VELVET-NOISE DECORRELATOR

Benoit Alary, Archontis Politis, and Vesa Välimäki *

Acoustics Lab, Dept. of Signal Processing and Acoustics
Aalto University
Espoo, Finland
Benoit.Alary@aalto.fi

ABSTRACT

Decorrelation of audio signals is an important process in the spatial reproduction of sounds. For instance, a mono signal that is spread on multiple loudspeakers should be decorrelated for each channel to avoid undesirable comb-filtering artifacts. The process of decorrelating the signal itself is a compromise aiming to reduce the correlation as much as possible while minimizing both the sound coloration and the computing cost. A popular decorrelation method, convolving a sound signal with a short sequence of exponentially decaying white noise which, however, requires the use of the FFT for fast convolution and may cause some latency. Here we propose a decorrelator based on a sparse random sequence called velvet noise, which achieves comparable results without latency and at a smaller computing cost. A segmented temporal decay envelope can also be implemented for further optimizations. Using the proposed method, we found that a decorrelation filter, of similar perceptual attributes to white noise, could be implemented using 87% less operations. Informal listening tests suggest that the resulting decorrelation filter performs comparably to an equivalent white-noise filter.

1. INTRODUCTION

Decorrelation is a useful operation in audio signal processing, as it can reduce correlation properties of partially correlated signals or can generate versions of a monophonic signal that are as little correlated as possible [1]. Leaving aside applications where it may be used internally by another method to improve its performance, such as echo cancellation, decorrelation is usually associated with multichannel processing and modification of the perceived spatial properties of sounds [1], [2], [3].

Reverberation, spatially extended sources, and ambiance with diffuse sound properties all produce binaural signals that have a low correlation. Conversely, when delivering two binaural signals to a listener in a spatial simulation environment the results can vary. The sound is perceived centrally when both signals are fully correlated but, when partially correlated, the spatial image is extended. Two fully incoherent signals may be perceived as two separated lateral events [4]. These properties make decorrelation a useful tool for tasks such as spatial source spreading, artificial reverberation, spatial audio coding, source distance rendering, and headphone externalization [1], [2], [5].

Various decorrelation methods have been proposed in the literature. Generally, most methods aim to a) minimize correlation between the input and the output signal, b) preserve the magnitude spectrum of the signal as much as possible, c) have a short decorrelation impulse response, if time-invariant, and d) be as computationally efficient as possible. Conditions b) and c) attempt to guarantee spatial modification of the sound without changing its spectral character or adding perceived reverberation.

A basic approach is to convolve the audio signal with a short white-noise sequence, in the range of 20–30 ms [1]. Variants of this apply an exponentially decaying envelope, or different decaying envelopes at different bandwidth decreasing time constants at higher frequencies, mimicking late reverberation properties. Such sequences have been studied by Hawksford and Harris [6]. The same work also proposes a *temporally diffuse impulse*, a sum of decaying sinusoids equalized with their minimum-phase equivalent to be flat. Xie *et al*. have studied the use of pseudo-random binary sequences for decorrelation [7]. Instead of synthetic signals, short late segments of room impulse responses are used directly by Faller in a spatial audio coding framework [5]. All these methods require convolution with the generated filters, making them costly in applications with multiple decorrelators applied in parallel, such as source spreading, or on the outputs of a large multichannel system. The fast convolution method is typically used to reduce the computational load, as it implements the convolution as a multiplication of spectra in the frequency domain. The efficiency then comes from the use of the fast Fourier transform (FFT) algorithm. Unfortunately, the fast convolution method causes latency, which is undesirable in real-time audio processing.

A more efficient method, which guarantees a flat magnitude response, is based on cascaded allpass filters [1], [8]. Alternatively, some methods operate in the time-frequency domain, using the short-time Fourier transform or a filter bank, and apply delays on each bin or subband. Such an approach was first introduced in [9], with fixed delays, drawn randomly, applied in perceptually motivated critical bands. This approach has been used with success in parametric spatial audio reproduction methods, e.g. in [10], to decorrelate diffuse/ambient components at the output channels. A more elaborate subband decorrelation approach has been presented by Vilkamo *et al*. [11] for artificial reverberation with convenient spectral control. Although more efficient, the filter-bank method also introduces latency.

This paper proposes the use of a velvet-noise sequence for decorrelating audio signals. Velvet noise consists of samples of values of either -1, 0, or 1 [12], [13]. The spacing between non-zero elements is also randomized to satisfy a density parameter. With sufficient density, these sparse sequences are perceived as smoother than Gaussian white noise [12], [13]. One of the useful applications of velvet-noise sequences is simulating room impulse responses. The late reverberation has been re-created using exponentially decaying filtered noise [14],[15]. Velvet-noise sequences offer a computationally efficient alternative for this purpose [12],

* This work was supported by the Academy of Finland (ICHO project, grant no. 296390).

[16], [17], [18], [19], [20], [21]. Whereas previous work on velvet noise has focused on its perceptual qualities and suitability to artificial reverberation, we investigate how to design a velvet-noise decorrelator (VND).

This paper is organized in the following way. In Section 2, we discuss velvet noise and how signals can be efficiently convolved with it. Section 3 introduces various decorrelation methods using velvet noise. Section 4 presents our results and compares the proposed method with a white-noise-based decorrelation method. Section 5 concludes this paper.

2. VELVET NOISE

2.1. Velvet-Noise Sequence

The main goal of using velvet-noise sequences (VNS) is to create spectrally neutral signals, comparable to white noise, while using as few non-zero values as possible. By taking advantage of the sparsity of the signal, we can efficiently convolve it in the time domain with another signal without any latency [19], [21]. The first step in the generation of velvet noise is to create a sequence of evenly spaced impulses at the desired density [12], as shown in Fig. 1(a). The sign of the impulses is then randomized (see Fig. 1(b)). Finally, the spacing between each impulse is also randomized within the space available to satisfy the density parameter, which is illustrated in Fig. 1(c).
For a given density ρ and sampling rate f_s, the average spacing between two impulses is

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

which is called the grid size. The total number of impulses is

$$M = L_\mathrm{s} T_\mathrm{d}, \tag{2}$$

where L_s is the total length in samples. The sign of each impulse is

$$s(m) = 2\,\mathrm{round}(r_1[m]) - 1, \tag{3}$$

where m is the impulse index, the round function is the rounding operation, and $r_1(m)$ is a random number between 0 and 1. The location of the impulse is calculated from

$$k(m) = \mathrm{round}[mT_\mathrm{d} + r_2[m](T_\mathrm{d} - 1)], \tag{4}$$

where $r_2(m)$ is also a random number between 0 and 1.

2.2. Velvet-Noise Convolution

To convolve a VNS with another signal x, we take advantage of the sparsity of the sequence. Indeed, by storing the velvet-noise sequence as a series of non-zero elements, all mathematical operations involving zero can be skipped. For further optimization, the location of the positive and negative impulses can be stored in separate arrays, k_+ and k_-, which removes the need for multipliers in the convolution [19], [21]

$$x * k = \sum_{m=0}^{M+} x[n - k_+[m]] - \sum_{m=0}^{M-} x[n - k_-[m]]. \tag{5}$$

For a sequence with a density of a 1000 impulses per second, which has been found sufficient for decorrelation, and a sample rate of 44.1 kHz, the zero elements represent 97.7% of the sequence. Therefore, given a sufficiently sparse sequence, time-domain convolution can be more efficient than a spectral domain convolution using an FFT for an equivalent white-noise sequence. Furthermore, this sparse time-domain convolution offers the benefit of being latency-free.

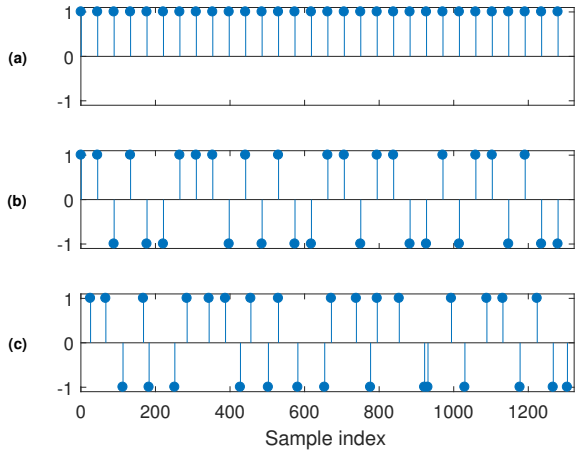

Figure 1: *Steps to create a velvet-noise sequence: (a) a sequence of evenly spaced impulses, (b) the values of impulses are randomized to either 1 or −1, and (c) the spacing between the impulses is randomized.*

3. VELVET-NOISE DECORRELATION

3.1. Exponential Decay

When using a sequence of white noise for decorrelation, adding an exponentially decaying amplitude envelope to the signal is recommended to minimize the potentially audible smearing of transients [6].

The decay characteristics of the envelope are designed to achieve a desired attenuation and target length. The envelope is given by

$$D(m) = e^{-\alpha m} \tag{6}$$

based on a decay constant α given by

$$\alpha = \frac{-\ln 10^{-L_\mathrm{dB}/20}}{M}, \tag{7}$$

where L_dB is the target total decay, while m represent the index of a specific impulse and M is the total number of non-zeros element in the sequence. The attenuation of an impulse is set by its index m. For a given set of parameters, every VNS generated has the same decay attenuation $D(m)$ for the same index m, only the sign varies. This distribution method ensures the energy is consistently distributed amongst the impulses regardless of their positions. The energy could also be distributed over time, depending on an impulse location. However, since the impulse locations vary based on a random value, a variation of distributed power would be obtained, which would lead to unbalanced decorrelation filters, unsuitable for multichannel purposes.

To generate an exponentially decaying velvet-noise filter (Figure 2), we combine the decay envelope of Eq. (6) to the random sign sequence:

$$s_\mathrm{e}(m) = e^{-\alpha m} s(m). \tag{8}$$

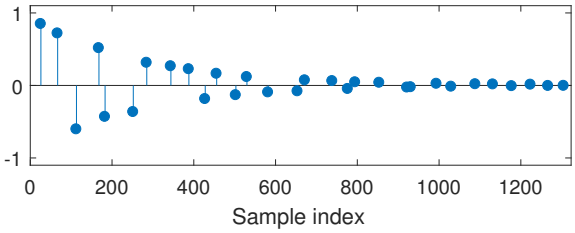

Figure 2: *Velvet-noise sequence with an exponentially decaying temporal envelope.*

The velvet-noise convolution equation also needs to be updated to include the attenuation factors:

$$x * k = \sum_{m=0}^{M} x[n - k[m]] s_{\mathrm{e}}(m). \tag{9}$$

3.2. Segmented Decay

Unfortunately, the decaying amplitude envelope requires multiplications and thus leads to more operations for the convolution. However, since the benefits of a smoothly decaying envelope are not really perceivable with this application, the process can be simplified while still attenuating impulses based on their location. Indeed, audible decorrelation artifacts can be caused by velvet-noise sequences that contain large impulses near the end of the sequence. Therefore, to prevent these artifacts while minimizing the number of multiplications required by the sparse convolution, we can simply limit the attenuation factors to a fixed set of attenuation coefficients, as illustrated in Fig. 3. Storing each segment in a separate list allows for a segmented convolution, applying each coefficient to a partial sum

$$x * k = \sum_{i=0}^{I} s_i \left(\sum_{m=0}^{M_i +} x[n - k_+[m]] - \sum_{m=0}^{M_i -} x[n - k_-[m]] \right). \tag{10}$$

The specific number of segments I and coefficient s_i values can be set manually. While testing with this approach, we have learned that a small number of segments can sound satisfying, e.g., $I = 4$.

3.3. Logarithmic Impulse Distribution

Since the loudness of later impulses needs to be minimized, the samples near the end of the sequence contribute very little to the convolved signal. To this end, the impulses can be concentrated at the beginning of the sequence, where loud impulses do not cause as much smearing of the transients in the decorrelated signals, could be beneficial. For this purpose, we can concentrate the impulses at the beginning of the sequence by distributing their location logarithmically. The grid size between each impulses is

$$T_{\mathrm{dl}}(m) = \frac{T_{\mathrm{T}}}{100} 10^{\frac{2 T_{\mathrm{d}} m}{M}} \tag{11}$$

by distributing the spacing logarithmically based on a total length T_{T} in samples and the density parameter T_{d}. This will also require the use of an updated location equation

$$k(m) = \mathrm{round}\Big(r_2[m] (T_{\mathrm{dl}}(m) - 1) \Big) + \sum_{i=0}^{m} T_{\mathrm{dl}}(i). \tag{12}$$

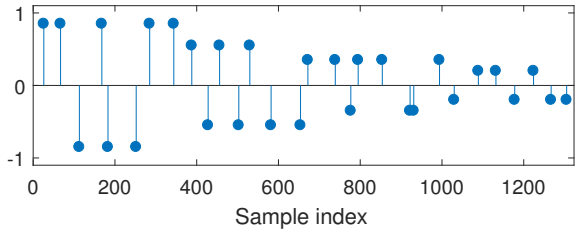

Figure 3: *Velvet-noise with a segmented, staircase-like decay envelope.*

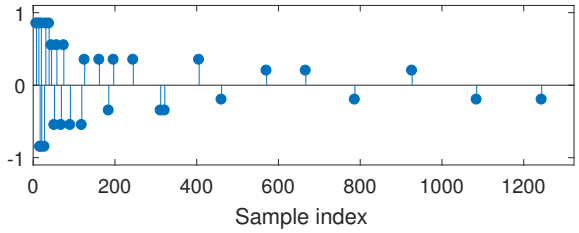

Figure 4: *Velvet-noise sequence having a logarithmically widening grid size combined with a segmented decay envelope.*

Figure 4 shows an example in which the impulse locations are distributed logarithmically and the segmented decay envelope is also applied. Since the impulses are not independent and identically distributed, this distribution may not preserve a flat power spectrum. However, given the low impulse density, this does not appear to have a significant impact on the spectral envelope in practice (see Section 4.2).

4. RESULTS

To evaluate the proposed methods, we first compared the computing cost of convolving velvet noise against the FFT-based white-noise method using different signal lengths and impulse density parameters. We analyzed the spectral envelope of the generated filters to ensure a spectral smoothness comparable to an exponentially decaying white-noise sequence and we calculated the cross-correlation after convolving a sine-sweep signal with the generated filters. The signal coherence was calculated over the audible frequency range. Finally, to complement the numerical evaluation methods, a preliminary perceptual test was conducted to validate the potential of the proposed method to externalize stereo sounds over headphones.

4.1. Computing Resources

We compared the total number of arithmetic operations required by the white-noise decorrelation to the VND method. To estimate the cost of the FFT-based convolution required by the white noise, we chose the radix-2 algorithm. However, a low-latency version would require extra operations to segment the input signal into multiple parts. The radix-2 algorithm has a complexity of $N \log_2(N)$ additions and $\frac{N}{2} \log_2(N)$ complex multiplications. We only counted the operations required to compute the FFT of the input signal, since the white-noise filter itself only needs to be converted into the spectral domain once [17]. The cost per sample was calculated using an $N/2$ window size for the input signal.

The chosen amplitude decay or the impulse distribution method do not have any impact on the required convolution, only the use of segments does. Therefore, the computational cost of the proposed methods can be regrouped into two categories: regular VND and segmented VND. In Fig. 5, the proposed methods are shown to outperform FFT-based convolution well below the required length of a decorrelation filter. Furthermore, for a length of 1024 samples, approximately 0.23 ms at a 44.1 kHz sampling rate, the proposed methods remain more efficient when the impulse density is kept below 6000 impulses per second or below 13000 impulses per second when using the segmented method with four segments.

In our sound examples, the decorrelation filters were chosen to be of length 30 ms, which leads to a signal length of 1323 samples for a sampling rate of 44.1 kHz. An impulse density of 1000 impulses per seconds was selected for the VND. The input signal was segmented into windows of 2048 samples. Since the convolution result has the length $2M - 1$, convolving two signals of length M requires a zero-padded FFT of 4096 samples to convolve with a filter of the same length. We can see from Table 1 that, for a length of 30 ms and a density of 1000 impulses per seconds, the VND with exponential decay envelope will require 76% less mathematical operations than fast convolution, whereas a segmented version of the algorithm with four segments will yield 87% fewer operations than fast convolution.

4.2. Spectral Envelope

In the context of decorrelation, a filter should be as spectrally flat as possible while randomizing the phase. Figure 6 compares the power spectra of a 30 ms white-noise signal with a VNS of the same length, both using an exponential decay envelope. Figure 7 compares the power spectra of a segmented VNS to another one that is similarly segmented but logarithmically distributed. The solid line shows the spectra of one randomly generated sequence. Since they are of short length, the randomization process does not lead to a perfectly flat spectrum. However, they both exhibit similar spectral behavior and they both have a comparable standard deviation over multiple randomized instances, as shown with the dashed lines.

4.3. Cross-Correlation

Although assessing the perceptual quality of a decorrelator numerically is difficult, the cross-correlation of two signals decorrelated with a VND should result in a similar plot when compared with a white-noise decorrelation. Assuming that the original input signal is a and its decorrelated version is $b = h * a$, where h is the decorrelation filter, the correlation between them is

$$r_{ab}(l) = \sum_n a(n+l)b(n), \tag{13}$$

where l is the correlation lag. For decorrelation purposes, the zero-lag case $r_{ab}(0)$ is of practical interest, since it corresponds to the maximum similarity between the two sequences. The normalized zero-lag correlation is given by

$$\rho_{ab} = \frac{\sum_n a(n)b(n)}{\sqrt{\sum_n a^2(n) \sum_n b^2(n)}}. \tag{14}$$

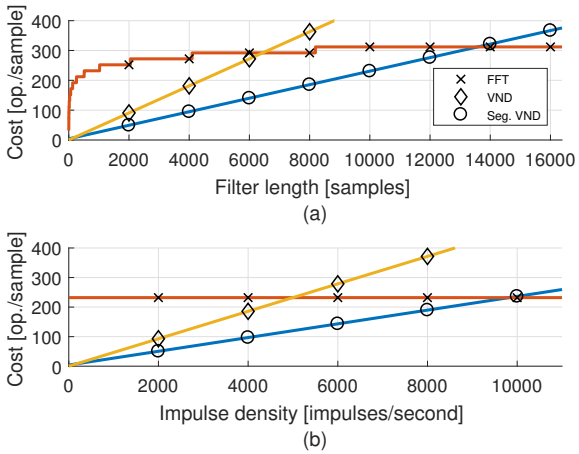

Figure 5: *Comparison of computational cost between block FFT-based convolution (cross), VND (diamond), and segmented VND using four segments (circle): (a) the number of operations required to process each input sample for different filter lengths when the impulse density ρ is set to 1000 impulses per second and (b) the cost for various impulse density settings when the filter length is set to $L_s = 1024$ samples.*

Table 1: *Computational cost per output sample in FFT-based fast convolution, in exponentially decaying velvet-noise convolution, and in segmented velvet-noise convolution. The VNS has 30 non-zero elements (30 ms at a density of 1000 impulses/seconds at 44.1 kHz).*

	Fast convolution	Exp. VND	Seg. VND
ADD	148	30	30
MUL	104	30	4
Total	252	60	34

The normalized cross-correlation coefficient is bounded between $\rho_{ab} \in [0, 1]$, and a lower maximum value indicates a more effective decorrelation.

We used a sine sweep to compare the correlation of two channels decorrelated either by a white-noise sequence or a segmented VNS (Fig. 8). Short windows of 20 ms were taken from both signals to calculate the cross-correlation values at different lag distances and generate a cross-correlogram. Figure 8(a) shows the auto-correlation patterns of the input signal, while Figs. 8(b) and (c) both display a comparable amount of decorrelation after applying the filters on each channels.

4.4. Coherence

Apart from the time-domain correlation, another useful metric is the cross-correlation in different frequency bands, called coherence. Normally, a decorrelator is more effective at higher frequencies than at lower, which is a result of the effective length of a decorrelation filter. Indeed, a longer filter will exhibit stronger decorrelation for longer wavelengths, but will also create potentially perceivable artifacts when the input signal contains transients. To study the decorrelation behavior on a frequency-dependent scale, we can use an octave or third-octave filterbank. The signals for the

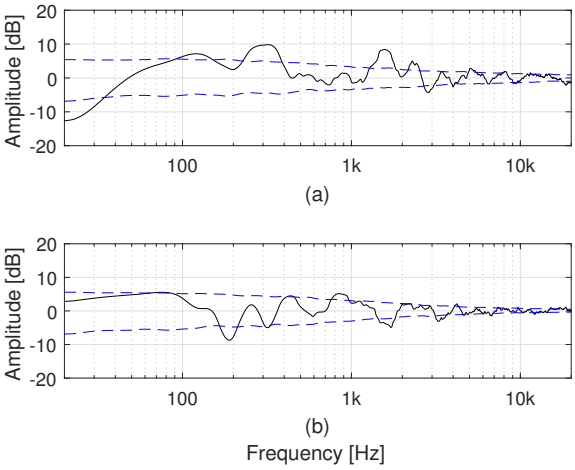

Figure 6: *Spectral envelope of (a) an exponentially decaying white-noise sequence and (b) an exponentially decaying VNS. A third-octave smoothing has been applied to these magnitude responses. The dashed lines represent the average standard deviation of randomly generating 500 sequences.*

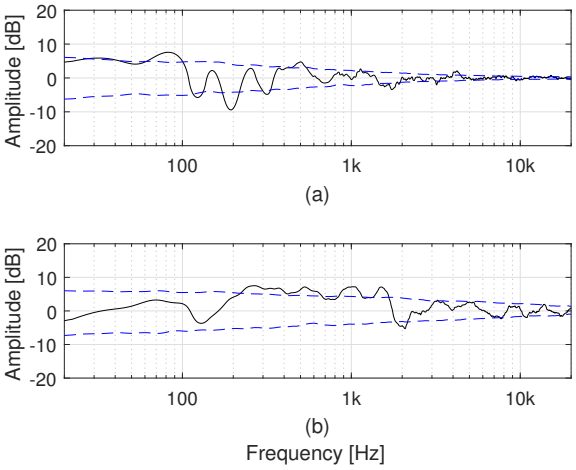

Figure 7: *Spectral envelope of (a) a segmented VNS and (b) a segmented and logarithmically distributed VNS. Both use four segments of equal length and with values 0.85, 0.55, 0.35, and 0.20, respectively. A third-octave smoothing has been applied to these magnitude responses. The dashed lines represent the average standard deviation of randomly generating 500 sequences.*

q^{th} band are denoted as a_q and b_q and the correlation coefficient as

$$\rho_{ab}^{(q)} = \frac{\sum_n a_q(n) b_q(n)}{\sqrt{\sum_n a_q^2(n) \sum_n b_q^2(n)}}. \tag{15}$$

Using this equation, we calculated the coherence of both white-noise decorrelation and VND methods. In Fig. 9, the decorrelation was applied to a white-noise input signal, and the tests were run five hundred times to accumulate the averaged results. Figure 10 shows the same calculation comparing a segmented VNS to another that is similarly segmented but logarithmically distributed. The exponentially decaying VNS shows similar results as the white-noise sequence. However, the segmented version shows better decorrelation when the impulses are linearly distributed in time (see Fig. 10(a)), whereas it shows less decorrelation when distributed logarithmically (see Figure 10(b)). Based on these results, the segmented version shows the most promising decorrelation potential.

4.5. Perceptual Testing

Perceptual testing of decorrelation filters is important, since it is difficult to numerically evaluate the impact of the decorrelator on the stereo image. During an informal perceptual test, several subjects found the proposed VND to lead to a similar or better decorrelation in comparison to the exponentially decaying white-noise decorrelator. The spectral coloration and spreading were comparable, and the transients did not have artifacts, provided that latter impulses were sufficiently attenuated in the VNS. However, a more thorough listening test is required before drawing definitive conclusions. The VND could also be compared to other decorrelation methods, such as the allpass filter and filter-bank methods.

The preliminary listening suggests that the VND can provide a satisfactory externalization of stereo sounds over headphones at a lower computational cost than the white-noise decorrelation method. No extra perceivable artifacts were detected when using the segmented decay envelope instead of the exponential decay. The tests subjects did detect some low-frequency filtering and a spectrally fluctuating stereo image. However, these are common by-products of decorrelation filters. Audio examples are available at http://research.spa.aalto.fi/publications/papers/dafx17-vnd/.

5. CONCLUSION

The decorrelation of audio signals by convolving them with a short VNS was proposed. Since a VNS is a spectrally flat signal, it can create decorrelated signals in the same way as white noise. The proposed VND method allows a latency-free, time-domain convolution at favorable computing costs. A VNS having an appropriate density and a decaying temporal envelope provides a comparable decorrelation with a smaller computational cost when compared to a white-noise sequence. When used to replace white noise in our test scenario, the VND reduces the number of arithmetic operations by 76% to 87%, depending on the configuration. Several parameters can be used to alter the decorrelation itself, such as density, decaying envelope, segmentation, and impulse distribution. According to this study, the VNS with a segmented decay envelope appears to be the best option, since it produces as good a decorrelation as white noise without latency and using 87% less operations. Future work may study the impact of these various parameters as well as conduct a formal perceptual evaluation and a thorough comparison with other well-established decorrelation methods.

6. ACKNOWLEDGMENT

The authors would like to thank Luis Costa for proofreading this paper.

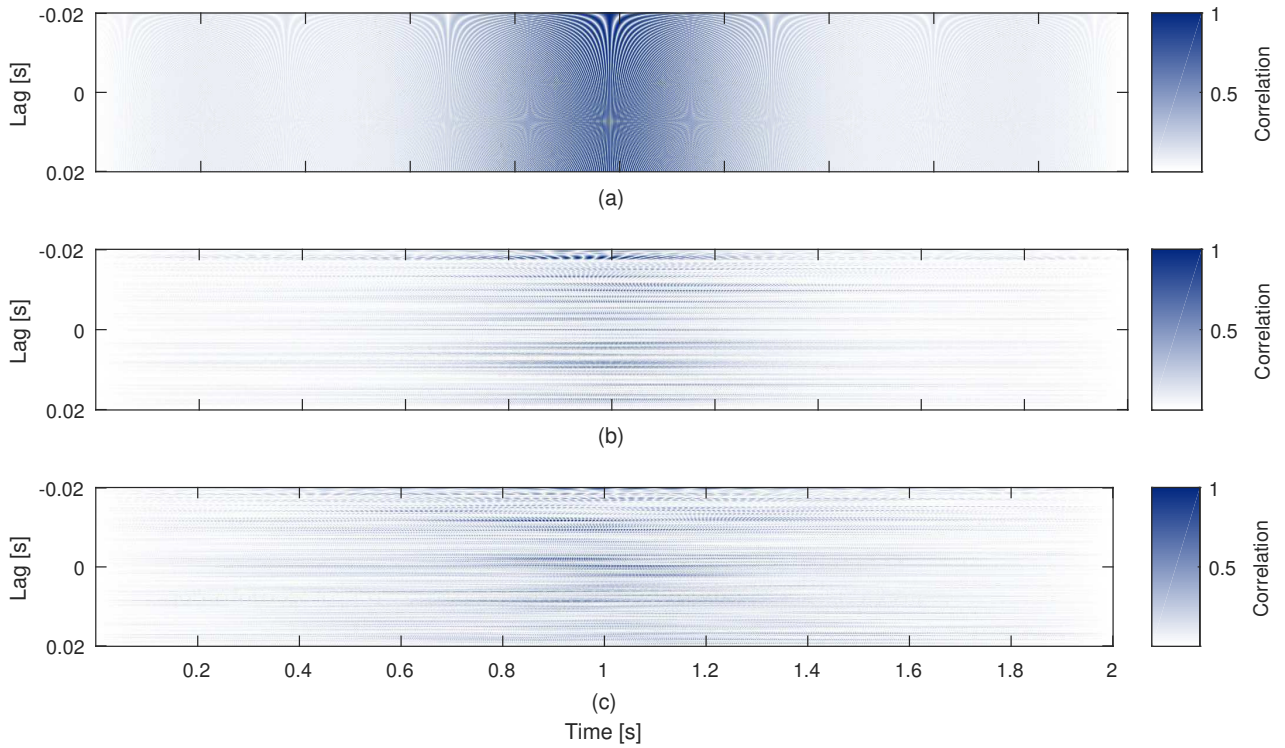

Figure 8: *(a) Auto-correlogram of the sine sweep signal. (b) Cross-correlogram of both channels using the white-noise sequence for decorrelation. (c) Cross-correlogram of both channels using the segmented VND for decorrelation. The correlation values were normalized for visualization purposes.*

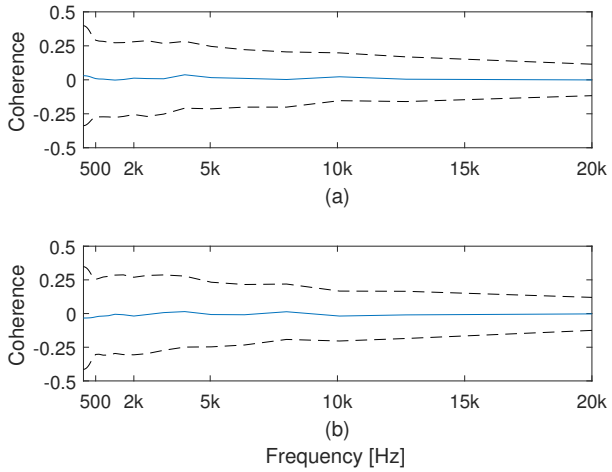

Figure 9: *Coherence graph, from 20 Hz to 20 kHz, between left and right channels when using (a) white noise and (b) velvet noise. The dashed lines represent the average standard deviation of the data after running the algorithm 500 times.*

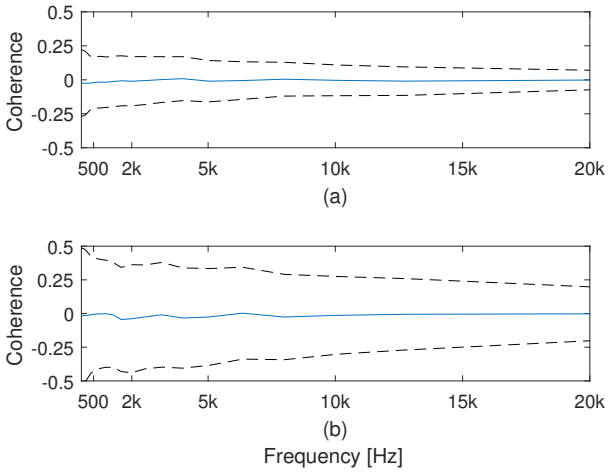

Figure 10: *Coherence graph, from 20 Hz to 20 kHz, between left and right channels when using (a) a segmented VNS and (b) a segmented and logarithmically distributed VNS. The dashed lines represent the average standard deviation of the data after running the algorithm 500 times.*

7. REFERENCES

[1] G. S. Kendall, "The decorrelation of audio signals and its impact on spatial imagery," *Computer Music J.*, vol. 19, no. 4, pp. 71–87, 1995.

[2] G. Potard and I. Burnett, "Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays," in *Proc. Int. Conf. Digital Audio Effects (DAFx-04)*, Naples, Italy, Oct. 2004, pp. 280–284.

[3] V. Pulkki and J. Merimaa, "Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests," *J. Audio Eng. Soc.*, vol. 54, no. 1/2, pp. 3–20, Jan./Feb. 2006.

[4] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT press, 1997.

[5] C. Faller, "Parametric multichannel audio coding: synthesis of coherence cues," *IEEE Trans. Audio, Speech, and Lang. Processing*, vol. 14, no. 1, pp. 299–310, Jan. 2006.

[6] M. J. Hawksford and N. Harris, "Diffuse signal processing and acoustic source characterization for applications in synthetic loudspeaker arrays," in *Proc. Audio Eng. Soc. 112nd Conv.*, Munich, Germany, May 2002.

[7] B. Xie, B. Shi, and N. Xiang, "Audio signal decorrelation based on reciprocal-maximal length sequence filters and its applications to spatial sound," in *Proc. Audio Eng. Soc. 133rd Conv.*, San Francisco, CA, USA, Oct. 2012.

[8] E. Kermit-Canfield and J. Abel, "Signal decorrelation using perceptually informed allpass filters," in *Proc. 19th Int. Conf. Digital Audio Effects (DAFx-16)*, Brno, Czech Republic, Sept. 2016, pp. 225–231.

[9] M. Bouéri and C. Kyriakakis, "Audio signal decorrelation based on a critical band approach," in *Proc. Audio Eng. Soc. 117th Conv.*, San Francisco, CA, USA, Oct. 2004.

[10] A. Politis, J. Vilkamo, and V. Pulkki, "Sector-based parametric sound field reproduction in the spherical harmonic domain," *IEEE J. Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 852–866, Aug. 2015.

[11] J. Vilkamo, B. Neugebauer, and J. Plogsties, "Sparse frequency-domain reverberator," *J. Audio Eng. Soc.*, vol. 59, no. 12, pp. 936–943, Dec. 2012.

[12] M. Karjalainen and H. Järveläinen, "Reverberation modeling using velvet noise," in *Proc. 30th Int. Conf. Audio Eng. Soc.: Intelligent Audio Environments*, Saariselkä, Finland, Mar. 2007.

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

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

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

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

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

[18] S. Oksanen, J. Parker, A. Politis, and V. Välimäki, "A directional diffuse reverberation model for excavated tunnels in rock," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013, pp. 644–648.

[19] B. Holm-Rasmussen, H.-M. Lehtonen, and V. Välimäki, "A new reverberator based on variable sparsity convolution," in *Proc. 16th Int. Conf. Digital Audio Effects (DAFx-13)*, Maynooth, Ireland, Sept. 2013, pp. 344–350.

[20] V. Välimäki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel, "More than 50 years of artificial reverberation," in *Proc. Audio Eng. Soc. 60th Int. Conf. Dereverberation and Reverberation of Audio, Music, and Speech*, Leuven, Belgium, Feb. 2016.

[21] V. Välimäki, B. Holm-Rasmussen, B. Alary, and H.-M. Lehtonen, "Late reverberation synthesis using filtered velvet noise," *Appl. Sci.*, vol. 7, no. 483, May 2017.