SHIMMER REVERBERATION WITH NONLINEAR FEEDBACK DELAY NETWORKS

*Gloria Dal Santo** , *Xiaojie Pi*
Acoustics Lab
Department of Information and Communications Engineering
Aalto University, Espoo, Finland

Karolina Prawda
AudioLab
School of Physics, Engineering and Technology
University of York, York, United Kingdom

Sebastian J. Schlecht
Multimedia Communications and Signal Processing
Friedrich-Alexander-Universität Erlangen-Nürnberg
Erlangen, Germany

Vesa Välimäki
Acoustics Lab
Department of Information and Communications Engineering
Aalto University, Espoo, Finland

ABSTRACT

Shimmer reverberation is an effect used in music production to deliver ethereal, pitch-shifted textures and evolving ambient soundscapes. This paper explores the synthesis of shimmer effects using the feedback delay network architecture, a popular real-time reverberator. We propose five distinct approaches for integrating nonlinear and time-varying operations into the feedback loop, focusing on expanding the harmonic content while adhering to energy-preservation and stability criteria. Our approach can generate a wide range of sonic characteristics, from harmonically rich distortions to musically coherent pitch-shifted reverberation, while maintaining stability and controllable decay behavior. This work bridges physically inspired structures with creative, non-physical audio effect design to support the exploration of novel music production tools.

1. INTRODUCTION

Modern music production, especially electronic music, frequently uses reverb plugins that produce unnatural pitch-shifted textures, such as the *shimmer reverberation* effect. This term is often used to describe a nonlinear artificial reverberator that generates new harmonics and frequency content beyond the spectral characteristics of the input signal. This paper studies five nonlinear signal processing techniques to realize unique shimmer reverb effects, comparing how the different approaches vary in their sonic characteristics as well as in the degree of parameter control they provide.

Pitch shifting and nonlinear audio effects are widely studied and ubiquitous in music production. Pitch shifters can be used to correct intonation or harmonize instruments and voices [1]. Meanwhile, nonlinearities are used for dynamic range control and distortion effects [1], to model analog circuits and components [2, 3, 4] and nonlinear physical behaviors in musical instruments [5, 6, 7], such as high-amplitude vibrations and string tension.

Among the many approaches to digital artificial reverberation, delay-based recursive structures are the earliest methods for real-time implementation [8]. Over time, such structures underwent various modifications, eventually evolving into the feedback delay network (FDN) [9], which has since the 1990s become a widely adopted algorithm for digital reverberation synthesis [8, 10, 11, 12]. Over the past decade, substantial research has continued to refine the analysis and design of FDNs, improving control over perceptually relevant properties such as reverberation time (T_{60}), echo density, and coloration [13, 14, 15, 16].

Despite these advancements, music production and artistic performance often prioritize non-physical reverb effects for creative expression. A wide range of plugins explores aesthetics described as "ethereal," "lush," and "dreamy," deliberately departing from physical plausibility and allowing reverberation to function as a secondary synthesizer. Sean Costello's ValhallaShimmer [17], for example, aims to reproduce the shimmer effect common in Brian Eno's productions by pitch-shifting a reverberated signal, with the option to reverse it in time. Costello based his algorithm on a feedback loop containing cascaded diffusion stages and pitch shifters [17]. Other popular nonlinear reverb effect plugins include Eventide's Blackhole [18] and ShimmerVerb [19], Soundtoys' Crystallizer [20], and Kentaro's Particle reverb [21].

Room reverberation is typically modeled as a linear time-invariant (LTI) system and characterized by the room impulse response [22]. Although atmospheric changes or surface absorption variations can introduce time variations, these effects are usually negligible for individual measurements [23]. Nonlinear behavior may only occur under extreme sound pressure levels or in the presence of resonant elements that are strongly excited by the sound source. For this reason, most artificial reverberation algorithms are based on an LTI system and research on intentionally non-physical reverberation effects has received relatively limited academic attention. One notable contribution is the work of Abel and Werner [24], who adapted a modal reverberator [25] to produce a wide range of pitch, distortion, and envelope manipulation effects.

A nonlinear reverberator based on an FDN structure could offer an efficient alternative that achieves a high modal and echo density. Besides room acoustics, the sonic capabilities of FDNs and FDN-like structures have been explored by replacing the delay lines with spring reverb models [26], to create filtering effects that range from reverb-like sounds to a self-oscillating system, or by modulating the delay lines to create a vibrato effect [27]. Delay modulation has been further explored by Norilo to create chorus, flanger, and pitch-shifter effects [28]. Norilo ensured stability through soft saturation, which introduces both harmonic and inharmonic aliasing components, and provided a limited evaluation

* Corresponding author, e-mail: gloria.dalsanto@aalto.fi

analysis. Finally, to break temporal periodicity, a common artifact in FDNs [29], Schlecht and Habets introduced a technique based on a modulated feedback matrix [30], producing effects analogous to delay modulation. In this study, we explore FDN capabilities of producing shimmer reverb through five techniques, and analyze them in terms of energy preservation and the spectral and temporal characteristics of the resulting sound.

The remainder of this paper is structured as follows. Section 2 defines the problem and introduces the nomenclature used throughout the paper. Section 3 describes the selected nonlinear and time-varying operations, and Section 4 analyzes their energy and stability properties when inserted into a recursive filter. The resulting shimmer reverberators are presented in Section 5, and their sonic performance is discussed in Section 6. Finally, Section 7 concludes the paper.

2. REVERB MODEL AND PROBLEM STATEMENT

An FDN is a recursive system consisting of delay lines, a set of gains, a feedback matrix coupling the delay-line outputs with the delay-line inputs, and attenuation filters controlling the decay rate. Its response is characterized by

$$\begin{aligned} y(n) &= \boldsymbol{c}^\top \boldsymbol{s}(n) \\ \boldsymbol{s}(n+\boldsymbol{m}) &= \boldsymbol{U}\boldsymbol{\Gamma}(n)\boldsymbol{s}(n) + \boldsymbol{b}x(n) \end{aligned} \tag{1}$$

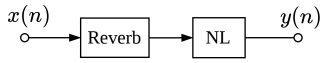
where $x(n)$ is the input signal, $y(n)$ is the output signal, $\boldsymbol{U}$ is the $N \times N$ mixing matrix, and $\boldsymbol{\Gamma}(n)$ is a diagonal matrix of N parallel filtering stages Γ_i. Here, N denotes the number of delay lines in $\boldsymbol{m}$ and i is the delay-line-specific index. Vectors $\boldsymbol{b}$ and $\boldsymbol{c}$ are $N \times 1$ column vectors of input and output gains, respectively, and the operator $(\cdot)^\top$ denotes the transpose. The vector $\boldsymbol{s}(n)$ is the $N \times 1$ output of the delay lines at time index n.

To introduce a shimmer-like effect in the FDN, we could explore the options depicted in Fig. 1, where NL denotes the nonlinear operation used to produce new harmonics and tones. Fig. 1a presents the most straightforward method, in which the nonlinearity is applied after the reverberator. Figure 1b presents the proposed approach, in which nonlinear operations are placed in the feedback loop of an FDN. In this configuration, the output of each nonlinearity feeds $\boldsymbol{U}$, which distributes its effect across the FDN channels, generating distinct sound textures depending on the chosen FDN parameters.

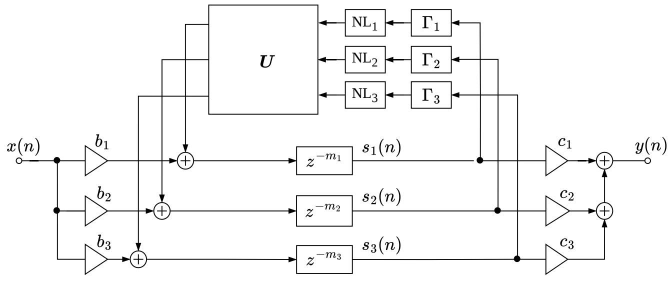
In this paper, we focus on the configuration in Fig. 1b to design a novel reverberator with a shimmering sonic texture. To ensure the stability of the FDN, we focus on operations that are nearly energy preserving, i.e.,

$$\sum_{n=0}^{L} |\mathrm{NL}(x(n))|^2 \approx \sum_{n=0}^{L} |x(n)|^2. \tag{2}$$

where L is the length of an arbitrary input signal $x(n)$ in samples. While lossy operations inherently guarantee stability, near energy preservation of NL allows the overall energy decay of the response to remain controllable. In the absence of attenuation filters, the degree of energy preservation must be further constrained to avoid any increase in the signal energy. Throughout the paper, we assume an FDN with an orthogonal feedback matrix $\boldsymbol{U}$ and attenuation filters satisfying $|\Gamma_i| < 1$. Under these conditions the system is stable. However, since the nonlinear operation recursively shifts energy from the original frequency bins to different regions of the spectrum, the T_{60} profile determined by $\boldsymbol{\Gamma}$ is no longer strictly preserved, and aliasing artifacts may arise.



(a) Series nonlinearity after reverberation.



(b) FDN with feedback nonlinearity.

Figure 1: Block diagrams of two options to introduce a shimmer-like effect using an FDN reverberator. The blocks labeled “NL” and “Reverb” are placeholders for the nonlinear operation and artificial reverberation algorithm (e.g., an FDN).

3. NONLINEAR AUDIO PROCESSING METHODS

In this section, we present the different nonlinear operations used within the feedback loop of the FDN (Fig. 1b) to produce a shimmer reverb effect. These techniques, some of which are commonly used in sound synthesis and audio processing, modify the spectral content in both harmonic and inharmonic ways. Figure 2 shows their response to a sinusoidal signal at 1.76 kHz (tone A6). For all experiments in this work, we use a sampling rate of $f_s = 96$ kHz.

3.1. Controllable Full-Wave Rectifier

We study a controllable full-wave rectifier (CFWR) [31], which is defined as:

$$y(n) = g_{\text{cfwr}}((1-\alpha)x(n) + \alpha|x(n)|), \tag{3}$$

where $0 \leq \alpha \leq 1$ controls the strength of the distortion effect and scaling factor $g_{\text{cfwr}} = \sqrt{2 - 2|\alpha - 1/2|}$ helps maintain the signal power when $\alpha < 1$. Figure 3 (left) shows the nonlinear transfer function of (3) and (right) the output for different values of α. For $\alpha < 1$, the positive portion of the input signal is amplified by g_{cfwr} to compensate for the attenuation incurred by the negative portion. The scaling factor g_{cfwr} reaches its maximum $\sqrt{2}$ at $\alpha = 0.5$.

The CFWR generates all even harmonics of the input signal. This introduces aliasing, as higher-frequency harmonics fold back at the Nyquist limit, producing mirrored spectral images that appear as discrete, staircase-like artifacts across the spectrum [32]. To reduce the resulting spectral aliasing, we apply first-order antiderivative antialiasing [33, 34] and obtain the following approximation to the full-wave rectifier:

$$|x(n)| \approx \frac{1}{2}\frac{x(n)|x(n)| - x(n-1)|x(n-1)|}{x(n) - x(n-1)}. \tag{4}$$

If the denominator of (4) is very small, such as when $x(n-1) \approx x(n)$, to avoid division by zero, we replace the antialiased approximation with $|x(n) + x(n-1)|/2$, as suggested by Parker *et al.* [33]. The effectiveness of this approach is illustrated in Fig. 4, which compares the spectrum of the standard full-wave rectifier with the antialiased approximation in (4), at $\alpha = 1$. The plot demonstrates a significant attenuation of the aliased spectral components between the primary harmonics.

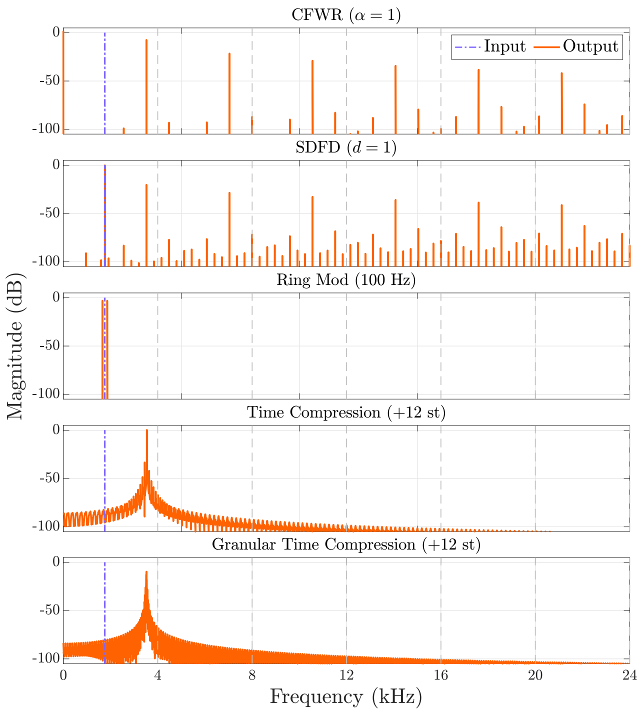


Figure 2: Spectrum of the nonlinearities excited by a 1.76 kHz sinusoidal input (tone A6) with $f_s = 96$ kHz.

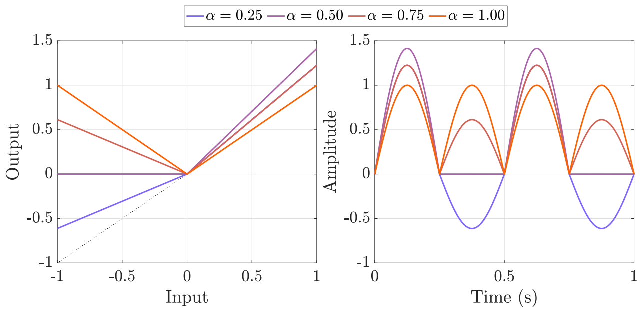


Figure 3: (Left) Transfer function of the CFWR nonlinearity, and (right) its output to a 2-Hz sinusoidal input with unitary amplitude. Different colors indicate different values of the control variable α.

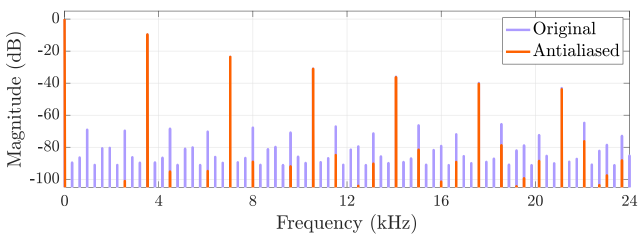


Figure 4: Spectrum comparison for a 1.76-kHz sinusoidal input processed via standard full-wave rectification, $y(n) = |x(n)|$, and its antiderivative antialiasing approximation defined in (4).

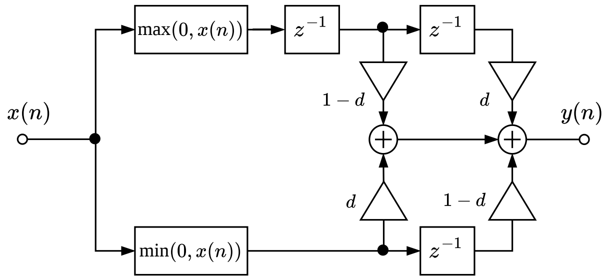


Figure 5: Block diagram of the SDFD nonlinear filter [6].

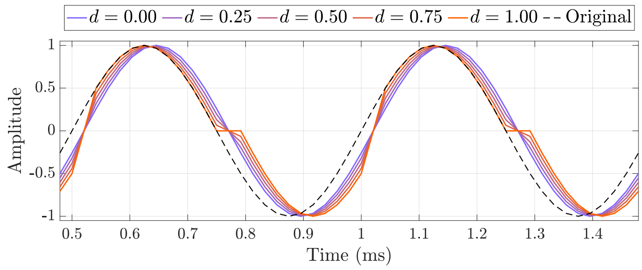


Figure 6: Output of the SDFD filter to a 2 kHz sinusoidal input with unitary amplitude. Different colors indicate different values of the control variable d.

3.2. Signal-Dependent Fractional Delay

Signal-dependent fractional delay (SDFD) filters were introduced to model amplitude-dependent nonlinear characteristics in musical instruments [5, 6]. We use the SDFD filter shown in Fig. 5, where the input signal is decomposed into positive and negative components using positive and negative half-wave rectifiers, respectively [6]. Each rectifier's component is then processed by a linear interpolator with complementary weights d and $1 - d$, where $0 \leq d \leq 1$. The positive component experiences an effective delay of approximately $1 + d$ samples, while the negative component is delayed by approximately $1 - d$ samples. The output signal is obtained by summing the two delayed components. The output waveform in Fig. 6 shows that, for a sinusoidal input, the SDFD retains the fundamental frequency and overall amplitude of the input waveform but exhibits distortions at the zero crossings. Similarly to the CFWR, the output spectrum of the SDFD contains even harmonics and some aliasing, as shown in Fig. 2, for $d = 1$.

3.3. Ring Modulation

Ring modulation (RM) is a sound processing technique that modulates the amplitude of the input signal by a sinusoidal signal with frequency f_{rm} [1]. Its output is given by

$$y(n) = g_{\mathrm{rm}} x(n) \sin(2\pi f_{\mathrm{rm}} n), \tag{5}$$

where g_{rm} is a scaling factor that maintains the input signal power. The multiplication modifies the spectrum of $x(n)$, producing two sidebands centered at the sum and difference frequencies relative to f_{rm}. As a result, the output spectrum consists of an upper sideband shifted by $+f_{\mathrm{rm}}$, and a mirrored lower sideband shifted by $-f_{\mathrm{rm}}$ (see Fig. 2). Unlike conventional amplitude modulation, the original input spectrum is suppressed, so that only the two sidebands remain in the output spectrum [1].

At low modulation frequencies, e.g., 10 Hz, the effect is perceived primarily as a tremolo due to slow amplitude variation. However, when the modulation frequency enters the audible range, the effect alters the timbre of the signal. If the frequency components of $x(n)$ and f_{rm} are related by an integer ratio, the generated sidebands are harmonic; otherwise, the resulting spectrum is inharmonic, often producing bell-like or metallic sounds [35].

3.4. Time Compression and Expansion

Pitch modification by time compression or expansion is achieved by driving the read pointer of a ring buffer at a different rate than the write pointer [36, 37, 38]. When the read pointer advances through the buffer faster than the write pointer, the output signal is time-compressed, resulting in a higher perceived pitch. Conversely, when the read pointer advances slower, the signal is time-expanded, producing a lower pitch. In this work, we use a single ring buffer with two read pointers, to avoid discontinuities whenever the read pointer wraps around to the beginning of the buffer.

Figure 7 illustrates the processing over two modulation periods. In the top panel, the normalized positions of the write and read pointers are shown. The read pointers always access past samples, and their positions are offset by half of the modulation period. This ensures that when one pointer approaches its reset point, the other continues reading from a stable region of the buffer. The signals from the read pointers are combined using an equal-power crossfade function, shown in the middle panel of Fig. 7. These crossfades are synchronized with the sawtooth modulation controlling the read pointers, allowing the gain to be smoothly transferred from one read pointer to the other before a reset occurs, assuming that the modulation period is shorter than the buffer length. In the bottom panel of Fig. 7, the output signal shows a frequency which is nearly double that of the input signal, corresponding to a pitch shift of one octave upward. Cancellation artifacts during crossfading, causing modulation in the amplitude, can be mitigated by incorporating a periodicity detection algorithm or by using an enhanced pitch-shifting method [38, 39].

For a desired transposition amount in semitones, T_{st}, the read pointer advances by $2^{T_{\text{st}}/12}$ samples at every processing step n. Unless exact octave transposition is desired, this value is non-integer. To interpolate between fractional sample positions, we use cubic polynomial interpolation.

3.5. Granular Time Compression and Expansion

In the granulation effect, the input signal is divided into grains [1], which are extracted by applying a window function w starting at an arbitrary time instance n_i. In this work, we combine the ring-buffer mechanism described in Sec. 3.4 with granulation by modifying the read pointers. The position of the read pointer is updated every L_w samples to a randomly selected index behind the write pointer by $n_i = \mathcal{U}(L_w, L_{\text{ring}})$ samples, where $\mathcal{U}$ denotes the discrete uniform distribution and L_{ring} is the size of the ring buffer. Randomness can be increased by choosing $L_{\text{ring}} \gg L_w$. Each grain $g_i(n)$ is then extracted by applying a window $w(n)$:

$$g_i(n) = x(n + n_i)w(n) \tag{6}$$

for $n = 0, \ldots, L_w$. In both ring-buffer-based pitch-shifting methods, crossfading is performed using two sinusoidal functions over the interval $[0, \pi]$, with a $\phi = \pi/2$ phase shift between them. The choice of L_w controls the effect such that shorter grains produce a more textured sound, whereas longer grains yield a more clearly

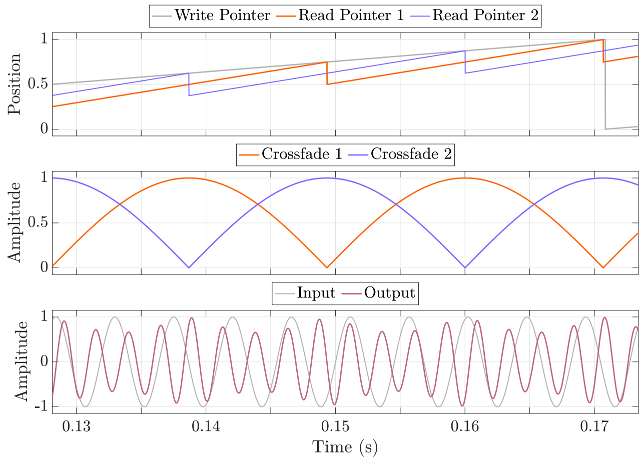


Figure 7: Operation of the pitch-shifter over two modulation periods: (top) normalized positions of the write and two read pointers; (middle) crossfade envelopes; (bottom) input and output signals. The input is a unit-amplitude sine wave at 22 Hz.

perceivable pitch-shifted effect. In the following, both the modulation period of the time modification in Sec. 3.4 and the grain size are denoted by L_w samples. Unless otherwise specified, the lengths in samples are $L_w = 2048$ and $L_{\text{ring}} = 8192$.

3.6. Operations Count

The computational cost of the proposed nonlinearities can be estimated by counting the number of floating-point multiplications required per processed sample. Additions, comparisons, and memory accesses are not considered in this analysis. The number of operations required by the FDN is reported in [15].

The CFWR nonlinearity requires eight operations, four for (3) and four for the antiderivative antialiasing approximation in (4). The SDFD nonlinearity requires four operations for the delays interpolation. Ring modulation is the least expensive method, requiring only two operations per sample. Time compression and expansion are computationally more demanding. The equal-power crossfade between the two read pointers requires two operations per sample. The cubic polynomial interpolation requires nine operations to compute the interpolation coefficients and three additional operations to evaluate the interpolator, yielding a total of 12 operations per sample. The dc-blocking filter (7), described in Section 4.1 and used together with CFWR and the ring-buffer-based methods, introduces two additional operations per sample.

4. ENERGY PRESERVATION AND STABILITY

In this section, we analyze the energy of the nonlinearities to assess whether and when they satisfy the energy preservation condition (2). This ensures that the system in Fig. 1b remains stable and identifies the conditions under which energy is lost, leading to a shortening of the reverberation time.

4.1. Energy at Zero Frequency

The response of the CFWR and, to a minor extent, of the ring-buffer-based time modifications exhibits a dc component (see Fig. 2), which can accumulate within the feedback structure and lead to instability. To prevent this, we follow the nonlinearities by a

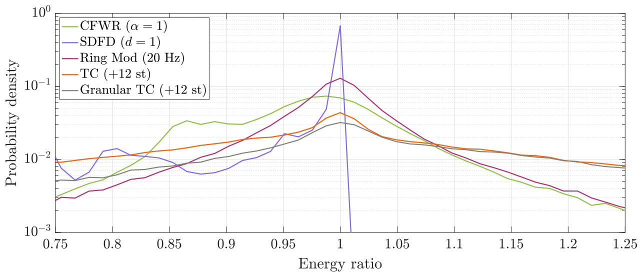


Figure 8: Distribution of the energy ratio of the nonlinearities.

dc-blocking filter characterized by the difference equation

$$y(n) = g_{\text{env}}(n)(x(n) - x(n-1) + Ry(n-1)), \tag{7}$$

where g_{env} is an energy compensation gain and R is a parameter that controls the pole location, and is typically $0.95 \leq R \leq 0.995$ [40]. The input signal $x(n)$ is the output of the nonlinearity. We use $R = 0.995$, corresponding to a cutoff frequency of 76 Hz.

Filtering out low-frequency components can lead to non-negligible energy loss, particularly for CFWR, causing the feedback signal to decay more rapidly than the desired T_{60}. To compensate for this loss, a slowly varying adaptive gain stage $g_{\text{env}}(n)$ is applied to the output of the dc blocker. The compensation gain is estimated from the instantaneous input–output energy ratio, $\frac{x(n)^2}{y(n)^2}$, using an exponentially smoothed envelope follower:

$$g_{\text{env}}(n) = \exp\left(-\frac{1}{f_s\tau}\right) g_{\text{env}}(n-1) + \left(1 - \exp\left(-\frac{1}{f_s\tau}\right)\right)\sqrt{\frac{x(n)^2}{y(n)^2}}, \tag{8}$$

where τ controls the envelope tracking time. To prevent rapid gain variations that could distort the waveform, we use $\tau = 50$ ms.

4.2. Energy Analysis

All of the nonlinearities presented in this work are level-invariant, except for the granular time modification, whose behavior is governed by the random process selecting the grains. Nonetheless, the total energy of its response scales linearly with the amplitude of the input signal. Therefore, the observations described below hold for input signals of any amplitude.

For the CFWR nonlinearity, the energy preservation property in (2) holds exactly for $\alpha = 0$ and $\alpha = 1$. For $0 < \alpha < 1$, the power varies quadratically, reaching a minimum at $\alpha = 0.5$. The compensation gain g_{cfwr} in (3) corrects this variation, ensuring nearly constant output energy across α.

In the SDFD nonlinearity, the positive and negative components of the input signals are shifted by at most one sample. As a result, zero-mean periodic signals can lose up to one sample due to the overlap between the delayed negative and positive components, as illustrated in Fig. 5. Thus, the SDFD cannot be considered strictly energy-preserving. The resulting attenuation depends on the input frequency and sampling rate, but it does not pose a risk to system stability.

The average power of a sinusoidal signal with unit amplitude is 0.5. Hence, to make the RM nonlinearity energy-preserving, the sinusoid must be scaled by $g_{\text{rm}} = \sqrt{2}$. However, this scaling does not guarantee stability in all cases. For fast and strongly recirculating delay lines, such as those with short lengths or coupled by a diagonally dominant feedback matrix, the instantaneous energy can grow rapidly, potentially leading to instabilities. In general, using lower amplitudes helps ensure the stability of the system.

The energy analysis of the time compression and expansion via ring buffer, Sec. 3.4 and 3.5, is non-trivial. With a power-complementary cross-fade envelope function combining the signals at the read positions, the energy is nearly preserved. For example, with a crossfade envelope based on a half-wave sinusoid (see middle panel of Fig. (7)), the two delay-line outputs are weighted over each modulation period, for $n = 0, \ldots, L_w - 1$, as

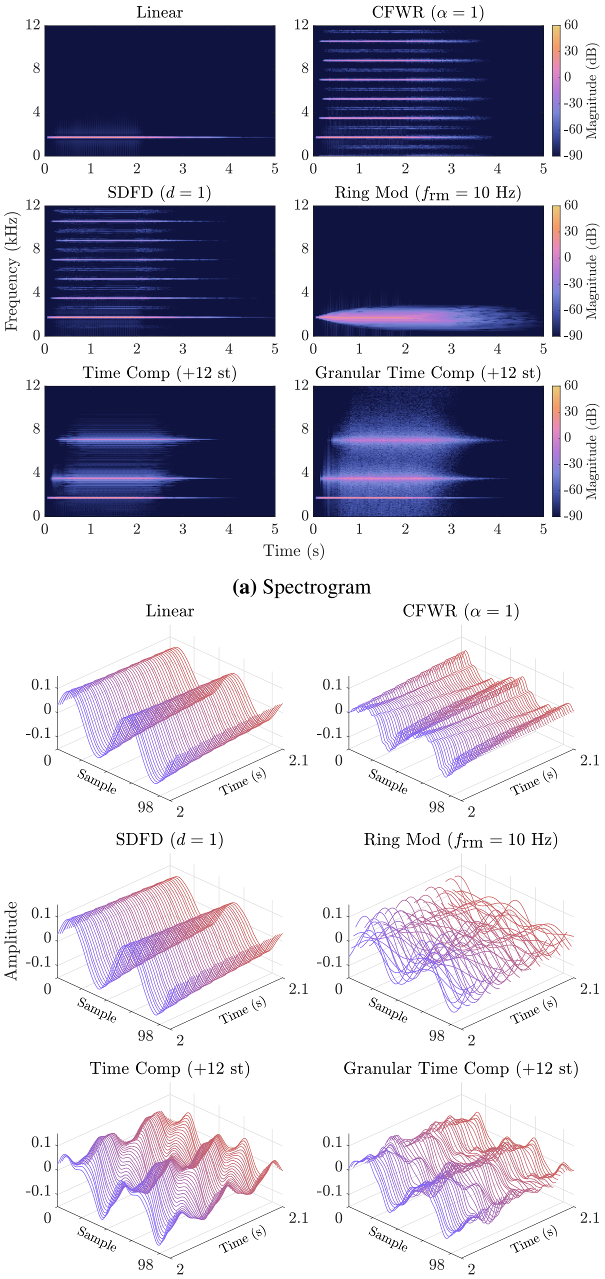
$$w_1(n) = \sin\left(\pi\frac{n}{L_w - 1}\right),$$
$$w_2(n) = \sin\left(\pi\frac{n}{L_w - 1} + \frac{\pi}{2}\right) = \cos\left(\pi\frac{n}{L_w - 1}\right),$$

which satisfy $w_1(n)^2 + w_2(n)^2 = 1$. However, reading the delay lines at a different rate alters the signal's energy density, breaking strict energy preservation. This is generally not problematic in typical FDN configurations, but in cases involving short and strongly recirculating delay lines, attenuation plays a more critical role in maintaining system stability.

To support the above discussion, Fig. 8 shows the distribution of the energy ratio $\sum_n |\text{NL}(x(n))|^2 / \sum_n |x(n)|^2$ across the considered nonlinearities. The resolution of the histogram from which the distribution was derived is 0.5 cents. The energy is computed over 100 ms windows extracted from a total of 336 minutes of music stems randomly sampled from the MoisesDB dataset [41]. Frames containing only silence were excluded. The energy ratio is generally concentrated around 1. Time compression exhibits the largest spread, which can be attributed to the inherent delay and the resulting mismatch in instantaneous energy. As discussed previously, SDFD attenuates high-frequency components, resulting in an asymmetric distribution skewed toward energy ratios below one. RM exhibit relatively broad but centered distributions, indicating approximate energy preservation with moderate variability. CFWR exhibits a pronounced peak around 0.87, indicating that the current energy compensation of the dc blocker is insufficient to achieve perfect energy preservation, and it could benefit from tuning τ to the dynamics of the input signal.

5. SHIMMER FEEDBACK DELAY NETWORK

We analyze the behavior of the nonlinearities within an FDN of size $N = 8$. We use a random orthogonal feedback matrix $\boldsymbol{U}$, and input and output gain vectors are defined as $\boldsymbol{b} = \boldsymbol{c} = \frac{1}{N}\boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix. The lengths of the delay lines are randomly chosen to be 71, 111, 235, 297, 307, 347, 381, and 400 ms and are sorted in ascending order, i.e., $m_1 = 71$ ms and $m_8 = 400$ ms. These delay lengths are relatively long compared to those typically used in realistic virtual acoustics applications, in order to visually and perceptually emphasize the progressive changes produced by successive recursions of the feedback loop. The attenuation filters are designed from a prototype first-order low-pass filter with crossover frequency $f_c = 10$ kHz, $T_{60}(0) = 2$ s, and $T_{60}(\pi) = 0.5$ s. CFWR, SDFD, and RM nonlinearities are applied to all delay lines. Ring-buffer-based time compressions corresponding to a one-octave pitch shift are applied to channels $i = 5, 6, 7, 8$. These channels have the longest delay lines, which results in a perceivable delay between successive octaves. CFWR and ring buffers are followed by the dc-blocker filter in (7).



(a) Spectrogram

(b) Temporal evolution

Figure 9: Response of the linear FDN and five nonlinear variations to a 1.76 kHz sinusoidal input with $f_s = 96$ kHz.

5.1. Spectrogram Analysis

Figure 9a shows the spectrograms of the responses to a 2-s 1.76 kHz sinusoid, with fade-in and fade-out envelopes generated using a Hann window of 0.4 s. The spectrograms are computed using 4096 frequency bins, a Hann window, and 75 % overlap. A large dynamic range is used to better visualize aliasing effects.

CFWR and SDFD exhibit similar responses: even harmonics appear after $2m_1$ samples, while odd harmonics, generated through recursion, appear after $3m_1$ samples. The harmonics generated by SDFD are generally weaker. As shown in Fig. 2, CFWR generates a stronger first harmonic, which leads to higher-amplitude subsequent harmonics in the FDN output. Aliasing artifacts, while visible between harmonics in the CFWR and SDFD responses, are considerably quieter than the main harmonics and can be partially mitigated by a higher-order attenuation filter, which reduces high-frequency content prior to nonlinear processing. The spectrogram also clearly illustrates the effect of RM in recursively spreading the input spectrum by $\pm f_{\mathrm{rm}}$. Both time-compression methods correctly generate higher octaves and do not cause significant aliasing artifacts. However, the pitch-shifting algorithm used in this work does not include a synchronization mechanism, and produces sidebands around the octave frequencies. The delay between octaves depends on the delay lines of the channels in which the operations are inserted. The ring-buffer introduces an additional initial delay of L_w samples before the input signal is read (see Fig. 7).

5.2. Temporal Evolution Analysis

Figure 9b shows the temporal evolution of the output waveforms, where the sample axis spans two periods of the input sinusoid. Only one in every two successive windows is displayed. The first 2 s are omitted to focus on the response after the input signal has decayed to silence.

Due to its less harmonically rich distortion, the SDFD better preserves the original waveform, producing an output that closely resembles that of the linear FDN outside the zero-crossing regions. In contrast, the stronger higher harmonics in the CFWR case have a more pronounced effect on the waveform, although the periodicity of the input signal remains clearly visible due to the feedforward path. The effect produced by RM is evident in the alternating amplitude pattern of the windows. The modulation period, $1/f_{\mathrm{rm}}$, is 176 times longer than the input period resulting in a repetition every 43 windows. However, this effect is not immediately apparent in Fig. 9b, as the visualization reflects the combined output of eight channels.

Time compression in the FDN preserves the period of the input signal while introducing higher octaves that add shorter-period modulations. Granular time compression, on the other hand, produces a noisier behavior due to its inherent stochastic nature, while still maintaining the original periodicity.

6. SHAPING THE SONIC CHARACTERISTICS

The sonic characteristics of the FDN and its nonlinear operations can be controlled through parameter selection, enabling a range of textures from smooth reverberation to harmonically rich, sustained echoes. This section provides intuition for the design and interpretation of these parameters. We categorize the effects into two groups: waveshaping and modulation (CFWR, SDFD, RM), and pitch shifting (time modification and granular time modification). Audio examples are available at the companion page [1].

6.1. Reverberation Control

The strength of the nonlinear effect is influenced by the mixing properties of the FDN. A large N and a dense feedback matrix produce rapid echo build-up and tend to excite nonlinearities more strongly, promoting a fast spread of the generated harmonics. This effect is particularly pronounced for short delay lines and may result in a distortion-like response. In contrast, if a more gradually

[1] http://research.spa.aalto.fi/publications/papers/dafx26-shimmer/

evolving effect is desired, longer delay lines and sparser feedback matrices should be used. Additionally, removing a subset of the nonlinear operations in Fig. 1b provides a simple means of controlling the overall strength of the effect. This should be considered in conjunction with the choice of delay-line length preceding the nonlinearity. The attenuation filters $\mathbf{\Gamma}$ can be designed to further accentuate the generated harmonics by assigning longer T_{60} values to the higher frequency range.

6.2. Wave Shaping and Modulation

CFWR and SDFD produce the most distortion-like effects due to the richness of their output spectra. A diagonally dominant feedback matrix, such as a Householder matrix, helps reduce inharmonicity and distortion, particularly for sustained input signals. For transient and percussive inputs, CFWR generates short, granular bursts with progressively increasing pitch. When $\alpha = 0.5$, CFWR behaves as a half-wave rectifier, producing more pronounced repetitions which, for percussive sounds, may be perceived as mechanical, e.g., resembling helicopter rotor noise. Values of $\alpha < 0.5$ generally result in a more subtle effect. SDFD typically produces a more moderate distortion and can therefore serve as an alternative to CFWR for sustained signals.

At low modulation frequencies, e.g., $f_{\text{rm}} \approx 10$ Hz, RM produces a tremolo-like effect, similar to amplitude modulation, which can result in pulsed or "chopped" reverberation for sustained signals. This effect can be further accentuated by using short delay lines, with lengths smaller than the modulation period, and a diagonally dominant $\boldsymbol{U}$. At higher f_{rm}, the effect no longer produces tremolo and generates audible inharmonic tones both above and below the original frequencies, creating a rough texture for sustained signals and metallic echoes for transient or short sounds.

6.3. Pitch Shifting

Pitch shifting via time modification produces the most "musical" effect among the considered nonlinearities. The current implementation allows each delay-line output to be shifted up or down by any number of semitones. In practice, musically meaningful results are safely obtained by octave shifts. Longer read-pointer modulation periods better preserve the original waveform and reduce distortion that would otherwise arise from more frequent pointer resets. However, this comes at the cost of increased latency before the effect becomes perceptually active and can produce undesired smearing when excited with short sounds. The delay-line lengths control the onset time of the shifted components. Shorter delays yield a response closer to a sweep, which may be desirable for sound design purposes. Granular time compression produces diffused, "blooming" textures, where longer buffer lengths increase randomness and spectral diffusion and add delay, extending the decay, resulting in a more sustained, reverberant response. Shorter buffer lengths, similar to large grain sizes, better preserve the original timbre in pitch-shifted echoes. For large L_w, the sinusoidal crossfade in Fig. 7 (middle), which is nonzero except at read-pointer resets, can be replaced with a piecewise-defined function. This reduces the transition duration between read pointers and allows for regions where only one pointer contributes, mitigating phasing.

6.4. Composite Nonlinearities

The proposed nonlinear FDNs can be extended by composition, i.e., by chaining operations of the same or different NL types within the feedback loop. This allows for a wider range of reverb textures while preserving overall energy. For example, RM combined with pitch shifting can introduce tremolo-like modulation to the shifted spectrum. Wave shaping can be applied before RM and pitch shifting to produce a warmer, richer sound with more spectrally smeared harmonics. Another degree of freedom lies in how these nonlinearities are arranged: instead of simple concatenation, different nonlinearities can be distributed within the same FDN, each applied to a subset of delay lines. These variations are also demonstrated on the companion webpage.

7. CONCLUSION

A shimmer reverberation using nonlinear operations directly into an FDN is proposed. The recursive structure of the FDN is used to generate a wide range of evolving harmonic and inharmonic textures. We explored five distinct nonlinear methods: controllable full-wave rectification, signal-dependent fractional delay, ring modulation, and two forms of time-based pitch shifting.

Through energy analysis, we established the conditions under which these nonlinearities are approximately energy-preserving and guarantee system stability. To mitigate potential instabilities and artifacts such as dc accumulation and aliasing, we implemented dc-blocking filters with energy compensation and first-order antiderivative antialiasing for the full-wave rectification.

We provide instructions to further tune the temporal evolution and texture of the reverb by adjusting FDN parameters such as delay-line lengths and feedback matrix density. The complementary audio examples help build intuition for designing the parameters governing the nonlinearities, and suggest the use of composite operation for more experimental sounds. Future work includes designing an adaptive attenuation filter that preserves a target T_{60} despite nonlinear energy redistribution, for example, through a differentiable implementation of the nonlinear FDN.

8. ACKNOWLEDGMENTS

The Aalto University School of Electrical Engineering funded the work of GDS. This work was supported by the HUCE infrastructure of the Aalto University School of Electrical Engineering.

9. REFERENCES

[1] U. Zölzer (Ed.), *DAFX Digital Audio Effects*. John Wiley & Sons, 2002.

[2] A. Huovilainen, "Non-linear digital implementation of the Moog ladder filter," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Oct. 2004, pp. 61–64.

[3] V. Välimäki and A. Huovilainen, "Oscillator and filter algorithms for virtual analog synthesis," *Computer Music J.*, vol. 30, no. 2, pp. 19–31, Jun. 2006.

[4] V. Lazzarini and J. Timoney, "A digital model for the Prologue voltage control filter," *J. Audio Eng. Soc.*, vol. 72, no. 1/2, pp. 79–88, Jan. 2024.

[5] J. R. Pierce and S. A. Van Duyne, "A passive nonlinear digital filter design which facilitates physics-based sound synthesis of highly nonlinear musical instruments," *J. Acoust. Soc. Amer.*, vol. 101, no. 2, pp. 1120–1126, Feb. 1997.

[6] V. Välimäki, T. Tolonen, and M. Karjalainen, "Signal-dependent nonlinearities for physical models using time-varying fractional delay filters," in *Proc. Int. Computer Music Conf. (ICMC)*, Ann Arbor, MI, USA, Oct. 1998.

[7] C. M. Cooper and J. S. Abel, "Digital simulation of "brassiness" and amplitude-dependent propagation speed in wind instruments," in *Proc. 13th Int. Conf. Digital Audio Effects (DAFx)*, Sep. 2010.

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

[9] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. AES Conv. 90*, Feb. 1991.

[10] S. J. Schlecht, "Feedback delay networks in artificial reverberation and reverberation enhancement," Ph.D. dissertation, FAU Erlangen-Nuernberg, Germany, 2018.

[11] R. Giampiccolo, A. I. Mezza, M. Pezzoli, S. Koyama, A. Bernardini, and F. Antonacci, "Modeling the impulse response of higher-order microphone arrays using differentiable feedback delay networks," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Sep. 2025, pp. 180–187.

[12] O. Das, G. Dal Santo, S. J. Schlecht, V. Välimäki, and Z. Cvetković, "Differentiable grouped feedback delay networks for learning coupled volume acoustics," *IEEE Trans. Audio, Speech and Lang. Process.*, vol. 34, Feb. 2026.

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

[14] V. Välimäki, K. Prawda, and S. J. Schlecht, "Two-stage attenuation filter for artificial reverberation," *IEEE Signal Process. Lett.*, Jan. 2024.

[15] G. Dal Santo, K. Prawda, S. J. Schlecht, and V. Välimäki, "Optimizing tiny colorless feedback delay networks," *EURASIP J. Audio Speech Music Process.*, vol. 2025, no. 13, Mar. 2025.

[16] P. Götz, G. Dal Santo, S. J. Schlecht, V. Välimäki, and E. A. P. Habets, "Matching reverberant speech through learned acoustic embeddings and feedback delay networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, May 2026.

[17] S. Costello, "ValhallaShimmer: The manual!" https://valhalladsp.com/2011/01/24/valhallashimmer-the-manual, [Online] Accessed: Jan. 15, 2026.

[18] Eventide, "Blackhole: Massive otherworldly reverb," www.eventideaudio.com/plug-ins/blackhole, [Online] Accessed: Jan. 15, 2026.

[19] ——, "ShimmerVerb: Ethereal reverb + pitch shifting," store.eventideaudio.com/products/shimmerverb, [Online] Accessed: Jan. 15, 2026.

[20] Soundtoys, "Crystallizer: Granular echo synthesizer," www.soundtoys.com/product/crystallizer/, [Online] Accessed: Jan. 15, 2026.

[21] Kentaro, "Particle-reverb," https://kentaro.tools/particle-reverb, [Online] Accessed: Mar. 10, 2026.

[22] H. Kuttruff, *Room Acoustics, Fifth Edition*. CRC Press, 2009.

[23] K. Prawda, "Sensitivity of room impulse responses in changing acoustic environment," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2025.

[24] J. S. Abel and K. J. Werner, "Distortion and pitch processing using a modal reverberator architecture," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Nov. 2015, pp. 95–102.

[25] J. S. Abel, S. Coffin, and K. Spratt, "A modal architecture for artificial reverberation with application to room acoustics modeling," in *AES Conv. 137*, Oct. 2014.

[26] J. Parker, S. Bilbao, H. Penttinen, and J. Abel, "Modelling methods for the highly dispersive slinky spring: A novel musical toy," in *Proc. Int. Conf. Digital Audio Effects*, Sep. 2010.

[27] S. Fan and S. Smith, "Integrating vibrato into artificial reverberation," in *Proc. AES Conv. 151*, Oct. 2021.

[28] V. Norilo, "Exploring the vectored time variant comb filter," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Sep. 2014, pp. 65–68.

[29] J. Dattorro, "Effect design, part 1: Reverberator and other filters," *J. Audio Eng. Soc.*, vol. 45, no. 9, pp. 660–684, Sep. 1997.

[30] S. J. Schlecht and E. A. Habets, "Time-varying feedback matrices in feedback delay networks and their application in artificial reverberation," *J. Acoust. Soc. Amer.*, vol. 138, no. 3, pp. 1389–1398, Sep. 2015.

[31] X. Pi, "Audio processing using nonlinear feedback systems," Master's thesis, Aalto University, Espoo, Finland, 2025.

[32] F. Esqueda, V. Välimäki, and S. Bilbao, "Rounding corners with BLAMP," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Sep. 2016, pp. 121–128.

[33] J. D. Parker, V. Zavalishin, and E. Le Bivic, "Reducing the aliasing of nonlinear waveshaping using continuous-time convolution," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Sep. 2016, pp. 137–144.

[34] S. Bilbao, F. Esqueda, J. D. Parker, and V. Välimäki, "Antiderivative antialiasing for memoryless nonlinearities," *IEEE Signal Processing Letters*, vol. 24, no. 7, pp. 1049–1053, Mar. 2017.

[35] C. Roads, *The Computer Music Tutorial*. MIT Press, 1996.

[36] F. F. Lee, "Time compression and expansion of speech by the sampling method," *J. Audio Eng. Soc.*, vol. 20, no. 9, pp. 738–742, Nov. 1972.

[37] M. Puckette, "Time shifts," in *The Theory and Technique of Electronic Music*. World Scientific, 2007.

[38] A. Haghparast, H. Penttinen, and V. Välimäki, "Real-time pitch-shifting of musical signals by a time-varying factor using normalized filtered correlation time-scale modification (NFC-TSM)," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Bordeaux, France, Sep. 2007, pp. 10–15.

[39] J. Laroche, "Time and pitch scale modification of audio signals," in *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, Eds. Springer, 1998, pp. 323–351.

[40] J. O. Smith, III, "DC blocker," in *Introduction to Digital Filters: With Audio Applications*. W3K Publishing, 2007, Accessed: Feb. 25, 2026. [Online]. Available: ccrma.stanford.edu/~jos/fp/DC_Blocker.html

[41] I. Pereira, F. Araújo, F. Korzeniowski, and R. Vogl, "MoisesDB: A dataset for source separation beyond 4-stems," 2023, arXiv preprint 2307.15913.