Fagerström, Jon; Meyer-Kahlen, Nils; Schlecht, Sebastian J.; Välimäki, Vesa

Dark Velvet Noise

Published in:
Proceedings of the 25th International Conference on Digital Audio Effects (DAFx20in22)

Published: 01/01/2022

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Fagerström, J., Meyer-Kahlen, N., Schlecht, S. J., & Välimäki, V. (2022). Dark Velvet Noise. In G. Evangelista, & N. Holighaus (Eds.), *Proceedings of the 25th International Conference on Digital Audio Effects (DAFx20in22)* (2022 ed., pp. 192-199). Article 31 (Proceedings of the International Conference on Digital Audio Effects). DAFx . https://dafx2020.mdw.ac.at/proceedings/papers/DAFx20in22_paper_31.pdf

DARK VELVET NOISE

Jon Fagerström[1], *Nils Meyer-Kahlen*[1], *Sebastian J. Schlecht*[1,2] *and Vesa Välimäki*[1]

[1]Acoustics Lab, Dept. of Signal Processing and Acoustics
[2]Media Lab, Dept. of Art and Media
Aalto University
Espoo, Finland
jon.fagerstrom@aalto.fi

ABSTRACT

This paper proposes dark velvet noise (DVN) as an extension of the original velvet noise with a lowpass spectrum. The lowpass spectrum is achieved by allowing each pulse in the sparse sequence to have a randomized pulse width. The cutoff frequency is controlled by the density of the sequence. The modulated pulse-width can be implemented efficiently utilizing a discrete set of recursive running-sum filters, one for each unique pulse width. DVN may be used in reverberation algorithms. Typical room reverberation has a frequency-dependent decay, where the high frequencies decay faster than the low ones. A similar effect is achieved by lowering the density and increasing the pulse-width of DVN in time, thereby making the DVN suitable for artificial reverberation.

1. INTRODUCTION

Velvet noise is a sparse, ternary pseudo-random sequence [1, 2]. It was originally proposed to model late reverberation, which resembles exponentially decaying filtered Gaussian noise [3]. This paper introduces a new type of sparse lowpass noise, that can also be used in artificial reverberation algorithms.

Different schemes to obtain the random impulse locations have been investigated, defining multiple variants of Original Velvet Noise (OVN) [2]. These include Totally Random Noise (TRN) [4, 5], which historically precedes OVN, and later additive random noise (ARN), Extended Velvet Noise (EVN), and Random Integer Noise (RIN) [2]. Of these alternatives, the OVN sequence has been found to be superior in terms of temporal smoothness compared to all the variants [2]. Another variant is frequency-domain velvet noise, in which the frequency bins are obtained using a similar scheme as used for obtaining the random impulse locations of OVN [6, 7].

Several different reverberation algorithms utilizing velvet noise have been proposed in the past [8, 9, 10, 11, 12]. Other applications include speech synthesis [13], where modified velvet noise has been used to generate unvoiced speech sounds. An extrapolation method utilizing velvet noise has also been proposed [14, 15]. The method convolves velvet noise with a stationary audio signal to create potentially endless sustain for the signal [14, 15]. Recently, a real-time implementation of the endless sustain technique was presented [16]. Short velvet-noise sequences have been used for efficient audio decorrelation [17, 18] as well as for humanizing sampling synthesis of percussive sounds [19].

While traditional velvet noise is white [20], colored spectra have been achieved via modifications in the sign and pulse-location distributions [21, 13, 20]. Werner [21] discovered crushed velvet noise, which has a highpass spectrum resembling that of low-shelf filtered white noise. The highpass characteristic is achieved by controlling the sign probability, i.e., the ratio of positive and negative impulses. At its extreme, a crushed-noise sequence consists of only non-negative sample values, 0 and 1, resulting in sparse binary noise. The density of crushed noise sequence controls the cutoff frequency, and the sign probability controls the amount of low-frequency attenuation [21, 20]. Similarly, other spectral characteristics are achievable by designing the gap distribution of a crushed additive random noise (CARN) sequence [20]. Morise [13] proposed another mechanism to obtain velvet noise with a highpass spectrum, by making sure the sign distribution is exactly balanced within each short segment of a longer sequence.

The previous colored velvet-noise sequences have shown mostly highpass characteristics, whereas lowpass characteristics offer more practical benefits in reverberation algorithms, which typically include some form of time-dependent lowpass filtering. The design of a sparse lowpass sequence was shown to be possible in theory [20]. However, the example provided therein does not warrant an efficient implementation. This paper proposes a practical algorithm for generating velvet noise with a lowpass spectrum, which we call Dark Velvet Noise (DVN) as opposed to the previous spectrally bright, highpass sequences [21, 13]. By its specific structure, DVN can be implemented efficiently with one multi-tap-delay line and a limited set of Recursive Running-Sum (RRS) [22, 23] blocks. This makes DVN ideal for efficient generation of late reverberation.

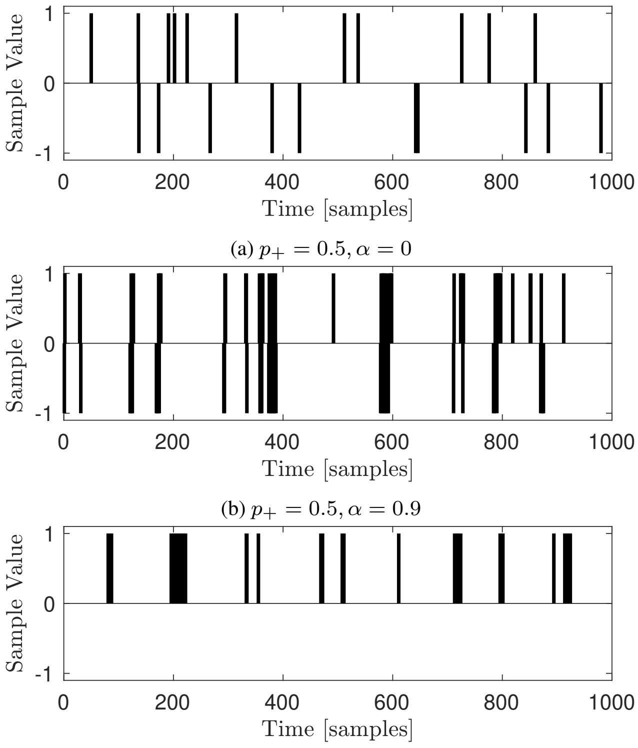
The rest of this paper is organized as follows. Section 2 gives a background on ternary sparse sequences and provides one option for creating such sequences with high-frequency attenuation. Section 3 introduces the novel DVN sequence, discusses its spectral and temporal characteristics, and proposes an efficient implementation for time-domain convolution with DVN. Section 4 gives an example of time-dependent DVN and its application to artificial reverberation. Section 5 discusses the strengths and drawbacks of DVN. Section 6 concludes the paper.

2. BACKGROUND

2.1. White Ternary-Noise Sequences

Sparse ternary sequences can be defined as

$$h(n) = \begin{cases} s(m) & \text{for } n = k(m), \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$



(a) $p_+ = 0.5, \alpha = 0$

(b) $p_+ = 0.5, \alpha = 0.9$

(c) $p_+ = 1, \alpha = 0.9$

Figure 1: *Crushed additive random noise with inflated uniform updates (see* α *in* (7)*), biased sign probability* p_+ *(see* (5)*), and pulse density* $\rho = 800$*. Only the non-zero samples are shown.*

where n denotes the time index in samples, $m \in \mathbb{Z}$ is the impulse index, $k(m) \in \mathbb{Z}$ determines the pulse locations in ascending order, and $s(m) = \pm 1$ is the sign of the pulse, created from an independent random process. Regardless of the process, we can define the mean spacing between pulses as

$$T_\mathrm{d} = \mathbb{E}[k(m) - k(m-1)], \tag{2}$$

where $\mathbb{E}[\cdot]$ is the expected value and denote pulse density $\rho = \frac{f_\mathrm{s}}{T_\mathrm{d}}$ in pulses per second.

Different ternary sequences are created by generating the location of the non-zero pulses with different processes [2]. One very general class of algorithms for determining the locations of the peaks is additive random sampling [24], where a random variable is drawn to determine the gap size from one non-zero peak to the next:

$$k(m) = k(m-1) + \Delta(m). \tag{3}$$

A common choice for the distribution of updates has been the uniform distribution [2], where the sequence is called ARN. An example of an ARN sequence with uniform updates is shown in Fig. 1a. All sequence generation processes in which the size of consecutive gaps is independent, can be expressed as ARN.

The OVN, however, which will be used below, is a ternary sparse sequence that cannot be generated through additive sampling. Instead, it is created by jittered sampling, where pulses are randomly displaced from a regular grid, as in

$$k(m) = \lfloor mT_\mathrm{d} + r(m)(T_\mathrm{d} - 1)\rceil, \tag{4}$$

where $r(m)$ is a uniform random number between 0 and 1 and $\lfloor\cdot\rceil$ denotes rounding to the closest integer.

Interestingly, all ternary sequences whose signs occur with equal probability converge to have a flat power spectral density (PSD) for an infinitely long sequence [20], no matter the generation process. Note that, for short sequences, deviations from the white spectrum occur [19], as is typical for all random noises.

2.2. Crushed Additive Random Noise

One option to create colored sequences is to employ crushed velvet noise, whose probability for the positive and negative signs $s(m)$ occurring is unequal. The probability to select a positive sign for a location of a non-zero sample is denoted as p_+. Werner [21] first observed that by modifying the sign probabilities of OVN, or ARN with uniform updates, a highpass character can be achieved. Morise independently used a similar principle to produce higpass filtered velvet noise [13].

Meyer-Kahlen *et al.* [20] recently showed that different choices of the index-update process (3), i.e., the ARN with an unequal sign probability, offer a much wider range of colored sequences. When the additive index update probability density function (PDF) and the sign probability are known, the PSD of the resulting sequence can be written analytically as

$$H(\omega) = \frac{1}{T_\mathrm{d}}\left(1 + Q(\omega)(1 - 2p_+)^2\right) \tag{5}$$

with

$$Q(\omega) = 2\Re\left(\frac{P_\Delta(\omega)}{1 - P_\Delta(\omega)}\right), \tag{6}$$

where P_Δ is the discrete-time Fourier transform (DTFT) of the additive update PDF p_Δ, i.e., its characteristic function, and $\Re$ denotes the real part of a complex number.

One option to create lowpass behavior is to "inflate" the uniform updates such that a proportion of pulses $0 \leq \alpha < 1$ are located next to the previous pulse so that there is no gap between them (see Fig. 1b). The additive index update can be expressed as

$$p_\Delta[k] = \begin{cases} \frac{1-\alpha}{2\mu_\alpha - 3} & \text{for } 2 \leq k \leq 2\mu_\alpha - 2, \\ \alpha & \text{for } k = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where the parameter μ_α is the mean gap size of all gaps that are larger than one. The overall mean gap size relates to the new parameters as $T_\mathrm{d} = \alpha + (1-\alpha)\mu_\alpha$.

By computing (7), (6), and (5), the resulting PSDs of such a sequence with different sign probabilities are shown in Fig. 2. Lowpass characteristics with the steepest slopes are achieved when setting the sign probability p_+ to one (or equivalently to zero). In this case, the sequence consists of randomly placed blocks of non-zero samples (see Fig. 1c).

Although the nice feature of colored-noise generation with CARN is that many different colors can be achieved and the PSD can be determined easily, it is not well suited for an efficient implementation. The main problem is that the size of the non-zero regions is unbounded, and the length of the gaps are independent. As is shown below, the former is inconvenient from a computational point of view, and the latter leads to a sequence that can be perceived as less smooth than OVN. Therefore, we propose a better way to produce lowpass velvet noise in the following.

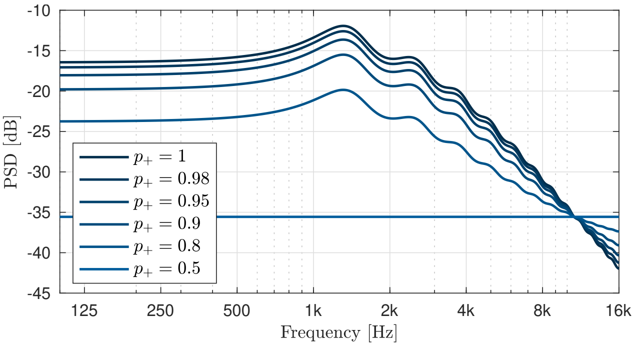


Figure 2: *PSD of CARN with lowpass behavior, generated using inflated uniform updates, $\rho = 800$, $\alpha = 0.9$, $f_s = 48$ kHz.*

3. DARK VELVET NOISE

This section, introduces a novel method for generating velvet noise with a lowpass spectrum. The lowpass spectrum is achieved by increasing the pulse width. We call this new type of sparse lowpass noise Dark Velvet Noise. As opposed to CARN sequences with lowpass behavior, DVN suits itself to an efficient implementation.

3.1. Generating DVN

Although an OVN sequence can have only one unit impulse within each grid segment, in a DVN sequence, we allow a single rectangular pulse of random length to occur in every grid segment. This can be conceptualized as randomized pulse-width modulation. The pulse width should be constrained between 1 and T_d to avoid any pulse exceeding the grid boundaries. The first eight pulses of a DVN sequence are shown in Fig. 3a and the corresponding magnitude spectrum of the full 1-s-long sequence in Fig. 3b.

The pulse width of the mth pulse in a DVN sequence is defined by

$$w(m) = \lfloor r_1(m)(w_\text{max} - w_\text{min}) + w_\text{min} \rceil, \tag{8}$$

where $m = 0, 1, 2, ..., M - 1$ is the pulse index, $r_1(m)$ is a uniform random number between 0 and 1, and w_max and w_min are the maximum and minimum pulse widths, respectively. The pulse location of DVN indicates the time index of the rising edge of each pulse and is defined by

$$k(m) = \lfloor mT_\text{d} + r_2(m)(T_\text{d} - w(m) \rceil, \tag{9}$$

where $r_2(m)$ is a uniform random number between 0 and 1. The sign of each pulse is

$$s(m) = 2 \lfloor r_3(m) \rceil - 1, \tag{10}$$

where $r_3(m)$ is a uniform random number between 0 and 1. The DVN sequence is then given as

$$h(n) = \begin{cases} s(m) & \text{for } k(m) \leq n < k(m) + w(m), \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

Note that for pulse width $w(m) \equiv 1$ the DVN sequence in (11) is equivalent to OVN.

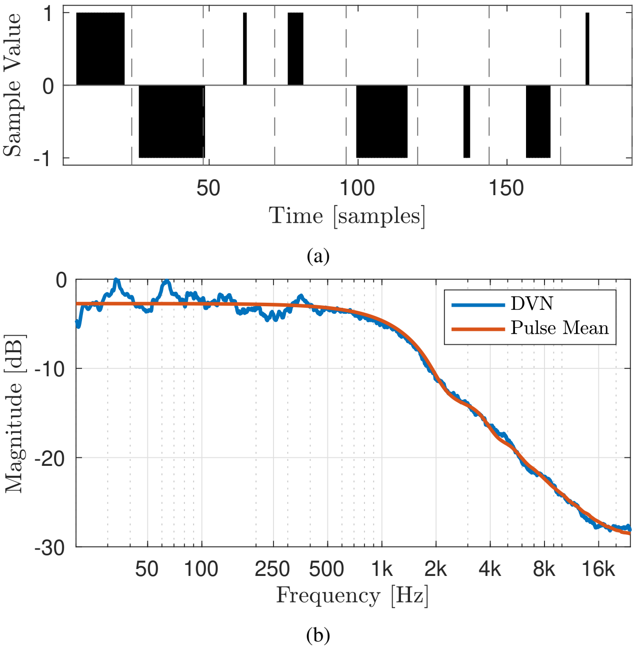


Figure 3: *(a) Beginning of a 1-s-long DVN sequence ($\rho = 2000$), (b) its third-octave smoothed magnitude spectrum, and the mean magnitude spectrum of the rectangular pulses used in the sequence.*

3.2. Spectral Shape

The spectral shape of DVN is determined by the rectangular pulses contained in the DVN sequence. Fig. 3a presents an example DVN sequence. Only the beginning of the sequence is shown for better visualization. Fig. 3b shows the third-octave-smoothed magnitude spectrum of the sequence (blue) and the corresponding mean pulse spectrum (red). The roll-off of the DVN spectrum is approximately 6 dB/octave, which resembles the roll-off of a first-order lowpass filter.

As seen in Fig. 3b, the spectral shape of the DVN follows closely the slope of the mean pulse spectrum. In the example sequence, each pulse width $w(m)$ is randomly drawn between $w_\text{min} = 1$ and $w_\text{max} = T_\text{d} = 24$. The corresponding mean pulse spectrum is computed as the mean spectrum of rectangular pulses in the same range. The mean pulse spectrum is given by

$$|H(\omega)| = \frac{\sum_{i=w_\text{min}}^{w_\text{max}} |W_i(e^{\imath\omega})|}{w_\text{max} - w_\text{min} + 1}, \tag{12}$$

where $|W_i(e^{\imath\omega})|$ is the magnitude spectrum of the rectangular pulse with pulse width $w = i$. Note that the spectrum is independent of the grid size T_d. Since the different pulse widths occur at random times with equal probability for each pulse width, they are uncorrelated, and thus the mean spectrum describes the overall spectrum of the full sequence.

Fig. 4 shows the magnitude spectrum of five different DVNs with varying pulse densities. The maximum pulse width is again set equal to the grid size, i.e., $w_\text{max} = T_\text{d}$, to avoid any overlapping pulses, and the set of pulse widths for each density is $w \in \{1, 2, ..., w_\text{max}\}$. Lower density allows wider pulses to fit

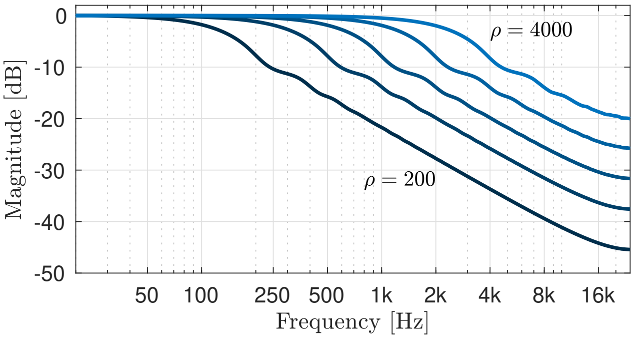


Figure 4: *DVN spectra with different pulse densities,* $\rho \in \{4000, 2000, 1000, 500, 200\}$ *pulses per second.*

within the grid segments of the DVN, thus resulting in a lower cutoff frequency, as seen in Fig. 4.

Fig. 5a shows the DVN spectrum with different maximum pulse widths w_{max}. The maximum pulse width is varied from the lightest blue $w \in \{1\}$, which corresponds to OVN, to 24, i.e., $w \in \{1, 2, ..., 24\}$. The resulting spectral change is from a white spectrum towards more and more lowpass filtered signal as the maximum pulse width increases.

The effect of varying the minimum pulse width w_{min} on the DVN spectrum is shown in Fig. 5b. The lightest color corresponds to DVN with pulse widths $w \in \{1, 2, ..., 24\}$ and the darkest with a single pulse width $w \in \{24\}$. Larger minimum pulse widths introduce a steeper roll-off, but at the same time increase the ripple in the DVN spectrum. Another problem arises with limited minimum pulse width, since the amount of jitter introduced in the pulse location decreases as the minimum pulse width increases. This is seen from (9); in the extreme case, setting $w(m) = T_{\text{d}}$ reduces it to $k(m) = \lfloor mT_{\text{d}} \rceil$.

Overall, Fig. 5 shows that the extreme cases: maximum pulse width close to the grid size and minimum pulse width close to 1 have a diminishing effect on the DVN spectrum. This suggests that the set of pulse widths can be further pruned to find the optimal set for a given application and desired spectral characteristic.

3.3. Perceptual Smoothness

OVN has been shown to sound smoother than Gaussian white noise when the pulse density is over 2000 pulses/s [1, 2]. The perceptual smoothness of OVN is due to its smooth auditory envelope [1]. Furthermore, Karjalainen and Järveläinen [1]showed that lowpass-filtered OVN needs less density than white OVN to sound smooth. Lowpass-filtered velvet noise ($F_c = 1.5\,\text{kHz}$) with an average pulse density of $\rho = 600$, was shown to sound smoother than full-band Gaussian white noise [1]. As each grid segment of DVN now contains a single rectangular pulse, resulting in a lowpass spectrum, the temporal smoothness of a DVN sequence may be at least as good as the smoothness of an OVN sequence with equal pulse density.

Perceptually, we cannot hear the steady DC level during each rectangular pulse. Instead we only. hear the rising and falling edges as the signal level abruptly changes from low to high and vice versa. The effect can be easily demonstrated by listening to a single rectangular pulse with long enough pulse width so that one clearly hears two distinct clicks.

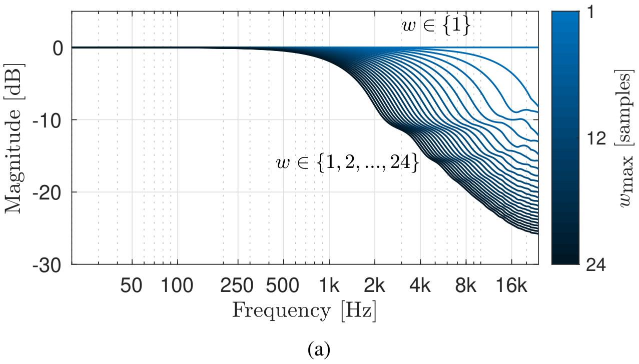


Figure 5: *(a) DVN spectra with varying maximum pulse width* w_{max} *and (b) with varying minimum pulse width* w_{min}.

Sound examples of DVN as well as filtered OVN are provided online [1], using the web audio player from [25]. The filtered OVN sequences were generated using linear prediction (LP) of order $N = 10$ on the DVN sequence. Filtering OVN with the inverse filter gives the OVN the same spectral characteristic as the DVN sequence. Filtered OVN sounds noticeably less smooth than a DVN sequence having the same density ρ. Doubling the OVN density ρ seems to result in similar smoothness. A formal listening test should be conducted to confirm these observations and learn about the perceptual characteristics of DVN.

3.4. Implementation

DVN consists of multiple rectangular pulses with varying width. The convolution with each of the pulses can be implemented efficiently with an RRS filter [22, 23]. The block diagram of the RRS filter is shown in Fig. 6. Each RRS filter requires only one addition, one subtraction, and two multiplications, or four operations in total. In theory, the moving average computing can be implemented with one addition, a delay line, and a subtraction. However, in floating-point computing, using an RRS filter based on the leaky integrator, as in Fig. 6, is safer. Otherwise, the hidden pole at DC can cause numerical instability. In this work, we use a conservative leakage factor $\epsilon = 2^{-12} \approx 2.441 \times 10^{-4}$ to avoid the numerical instability.

Fig. 7a shows the impulse response (IR) of an RRS filter with $M = 95$ for the perfectly rectangular RRS having $\epsilon = 0$ and for the leaky implementation with $\epsilon = 2^{-12}$. The falling edge of both

[1] `http://research.spa.aalto.fi/publications/papers/dafx22-dvn/`

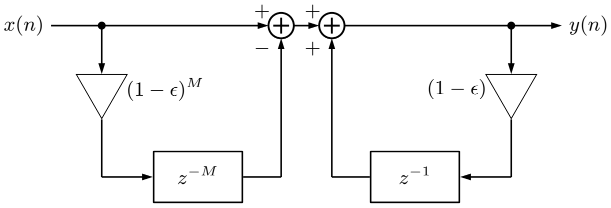
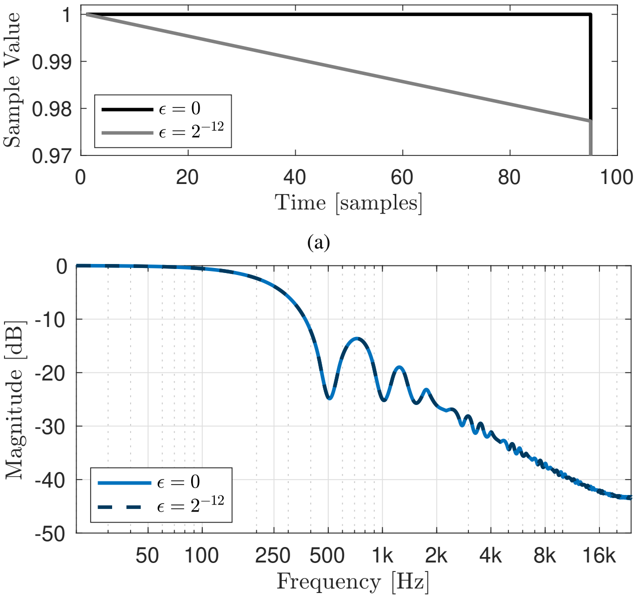


Figure 6: *Block diagram of the RRS filter using a leaky integrator (on the right-hand side) and a feedforward comb filter (on the left). When ϵ is very small, the impulse response of the RRS filter resembles a rectangular pulse, that is M samples long.*



(b)

Figure 7: *(a) RRS filter IR with $M = 95$ and (b) the corresponding third-octave smoothed magnitude response.*

IRs goes to zero, but the y-axis range is limited for better visualization of the difference. The leaky implementation results in a slow exponential decay along the IR, which looks almost linear in 7a due to ϵ and M being relatively small. The corresponding magnitude responses are shown in Fig. 7b, where the deviation from a perfect rectangular pulse clearly has a negligible effect on the spectrum.

The DVN filtering structure needs as many RRS filters as there are unique pulse widths $w(m)$ present in the DVN sequence. For instance, with the sampling frequency $f_s = 48$ kHz and the pulse density $\rho = 2000$, the corresponding grid size is $T_d = 24$, which is also the maximum pulse width. Thus, we need only up to 24 different RRS filters for a DVN convolution of any length. In this example, the RRS filters require $4 \times 24 = 96$ floating-point operations (FLOPS) per output sample.

The block diagram of an example DVN convolution implementation is shown in Fig. 8. The block diagram shows only the connections for a particular example sequence. The pulse width is limited to a maximum of three for better visualization of the structure. The taps of the multi-tap delay line indicate the randomized pulse locations of the DVN. The total length of the delay line L determines the length of the corresponding DVN sequence. Each z^{-T_d} block of the delay line has a single jittered output tap. The ±1 gain blocks randomize the sign $s(m)$ of each pulse. Each of the pulses is then randomly routed to one of the three RRS filters. Finally, the convolution result between signal $x(n)$ and a DVN sequence $h(n)$ is the sum of the RRS filter outputs. The total computational cost of the DVN convolution is still dependent on the pulse density ρ and length L of the DVN sequence, which dictate the total number of output taps in the delay line of Fig. 8.

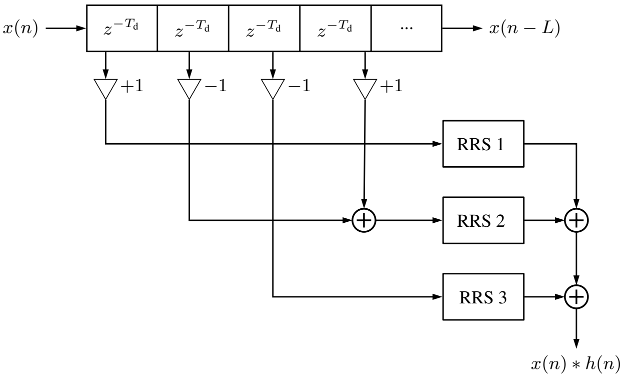


Figure 8: *Example of a DVN convolution structure with three shared RRS filters.*

4. APPLICATION TO ARTIFICIAL REVERBERATION

In this section, application of DVN as an artificial reverberator is discussed. A room impulse response (RIR) usually has a frequency-dependent reverberation time ($T60$), where the high frequencies decay faster than the low frequencies. The frequency-dependent decay is traditionally implemented with a loop filter in structures with feedback loops [26] or with a sub-band structure [27], where each sub-band is processed separately.

As explained in Sec. 3.2, the pulse widths present in the DVN determine its spectral shape. When the maximum pulse width is coupled to the grid size, decreasing the density reduces the bandwidth of the DVN, as seen in Fig. 4. This property lends itself particularly well to applications in artificial reverberation. Previously the density of a velvet noise based RIR model has been shown to be reduced along its length as the bandwidth also decreases without sacrificing temporal smoothness [1, 9].

Fig. 9a shows the spectrogram of an example IR of the DVN reverberator. The parameter values of the maximum pulse width and the pulse density are shown as functions of time in Figs. 9c and 9d, respectively. The IR in Fig. 9a starts as a white spectrum since the maximum pulse width starts at 1. As the maximum pulse width is increased and the pulse density is decreased the spectrum darkens towards the end of the response. The corresponding $T60$ values were estimated from the DVN IR in octave bands, using the decay fit net [28], and are shown in Fig. 9b. Fig. 10a shows the example IR. Additionally, short 20 ms-long segments at the beginning of the IR and at 1 s of the IR are shown in Fig. 10b and Fig. 10c, respectively. Wider and fewer pulses are seen in the latter segment (Fig. 10c), whereas the first segment contains (Fig. 10b) a larger amount of narrower pulses.

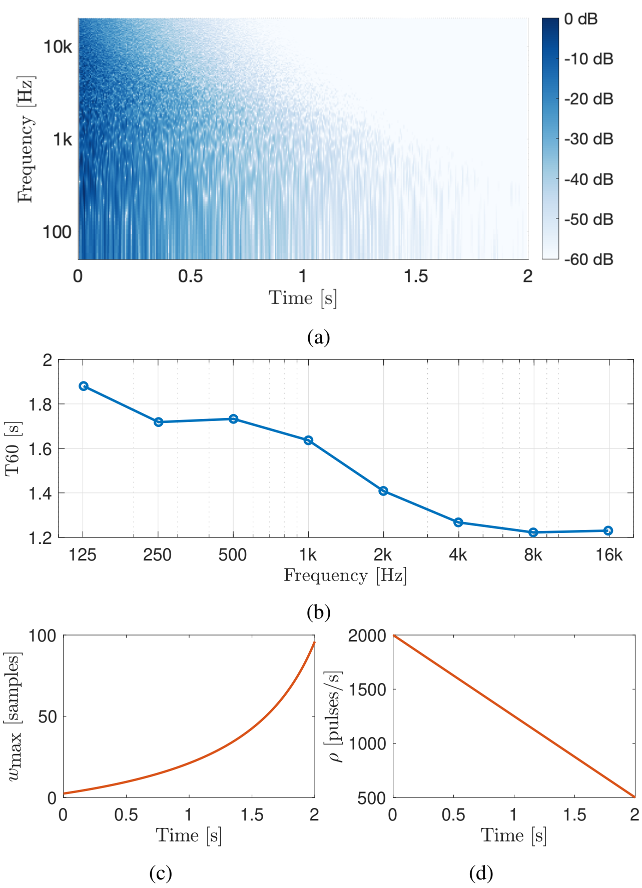


Figure 9: *(a) DVN reverb spectrogram, (b) the corresponding estimated $T60$ values at octave bands, and (c) the maximum pulse width w_{max}, and (d) pulse density ρ as functions of time. The $T60$ values were estimated using the decay fit net [28].*

Since changing the maximum pulse width and the density affect the energy of the sequence, one should apply normalization prior to applying the designed envelope to the DVN reverb. In the provided design example, the generated IR was normalized with both the time-dependent maximum pulse width (Fig. 9c) and the pulse-density (Fig. 9d) parameter values to ensure constant energy throughout the sequence. In practice, the normalization can be baked into the decay envelope as

$$s_{\mathrm{e}}(m) = \frac{1}{w_{\max}(m)\rho(m)} e^{-\alpha m} s(m), \tag{13}$$

where $\alpha > 0$ is the decay rate, $w_{\max}(m)$ and $\rho(m)$ are the time-dependent maximum pulse width and pulse density, respectively, at the pulse locations $k(m)$. Sound examples of the DVN reverb are available in the companion web page of this paper 1.

Table 1 shows the number of arithmetic operations (in FLOPs) per output sample and the required delay-line memory to compute the DVN convolution of the example RIR model in Fig. 9. The example IR has $M = 2502$, pulses resulting in 5003 FLOPs from the decay envelope multipliers $s_{\mathrm{e}}(m)$ and additions of each output tap of the delay line. The remaining part of the computational cost comes from the 95 RRS filters requiring 380 FLOPs. Thus, the total number of FLOPs is $5003 + 380 = 5383$. For comparison, the corresponding costs for direct convolution as well as the partioned fast convolution [29] method are shown in Table 1. The DVN convolution is more efficient in terms of required memory, whereas the partitioned fast convolution is the more efficient method in terms of FLOPs per output sample.

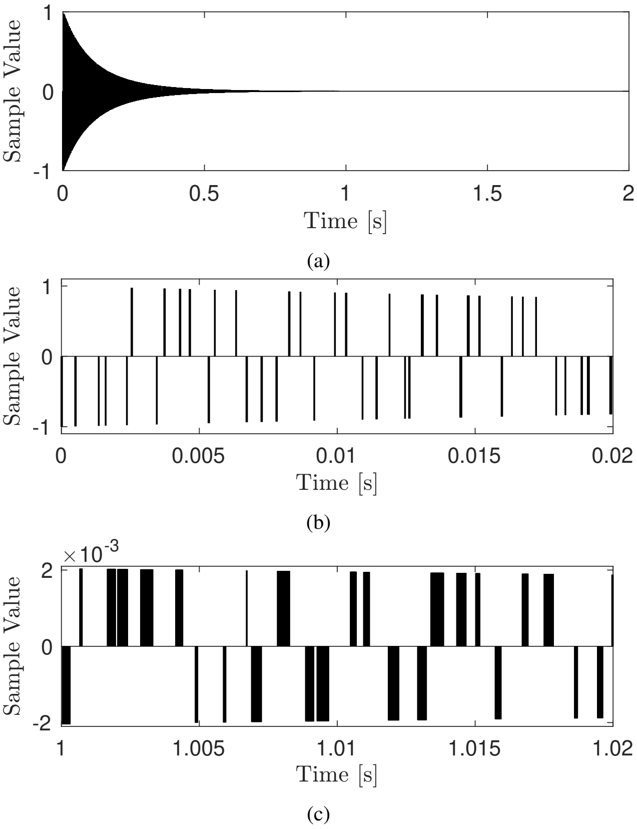


Figure 10: *(a) DVN reverb IR, (b) the first* 20 *ms of the IR, and (c) the* 20 *ms-long segment of the IR starting at* 1 *s.*

Table 1: *Operation count per output sample and memory cost for modeling a 2-s RIR at $f_s = 48$ kHz.*

Algorithm	FLOPs	Delay-line Memory
Direct convolution	191,999	96,000
Partioned fast convolution	451	192,000
DVN convolution	5383	96,000

5. DISCUSSION

In this section, the strengths and drawbacks of the proposed DVN are discussed and compared to other alternative techniques, with the main emphasis on the artificial reverberation application. The CARN sequences discussed in Sec. 2 provided the theoretical background for generating sparse noise with a lowpass spectrum. Now that DVN is introduced as an alternative, the two options can be compared. The lowpass characteristic of the shown CARN with inflated uniform updates is clearly due to the same principle as the lowpass behaviour of DVN: both have segments containing neighboring impulses that act as lowpass-filter responses.

However, the CARN sequences lack the temporal smoothness since the pulse locations are based on the additive random process, which has been shown to require more density for smooth sound than the grid-based pulse location distribution of OVN [2]. Moreover, the pulse width of CARN is unbounded, making it unsuitable for the efficient implementation of the convolution structure in Fig. 8 based on a fixed set of RRS filters. The unbounded pulse width can influence the smoothness as well, since long rectangular pulses create long perceptual gaps.

A further advantage of DVN, in the context of artificial reverberation, is that its cutoff frequency can be freely modulated in time to achieve a time-dependent spectrum of varying brightness. Compared to the feedback delay network (FDN) algorithm [26], the temporal envelope of DVN can be designed to have an arbitrary shape, whereas an FDN can only create a decaying response. In comparison with the partioned fast convolution algorithm [29], the DVN reverb is less effcient in terms of FLOPs for modeling a 2-s long RIR. However, the DVN has the benefit of requiring less memory, and by providing a parametric framework where the parameters of the reverberation, e.g., frequency-dependent $T60$ and the decay profile, can be freely modified.

For the 2-s-long example RIR, the main cost of the DVN convolution comes from the delay line output taps ($M = 2502$) with the decay envelope gains for each pulse $s_e(m)$. Future research may aim to improve the computational efficiency and to investigate further the perceptual quality of the method for applications in artificial reverberation. An efficient parametric reverberation algorithm may be developed based on the DVN convolution.

6. CONCLUSIONS

DVN is proposed in this paper as an extension of velvet noise with a lowpass spectrum. The DVN sounds perceptually dark as opposed to bright, hence the name. The lowpass spectrum is achieved by replacing the unit impulses of a velvet-noise sequence with rectangular pulses of random width. Convolution with the DVN can be implemented efficiently utilizing one RRS filter per unique pulse width. The RRS filter realizes the convolution with a rectangular pulse using only four operations regardless of the pulse width.

The cutoff frequency of the DVN is shown to be controlled by changing the maximum pulse width, which in turn is coupled to the pulse density of the sequence. Decreasing the density lowers the cutoff frequency, since the larger grid can fit in wider pulses. By modulating the maximum pulse width and density over time, the DVN method can produce characteristic artificial reverberation, where the high frequencies decay faster than the low ones. The DVN method is also more flexible in terms of the design of its temporal envelope and modulation of the frequency response than feedback-based reverberation algorithms. Future research may investigate the perceptual smoothness of the DVN in a formal listening test.

7. ACKNOWLEDGMENTS

This research is part of the activities of the Nordic Sound and Music Computing Network— NordicSMC (NordForsk project no. 86892) and has received funding from the European Union Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 812719.

8. REFERENCES

[1] M. Karjalainen and H. Järveläinen, “Reverberation modeling using velvet noise,” in *Proc. Audio Eng. Soc. 30th Int. Conf. Intell. Audio Environ.*, Saariselkä, Finland, Mar. 2007.

[2] 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, Language Process.*, vol. 21, no. 7, pp. 1481–1488, Jul. 2013.

[3] J. A. Moorer, “About this reverberation business,” *Computer Music J.*, vol. 3, no. 2, pp. 13–28, Jun. 1979.

[4] P. Rubak and L. G. Johansen, “Artificial reverberation based on a pseudo-random impulse response: Part I,” in *Proc. 104th Conv. Audio Eng. Soc.*, Amsterdam, The Netherlands, May 1998.

[5] P. Rubak and L. G. Johansen, “Artificial reverberation based on a pseudo-random impulse response: Part II,” in *Proc. 106th Conv. Audio Eng. Soc.*, Munich, Germany, May 1999.

[6] 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.

[7] H. Kawahara, K. I. Sakakibara, M. Mizumachi, M. Morise, and H. Banno, “Simultaneous measurement of time-invariant linear and nonlinear, and random and extra responses using frequency domain variant of velvet noise,” in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Dec. 2020, pp. 174–183.

[8] B. Holm-Rasmussen, H.-M. Lehtonen, and V. Välimäki, “A new reverberator based on variable sparsity convolution,” in *Proc. Int. Conf. Digital Audio Effects (DAFx-13)*, Maynooth, Ireland, Sep. 2013, pp. 344–350.

[9] V. Välimäki, B. Holm-Rasmussen, B. Alary, and H-M. Lehtonen, “Late reverberation synthesis using filtered velvet noise,” *Applied Sciences*, vol. 7, no. 5, 2017.

[10] S. J. Schlecht and E. A. P. Habets, “Scattering in feedback delay networks,” *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 28, pp. 1915–1924, Jun. 2020.

[11] J. Fagerström, B. Alary, S. J. Schlecht, and V. Välimäki, “Velvet-noise feedback delay network,” in *Proc. Int. Conf. Digital Audio Effects (DAFx-20)*, Wien, Austria (remote), Sep. 2020, pp. 219–226.

[12] V. Välimäki and K. Prawda, “Late-reverberation synthesis using interleaved velvet-noise sequences,” *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 29, pp. 1149–1160, Feb. 2021.

[13] M. Morise, “Modification of velvet noise for speech waveform generation by using vocoder-based speech synthesizer,” *IEICE Transactions on Information and Systems*, vol. E102D, no. 3, 2019.

[14] S. D’Angelo and L. Gabrielli, “Efficient signal extrapolation by granulation and convolution with velvet noise,” in *Proc. 21st Int. Conf. Digital Audio Effects (DAFx-18)*, Aveiro, Portugal, Sep. 2018, pp. 107–112.

[15] V. Välimäki, J. Rämö, and F. Esqueda, "Creating endless sounds," in *Proc. 21st Int. Conf. Digital Audio Effects (DAFx-18)*, Aveiro, Portugal, Sep. 2018, pp. 219–226.

[16] C. Muñoz Lázaro, "The Velvet Cloud: Infinite sustain and reverberation using velvet noise," M.S. thesis, Aalborg University, Copenhagen, Dermark, Dec. 2020, `https://projekter.aau.dk/projekter/files/402531233/Master_ThesisV2_10_.pdf`.

[17] B. Alary, A. Politis, and V. Välimäki, "Velvet-noise decorrelator," in *Proc. Int. Conf. Digital Audio Effects (DAFx-17)*, Edinburgh, UK, Sep. 2017, pp. 405–411.

[18] S. J. Schlecht, B. Alary, V. Välimäki, and E. A. P. Habets, "Optimized velvet-noise decorrelator," in *Proc. Int. Conf. Digital Audio Effects (DAFx-18)*, Aveiro, Portugal, Sep. 2018, pp. 87–94.

[19] J. Fagerström, S. J. Schlecht, and V. Välimäki, "One-to-many conversion for percussive samples," in *Proc. Int. Conf. Digital Audio Effects (DAFx-21)*, Vienna, Austria, Sep. 2021, pp. 129–135.

[20] N. Meyer-Kahlen, S. J. Schlecht, and V. Välimäki, "Colours of velvet noise," *Electronics Letters*, vol. 58, no. 12, pp. 495–497, Jun. 2022.

[21] K. J. Werner, "Generalizations of velvet noise and their use in 1-bit music," in *Proc. Int. Conf. Digital Audio Effects (DAFx-19)*, Birmingham, UK, 2019.

[22] T. Saramäki, "Narrowband linear-phase FIR filters requiring a small number of multipliers," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Paris, France, May 1982, pp. 278–281.

[23] J. W. Adams and A. N. Willson, Jr., "A new approach to FIR digital filters with fewer multipliers and reduced sensitivity," *IEEE Trans. Circ. Syst.*, vol. 30, no. 5, pp. 277–283, May 1983.

[24] I. Bilinskis and A. K. Mikelson, *Randomized Signal Processing*, Prentice-Hall, 1992.

[25] N. Werner, S. Balke, F.-R. Stöter, M. Müller, and B. Edler, "trackswitch.js: A versatile web-based audio player for presenting scientific results," in *Proc. 3rd Web Audio Conf.*, London, UK, Aug. 2017.

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

[27] J. Vilkamo, B. Neugebauer, and J. Plogsties, "Sparse frequency-domain reverberator," *J. Audio Eng. Soc.*, vol. 59, no. 12, pp. 752–762, Dec. 2011.

[28] G. Götz, R. Falcón Pérez, S. J. Schlecht, and V. Pulkki, "Neural network for multi-exponential sound energy decay analysis," submitted to J. Acoust. Soc. Am, preprint accessible at `https://arxiv.org/abs/2205.09644v1`, 2022.

[29] F. Wefers and M. Vorländer, "Optimal filter partitions for non-uniformly partioned convolution," in *Proc. Audio Eng. Soc. 45th Conf. Applications of Time-Frequency Processing in Audio*, Helsinki, Finland, Mar. 2012.