OPEN ACCESS Freely available online
J. Fagerström, S. J. Schlecht, And V. Välimäki,
"Non-Exponential Reverberation Modeling Using Dark Velvet Noise,"
J. Audio Eng. Soc., vol. 72, no. 6, pp. 370–382 (2024 Jun.).
https://doi.org/10.17743/jaes.2022.0138.

Non-Exponential Reverberation Modeling Using Dark Velvet Noise

JON FAGERSTRÖM,[1,*] **SEBASTIAN J. SCHLECHT,**[1,2] *AES Member,* **AND VESA VÄLIMÄKI,**[1] *AES Fellow*
(jon.fagerstrom@aalto.fi) (sebastian.schlecht@aalto.fi) (vesa.valimaki@aalto.fi)

[1]*Acoustics Lab, Department of Information and Communications Engineering, Aalto University, Espoo, Finland*
[2]*Media Lab, Department of Art and Media, Aalto University, Espoo, Finland*

Previous research on late-reverberation modeling has mainly focused on exponentially decaying room impulse responses, whereas methods for accurately modeling non-exponential reverberation remain challenging. This paper extends the previously proposed basic dark-velvet-noise reverberation algorithm and proposes a parametrization scheme for modeling late reverberation with arbitrary temporal energy decay. Each pulse in the velvet-noise sequence is routed to a single dictionary filter that is selected from a set of filters based on weighted probabilities. The probabilities control the spectral evolution of the late-reverberation model and are optimized to fit a target impulse response via non-negative least-squares optimization. In this way, the frequency-dependent energy decay of a target late-reverberation impulse response can be fitted with mean and maximum reverberation-time errors of 4% and 8%, respectively, requiring about 50% less coloration filters than a previously proposed filtered-velvet-noise algorithm. Furthermore, the extended dark-velvet-noise reverberation algorithm allows the modeled impulse response to be gated, the frequency-dependent reverberation time to be modified, and the model's spectral evolution and broadband decay to be decoupled. The proposed method is suitable for the parametric late-reverberation synthesis of various acoustic environments, especially spaces that exhibit a non-exponential energy decay, motivating its use in musical audio and virtual reality.

0 INTRODUCTION

Artificial reverberation algorithms have been developed since the 1960s, starting with Schroeder's original algorithm [1, 2]. Schroeder's algorithm, as well as many that followed, are based on the assumption that the late reverberation part of a room impulse response (IR) can be modeled with exponentially decaying filtered white noise [2−5]. However, non-exponentially decaying reverberation can be observed in forests [6, 7], coupled rooms [8−10], and the famous gated reverb sound from the 1980s [11].

A room IR can be divided into three perceptually motivated parts, the direct sound, early reflections, and late reverberation [12]. The study presented in this paper focuses on the modeling of the late-reverberation part. In this paper, the authors propose a novel artificial reverberator capable of modeling target late reverberation with arbitrary energy decay and spectral evolution.

A notable branch of artificial reverberation algorithms is based on pseudo-random noise. Rubak and Johansen proposed using sparse random noise to model exponentially decaying Gaussian noise [4, 13]. However, the resulting sparse finite-impulse-response (FIR) filters placed inside a feedback loop would still need over 10,000 filter coefficients to produce a smooth reverberation IR, i.e. reverberation IR that does not sound rough. Later, Karjalainen and Järveläinen introduced velvet noise [5], which becomes smooth broadband noise with a pulse density of 1,500−2,000 pulses/s [5, 14]. To reduce computational costs, several velvet-noise−based algorithms employ a feedback structure, limiting them to generating only exponential decay [5, 15, 16].

It is important to make a distinction between the psychoacoustic quantity roughness, measured in asper [17], and the term "temporal roughness" used in this work to describe the perceived quality of sparse noise. The former is defined as the auditory sensation caused by amplitude-modulated pure tones, with modulation frequencies within the range of 15 to −300 Hz [17]. The latter is only loosely defined in previous literature as the sensation when a sparse noise sequence is not perceived as sounding smooth [5, 14]. As Meyer-Kahlen et al. [18] pointed out, the random assignment of the pulses of a sparse noise sequence can be interpreted as pseudo-random amplitude modulation.

*To whom correspondence should be addressed, email: jon.fagerstrom@aalto.fi, Last updated: April 11, 2024

The feedback delay network (FDN) [19] is a generalization of the comb-filter−based reverberator, which is still actively studied today [20−23]. Hybrid reverberators combining an FDN and velvet noise have also been proposed, which place the velvet-noise filters at the inputs and outputs [24] or within the feedback matrix of an FDN [22] to increase the echo density. However, in its basic form, the FDN can produce only an exponential decay. Combining two FDNs with different parameters allows for generating various non-exponential attenuation patterns, such as fade-in control or two-stage decay [25−27, 10]. An extended method based on the FDN has been proposed to synthesize double-slope decays of coupled rooms [23, 9]. However, no FDN-based method is capable of synthesizing reverberation, which has an arbitrary and non-exponential energy decay.

Karjalainen and Järveläinen proposed a modal reverberator structure for modeling late reverberation [28], and the idea was later refined by Abel et al. [29]. The modal reverberator is implemented with a parallel combination of mode filters, whose resonant frequencies and damping coefficients are tuned to match those of the target space. The number of modes to produce high-quality reverberation is suggested to be between 1,000 and 2,000 modes, based on informal listening. Wells recommends the use of a much larger number of modes [30]. The modal reverberators are best suited for exponentially decaying reverberation. However, implementing two-stage decays and fade-ins is also possible by tuning the damping of a portion of the mode filters [26]. Recently, modal synthesis has been proposed for the resynthesis of denoised anisotropic late reverberation with multi-slope decays by Hold et al. [31]. The multi-slope decays, however, are still a linear combination of exponential decays.

Holm et al. introduced an FIR-filter–based algorithm called the filtered-velvet-noise (FVN) reverberator [32] that Välimäki et al. refined later [33]. The FVN models a target late-reverberation IR with concatenated filtered-noise segments of different lengths. The filters are designed based on the time-frequency analysis of a target IR. The variable-length windowing mitigates audible transitions between the consecutive filters. The FIR-based FVN structure allows manipulating the IR by changing the lengths of the VN block, e.g., lengthening or shortening the decay time [33].

In this paper, the authors' previous work on the dark-velvet-noise (DVN) reverberator [34] is extended to fit its IR to a measured target late-reverberation IR. An extension to the original DVN algorithm is introduced, which replaces its recursive running-sum filters with arbitrary dictionary filters. Furthermore, the uniform probability for a pulse to be connected to a certain filter is set as a free parameter. The proposed method fits the extended DVN model to a target response via non-negative least-squares (NNLS) optimization [35], which is applied for the first time to model reverberation. The resulting model is parametric and facilitates various perceptually relevant modifications. Additionally, the authors investigate the temporal roughness properties of the proposed extended DVN and propose a scheme for mitigating it.

The proposed extended DVN model IR is subjected to an objective evaluation: First, the original target IRs are compared in terms of spectro-temporal fit and reverberation time (T60) estimation in the case where the target IR has an exponential energy decay, with IRs synthesized with optimized extended DVN model instances. It is demonstrated that the new method provides a good objective fit in parametrizing two distinctly different target reverberation IRs, those of a concert hall and outdoor space coupled to a cave opening. Next, the proposed method is compared to the previously proposed FVN method [33], and the FVN method is formulated as a special case of the proposed method.

The rest of this paper is organized as follows. SEC. 1 summarizes the previously proposed DVN algorithm [34]. SEC. 2 proposes the novel extension of the DVN structure. SEC. 3 discusses the proposed reverberation modeling scheme, step by step, including the NNLS optimization scheme. SEC. 4 presents an objective performance evaluation when using the proposed method for modeling a target IR and discusses the flexibility of the extended DVN in generating parametric modifications of the modeled IR. SEC. 5 concludes the paper.

1 DARK VELVET NOISE

This section provides the relevant background on the previously proposed DVN algorithm [34]. The basics of velvet noise and dark velvet noise serve as the basis for developing the proposed method for late-reverberation synthesis.

Original velvet noise [5] is a sparse pseudo-random noise, which consists of sparsely placed unit impulses with uniformly distributed signs. The main design parameter of a velvet-noise sequence is its pulse density ρ in pulses per second. Based on the desired pulse density, the grid size T_d can be computed as

$$T_\mathrm{d} = \frac{f_\mathrm{s}}{\rho}, \tag{1}$$

where f_s is the sample rate in hertz. Within each grid segment, a single unit impulse occurs.

Whereas the original velvet noise has a white magnitude spectrum [36], DVN is an extension of velvet noise that has a low-pass spectrum [34]. The low-pass spectrum is achieved by a random modulation of the pulse width w along the DVN sequence, i.e., the unit impulses of the original velvet noise are replaced by square pulses of varying width. In practice, the pulse-width modulation is implemented using a discrete set of recursive running-sum filters, one for each required pulse width [34]. The DVN sequence is given as

$$h(n) = \begin{cases} s(m) & \text{for } k(m) \le n < k(m) + w(m), \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where n is the sample index, m is the pulse index, $s(m)$ is the sign of the mth pulse, and $k(m)$ is the location of the mth pulse. When $w(m) \equiv 1$, Eq. (2) gives the original velvet noise sequence where each pulse is a unit impulse.

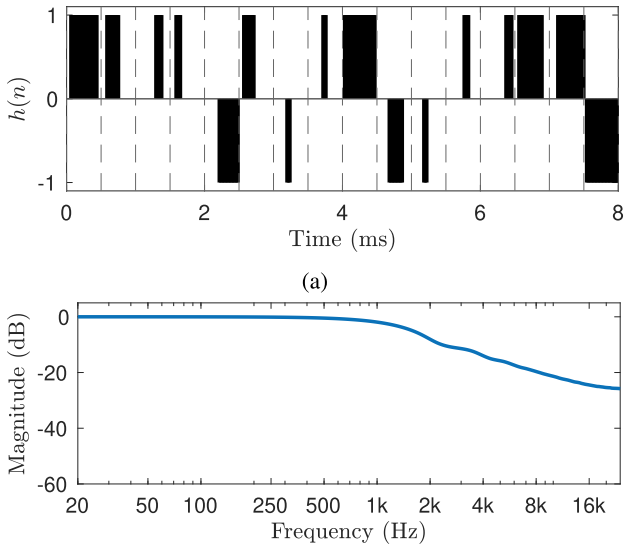


Fig. 1. (a) Example of a basic DVN sequence, and (b) its PSD normalized to 0 dB. The dashed lines show the grid spacing mT_d at the sample rate $f_\mathrm{s} = 48$ kHz.

The pulse locations of DVN are computed as

$$k(m) = \lfloor mT_\mathrm{d} + r_1(m)(T_\mathrm{d} - w(m)) \rceil, \tag{3}$$

where $\lfloor \cdot \rceil$ is the rounding operator and $r_1(m)$ is a uniform random number between 0 and 1. This formulation assures that the pulses do not overlap each other, given that $1 \leq w(m) \leq T_\mathrm{d}$.

The pulse widths of DVN (in samples) are computed as

$$w(m) = \lfloor r_2(m)(w_\mathrm{max} - w_\mathrm{min}) + w_\mathrm{min} \rceil, \tag{4}$$

where $r_2(m)$ is an uniform random number between 0 and 1, and w_max and w_min are the maximum and minimum pulse widths, respectively. The sign of each pulse is computed by [5]:

$$s(m) = 2\lfloor r_3(m) - 1 \rceil, \tag{5}$$

where $r_3(m)$ is a uniform random number between 0 and 1.

The first 8 ms ($f_\mathrm{s} = 48$ kHz[1]) of an example basic DVN sequence and the power spectral density (PSD) of the corresponding infinitely long DVN sequence are shown in Figs. 1(a) and Fig. 1(b), respectively.

2 EXTENDED DARK VELVET NOISE

In this section, an extension to the previously proposed DVN [34] is presented. The proposed extension replaces the previously used recursive running-sum filters with arbitrary dictionary filters, whose probabilities are set as a free parameter. Additionally, the temporal roughness of the proposed extension is discussed.

[1]The authors used the sample rate of $f_\mathrm{s} = 48$ kHz throughout this work.

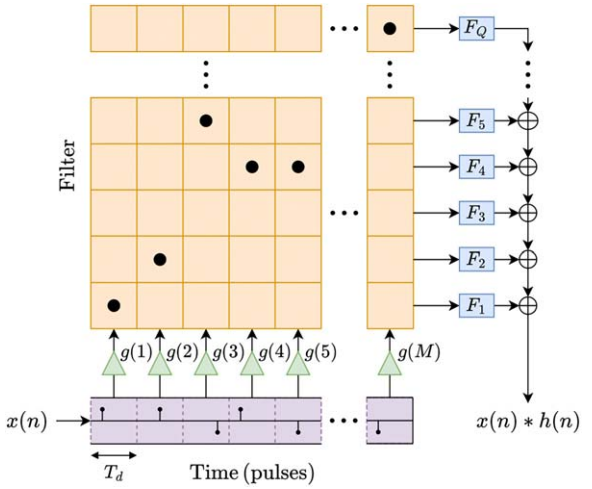


Fig. 2. Structure of the proposed extended DVN convolution with Q dictionary filters and M pulses. Each pulse is multiplied by a gain $g(m)$ and routed to the input of one dictionary filter, as indicated by the matrix. Each dot in the routing matrix represents a connection path.

2.1 Generalization of Dark Velvet Noise

As the first step of the generalization, the recursive running-sum filters are replaced with arbitrary dictionary filters $\boldsymbol{F}(z)$, allowing the generation of colored noise of desired spectral shape. Fig. 2 shows the block diagram of the proposed extended DVN convolution structure with Q dictionary filters $\boldsymbol{F}(z)$ and $M \gg Q$ pulses. The decay-envelope gains $g(m)$ parametrize the broadband energy decay. Finally, in contrast to the basic DVN, the uniform probabilities of each dictionary filter were set as a free parameter.

The filter probabilities for a single pulse are denoted by the vector

$$\boldsymbol{p} = \left[p_1, p_2, \ldots, p_Q\right]^\mathrm{T} \geq 0, \text{ with } \sum_{q=1}^{Q} p_q = 1, \tag{6}$$

where $[\,\cdot\,]^\mathrm{T}$ is the transpose operation.

A list of filter indices for each pulse is determined based on the pulse-filter probabilities $\boldsymbol{p}$ as

$$\phi(m) = f(\boldsymbol{p}(m)) \in \{1, 2, ..., Q\}, \tag{7}$$

where $f(\,\cdot\,)$ is any function that selects the pulse filter based on the probabilities $\boldsymbol{p}$. The resolved list of filter indices ϕ is visualized in matrix form in Fig. 2, where the one-hot column vectors show the filter selection for each pulse.

Eq. (2) can be reformulated so that the velvet-noise sequence is split into Q sub-sequences. Each sub-sequence includes the pulses routed to one of the dictionary filters, corresponding to a single row of the connection matrix of Fig. 2. The qth sub-sequence is then given as

$$v_q(n) = \begin{cases} s(m)g(m) & \text{for } n = k(m) \wedge \phi(m) = q, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

The transfer function of the extended DVN can now be written as

$$H(z) = \sum_{q=1}^{Q} V_q(z) F_q(z), \tag{9}$$

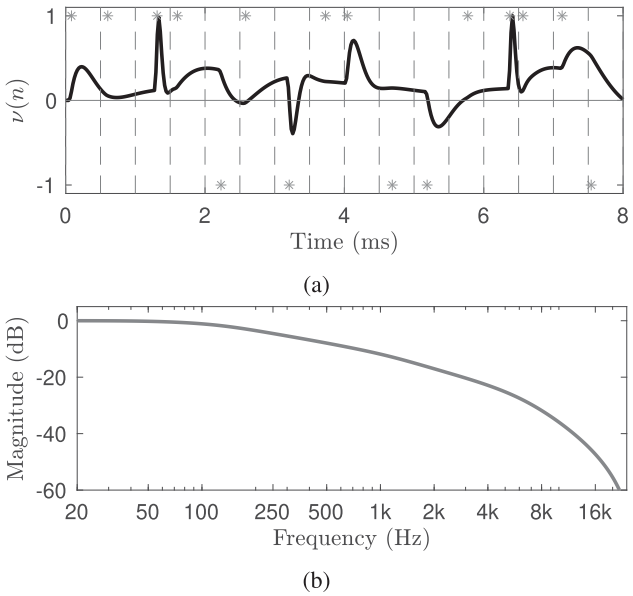


Fig. 3. (a) Extended DVN IR with the underlying pulse locations and signs indicated by asterisks. (b) The corresponding PSD, computed using Eq. (10) and normalized to 0 dB. The used dictionary filters correspond to the ones presented in Fig. 4(a).

where $V_q(z)$ is the velvet-noise subsequence routed to the qth dictionary filter $F_q(z)$. The time-dependent PSD of the extended DVN sequence is described by the weighted mean magnitude response of the dictionary filters:

$$|H(m,\omega)| = \sum_{q=1}^{Q} F_q(e^{j\omega}) p_q(m), \tag{10}$$

where $F_q(e^{j\omega})$ is the magnitude response of the qth dictionary filter, and the probabilities p_q are defined in Eq. (6). The PSD of the extended DVN sequence, Eq. (10), is independent of the pulse density ρ, and the equation holds true as the occurrences of dictionary filters are uncorrelated due to the randomized selection of the filter and placement of pulses.

Figs. 3(a) and 3(b) show an example of an extended DVN sequence and its PSD, respectively. The pulse locations are the same as in the basic DVN sequence in Fig. 1(a). Note that the pulses are no longer located within the grid segments of the basic DVN and are smeared in time.

2.2 Mitigating Temporal Roughness

In this work, and in previous studies on velvet noise [5, 14], temporal roughness describes the perceived quality of sparse noise. For the original velvet noise, which consists of randomly placed unit impulses, the perceived temporal roughness is simply inversely proportional to its pulse density ρ. Based on informal listening experiments the basic DVN method, which uses the recursive running-sum filters, shows similar behavior in terms of temporal roughness [34].

In this work, however, the authors used high-order filters with a potentially steep spectral decay. In combination with the naive implementation of the pulse-filter selection [see Eq. (7)], this could lead to perceivable temporal roughness. In particular, roughness issues may originate from the random switching between filters that have large energy differences within certain frequency bands, as seen in the spectrogram of Fig. 4(b).

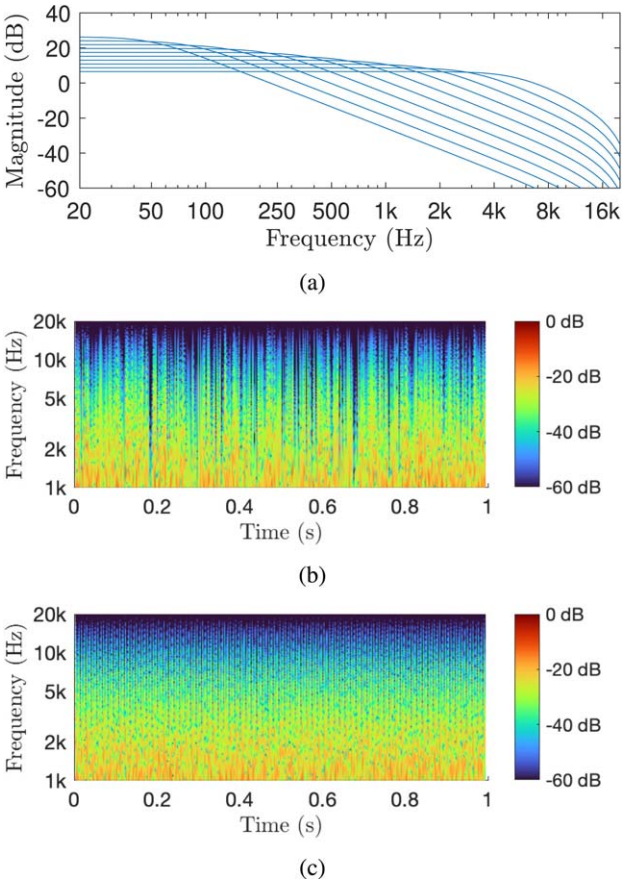


Fig. 4. (a) Magnitude responses of the example second-order dictionary filters $F_q(z)$ with $q = 1, 2, ..., 10$, and the spectrograms of the resulting extended DVN sequence with (b) a totally random and (c) greedy filter assignment.

To mitigate issues related to temporal roughness, a two-step solution is proposed. First, the filter energies are normalized to minimize the temporal roughness, by ensuring the broadband filter energy does not fluctuate when switching between different filters. Next, a greedy filter assignment is proposed. The naive pulse filter selection is based on a weighted uniform random number and is given as

$$\phi(m) = \arg\max_q \{r_q(m) p_q(m)\}, \tag{11}$$

where $r_q(m)$ is a uniform random number in the range [0, 1].

The following greedy filter assignment is proposed instead:

$$\phi(m) = \arg\max_q \{(\tau_q(m) + \epsilon r_q(m)) p_q(m)\}, \tag{12}$$

where ϵ is a free parameter controlling the amount of randomization and τ_q is a sample index (i.e., time) when the qth dictionary filter was last selected. The value τ_q is up-

dated sequentially based on the previously selected pulse filter with

$$\tau_q(m+1) = \begin{cases} 0 & \text{for } \phi(m) = q, \\ \tau_q(m) + 1 & \text{otherwise.} \end{cases} \quad (13)$$

When the qth dictionary filter stays inactive, its weight τ_q grows, thus making it more likely for the greedy assignment to pick that filter.

The roughness problem is best visualized by synthesizing stationary noise with uniform probabilities $p_q \equiv 1/Q$, i.e., a pulse routing to any of the defined dictionary filters with uniform probability. Fig. 4(a) shows the magnitude responses of an example set of ten second-order dictionary filters. However, the spectrogram in Fig. 4(b) highlights the frequency-dependent roughness problem that remains after the normalization when the naive pulse filter selection of Eq. (11) is used. The spectrograms were computed using Fast Fourier Transform of length 2,048, with a 256-sample Hann window with 50% overlap. The frequency axis in Fig. 4(b) was limited from 1 to 20 kHz for better visualization of the sparsity, which creates audible temporal roughness in the noise sequence. The naive filter assignment is based on uniform probabilities, and thus, the resulting sub-sequences defined by Eq. (8) resemble totally random noise, which was shown to sound rougher than velvet noise [5, 14].

Fig. 4(c) shows the spectrogram of the extended DVN sequence generated based on the same uniform probabilities, but now applying the greedy filter assignment of Eq. (12) instead of the random assignment. The resulting extended DVN sequence is visually and audibly much smoother. However, there is a trade-off since the greedy assignment creates some periodicity in the sequence. Note that the PSD of the sequences in Figs. 4(b) and 4(c) is the same. The noise sequences in Figs. 4(b) and 4(c) can be listened to on the companion web page of this paper.[2]

3 REVERBERATION MODELING

The proposed framework for parametrizing a target late-reverberation IR using the novel extended DVN structure is presented in this section. The previously proposed FVN method is shown to be a special case of the extended DVN algorithm.

3.1 Preprocessing and Analysis

To parameterize a late-reverberation IR with the extended DVN algorithm, the short-time Fourier transform (STFT) was used to obtain a time-frequency representation of the target IR. The late-reverberation IR of the Promenadi Hall [37], Pori, Finland, serves as an example target IR throughout this section. The T60 of the target IR, evaluated at octave bands between 125 and 8,000 Hz, ranges from 2.7 s at the lowest band to 1.2 s at the highest band [37].

In the following, the number of STFT time frames is denoted with T. A Fast Fourier Transform length of 2,048 was used. The design choice in the analysis step is to choose a suitable frame size and overlap amount to capture the temporal and spectral changes in sufficient detail. The choice may vary depending on the characteristics of the target IR. Typically, larger rooms with longer decay times will tend to have smooth exponentially decaying time envelopes for which a shorter window size yields good results. Small rooms or outdoor spaces with short reverb time tend to have a more detailed envelope that heavily affects the timbre of the IR, thus requiring a denser framing for the analysis to yield a more detailed decay envelope for the extended DVN model.

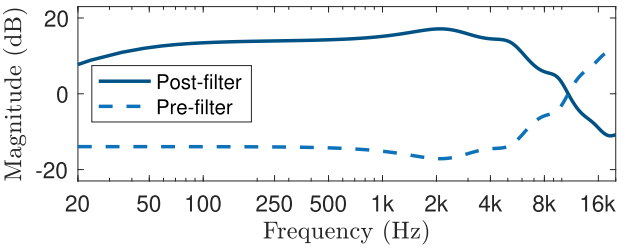


Fig. 5. Example magnitude responses of the pre-filter and complementary post-filter of an extended DVN model instance tuned to the Promenadi Hall IR.

As a preprocessing step, linear prediction (LP) was applied on the first time frame of the STFT representation of the target IR. The solution of the LP gives the allpole post-filter $1/R(z)$ coefficients. The coefficients of the LP filters were obtained using MATLAB's `lpc` function. A similar LP modeling approach was applied previously by Holm et al. [32]. A filter order of 10 was used in this case for LP modeling since this was found to be sufficient for capturing the low-pass characteristic of late reverberation in a previous study [33].

However, the low-order LP filter cannot model any possible low-frequency rolloff present in the target IR. Thus, in this work, an additional first-order DC-blocker filter [38] was fitted to the target IR to improve the model's spectral accuracy at low frequencies. The inverse filter of the allpole filter, i.e., the pre-filter, was applied to the whole target IR to whiten it. The whitening ensures that the target IR starts with a flat magnitude spectrum. Thus, the post-filter takes care of matching the initial coloration of the target IR, whereas the dictionary filters concentrate on modeling the time-dependent relative spectral change.

The magnitude response of the post-filter is shown in Fig. 5 (solid line). The response is the combined response of the tenth-order allpole filter and first-order DC-blocker filter. In this example, the effect of the DC blocker is visible as a slight cut below 100 Hz. The corresponding pre-filter magnitude response of the target IR is shown as a dashed line in Fig. 5.

3.2 Dictionary-Filter Design

The main design question was to define the dictionary filters to be used in the extended DVN structure shown in Fig. 2. In this work, the dictionary filters were obtained directly from the analysis stage, reminiscent of the FVN algo-

[2]https://github.com/Ion3rik/dark-velvet-noise-reverb.

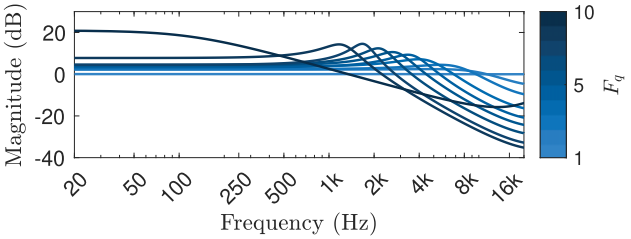


Fig. 6. Normalized magnitude responses of the dictionary filters used for modeling the late-reverberation IR of the Promenadi Hall. Here, the dictionary consists of ten second-order allpole filters extracted from the analyzed target reverberation. The filter energies are normalized to one.

rithm [33], by approximating each analysis frame response with a second-order allpole filter. The coefficients of the second-order allpole filter were estimated using an LP. The single tenth-order pre-filter allows applying second-order filters in the dictionary instead of the tenth-order filters used in the previously proposed FVN method [33].

Applying a subset of the analyzed filters that were logarithmically spaced in time was found to yield good results. Due to the logarithmic spacing, closely spaced frame filters are picked from the beginning of the analyzed IR, and more sparsely spaced filters toward the end. The justification for the logarithmic spacing was inspired by the results of Välimäki et al. [33]. They observed that the magnitude spectrum of late reverberation typically varies rapidly at the beginning of the IR and slower toward the end of the IR [33].

Fig. 6 shows example magnitude responses of the set of dictionary filters used to model the Promenadi Hall late-reverberation IR. In this example, the target IR was analyzed in 50 frames, and the logarithmically spaced subset of ten dictionary filters included the filters 1, 2, 3, 5, 7, 10, 15, 23, 34, and 50. The filter energies were normalized to one. Note that the first dictionary filter in Fig. 6, which is estimated from the first frame of the pre-whitened target IR, has a practically flat magnitude response and can be omitted in the implementation without causing an audible effect.

3.3 Solving the Filter Probabilities

After designing the dictionary filters, their magnitude responses need to be fit to the target magnitude response by solving an NNLS problem. NNLS is a constrained version of the least-squares optimization problem, which is a convex problem with linear constraints [35].

The dictionary filter magnitude responses are denoted by the matrix $|\boldsymbol{F}(e^{j\omega})|$, where the columns contain the magnitude response of each dictionary filter at discrete frequencies. The NNLS solution yields the activation vector $\boldsymbol{z}$. The values of coefficients $\boldsymbol{z}$ are constrained to be positive or zero. The optimization problem has the form

$$\min_{\boldsymbol{z}} \left\| \left|\mathbf{F}(e^{j\omega})\right| \boldsymbol{z} - |\boldsymbol{h}(\omega_k)| \right\|_2^2, \text{ subject to } \boldsymbol{z} \geq 0, \tag{14}$$

where ω_k are the discrete frequencies and $|\boldsymbol{h}(\omega_k)|$ is the target magnitude response. In this work, MATLAB's `lsqnonneg` function was used to solve the optimization

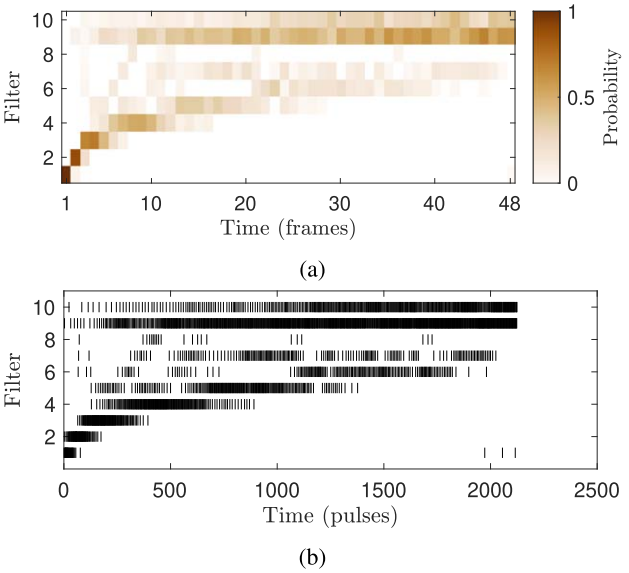


Fig. 7. (a) Probability matrix $\boldsymbol{P}$ and (b) the corresponding resolved pulse filters for modeling the Promenadi Hall late-reverberation IR.

problem. To compose the activation matrix for the whole time-frequency representation of the target IR, the optimization problem, cf. Eq. (14), was solved for each time step separately. The activation matrix has the form

$$\boldsymbol{Z} = \big[z(1), z(2), \ldots, z(T)\big], \tag{15}$$

where $z(t)$ are the column vectors containing the solved activations for each analysis frame t.

Due to the formalization of the proposed extended DVN structure, the solved activation matrix $\boldsymbol{Z}$ was normalized to obtain the probability of each dictionary filter for each time step. The normalization gains are computed as

$$\gamma(t) = ||z(t)||_1 = \sum_{i=1}^{Q} |z(t)|. \tag{16}$$

The probability matrix is then given as

$$\boldsymbol{P} = \boldsymbol{Z} \odot \frac{1}{\boldsymbol{\gamma}}, \tag{17}$$

where $\odot$ is the Hadamard product and $\boldsymbol{\gamma} = [\gamma(1), \gamma(2), \ldots, \gamma(T)]^{\mathrm{T}}$. After the normalization, each column vector $\boldsymbol{p}(t)$ of the probability matrix fulfills Eq. (6).

Fig. 7(a) shows the probability matrix $\boldsymbol{P}$ obtained via normalization of the activation matrix solution for the target IR. Dark areas indicate the most probable filter routings for each analysis time frame. The general trend shows that the higher probabilities shift from the first dictionary filter toward the last dictionary filter.

3.4 Sparse Synthesis with Dark Velvet Noise

The first step of the synthesis is to generate a velvet-noise sequence that has the same length as the target IR with the desired pulse density ρ. In this work, the authors used a time-dependent density starting from $\rho = 2{,}000$ pulses/s

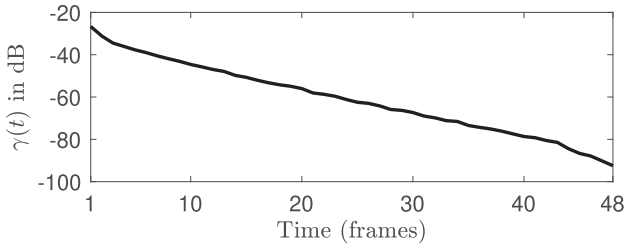


Fig. 8. Normalization gains $\gamma(t)$ per frame.

and linearly decreasing toward $\rho = 500$. The decreasing density allows maintaining a low computational load of the algorithm without introducing temporal roughness to the reverberation [33]. The pulse locations and signs of the sequence were computed using Eqs. (3) and (5), respectively.

Conveniently, the normalization gains $\gamma(t)$ defined in Eq. (16) describe the broadband energy decay envelope of the target IR. Fig. 8 shows the normalization gains $\gamma(t)$ corresponding to the probability matrix $\boldsymbol{P}$ of Fig. 7. On a logarithmic scale, the normalization gains $\gamma(t)$ decrease approximately linearly over time, except for their beginning and end—a trend that corresponds to the exponential decay of the target IR.

Because the synthesis requires a discrete number of pulses, the probability matrix $\boldsymbol{P}$ and the normalization gains $\boldsymbol{\gamma}$ were interpolated to match the number of pulses M. The pre-computed pulse locations k were used as query points, and linear interpolation was applied. The interpolated normalization gains are denoted by $\widehat{\boldsymbol{\gamma}}$. The per-pulse decay gains of the extended DVN convolution structure shown in Fig. 2 are obtained as

$$g(m) = \widehat{\gamma}(m)\sqrt{T_\mathrm{d}(m)}, \tag{18}$$

where the multiplication with the square root of the grid size $T_\mathrm{d}(m)$ compensates for the change in energy due to the time-dependent density.

Finally, the pulse filters are resolved using Eq. (12), based on the interpolated probability matrix. Fig. 7(b) shows the resolved pulse filters computed for the probability matrix of Fig. 7(a). High concentrations of pulses [see Fig. 7(b)] are likely routed to filters with high probabilities [see Fig. 7(a)], whereas areas of low probability, i.e., the areas of lighter color, show sparser patterns of filter selection.

3.5 Comparison to Filtered Velvet Noise

The reverberation modeling framework presented above resembles the previously proposed FVN algorithm [32, 33]. In this section, a comparison between the two methods is provided.

In the FVN algorithm, the target IR is segmented into non-overlapping frames of different lengths, and an allpole LP filter is extracted from each segment [33]. In the extended DVN method, the extracted LP filters can be used as both the target and dictionary responses. In this special case, the solution of the NNLS problem returns a diagonal probability matrix, since the best fit is obtained by simply activating the extracted segment filters in order.

Fig. 9(a) shows the first ten segments of the probability matrix $\boldsymbol{P}$ obtained using the matched dictionary filters extracted from the target IR. Fig. 9(b) shows the corresponding filter assignment for each pulse when the filters are assigned using the per-frame probabilities and the resulting filter list is interpolated to yield a filter routing for each pulse. In this configuration, the extended DVN essentially implements the FVN algorithm, where the filter is switched at the segment boundaries. However, the constant segment length results in periodic switching of the dictionary filter, which can cause audible disturbance in the synthesized IR [33]. The switching periodicity was mitigated in the FVN algorithm by using a non-uniform segment length.

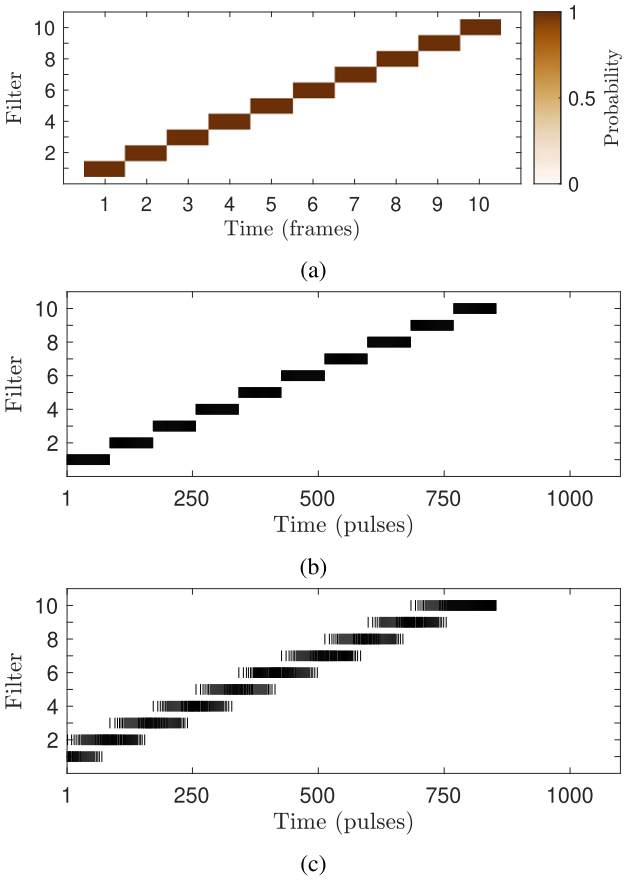


Fig. 9. First ten frames of the (a) probability matrix $\boldsymbol{P}$ and corresponding resolved pulse filters (b) with filter interpolation, resembling uniformly segmented FVN, and (c) with probability interpolation, unique to the extended DVN. The presented dictionary filters are the LP allpole target filters of each frame.

The extended DVN presents another option for the filter assignment, where the probability matrix is first interpolated, and the filters are then assigned per pulse based on the interpolated probability matrix. Fig. 9(c) shows the corresponding filter assignment for each pulse. In this case, the switching is not discrete at each segment boundary. Instead, a smoother mix is obtained between the consecutive filters. Informal listening experiments suggested that the mixing mitigates the problem of periodic disturbances even when using uniform segmentation. Sound examples of the noise sequences shown in Fig. 9(b) and 9(c) are available online.[2]

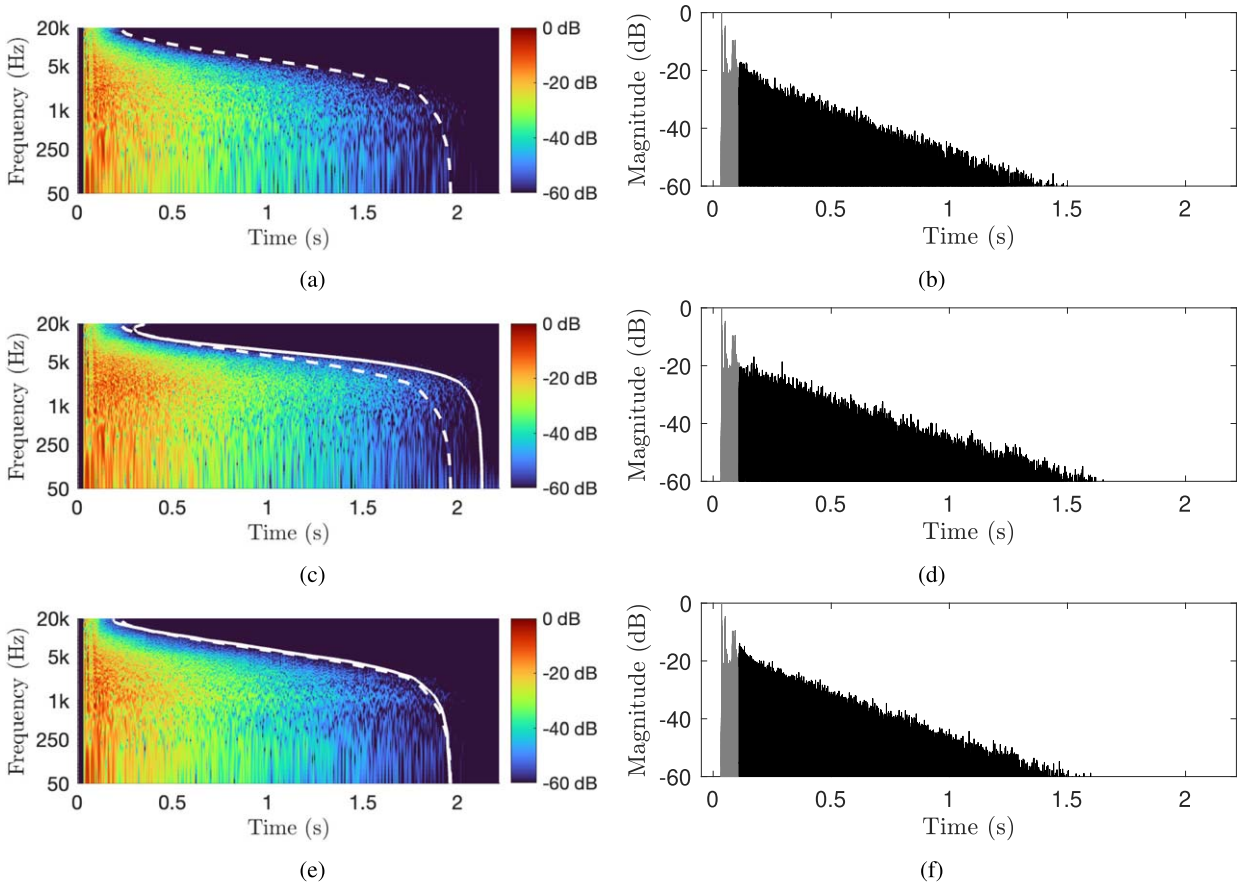


Fig. 10. The left-column sub-figures show the spectrograms of the Promenadi Hall and their modeled counterparts: (a) target IR and estimated IRs based on (c) the FVN model and (e) extended DVN model. The T60 estimates of the target IR (dashed) and modeled IRs (solid) are overlayed on the spectrograms, indicating a better fit using the proposed method. The right-column sub-figures present (b) the target IR, (d) FVN model instance IR, and (f) extended DVN model instance IR. The direct part and early reflections of the IR, both of which are not modeled, are shown in gray.

4 EVALUATION

In this section, an objective evaluation of the extended DVN reverberation algorithm is presented. The method was applied to model the late-reverberation part of two target IRs from different spaces. The early IR parts containing the direct sound and early reflections were not modeled but directly adopted from the original data. The two spaces have distinctive acoustical characteristics and were selected to provide challenging conditions to test the generalizability of the proposed reverberation algorithm.

4.1 Target Impulse Responses

The first of the two target IRs is the high-quality measurement of the IR of the Promenadi concert hall in Pori, Finland, conducted by Merimaa et al. [37]. The late part of the IR was modeled after 110 ms as a test case to allow a direct comparison with the previously proposed advanced FVN model [33]. Välimäki et al. [33] determined the late reverberation of the Pori Hall IR to start after 110 ms based on preliminary testing. The second target IR has been recorded in Creswell Crags [39], where the IR has a strong second echo. The specific IRs the authors chose from the two datasets are "s1_r3_o.wav" and "8_r_grundymouth_s_grundypath.wav."

4.2 Modeling Concert-Hall Reverberation

In this section, the authors compare IRs synthesized with their extended DVN and the previously proposed FVN [33] objectively to the target IR of the Promenadi Hall. The objective accuracy was analyzed in terms of spectrogram comparison and T60 estimation. The extended DVN model of the Promenadi Hall was parametrized using the configuration described in Sec. 3.

Figs. 10(b) and 10(f) show the target IR of the Promenadi Hall and the IR synthesized with an optimized instance of the extended DVN model, respectively. In Fig. 10, the early parts of the IR are shown in gray and the late reverberation in black. The target IR envelope is visually highly similar to the envelope of the extended DVN model instance IR. The temporal resolution of the extended DVN model envelope is determined by the STFT window length, which was selected to be 85 ms with 50% overlap. Lengthening the STFT window would result in a smoother envelope. The previously proposed advanced FVN [33] model instance IR, shown in Fig. 10(d), has a slightly longer decay compared to the target IR envelope.

The estimated T60 curves are overlayed on the spectrograms in Fig. 10, where the solid line shows the T60 of the models and the dashed line shows the target T60. The mean and maximum T60 errors measured in the frequency

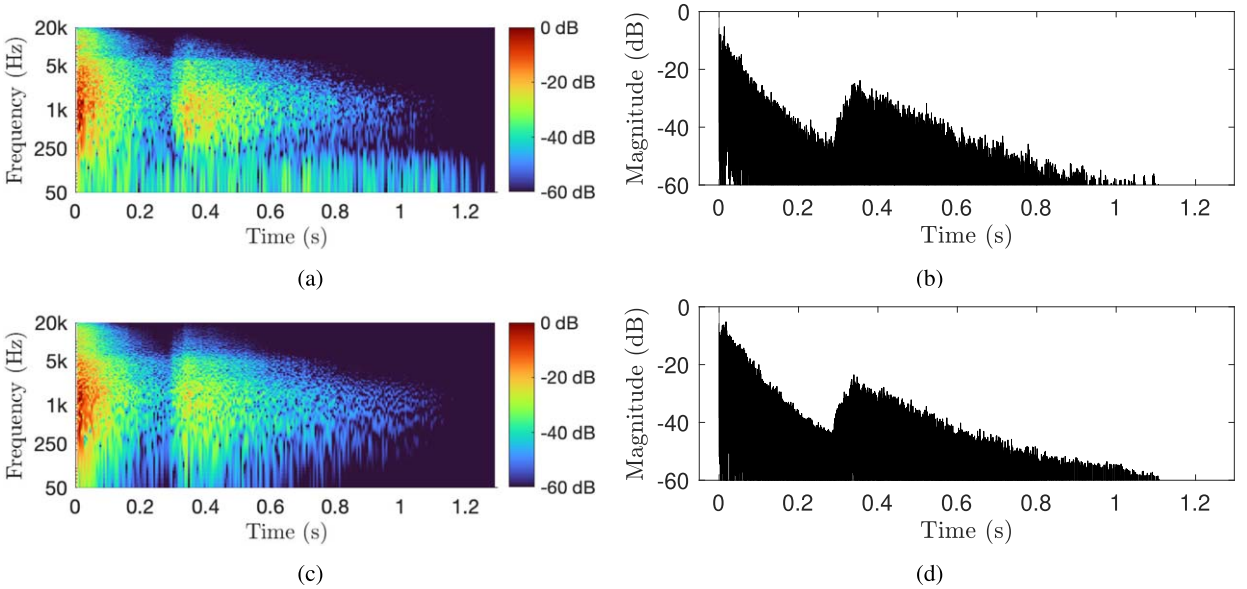


Fig. 11. Spectrogram of the Creswell Crags (a) target IR and (c) the extended DVN model. Energy (in decibels) of the Cresswell Crags (b) target IR and (d) the corresponding extended DVN model. The late and early reflections, after the direct sound (1 ms), are modeled.

range of 20–16,000 Hz of the FVN model instance are 14% and 28%, respectively. The extended DVN model instance in Fig. 10(e) follows the frequency-dependent decay of the target IR well across all frequencies with mean and maximum T60 errors of 4% and 8%, respectively.

Overall, the extended DVN model instance achieved a better spectro-temporal fit than the FVN model instance using only ten dictionary filters, whereas the advanced FVN model uses 20 coloration filters. This implies that the proposed method is both more accurate and more efficient in modeling a late-reverberation IR than the previously proposed FVN method.

4.3 Modeling Non-Exponential Reverberation

The authors evaluate objectively the accuracy of the synthesized IR of the extended DVN model instance in modeling the Cresswell Crags target IR. The target IR and its spectrogram are shown in Figs. 11(b) and 11(a), respectively. There is no other parametric method that can successfully model target IRs containing a single echo that results in two exponential energy decays. The analysis presented in this section is intended to demonstrate the variable capabilities of the proposed method.

The extended DVN model uses ten second-order allpole dictionary filters, a single 12th-order post-filter, and two cascaded first-order DC blockers. The order of the post-filter components was increased to obtain a better spectral fit to the target IR. The analysis methods are identical to the ones used in SEC. 4.2, except for the T60 analysis, which is not a meaningful metric for the target IR in question. For this example, the extended DVN was used to synthesize the entire target IR except for the direct sound (first 1 ms), since the target IR does not show any prominent early reflections.

Fig. 4(c) shows the spectrogram of the extended DVN model instance IR. The DVN model provides an accurate approximation of the overall shape of the target spectrogram. However, a slightly longer decay is seen in Fig. 11(c) around 1 kHz. The largest difference between the target and model instance spectrograms is seen at low frequencies, at 100 Hz and just above it, where the DC-blocker filters out the low-frequency noise visible in the target spectrogram of Fig. 11(a).

The IR of the extended DVN model instance in Fig. 11(d) now features a more detailed energy-decay envelope compared to that of the concert hall reverberation DVN model instance in Fig. 10(f). This is due to the smaller frame length (5.3 ms) with 50% overlap for the STFT analysis. The smaller frame length and hop size are beneficial in modeling the more complex double-stage decay of the Crasswell Crags target reverberation. The envelope shape of the extended DVN model instance in Fig. 10(d) follows well that of the target IR in Fig. 11(b). In summary, this design example demonstrates the ability of the proposed method to model a non-exponential IR.

4.4 Modification of the Proposed Model

One of the benefits of the parametric extended DVN is its flexibility concerning modification possibilities of the synthesized IR. The direct modification of IRs used in convolution reverberation has been proposed by Canfield-Dafilou and Abel [40] to change the perceived room size of the original IR. Modifications such as time stretching and temporal envelope modification have been demonstrated already using the previously proposed FVN algorithm [32, 33].

The extended DVN lends itself just as well to time stretching and envelope modifications. Furthermore, the achievable level of detail in envelope modifications with the extended DVN algorithm is vaster than with the previously proposed FVN algorithm, since the former method relies on pulses scaled by individual gains. The choice of

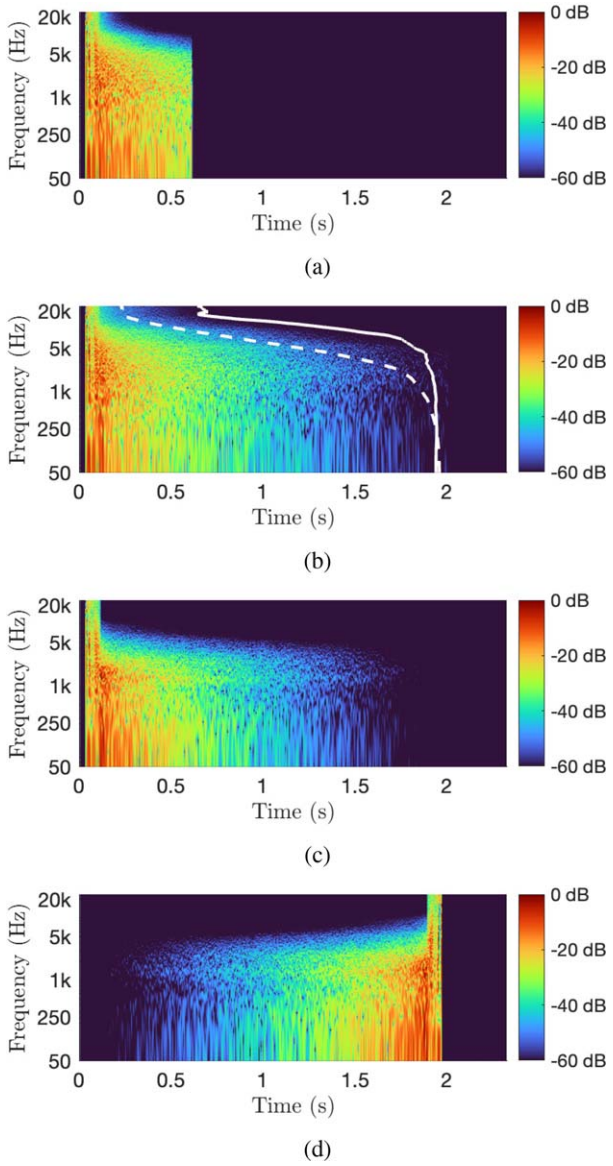


Fig. 12. Spectrograms of the modified Promenadi Hall extended DVN model instance with (a) gated reverb, (b) modified T60 [compared to the original in Fig. 10(e)], (c) reversed spectral evolution, and (d) reversed decay. The modifications are applied only to the late part, which is modeled, except for the reversed decay, where the early parts are also reversed and placed after the modeled part. The T60 estimates of the original target IR (dashed) and modified model IR (solid) are overlayed on top of the spectrogram in (b).

envelope resolution is a trade-off between computational cost and envelope detail. Time stretching with the extended DVN is implemented by changing the length of the generated velvet-noise sequence and interpolating the probability matrix $\boldsymbol{P}$ to match the modified pulse locations. In the following examples, three more modifications achievable with the extended DVN algorithm are highlighted.

Fig. 12 shows the spectrograms of three modifications of the Promenadi Hall extended DVN model instance. Fig. 12(a) illustrates the effect of the gated reverb modification made to the extended DVN model instance IR. Since the extended DVN is implemented as an FIR structure, generating a gated reverb is achieved simply by truncating part of the delay-line coefficients from the end of the model instance IR.

In Fig. 12(b), the authors slowed down the spectral change relative to the original model instance of Fig. 10(e). With the proposed method, the relative spectral change of the late reverberation can be adjusted by manipulating the probability matrix $\boldsymbol{P}$, cf. Eq. (17). The slowing down of the decay was achieved by taking a subset of columns from the beginning of the original probability matrix, i.e., the matrix

$$\boldsymbol{P}_\alpha = \left[\, p(1) \quad p(2) \quad \ldots p(\hat{T}) \,\right], \tag{19}$$

where $\hat{T} = \lfloor T\alpha \rceil$ and $0 \leq \alpha \leq 1$. The new matrix $\boldsymbol{P}_\alpha$ was then interpolated to fit the original pulse locations. In Fig. 12(b), the resulting T60 estimate is overlayed on the spectrogram, where the dashed line shows the original target T60 and the solid line shows the modified T60, where the high frequencies ring considerably longer than in the target.

The final modification example is the time-reversed reverb effect, which has two versions. Since the spectral change is parametrized by the probability matrix $\boldsymbol{P}$ and the energy decay with $g(m)$ of Eq. (18), they are decoupled. In Fig. 12(c), the authors have time reversed the filter list ϕ, cf. Eq. (7), of the extended DVN model instance. The resulting IR retains its overall energy decay, but the spectral change is now from dark to bright. Implementing the opposite version of the two flip operations is also a possibility. Flipping the decay envelope while using the original filter list results in an IR whose energy rises while retaining the original spectral change, as shown in Fig. 12(d). Sound examples of the modifications presented in Fig. 12 are provided together with the MATLAB application that can recreate them.[2]

The brightening effect and gated effect can be implemented also with recursive algorithms given that an accurate model is first obtained. However, the gated reverb requires either a second parallel recursive algorithm to be run to cancel the remaining part of the IR [25] or, alternatively, an additional noise gate to mute the output of the reverberator. On the other hand, the reversed energy decay is impossible to implement in a recursive manner, since it requires growth of energy over time, which would result in instability in a recursive implementation.

5 CONCLUSION

In this paper, an extension of the previously proposed DVN algorithm was presented. A parametric reverberator was developed based on the extended DVN structure utilizing sparse velvet-noise convolution. By replacing the square pulses of the previously proposed DVN method with arbitrary dictionary filters and setting the probability of each dictionary filter as a free parameter, the extended DVN model can be fitted to a target IR via NNLS optimization. The authors additionally showed that a previously proposed

FVN reverberation algorithm [33] is a special case of the proposed extended DVN reverberator.

The authors evaluated the proposed method objectively and demonstrated its capability to accurately model the late reverberation of two distinctly different spaces, a large concert hall and a coupled space in which sound decays in a non-exponential manner. No previously proposed parametric reverberation algorithm exists for modeling the latter target IR. The authors assessed the spectro-temporal fit of optimized instances of the proposed extended DVN model to the two target IRs objectively, and the proposed method produced an accurate model in both cases. For the concert hall target IR, the extended DVN yielded a better spectro-temporal fit than the previously proposed FVN algorithm, when using half the number of filters. The method's capability was demonstrated by providing practically relevant examples of model IR modification, such as slowing down the decay in a frequency-dependent way and gating the model IR.

All in all, the proposed extended DVN reverberator is applicable to synthesizing late reverberation of various spaces while providing perceptually meaningful parametric control. Future work could investigate a more general approach to specify a broad set of dictionary filters and find a sparse solution, instead of using the empirically designed filters.

6 ACKNOWLEDGMENT

This research is part of the activities of the Nordic Sound and Music Computing Network (NordicSMC; NordForsk project no. 86892).

7 REFERENCES

[1] M. R. Schroeder, "Natural Sounding Artificial Reverberation," *J. Audio Eng. Soc.*, vol. 10, no. 3, pp. 219–223 (1962 Jul.).

[2] 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 (2012 Jul.). https://doi.org/10.1109/TASL.2012.2189567.

[3] J. A. Moorer, "About This Reverberation Business," *Comput. Music J.*, vol. 3, no. 2, pp. 13–28 (1979 Jun.). https://doi.org/10.2307/3680280.

[4] P. Rubak and L. G. Johansen, "Artificial Reverberation Based on a Pseudo-Random Impulse Response: Part I," presented at the *104th Convention of the Audio Engineering Society* (1998 May), paper 4725.

[5] M. Karjalainen and H. Järveläinen, "Reverberation Modeling Using Velvet Noise," in *Proceedings of the 30th AES International Conference on Intelligent Audio* (Saariselkä, Finland) (2007 Mar.), paper 7.

[6] K. Spratt and J. S. Abel, "A Digital Reverberator Modeled After the Scattering of Acoustic Waves by Trees in a Forrest," presented at the *125th Convention of the Audio Engineering Society* (2008 Oct.), paper 7650.

[7] F. Stevens, D. T. Murphy, L. Savioja, and V. Välimäki, "Modeling Sparsely Reflecting Outdoor Acoustic Scenes Using the Waveguide Web," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 8, pp. 1566–1578 (2017 Aug.). https://doi.org/10.1109/TASLP.2017.2699424.

[8] C. F. Eyring, "Reverberation Time Measurements in Coupled Rooms," *J. Acoust. Soc. Am.*, vol. 3, no. 2, pp. 181–206 (1931 Oct.).

[9] O. Das and J. S. Abel, "Grouped Feedback Delay Networks for Modeling of Coupled Spaces," *J. Audio Eng. Soc.*, vol. 69, no. 7/8, pp. 486–496 (2021 Jul.). https://doi.org/10.17743/jaes.2021.0026.

[10] C. Kirsch, T. Wendt, S. Van De Par, H. Hu, and S. D. Ewert, "Computationally-Efficient Simulation of Late Reverberation for Inhomogeneous Boundary Conditions and Coupled Rooms," *J. Audio Eng. Soc.*, vol. 71, no. 4, pp. 186–201 (2023 Apr.). https://doi.org/10.17743/jaes.2022.0053.

[11] R. Fink, M. Latour, and Z. Wallmark (Eds.), *The Relentless Pursuit of Tone: Timbre in Popular Music* (Oxford University Press, New York, NY, 2018).

[12] T. Rossing, R. Moore, and P. Wheeler, "Auditorium Acoustics," in *The Science of Sound*, pp. 525–545 (Addison-Wesley, London, UK, 2002), 3rd ed.

[13] P. Rubak and L. G. Johansen, "Artificial Reverberation Based on a Pseudo-Random Impulse Response: Part II," presented at the *106th Convention of the Audio Engineering Society* (1999 May), paper 4900.

[14] V. Välimäki, H.-M. Lehtonen, and M. Takanen, "A Perceptual Study on Velvet Noise and Its Variants at Different Pulse Densities," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 21, no. 7, pp. 1481–1488 (2013 Jul.). https://doi.org/10.1109/TASL.2013.2255281.

[15] K. Lee, J. Abel, V. Välimäki, T. Stilson, and D. P. Berners, "The Switched Convolution Reverberator," *J. Audio Eng. Soc.*, vol. 60, no. 4, pp. 227–236 (2012 Apr.).

[16] V. Välimäki and K. Prawda, "Late-Reverberation Synthesis Using Interleaved Velvet-Noise Sequences," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 29, pp. 1149–1160 (2021 Feb.). https://doi.org/10.1109/TASLP.2021.3060165.

[17] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models* (Springer, Berlin, Germany, 2007), 3rd ed.

[18] N. Meyer-Kahlen, S. J. Schlecht, and T. Lokki, "Perceptual Roughness of Spatially Assigned Sparse Noise for Rendering Reverberation," *J. Acoust. Soc. Am.*, vol. 150, no. 5, pp. 3521–3531 (2021 Nov.). https://doi.org/10.1121/10.0007048.

[19] J.-M. Jot, "An Analysis/Synthesis Approach to Real-Time Artificial Reverberation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 221–224 (San Francisco, CA) (1992 Mar.). https://doi.org/10.1109/ICASSP.1992.226080.

[20] B. Alary, A. Politis, S. Schlecht, and V. Välimäki, "Directional Feedback Delay Network," *J. Audio Eng. Soc.*, vol. 67, no. 10, pp. 752–762 (2019 Oct.). https://doi.org/10.17743/jaes.2019.0026.

[21] K. Prawda, V. Välimäki, and S. J. Schlecht, "Improved Reverberation Time Control for Feedback Delay Networks," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pp. 1–8 (Birmingham, UK) (2019 Sep.).

[22] S. J. Schlecht and E. A. P. Habets, "Scattering in Feedback Delay Networks," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 28, pp. 1915–1924 (2020 Jun.). https://doi.org/10.1109/TASLP.2020.3001395.

[23] O. Das, J. S. Abel, and E. K. Canfield-Dafilou, "Delay Network Architectures for Room and Coupled Space Modeling," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pp. 234–241 (Vienna, Austria) (2020 Sep.).

[24] J. Fagerström, B. Alary, S. J. Schlecht, and V. Välimäki, "Velvet-Noise Feedback Delay Network," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pp. 219–226 (Vienna, Austria) (2020 Sep.).

[25] E. Piirilä, T. Lokki, and V. Välimäki, "Digital Signal Processing Techniques for Non-Exponentially Decaying Reverberation," in *Proceedings of the COST-G6 Workshop on Digital Audio Effects*, pp. 21–24 (Barcelona, Spain) (1998 Nov.).

[26] K.-S. Lee and J. S. Abel, "A Reverberator With Two-Stage Decay and Onset Time Controls," presented at the *129th Convention of the Audio Engineering Society* (2010 Nov.), paper 10287.

[27] N. Meyer-Kahlen, S. J. Schlecht, and T. Lokki, "Fade-In Control for Feedback Delay Networks," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pp. 227–233 (Vienna, Austria) (2020 Sep.).

[28] M. Karjalainen and H. Järveläinen, "More About This Reverberation Science: Perceptually Good Late Reverberation," presented at the *111th Convention of the Audio Engineering Society* (2001 Sep.), paper 5415.

[29] J. S. Abel, S. Coffin, and K. Spratt, "A Modal Architecture for Artificial Reverberation With Application to Room Acoustics Modeling," presented at the *137th Convention of the Audio Engineering Society* (2014 Oct.), paper 9208.

[30] J. J. Wells, "Modal Decompositions of Impulse Responses for Parametric Interaction," *J. Audio Eng. Soc.*, vol. 69, no. 7/8, pp. 530–541 (2021 Jul.). https://doi.org/10.17743/jaes.2021.0027.

[31] C. Hold, T. McKenzie, G. Götz, S. J. Schlecht, and V. Pulkki, "Resynthesis of Spatial Room Impulse Response Tails With Anisotropic Multi-Slope Decays," *J. Audio Eng. Soc.*, vol. 70, no. 6, pp. 526–538 (2022 Jun.) https://doi.org/10.17743/jaes.2022.0017.

[32] B. Holm-Rasmussen, H.-M. Lehtonen, and V. Välimäki, "A New Reverberator Based on Variable Sparsity Convolution," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pp. 344–350 (Maynooth, Ireland) (2013 Sep.).

[33] V. Välimäki, B. Holm-Rasmussen, B. Alary, and H.-M. Lehtonen, "Late Reverberation Synthesis Using Filtered Velvet Noise," *Appl. Sci.*, vol. 7, no. 5, paper 483 (2017 May). https://doi.org/10.3390/app7050483.

[34] J. Fagerström, N. Meyer-Kahlen, S. J. Schlecht, and V. Välimäki, "Dark Velvet Noise," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pp. 192–199 (Vienna, Austria) (2022 Sep.).

[35] C. L. Lawson and R. J. Hanson, "Linear Least Squares With Linear Inequality Constraints," in *Solving Least-Squares Problems*, pp. 158–173 (Prentice Hall, Upper Saddle River, NJ, 1974).

[36] N. Meyer-Kahlen, S. J. Schlecht, and V. Välimäki, "Colours of Velvet Noise," *Electron. Lett.*, vol. 58, no. 12, pp. 495–497 (2022 Jun.). https://doi.org/10.1049/ell2.12501.

[37] J. Merimaa, T. Peltonen, and T. Lokki, "Concert Hall Impulse Responses—Pori, Finland," http://legacy.spa.aalto.fi/projects/poririrs/ (accessed Jun. 1, 2023).

[38] J. Pekonen and V. Välimäki, "Filter-Based Alias Reduction for Digital Classical Waveform Synthesis," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 133–136 (Las Vegas, NV) (2008 May). https://doi.org/10.1109/ICASSP.2008.4517564.

[39] D. T. Murphy and S. Shelley, "OpenAIR: An Interactive Auralization Web Resource and Database," presented at the *129th Convention of the Audio Engineering Society* (2010 Nov.), paper 8226.

[40] E. K. Canfield-Dafilou and J. S. Abel, "Resizing Rooms in Convolution, Delay Network, and Modal Reverberators," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pp. 229–236 (Aveiro, Portugal) (2018 Sep.).

THE AUTHORS


Jon Fagerström


Sebastian J. Schlecht


Vesa Välimäki

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

•

Sebastian J. Schlecht is a Professor of Practice for Sound in Virtual Reality at the Acoustics Lab, Department of Information and Communications Engineering and Media Labs, Department of Art and Media, of Aalto University, Finland. He received the Diploma in Applied Mathematics from the University of Trier, Germany, in 2010 and an M.Sc. degree in Digital Music Processing from the School of Electronic Engineering and Computer Science at Queen Mary University of London, UK, in 2011. In 2017, he received a doctoral degree at the International Audio Laboratories Erlangen, Germany, on artificial spatial reverberation and reverberation enhancement systems. From 2012 to 2019, Dr. Schlecht was also external research and development consultant and lead developer of the 3D Reverb algorithm at the Fraunhofer IIS, Erlangen, Germany.

•

Vesa Välimäki received his D.Sc. degree in electrical engineering from the Helsinki University of Technology, Espoo, Finland, in 1995. In 1996, he was a Postdoctoral Researcher at the University of Westminster, London, UK. In 2001–2002, he was a Professor of signal processing at the Pori unit of the Tampere University of Technology, Finland. In 2008–2009, he was a Visiting Scholar at Stanford University. He is a Full Professor of audio signal processing and the Vice Dean for Research in electrical engineering at Aalto University. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE). In 2015–2020, he was a Senior Area Editor of the *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. Since 2020, Prof. Välimäki has been the Editor-in-Chief of the *Journal of the Audio Engineering Society*.