SPECTRAL ANALYSIS OF STOCHASTIC WAVETABLE SYNTHESIS

Nicholas Boyko

New York University
New York, USA
nboyko@nyu.edu

Elliot K. Canfield-Dafilou

New York University
New York, USA
e.canfield.dafilou@nyu.edu

ABSTRACT

Dynamic Stochastic Wavetable Synthesis (DSWS) is a sound synthesis and processing technique that uses probabilistic waveform synthesis techniques invented by Iannis Xenakis as a modulation/ distortion effect applied to a wavetable oscillator. The stochastic manipulation of the wavetable provides a means to creating signals with rich, dynamic spectra. In the present work, the DSWS technique is compared to other fundamental sound synthesis techniques such as frequency modulation synthesis. Additionally, several extensions of the DSWS technique are proposed.

1. INTRODUCTION

Wavetable and direct synthesis are two popular approaches to digital sound generation [1]. Direct synthesis utilizes a phase accumulator to directly construct a digital signal, while wavetable synthesis utilizes a lookup table holding a single cycle of a periodic waveform where the rate of indexing into the table controls the fundamental frequency of the output signal. Recently, Radna proposed Dynamic Stochastic Wavetable Synthesis (DSWS) as an extension to Iannis Xenakis's direct synthesis technique Dynamic Stochastic Synthesis (DSS) [2, 3]. It extends the basic principles of DSS by applying dynamic, probabilistic alterations of audio sample values to a table-lookup oscillator, effectively transforming the DSS method from a pure synthesis technique into an audio processing algorithm for arbitrary waveforms. DSWS uses the principles of DSS to apply pitch and amplitude deviations to a signal by altering the table-lookup process utilized in wavetable synthesis.

The compositional utility of such a tool then lies in its dynamic behavior across time, which is defined by the chosen parameters and probability distributions. As such, it is useful to model and analyze the expected change in power spectra and bandwidth of a DSWS signal with respect to its unique parameters and probability distributions. Because DSWS employs the use of stochastic processes to manipulate the sound, non-stationarities are introduced to an audio signal through the parameters of the technique. There has been some analysis on the behavior of the original DSS procedure itself [4–7], but its extension to a more general audio processing approach warrants further exploration and characterization. To model the characteristics of the DSWS procedure, we compare power spectral density (PSD) across parameters and approaches.

Copyright: © 2024 Nicholas Boyko et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, adaptation, and reproduction in any medium, provided the original author and source are credited.

2. PROBABILISTIC SOUND SYNTHESIS

Probabilistic techniques have been used in sound synthesis and computer music composition for several decades. Notable composers such as John Cage and Iannis Xenakis were at the forefront, employing indeterminate and stochastic processes in their compositions. For example, Cage's *Music of Changes* used the *I Ching* in the compositional process to determine aspects of the sound such as pitch and duration [8]. In his composition *HPSCHD*, a FORTRAN program generated random changes for parameters of synthesized audio [9]. Xenakis often drew inspiration from natural processes, coining the term "Stochastic music" and using mathematical models of natural systems, such as the Wiener process for modeling Brownian motion or the Maxwell-Boltzmann kinetic theory of gases [3]. Xenakis was concerned with the concept of musical "density," across both time and frequency domains. This concept of density led to the development of granular synthesis, a technique that segments an audio signal into small "grains" to create signals with varied timbre while maintaining some spectral characteristics of the source audio [10].

2.1. Dynamic Stochastic Synthesis

Dynamic Stochastic Synthesis, described in Xenakis's book *Formalized Music*, generates a single cycle of a periodic waveform by linearly interpolating a set of breakpoints generated randomly in the time-domain [3]. The position of the breakpoints in amplitude-time space is altered in each successive period of the wave, creating a dynamic, varied signal with fluctuating timbre and spectra. DSS has been extended in recent years, encompassing both real-time implementations [4, 6, 11] and theoretical and compositional analyses of the technique [5, 7].

2.2. Dynamic Stochastic Wavetable Synthesis

The DSWS procedure proposed by Radna is summarised here, while a more detailed treatment can be found in [2].

2.2.1. Table Lookup and Deviations

In DSWS, audio samples corresponding to one period of a waveform are stored in a lookup table and indexed circularly with a phasor to generate an audio stream. With each iteration reading through the wavetable, the amplitude values of the stored sample data and the speed of the read phasor are modified according with stochastic processes, affecting the pitch and and spectral content of the output. In the present work, sinusoidal and bandlimited sawtooth waves are used in the wavetable, however arbitrary initial waveforms can also be used.

The DSWS approach divides the wavetable into M segments. Pitch deviations are applied by altering the lookup phase increment φ according to a random walk at each segment index m and linearly interpolating between existing sample values. As in the original implementation, pitch deviation values are generated with a MIDI value f_{MIDI}, and converted to frequency according to

$$f_{\text{Hz}} = 440\,\text{Hz} \cdot 2^{(f_{\text{MIDI}}+P[m]-69)/12} \tag{1}$$

where $P[m]$ is the pitch modulation value at index m. Because of this, pitch deviations are applied in a perceptually symmetric manner around the center frequency, which may subjectively be considered preferable.

The amplitude deviations use an anti-clipping wavefolder approach, increasing or decreasing output values $y[n]$ by the amount d by which it lies outside of the range $[-1, 1]$, according to

$$y[n] = \begin{cases} 1-d, & x[n]+a > 1 \\ x[n]+a, & -1 \leq x[n]+a \leq 1 \\ -1+d, & x[n]+a < -1 \end{cases} \tag{2}$$

where d is given by

$$d = |x[n]+a| - 1. \tag{3}$$

The amplitude deviations a are also set using a random walk, and applied for each segment m.

2.2.2. *Random Walks and Probability Distributions*

The pitch and amplitude deviations are generated by second-order random walks with uniform probability distribution and stored in lists of length M. The generation of these random walks is a key component of Xenakis's original DSS algorithm, as utilized compositions such as GENDY3 [12]. Values are generated within some variable bounds $[-\alpha, \alpha]$, with each deviation value $x[n]$ taking the form

$$x[n] = x[n-1] + \Delta[n] \tag{4}$$

where $\Delta[n]$ is given in the range $[-\beta, \beta]$ with $\beta \leq \alpha$, according to the probability distribution $P(\Delta)$. Thus β defines the step size of the random walk. We utilize both continuous and discrete probability distribution functions, namely normal, uniform, and Cauchy continuous distributions (see Fig. 1); and Poisson and "drunk walk" distributions. The "drunk walk" is a special case of a Bernoulli distribution, with equal probability of a value of ± 1.

In our implementation, the uniform distribution was bounded on the range $[-1, 1]$, and the other functions were designed so that $\geq$99.9% of values fell within the range of $[-1, 1]$. These probability distribution functions are then utilized for the generation of both first-order and second-order random walks, which are stored and read as deviation values.

3. ANALYSIS

To analyze DSWS, we utilized the power spectral density computed via Welch's method for several output signals [13].[1] This calculates an average of periodograms across time, which helps to smooth non-stationarities and account for dynamic random behavior in the signal. Power spectral density is thus a natural metric for the analysis for probabilistically altered or generated signals, such

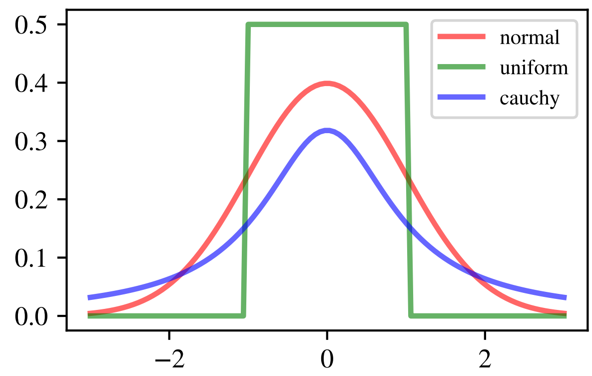

Figure 1: *The probability density functions of normal, uniform, and Cauchy probability distributions.*

as the output of DSWS, as it allows for a more holistic view of the expected value of a stochastic process. It is important to note that PSD is not useful for temporal analysis. Several parameters were evaluated. We first characterize the stochastic alterations afforded by DSWS. In Fig. 2, the Bernoulli distribution "drunk walk" is used to generate the pitch deviations for a sinusoidal signal at a frequency of 10 kHz.[2] The center frequency has constant power across variable pitch deviation values, while the outer bandwidth increases in direct correlation to the pitch deviation values.

We compare this behavior to a pitch deviation table generated with a regular frequency deviation (i.e., sinusoidally modulating the frequency parameter) in Fig. 3. A similar spectral envelope is apparent in this figure, with side lobes appearing at integer multiples of the modulation rate above and below the center frequency. Greater bandwidth expansion is again apparent with larger pitch deviation values. This behavior can also be compared to the spectrum created through direct frequency modulation (FM) synthesis.

The output $y[n]$ of a signal generated via frequency modulation synthesis can be written

$$y[n] = \sin(\omega_c n + A_m \sin(\omega_m n)) \tag{5}$$

with ω_c, ω_m the carrier and modulator frequencies, and A_m the amplitude of the modulator, or modulation index. In the case of both FM synthesis and sinusoidal DSWS pitch deviation, a constant sinusoidal signal modulates the pitch of the input signal, causing bandwidth expansion with peaks at integer multiples of the modulator frequency ω_m about the carrier frequency ω_c:

$$\omega_c + k\omega_m, \quad k \in \mathbb{Z}. \tag{6}$$

With frequency modulation, the amplitudes of the side lobes are related to Bessel functions of the first kind [14], and are affected by the modulation index A_m, as seen in Fig. 4. In the sinusoidal DSWS pitch deviation, the bandwidth expansion is more concentrated around the center frequency, but contains frequency content across the entire spectrum, as seen in Fig. 3.

The bandwidth expansion effect of the pitch deviations is also visible on complex waveforms, as in Fig. 5, in which a band-limited sawtooth wave has pitch deviations applied according to

[1] Audio examples and Python code associated with the figures from this paper can be found online at https://ccrma.stanford.edu/∼kermit/website/dsws.html.

[2] This frequency was chosen for visualisation purposes, to be far from both DC and the Nyquist limit for a sample rate of 44.1 kHz.

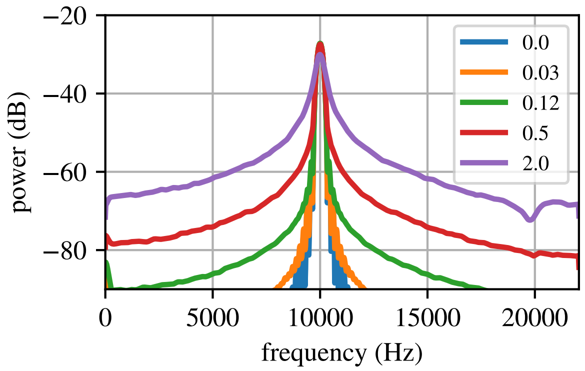

Figure 2: *PSD for DSWS using a* 10 kHz *sinusoidal wavetable and stochastic pitch deviations according to a Bernoulli probability distribution and variable pitch deviation amount.*

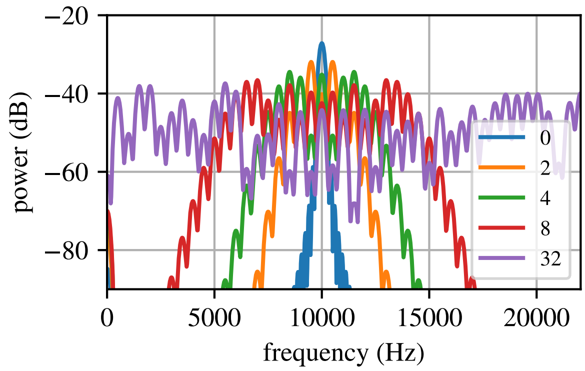

Figure 4: *PSD for FM synthesis with a carrier frequency of* 10 kHz*, modulation frequency of* 500 Hz*, and modulation index (controlling the bandwidth expansion) ranging from* 0 *to* 32*.*

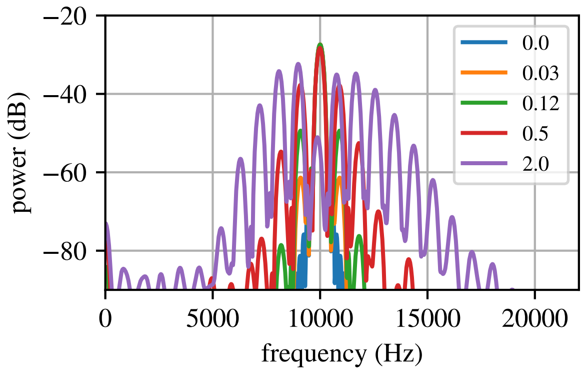

Figure 3: *PSD for DSWS using a* 10 kHz *sinusoidal wavetable and pitch deviations set through sinusoidal modulation with a frequency of* 500 Hz *and variable pitch deviation amount.*

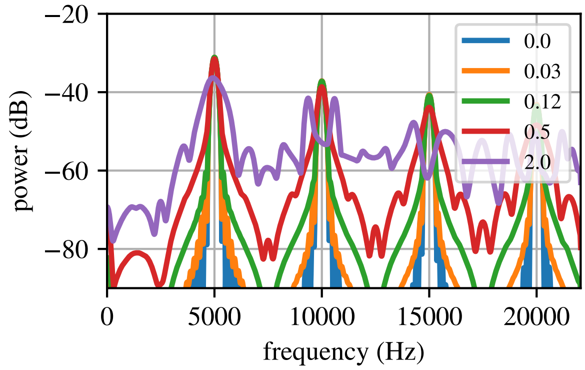

Figure 5: *PSD for DSWS using a* 5 kHz *band-limited sawtooth wavetable and stochastic pitch deviations according to a Bernoulli probability distribution and variable pitch deviation amount.*

the same Bernoulli probability distribution. In the harmonically rich spectrum case, the DSWS causes bandwidth expansion around each spectral peak. As the pitch modulation amount increases, we see an increase in low-frequency content in addition to the blending of adjacent spectral peaks.

The amplitude deviation parameter behaves similarly across probability distributions, with the resultant signals visualized in Fig. 6. A larger maximum amplitude deviation value results in a near-constant increase in power at frequencies outside of the main lobe, with little effect on the center frequency itself. Additionally, the increase in power is directly correlated with an increase in the amplitude deviation amount itself. Due to the fold-over distortion, additional spectral peaks are introduced.

The spectra of each probability distribution are also compared in single-segment mode with respect to pitch deviation at a maximum of one semitone above and below the center frequency, as seen in Fig. 7. The bandwidths of each distribution are mostly similar, with slight differences in the extreme spectra, likely associated with the heaviness of the tails of each probability distribution. In addition, the Poisson distribution stands out due to its skewness, shifting the main lobe towards lower frequencies. The two independent lobes on either side of the center frequency are due to the use of a bounded second-order random walk for the pitch deviations, whose values spend most of their time towards the boundaries. As the maximum allowed pitch deviation is ± 1 semitone, the lobes have peaks at roughly 1 kHz above and below the center frequency of 10 kHz.

4. FUTURE WORK

There are several potential avenues for further extension of the DSWS technique. For instance, the implementation and choice of probability distributions could be further considered and extended. One example of another stochastic process favored by Xenakis is the Wiener process for modeling Brownian motion, which has already been used in an extension to the DSS algorithm [15]. Additionally, the skewness of certain probability distributions (e.g. logistic, Poisson) is another potential variable for altering the spectral content of the output signal. The analyses presented in this work suggest that a skewed distribution would present asymmetric

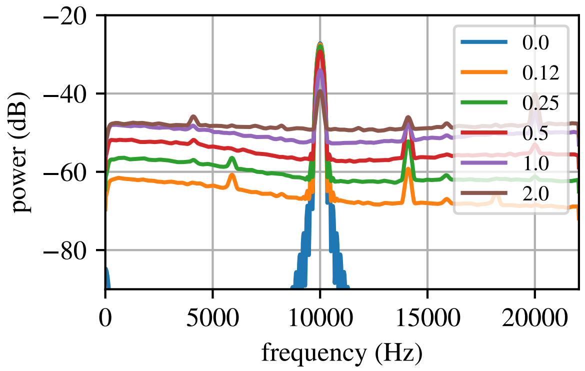

Figure 6: *PSD for DSWS using a* 10 kHz *sinusoidal wavetable and stochastic amplitude deviations with variable amplitude deviation amount.*

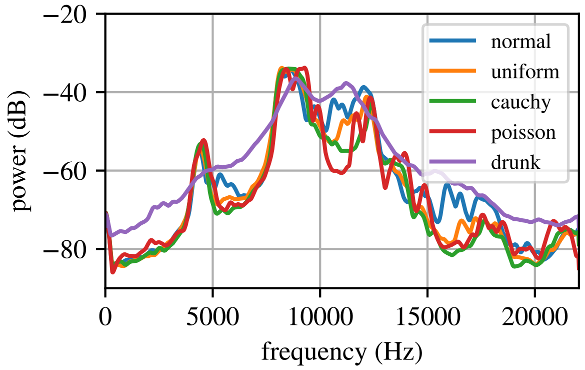

Figure 7: *PSD for DSWS using a* 10 kHz *sinusoidal wavetable and stochastic pitch deviations according to each probability distribution, with a maximum deviation of one semitone.*

bandwidth expansion in the output signal.

As described by Radna, segmenting the wavetable also affects the bandwidth of the signal, with a larger number of segments producing a brighter sound. In our implementation, segments are generated with evenly distributed divisions of the wavetable, but generating segments randomly or according to points of large differential in the slope of the signal could provide unique behavior.

Lastly, further steps to prevent aliasing could be investigated. A naive 8x oversampling approach with an anti-aliasing lowpass filter was able to remove aliased peaks in the spectrum created by amplitude deviation, such as those seen in Fig. 6. Rather than over sampling, it could be beneficial to smooth the pertubed wavetable using the method presented in [16].

5. CONCLUSIONS

In the present work, we have evaluated the characteristic bandwidth expansion of the Dynamic Stochastic Wavetable Synthesis technique with respect to several probability distribution functions. The behavior of DSWS presents a novel alternative to classical non-linear synthesis techniques such as frequency modulation (FM) synthesis as well as direct digital synthesis techniques, while maintaining real-time performance and efficiency. Additionally, as an audio processing unit, DSWS presents itself as a versatile method for introducing stochastic volatility into an audio stream while maintaining timbral characteristics of the source audio. Finally, we hope this analysis of the DSWS algorithm parameter space will inspire others when describing new sound synthesis and processing techniques.

6. REFERENCES

[1] Robert Bristow-Johnson, "Wavetable synthesis 101, a fundamental perspective," in *Proc. Audio Eng. Conv.*, 1996.

[2] Raphael Radna, "Dynamic stochastic wavetable synthesis," in *Proc. Int. Conf. Digital Audio Effects*, 2023.

[3] Iannis Xenakis, *Formalized music: thought and mathematics in composition*, Pendragon Press, revised edition, 1992.

[4] Peter Hoffmann, "Implementing the dynamic stochastic synthesis," in *Journées d'Informatique Musicale*, 1996.

[5] Peter Hoffmann, ""something rich and strange": Exploring the pitch structure of GENDY3," *J. New Music Research*, vol. 33, no. 2, pp. 137–144, 2004.

[6] Andrew Brown, "Extending dynamic stochastic synthesis," in *Proc. Int. Comp. Music Conf.*, 2005, pp. 111–114.

[7] Sergio Luque, "Stochastic synthesis: Origins and extensions," *M.M. thesis, Inst. Sonology, Roy. Conservatory, The Netherlands*, 2006.

[8] John Cage, "To describe the process of composition used in music of changes and imaginary landscape no. 4," *Silence: Lectures and Writings*, pp. 57–59, 1961.

[9] Larry Austin, "HPSCHD," *Comp. Music J.*, vol. 28, no. 3, pp. 83–85, 2004.

[10] Sergio Cavaliere and Aldo Piccialli, "Granular synthesis of musical signals," in *Musical Signal Process.*, pp. 155–186. Routledge, 2013.

[11] Nick Collins, "Implementing stochastic synthesis for supercollider and iphone," in *Proc. Xenakis Int. Symp.*, 2011.

[12] Marie-Hélène Serra, "Stochastic composition and stochastic timbre: Gendy3 by Iannis Xenakis," *Perspectives of New Music*, pp. 236–257, 1993.

[13] Peter Welch, "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, 1967.

[14] John M Chowning, "The synthesis of complex audio spectra by means of frequency modulation," *J. Audio Eng. Soc.*, vol. 21, no. 7, pp. 526–534, 1973.

[15] Emilio L Rojas and Rodrigo F Cádiz, "A physically inspired implementation of Xenakis's stochastic synthesis: Diffusion dynamic stochastic synthesis," *Comp. Music J.*, vol. 45, no. 2, pp. 48–66, 2021.

[16] Kurt J. Werner and Emma Azelborn, "Antialiasing piecewise polynomial waveshapers," in *Proc. Int. Conf. Digital Audio Effects*, 2023.