A DIRECTIONAL DIFFUSE REVERBERATION MODEL FOR EXCAVATED TUNNELS IN ROCK

Sami Oksanen, Julian Parker, Archontis Politis, and Vesa Välimäki

Aalto Univ., School of Electrical Engineering
Department of Signal Processing and Acoustics
P.O. Box 13000, FI-00076 AALTO, Espoo, Finland

ABSTRACT

Acoustic impulse responses of an excavated tunnel were measured. Analysis of the impulse responses shows that they are very diffuse from the start. A reverberator suitable for reproducing this type of response is proposed. The input signal is first comb-filtered and then convolved with a sparse noise sequence of the same length as the filter's delay line. An IIR loop filter inside the comb filter determines the decay rate of the response and is derived from the Yule-Walker approximation of the measured frequency-dependent reverberation time. The particular sparse noise sequence proposed in this work combines three velvet noise sequences, two of which have time-varying weights. To simulate the directional soundfield in a tunnel, the use of multiple such reverberators, each associated with a virtual source distributed evenly around the listener, is suggested. The proposed tunnel acoustics simulation can be employed in gaming, in film sound, or in working machine simulators.

Index Terms—Acoustic measurements, acoustic signal processing, convolution, reverberation

1. INTRODUCTION

Recorded impulse responses of spaces have been directly used as a means to add natural reverberation of a specific character to a dry recording, or as a way to listen how a specific space would affect a source signal, a process that in the context of architectural and virtual acoustics is termed auralisation [1, 2]. A straightforward way to achieve that is to convolve a source signal with a recorded room impulse response (RIR), which herein is termed as convolution reverberation [1]. An alternative approach, which provides more flexibility and computational efficiency, is to estimate the most relevant parameters from a recorded RIR and then map these parameters to common reverberation algorithms, such as combinations of delay-lines and feedback delay networks [3, 4, 5, 1]. Common parameters are mixing times between the early and late reverberant part, arrival times and amplitudes of discrete early reflections and reverberation times of the late diffuse part [6].

This work was funded by the Finnish Work Environment Fund, grant no. 111244, GETA, Sandvik Mining and Construction Corp., Finland. The authors would like to thank Mr. Otto Hedström from Aalto Univ. Dept. Civil and Environmental Engineering for assistance provided during the measurements and for the photograph presented in Fig. 1, Dr. Heidi-Maria Lehtonen and Mr. Bo Holm-Rasmussen for helpful comments.

Fig. 1. Overall view of the main practice drilling tunnel. The sound source was located at the far end of the tunnel and receiver locations varied along the tunnel.

For surrounding multichannel rendering of reverberation, its directional characteristics should be taken into account, including direction of incidence of early echoes and overall directional distribution of energy and decay times of the late reverberant part [7]. A direct approach to generate multichannel RIRs for convolution reverberation from a directional recording is the ambisonic method [8], [9], in which a 4-channel B-format RIR is distributed to a multichannel loudspeaker setup by means of a least-squares inversion. For an overall description on the ambisonic methods and the B-format see [10]. However, due to the very limited angular resolution of first-order ambisonics, inter-channel coherence is high resulting in strong coloration of the reproduced reverberation in reproduction, position-dependent comb filtering effects and a reduced sense of diffusion.

A more recent method which utilizes the information encoded in the B-format in a more perceptually-relevant way is the Spatial Impulse Response Rendering (SIRR) [11]. SIRR performs a time-frequency energetic analysis on the B-format RIR and decomposes it in a directional part, which contains the direct sound and strong reflections, panned to the appropriate speakers by means of vector-base amplitude panning (VBAP) [12], and a diffuse part which is reproduced by all loudspeakers by means of decorrelating filters. Another recent approach similar in spirit to SIRR, but for arbitrarily spaced arrays instead of B-format input, is the one presented in [13]. Both of these methods achieve high-quality rendering of recorded responses, however the end product is meant for convolution reverberators, mainly for auralisation of concert halls. In addition, the time-frequency analysis and synthesis of the RIRs can be computationally intensive for many applications.

This work's relation to prior research is that the use of artificial reverberation is expanded from concert halls and rooms to tunnels, and a novel directional reverberation algorithm is introduced. This study presents an efficient algorithm for surrounding reverberation also based on a B-format RIR, which avoids the computational load of convolution reverbs while retaining the basic directional characteristics of the reverberation. The algorithm is suitable for diffusive spaces without dominant reflections, such as the cave space under study, or as a processing block in a system that handles the early reflections separately. Prior work of modeling tunnel reverbs includes the study by Collecchia et al. [14] where a method to model tunnel systems using digital waveguide networks is presented. However, it does not cover directional reverberation modeling.

In Sec. 2, we describe the measurement of the B-Format impulse response of an excavated rock tunnel, and present some analysis of its main characteristics. In Sec. 3 we present a method of approximating this reverberation for arbitrary input through a novel multichannel sound reproduction system. In Sec. 4, we conclude.

2. TUNNEL IMPULSE RESPONSE MEASUREMENTS

The measured space is a teaching and test tunnel located at the Aalto University campus, see Fig. 1. The characteristics of the tunnel match closely the ones at excavation and underground construction sites. The tunnel floor consists of a layer of gravel on top of the bedrock. Parts of the tunnel walls and ceiling are reinforced with sprayed plaster coating, but the majority of the wall area is bare and extremely uneven rock.

The acoustic properties of the space were determined using the standard log-sweep method [15, 16]. A sound source is located at a certain position within the tunnel, and a set of logarithmic sine sweeps are played. The acoustic response to this sweep is measured using a microphone positioned at a pre-determined location, at a great enough distance to be well within the reverberant soundfield. Sound signal playback and recording were carried out with custom made software in Matlab. Log-sweep duration was 10 s and frequency range was 20 Hz to 20 kHz. The excitation signal was repeated five times in each speaker and receiver position to help avoid sources of random error.

The excitation signals were produced using a stand-mounted active loudspeaker (Genelec 1032A). To approximate an omnidirectional acoustic excitation with the directional speaker, the measurements were conducted in seven different speaker orientations and averaged. The set of speaker orientations consisted of six positions in the horizontal plane 60 degrees apart from each other and one where loudspeaker was pointing upwards. The response measurements were also carried out using a passive omnidirectional loudspeaker. The omnidirectional loudspeaker was limited in efficiency and was not capable of producing the high SPLs required at greater distances, so the omnidirectional measurements were used to verify the results produced by the directional speaker.

Acoustic responses were measured using a pressure-field microphone (B&K 4192) and a B-format Soundfield microphone (ST 350) simultaneously. The microphones were placed approximately 1.5 m above the floor and close to the axial line of the tunnel. The distance between microphones was approximately 30 cm. The distance of the receiver point from the sound source was varied from 5 m to 20 m.

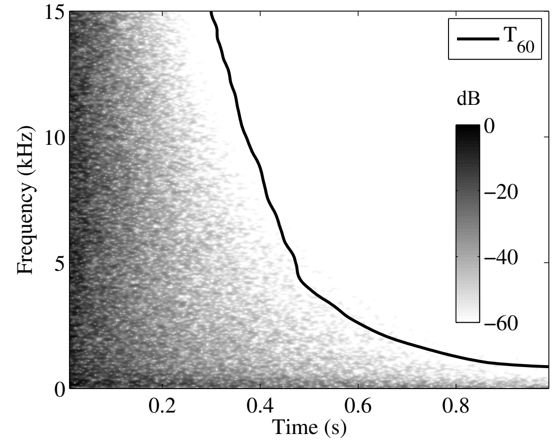

Fig. 2. Spectrogram of the measured impulse response and estimated T_{60} value.

The impulse response of the space is derived from the measured responses by convolving it with the time reversed version of the excitation signal. The impulse responses are then time-aligned and averaged with the other appropriate measurements. In the case of the directional speaker, this is all repeated for measurements at the seven different angles for a particular distance between sound source and receiver. In the omnidirectional speaker case, this is just the repeated measurements made at that distance.

2.1. Analysis of measured responses

The first step of analysis was to compare the responses produced by the average of directional speaker responses with those produced by the omnidirectional speaker. The results produced agree closely enough to suggest that the average of directional speaker responses is a good approximation, especially within the horizontal plane. The omnidirectional speaker has a poor response below around 100 Hz, and so verification is not possible below that frequency. However, at this frequency the directionality of the Genelec 1032A is less pronounced, so we assume that the approximation holds.

Figure 2 shows a spectrogram of the response produced using the average of directional speaker responses, taken at a distance of 13m between source and receiver. The response is created from the B-format signal recorded at the Soundfield microphone by creating a virtual microphone pointing towards the source. Overlayed on the spectrogram is the estimated T_{60} with respect to frequency. Inspection of the spectrogram shows that the response consists of only diffuse sound, with no clear discrete echoes visible.

Figure 3 shows relative T_{60} times of the responses produced by calculating eight different virtual microphones from the B-format signal, pointing in eight directions distributed evenly around the circle. It should be clear from inspection that there is a fairly significant variation of frequency dependent reverberation time with direction.

3. A DIRECTIONAL DIFFUSE REVERBERATION MODEL

As can be seen in the analysis of the measured responses of the tunnel given above in Sec. 2, the defining characteristic of reverberation in an irregular rock tunnel is that no strong early reflections are

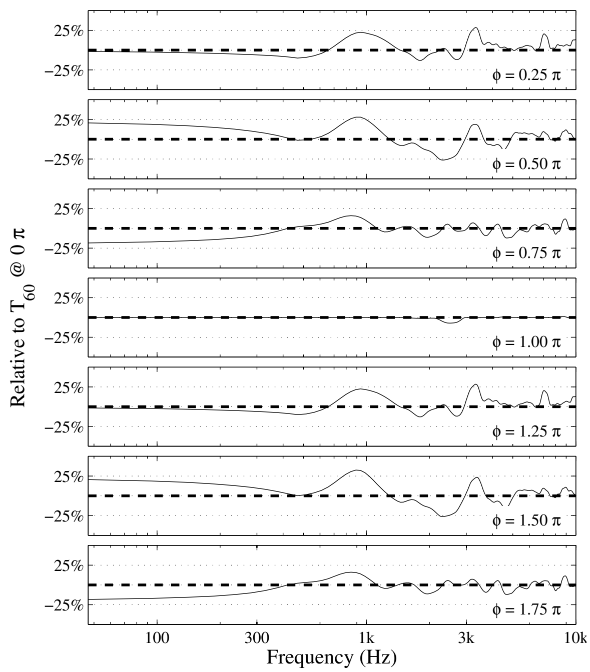

Fig. 3. Relative difference in T_{60} with respect to direction ϕ.

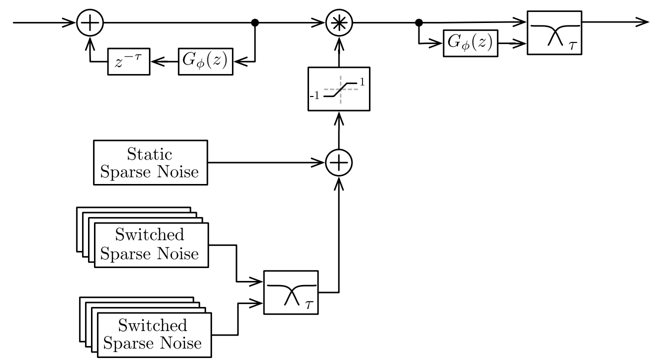

Fig. 4. Block diagram showing structure of a single channel of the novel reverberator.

present, only a diffuse tail. The density of echoes is also extremely high from the start of the response. The response is similar to that of a plate reverberator in these respects [17, 18, 19]. The other interesting aspect of the measured reverberation is the frequency-dependent variation of T_{60} with angle. A successful model should be able to reproduce this fast onset of diffuse sound, as well as being computationally efficient enough to be used in multi-channel or spatial audio context to reproduce the directional nature of the sound field.

The first task in the production of the model was to choose an appropriate modeling approach. Physical modeling of the space via FDTD or some form of acoustic ray-tracing is not ideal, due to the complex irregular geometry of the tunnel's rock walls. Convolution with a measured impulse response is an option, but is inflexible and computationally inefficient if we want to produce a model which accurately reproduces the frequency dependent T_{60} in multiple directions. Traditional algorithmic reverberation methods of the Schroeder [20], Moorer [21] or Gardner [22] type have difficulty producing the immediate onset of dense echoes necessary [1]. A Feedback Delay Network (FDN) [3] is reasonably appropriate, although again there can be difficulty in easily producing the immediate density of reflections necessary. The most appropriate solution would appear to be one based on convolution with a sparse noise sequence [23, 24]. This method has been shown to be more computationally efficient than a comparable FDN [24], and by its nature is entirely diffuse for the whole duration.

3.1. Single-channel reverberation structure

The basic reverberation structure is similar to those presented by Lee et al. [24], with some modifications. The core idea of the approach is that the diffuse tail of an acoustic response can be approximated by convolution with exponentially decaying white noise. However, direct convolution with exponentially decaying white noise has some drawbacks – namely that it is inefficient and lacks the possibility of varying reverberation time with frequency. Instead, we can use a comb filter to produce a series of exponentially decaying repeats of the input signal, and then convolve with a shorter sequence of noise to fill the gaps. The result is very close to the direct convolution, but computationally much less expensive. Frequency dependent reverberation time can be obtained by the addition of a damping filter into the feedback loop of the comb filter.

The efficiency of such a structure can be improved further by utilizing a form of sparse noise (noise that contains many zeros) instead of Gaussian white noise. Karjalainen and Järveläinen [23] propose one such type of noise, which they call 'velvet noise'. This particular sparse noise possesses the desirable quality of a constant, non-lumpy distribution in time. This quality manifests itself as an audible smoothness. This density of this form of noise may be greatly reduced (down to as low as several thousand impulses per second) whilst still being audibly equivalent to Gaussian noise [23, 25].

The disadvantage of this approach to reverberation is that depending on the exact structure of the reverberator, unwanted artifacts may be produced for certain types of signal. For example, if a static noise sequence is calculated once and used constantly, a periodic element is clearly audible in the response of the system to transients [23, 24]. Conversely, this periodicity is hidden during the systems response to steady-state signals, which contain no audible artifacts. To mitigate this periodicity it is necessary to vary the sparse noise sequence over time. The most basic form of this variation is to replace the noise sequence each time one cycle through the comb-filter has been completed. This modification removes the periodicity from the response to transient signals completely, but when a steady-state signal is applied heavy modulation of the sound is audible. Therefore, much of the difficulty in designing this type of reverberator is to find the correct form and type of variation of the noise sequence which strikes a balance between audible periodicity and unwanted modulation, and ideally makes both of these artifacts inaudible. Lee et al. [24] suggest several such types of variation. The method used in this work is a new hybrid of these methods, with some extension.

The other major challenge of this form of reverberation is modifying the sparse noise sequence so that continuity is maintained between the sections delineated by the echoes of the comb-filter. In the case of frequency-independent reverberation time, this is simply a case of providing the noise sequence with an exponentially decaying envelope dependent on the length and damping coefficient

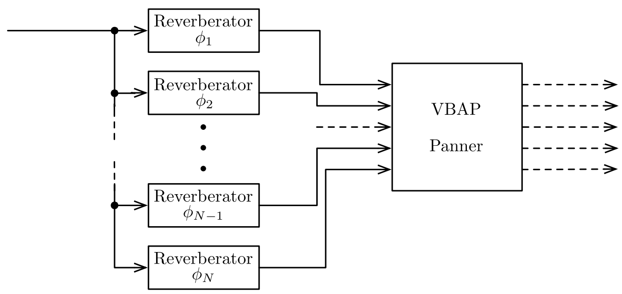

Fig. 5. Block diagram showing structure of complete multi-channel reverberator.

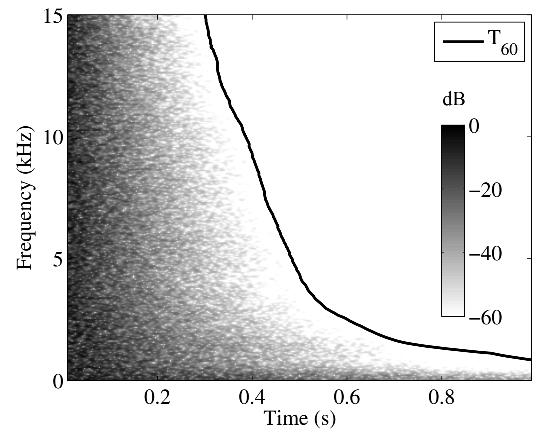

Fig. 6. Spectrogram of the impulse response of the novel reverberator and its T_{60} estimate. Compare with Fig. 2.

of the comb filter. In the case of frequency-dependent reverberation time, Lee et al. [24] suggest crossfading between an unfiltered noise-sequence and the same noise sequence filtered with the damping filter.

Figure 4 shows the proposed structure of the reverberator. It consists of a comb filter of length τ samples, containing a damping filter $G_\phi(z)$. This damping filter is an IIR filter derived from Yule-Walker approximation of the frequency-dependent T_{60} of the measured tunnel response for a particular direction. The output of this comb filter is convolved with a sparse noise sequence of length τ. This sparse noise sequence is a new hybrid of the non-overlapping sequence and overlapping sequence methods described by Lee et al. [24]. Three noise sequences are generated. Firstly, a static noise sequence of low density (500 impulses per second). Secondly, two denser noise sequences (2000 impulses per second) are generated, and crossfading is performed between them over each comb-filter period τ. At the end of a comb-filter period, the first noise-sequence (where the crossfading started) is replaced by the second noise sequence, which is then in turn replaced by a newly generated noise sequence. This process repeats for every comb-filter period τ. The crossfaded noise sequence is added to the static noise sequence, passed through a hard clipper to remove the occasional occurrences of co-incident (and hence double amplitude) impulses, and then convolved with the output of the comb-filter. This combination of static and varying sparse noise was chosen as it successfully suppressed artifacts whilst still being sparse enough to keep the convolution efficient. Finally, the convolved signal is crossfaded over each period τ with a version of itself filtered by the damping filter $G_\phi(z)$, in order to approximate the correct envelope. Experimentation showed that best results were obtained when the comb-filter period is $\tau \approx 30$ ms.

3.2. Extending the model for directional reverberation

To extend the reverberation structure to a directional model, we make the assumption that due to the highly diffuse nature of the reflections, the sound arriving from different directions is essentially uncorrelated. We can then approximate the directional response by the use of a number of separate reverberators, each of which is treated as a virtual source within the space, and spatialized according to some established system for distribution to a loudspeaker system. In this case, we use VBAP [12] to perform this spatialization. For headphone audio applications, a spatialization scheme making use of HRTFs could replace VBAP. Each reverberator ϕ_n has its damping filter G_{ϕ_n} derived from the directional T_{60} time measured in the same direction as the corresponding virtual source. For the purpose of this work we employed 8 reverberators and hence 8 virtual sources, distributed evenly at increments of $\frac{\pi}{4}$ in a plane around the listener. The dry signal is also treated as a virtual source, and placed at the front of the space, to be consistent with the position of the source during the measurements. It is possible that a smaller number of virtual sources and reverberators could be used without much loss of directional information due to the limited angular resolution of the B-Format microphone. Figure 5 shows this structure.

3.3. Results

Figure 6 shows spectrogram of the impulse response of the first reverberator (and hence first virtual source) in the model described above. The first reverberator is placed at zero angle (i.e. straight ahead). This response can be compared to the measured response given in Figure 2, which is the response to which the frequency dependent reverberation of this particular reverberator has been fitted. The diffuse nature of the reverberation seems to have been captured correctly, and the T_{60} is approximated well.

The model was tested with a variety of input sounds, on an 8.1 surround audio system in a listening room conforming to the ITU-R BS.1116 standard. The resulting sound was consistent with what we experienced in the tunnel during the measurement process, and compared favorably to the sound produced by a first-order ambisonic decomposition of the B-format impulse response. Sound samples are available for download.[1]

4. CONCLUSIONS

In this work we have presented measurements of the acoustic response of an excavated rock tunnel, and proposed a reverberation structure which can replicate these results. The reverberator could be applied in any situation in which the acoustic environment of a rock tunnel needs to be replicated. This could include game sound, film sound, and sound for simulators used for the training of machine operators. The reverberator is computationally efficient, and can additionally be applied in other applications where directional reverberation consisting of only diffuse sound is desired – for example as the late-reverberation portion of a more general reverberator.

[1] `http://www.acoustics.hut.fi/go/icassp13-caverev/`

5. REFERENCES

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

[2] M. Vorländer, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, Springer London, London, UK, 2008.

[3] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *90th AES Convention*, Paris, France, 1991.

[4] R. Stewart and D. Murphy, "A hybrid artificial reverberation algorithm," in *122nd AES Convention*, Vienna, Austria, 2007.

[5] A. Primavera, S. Cecchi, L. Romoli, P. Peretti, and F. Piazza, "An advanced implementation of a digital artificial reverberator," in *130th AES Convention*, London, UK, 2011.

[6] J.-M. Jot, L. Cerveau, and O. Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in *103rd AES Convention*, New York, NY, USA, 1997.

[7] M. Kuster, "Multichannel room impulse response rendering on the basis of underdetermined data," *Journal of the Audio Engineering Society*, vol. 57, no. 6, pp. 403–412, June 2009.

[8] M. A. Gerzon, "Recording concert hall acoustics for posterity," *Journal of the Audio Engineering Society*, vol. 23, no. 7, pp. 569–571, 1975.

[9] A. Farina and R. Ayalon, "Recording concert hall acoustics for posterity," in *Proc. 24th International Conference of AES: Multichannel Audio, The New Reality*, Banff, Alberta, Canada, 2003.

[10] M. A. Gerzon, "Ambisonics in multichannel broadcasting and video," *Journal of the Audio Engineering Society*, vol. 33, no. 11, pp. 859–871, 1985.

[11] J. Merimaa and V. Pulkki, "Spatial impulse response rendering I: Analysis and synthesis," *Journal of the Audio Engineering Society*, vol. 53, no. 12, pp. 1115–1127, 2005.

[12] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, Jun. 1997.

[13] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial decomposition method for room impulse responses," *Journal of the Audio Engineering Society*, vol. 61, no. 1/2, pp. 16–27, 2013.

[14] R. E. Collecchia, M. A. Kolar, and J. S. Abel, "A computational acoustic model of the coupled interior architecture of ancient Chavín," in *133th AES Convention*, San Francisco, CA, USA, Oct. 2012.

[15] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *108th AES Convention*, Paris, France, Feb. 2000.

[16] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 443–471, 2001.

[17] K. Arcas and A. Chaigne, "On the quality of plate reverberation," *Applied Acoustics*, vol. 71, no. 2, pp. 147–156, 2010.

[18] S. Bilbao, K. Arcas, and A. Chaigne, "A physical model for plate reverberation," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toulouse, France, May 2006, vol. 5, pp. 165–168.

[19] S. Bilbao, "A digital plate reverberation algorithm," *Journal of the Audio Engineering Society*, vol. 55, no. 3, pp. 135–144, Mar. 2007.

[20] M. R. Schroeder, "Natural-sounding artificial reverberation," *Journal of the Audio Engineering Society*, vol. 10, no. 3, pp. 219–223, July 1962.

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

[22] W. Gardner, "A realtime multichannel room simulator," in *Proc. 124th meeting of the Acoustical Society of America*, New Orleans, LA, USA, Nov. 1992.

[23] M. Karjalainen and H. Järveläinen, "Reverberation modeling using velvet noise," in *Proc. 30th International Conference of AES: Intelligent Audio Environments*, Saariselkä, Finland, Mar. 2007, pp. 1–9.

[24] K. S. Lee, J. S. Abel, V. Välimäki, T. Stilson, and D. P. Berners, "The switched convolution reverberator," *Journal of the Audio Engineering Society*, vol. 60, no. 4, pp. 227–236, Apr. 2012.

[25] V. Välimäki, H.-M. Lehtonen, and M. Takanen, "A perceptual study on velvet noise and its variants at different pulse densities," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, 2013.