-taslp-ivn/

[27] V. Välimäki, B. Holm-Rasmussen, B. Alary, and H.-M. Lehtonen, "Late reverberation synthesis using filtered velvet noise," *Appl. Sci.*, vol. 7, no. 5, pp. 1–17, May 2017.

[28] B. Holm-Rasmussen, H.-M. Lehtonen, and V. Välimäki, "A new reverberator based on variable sparsity convolution," in *Proc. Int. Conf. Digit. Audio Effects*, Maynooth, Ireland, Sep. 2013, pp. 344–350.

[29] E. Piirilä, T. Lokki, and V. Välimäki, "Digital signal processing techniques for non-exponentially decaying reverberation," in *Proc. 1st COST-G6 Workshop Digit. Audio Effects*, Barcelona, Spain, Nov. 1998, pp. 21–24.

[30] K.-S. Lee and J. S. Abel, "A reverberator with two-stage decay and onset time controls," in *Proc. Audio Eng. Soc. 129th Conv.*, San Francisco, CA, USA, Nov. 2010, pp. 1–6.

[31] N. Meyer-Kahlen, S. J. Schlecht, and T. Lokki, "Fade-in control for feedback delay networks," in *Proc. Int. Conf. Digit. Audio Effects*, Vienna, Austria, Sep. 2020, pp. 227–233.

[32] 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, Sep. 1997, pp. 1–8.

[33] T. Wendt, S. van de Par, and S. D. Ewert, "A computationally-efficient and perceptually-plausible algorithm for binaural room impulse response simulation," *J. Audio Eng. Soc.*, vol. 62, no. 11, pp. 748–766, Nov. 2014.

[34] J.-M. Jot, "Proportional parametric equalizers—Application to digital reverberation and environmental audio processing," in *Proc. Audio Eng. Soc. 139th Conv.*, New York, NY, USA, Oct. 2015, pp. 1–8, Paper 9358.

[35] S. J. Schlecht and E. A. P. Habets, "Accurate reberberation time control in feedback delay networks," in *Proc. Int. Conf. Digit. Audio Effects*, Edinburgh, U.K., Sep. 2017, pp. 337–344.

[36] K. Prawda, S. J. Schlecht, and V. Välimäki, "Improved reverberation time control for feedback delay networks," in *Proc. Int. Conf. Digit. Audio Effects*, Birmingham, U.K., Sep. 2019, pp. 1–8.

[37] V. Välimäki and J. Liski, "Accurate cascade graphic equalizer," *IEEE Signal Process. Lett.*, vol. 24, no. 2, pp. 176–180, Feb. 2017.

[38] T. Cox and A. Kilpatrick, "A record longest echo within the Inchindown oil despository (L)," *J. Acoust. Soc. Amer.*, vol. 137, no. 3, pp. 1602–1604, Mar. 2015.

[39] T. Cox. World's 'longest-Echo' Fifth Impulse. Jan. 2014. [Online]. Available: http://freesound.org/people/acs272/sounds/214221/

[40] ISO, "ISO 3382-2:2008, Acoustics - Measurement of room acoustic parameters - part 2: Reverberation time in ordinary rooms," Int. Org. Standardization, Geneva, Switzerland, Tech. Rep., 2009.

[41] ISO, "ISO 3382-1:2009, Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces," Int. Org. Standardization, Geneva, Switzerland, Tech. Rep., 2009.

[42] J. S. Abel and P. Huang, "A simple, robust measure of reverberation echo density," in *Proc. Audio Eng. Soc. 121st Conv.*, San Francisco, CA, USA, Oct. 2006, pp. 1–10.

[43] P. Huang and J. S. Abel, "Aspects of reverberation echo density," in *Proc. Audio Eng. Soc. 123 rd Conv.*, New York, NY, USA, Oct. 2007, pp. 1–7.

[44] P. Huang, J. S. Abel, H. Terasawa, and J. Berger, "Reverberation echo density psychoacoustics," in *Proc. Audio Eng. Soc. 125th Conv.*, San Francisco, CA, USA, Oct. 2009, pp. 1–10.

[45] S. J. Schlecht and E. A. P. Habets, "Feedback delay networks: Echo density and mixing time," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 2, pp. 374–383, Feb. 2016.

[46] E. De Sena, H. Hachabiboglu, Z. Cvetkovic, and J. O. Smith, "Efficient synthesis of room acoustics via scattering delay networks," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 23, no. 9, pp. 1478–1492, Sep. 2015.

[47] K. Prawda, S. Willemsen, S. Serafin, and V. Välimäki, "Flexible real-time reverberation synthesis with accurate parameter control," in *Proc. Int. Conf. Digit. Audio Effects*, Vienna, Austria, Sep. 2020, pp. 16–23.

[48] 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.

[49] J. Fagerström, B. Alary, S. J. Schlecht, and V. Välimäki, "Velvet-noise feedback delay network," in *Proc. Int. Conf. Digit. Audio Effects*, Vienna, Austria, Sep. 2020, pp. 219–226.

[50] A. Czyżewski, "A method of artificial reverberation quality testing," *J. Audio Eng. Soc.*, vol. 38, no. 3, pp. 129–141, Mar. 1990.

[51] L. Cremer and H. A. Müller, *Principles and Applications of Room Acoustics*. Barkiny, Essex, England: Applied Science, 1982.

[52] P. Malecki, K. Sochaczewska, and J. Wiciak, "Settings of reverb processors from the perspective of room acoustics," *J. Audio Eng. Soc.*, vol. 68, no. 4, pp. 292–301, Apr. 2020.

[53] W. C. Sabine, *Collected Papers on Acoustics*. Cambridge, MA, USA: Harvard Univ. Press, 1922.

[54] N. Kaplanis, S. Bech, T. Lokki, T. van Waterschoot, and S. Holdt Jensen, "Perception and preference of reverberation in small listening rooms for multi-loudspeaker reproduction," *J. Acoust. Soc. Amer.*, vol. 146, no. 5, pp. 3562–3576, Nov. 2019.

[55] N. Kaplanis, S. Bech, S. H. Jensen, and T. van Waterschoot, "Perception of reverberation in small rooms: A literature study," in *Proc. Audio Eng. Soc. 55th Int. Conf. Spatial Audio*, Helsinki, Finland, Aug. 2014, pp. 1–14.

[56] A. Kuusinen and T. Lokki, "Wheel of concert hall acoustics," *Acta Acoust. United Acustica.*, vol. 103, no. 2, pp. 185–188, Mar./Apr. 2017.

[57] T. Lokki, J. Pätynen, A. Kuusinen, H. Vertanen, and S. Tervo, "Concert hall acoustics assessment with individually elicited attributes," *J. Acoust. Soc. Amer.*, vol. 130, no. 2, pp. 835–849, Aug. 2011.

[58] T. Lokki, J. Pätynen, A. Kuusinen, and S. Tervo, "Disentangling preference ratings of concert hall acoustics using subjective sensory profiles," *J. Acoust. Soc. Amer.*, vol. 132, no. 5, pp. 3148–3161, Nov. 2012.

[59] T. Wendt, S. van de Par, and S. D. Ewert, "A computationally-efficient and perceptually-plausible algorithm for binaural room impulse response simulation," *J. Audio Eng. Soc.*, vol. 62, no. 11, pp. 748–766, Nov. 2014.

[60] M. Steimel, "Implementation of a hybrid reverb algorithm-parameterizing synthetic late reverberation from impulse responses," Master's thesis, Aalborg Univ., Aalborg, Denmark, Aug. 2019.

[61] S. Djordjevic, H. Hacihabiboglu, Z. Cvetkovic, and E. De Sena, "Evaluation of the perceived naturalness of artificial reverberation algorithms," in *Proc. Audio Eng. Soc. 148th Conv.*, May 2020, pp. 1–10.

[62] K. Prawda, V. Välimäki, and S. Serafin, "Evaluation of accurate artificial reverberation algorithm," in *Proc. 17th Sound Music Comput. Conf.*, Turin, Italy, Jun. 2020, pp. 247–254.

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 Pori unit, Tampere University of Technology, Pori, Finland. From 2006 to 2007, he was the Head with the Laboratory of Acoustics and Audio Signal Processing, TKK. From 2008 to 2009, he was a Visiting Scholar with Stanford University, Stanford, CA, USA. He is currently a Full Professor of audio signal processing and the Vice Dean for research in electrical engineering with Aalto University, Espoo, Finland. His research interests include artificial reverberation, digital filter design, audio effects processing, and sound synthesis.

Prof. Välimäki is a Fellow of the Audio Engineering Society and a Life Member of the Acoustical Society of Finland. 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 and from 2007 to 2011, he was an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS and the IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING. From 2015 to 2020, he was a Senior Area Editor of the IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING. He was on the Editorial Board of the *Research Letters in Signal Processing* and the *Journal of Electrical and Computer Engineering*. He was the Lead Guest Editor of a Special Issue of the IEEE *Signal Processing Magazine* in 2007 and of a Special Issue of the IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING in 2010. He was the Guest Editor of Special Issues of the IEEE *Signal Processing Magazine* in 2015 and 2019. He has been the Guest Editor of Special Issues published in *Applied Sciences* in 2016, 2018, and 2020. He was the Chair of the 2008 International Conference on Digital Audio Effects (DAFX) and the Chair of the 2017 International Conference on Sound and Music Computing. He is currently the Editor-in-Chief of the *Journal of the Audio Engineering Society*.

Karolina Prawda received the M.Sc. degree in acoustic engineering from the AGH University of Science and Technology, Kraków, Poland, in 2017.

She is currently working toward the Doctoral degree with the Acoustics Lab, Aalto University, Espoo, Finland. Her research interests include artificial reverberation and variable acoustics.

Since 2018, she has been a Member of the Polish Section of the Audio Engineering Society.