

Sonically Optimized Noise Shaping Techniques

Second edition

Alexey Lukin
lukin@ixbt.com

2002

In this article we compare different commercially available noise shaping algorithms and introduce a new highly optimized frequency response curve for noise shaping filter implemented in our ExtraBit Mastering Processor.

Contents

INTRODUCTION	2
TESTED NOISE SHAPERS	2
CRITERIONS OF COMPARISON.....	3
RESULTS.....	4
NOISE STATISTICS & PERCEIVED LOUDNESS	4
NOISE SPECTRUMS	5
No noise shaping (TPDF dithering).....	5
Cool Edit (C1 curve).....	5
Cool Edit (C3 curve).....	6
POW-R (pow-r1 mode).....	6
UV22	6
Sound Forge (hi-pass triangular dither + equal loudness noise shaping contour)	6
L2 (“Normal” preset).....	7
L1-Ultramaximizer (“Normal” preset).....	7
POW-R (pow-r2 mode).....	7
Sony Super Bit Mapping.....	7
WaveLab (Type3 curve)	8
Cool Edit (44.1 KHz curve).....	8
L2 (“Ultra” preset).....	8
POW-R (pow-r3 mode).....	8
ExtraBit (“Medium+LPDN” preset).....	9
L1-Ultramaximizer (“Ultra” preset).....	9
ExtraBit (“Ultra+LPDN” preset).....	9
NOISE SAMPLES	9
RESULT ANALYSIS.....	10
SUMMARY.....	11
FEEDBACK	11
ACKNOWLEDGEMENTS.....	11

Introduction

A compact disk has become the most common medium for distribution of a high-quality audio. In spite of many drawbacks of its 44100 Hz, 16-bit digital audio format, there are various techniques to improve its performance. A quantization operation is frequently used to reduce the word length of digital audio data before placing it on a compact disk from original 24-bit master recording. A common practice in CD mastering is to use dithering and noise shaping processes to prevent correlation of quantization noise with audio waveform and to reduce the perceived amount of added noise respectively. In general, the noise shaping algorithm is standardized, so the differences between various techniques are in the curve of noise shaping (which is the frequency response of feedback filter). In this paper we analyze some popular commercially available noise shapers and compare their performance using several simple criteria.

Tested noise shapers

We have tested the following popular commercially available noise shapers:

- **L1-Ultramaximizer** by Waves,
- **L2** processor by Waves,
- **UV22** algorithm by Apogee (which is not a traditional noise shaper in fact. It uses dithering with a specially generated HF-noise and doesn't employ standard noise shaping),
- **POW-R** algorithm by POW-R consortium LLC (including POW-r1, POW-r2, and POW-r3 modes),
- **Sony Super Bit Mapping** system by Sony,
- Several noise shaping curves integrated into **Cool Edit Pro** v1.2a by Syntrillium,
- Several noise shaping curves integrated into **Sound Forge** v5.0 by Sonic Foundry,
- Several noise shaping curves integrated into **WaveLab** v3.04 by Steinberg,
- Our **ExtraBit Mastering Processor**. We propose it as optimal solution for noise shaping.

Criteria of comparison

There are several criteria for estimating performance of noise shapers. First of them is the **degree of distortion elimination** from 16-bit quantized recording. This parameter depends on amount of dither noise added to the recording before noise shaping and on correctness of noise shaping feedback filter (it must preserve the average spectral power of noise). We assume that a standard TPDF dither noise with a maximal peak-to-peak value of 2 LSB is suitable for high-quality audio purposes. It eliminates distortion and noise modulation by making the first two statistical moments of the error signal independent from the input audio recording. Lower amplitudes of dither noise usually produce unwanted noise modulation (noise level depends on the input signal) or even harmonic distortion. If some noise shaping system meets our criterion for amount of dither noise, then we can analyze it just by examining the properties of noise produced by noise shaper at the absence of audio signal (because the output signal of noise shaper is a sum of a high-resolution input signal and a constant quantization noise, which doesn't sonically depend on input signal).

The second criterion is the **perceived loudness of quantization noise**. The primary task for noise shaper is to reduce this loudness by shifting the spectrum of dithering noise to the frequency bands where the ear's sensitivity is relatively low. It is important to notice that the loudness must be estimated at the lowest listening volume, when the noise becomes barely audible because we intend to make the noise completely inaudible at reasonable listening levels. We must emphasize that the loudness is a subjective value because it depends not only on parameters of noise and listening conditions but also on individual threshold-of-hearing curves. We used various listening conditions, various playback devices (such as headphones, hi-fi stereo systems, and studio monitors), and a group of people of different ages to evaluate the loudness of noise samples.

The third criterion is the **power of output shaped quantization noise**. Each noise shaper changes the white spectrum of TPDF dithering noise at the expense of increasing the total noise power. Usually the noise is shifted into high-frequency bands, so at the output we get a high-frequency noise with a significant power. If the power of high-frequency noise becomes too high, it can damage some audio gear (e.g. tweeters at high listening levels). Also high power HF-noise can degrade the quality of interpolation of audio data in CD-players, which cannot correctly read digital data from CD medium and try to interpolate invalid samples. There is no standard for the maximal dithering noise power but we propose to use a reasonable threshold of -60 dB FS for noise peak values. It prevents most VU meters with an amplitude range of 60 dB from flickering when playing noise shaped recordings.

Results

Here we present spectrums of quantization noise from different noise shapers and some other properties of noise samples. At the next paragraph we will analyze the results and give some comments on the performance of tested noise shapers.

Noise statistics & perceived loudness

At the table below you can see some measured properties of dither noise samples produced by different tested word length reduction systems. The algorithms are sorted according to their sonic performance that was estimated from the perceived loudness of dither noise (compared to the loudness of a standard TPDF dither noise). We should emphasize that loudness of dither noise is a subjective value and depends on individual hearing thresholds and listening conditions. The figures below show some average loudness obtained from a group of people.

Algorithm	Noise peaks, dB FS	Noise RMS, dB	Noise loudness, dB
ExtraBit (<i>“Ultra+LPND” preset</i>)	-61	-71	-12
L1-Ultramaximizer (<i>“Ultra” preset</i>)	-67	-79	-8
ExtraBit (<i>“Medum+LPDN” preset</i>)	-84	-89	-7
POW-R (<i>pow-r3 mode</i>)	-66	-78	-7
L2 (<i>“Ultra” preset</i>)	-65	-78	-7
Cool Edit (<i>44.1 KHz curve</i>)	-62	-75	-7
WaveLab (<i>Type3 curve</i>)	-60	-73	-7
Sony Super Bit Mapping	-81	-92	-6
POW-R (<i>pow-r2 mode</i>)	-77	-89	-6
L1-Ultramaximizer (<i>“Normal” preset</i>)	-75	-87	-6
L2 (<i>“Normal” preset</i>)	-74	-88	-6
Sound Forge (<i>hi-pass triangular dither + equal loudness noise shaping contour</i>)	-68	-80	-6 [*]

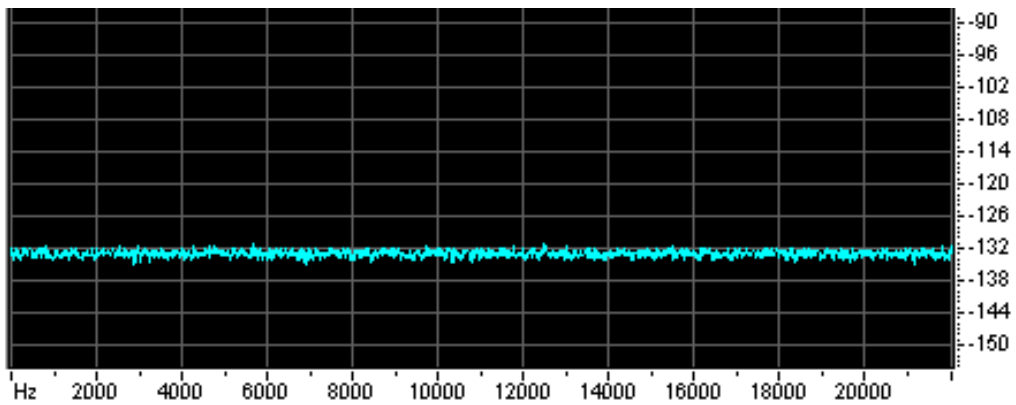
UV22	-78	-84	-5
POW-R (<i>pow-r1 mode</i>)	-75	-83	-5
Cool Edit (<i>C3 curve</i>)	-75	-87	-4
Cool Edit (<i>C1 curve</i>)	-75	-86	-2
No noise shaping (<i>TPDF dithering</i>)	-90	-96	0

[*] - This noise shaper does not produce correct result because of insufficient dithering amount. The resulting noise appears modulated by audio signal. All the noise shapers from Sound Forge have this problem, so for our testing we chose only the most powerful preset.

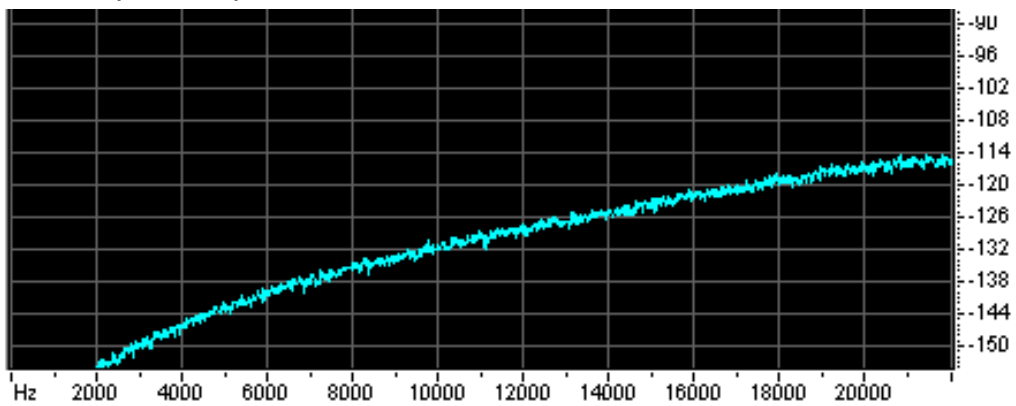
Noise spectrums

Spectrums of noise generated by these noise shapers look as following:

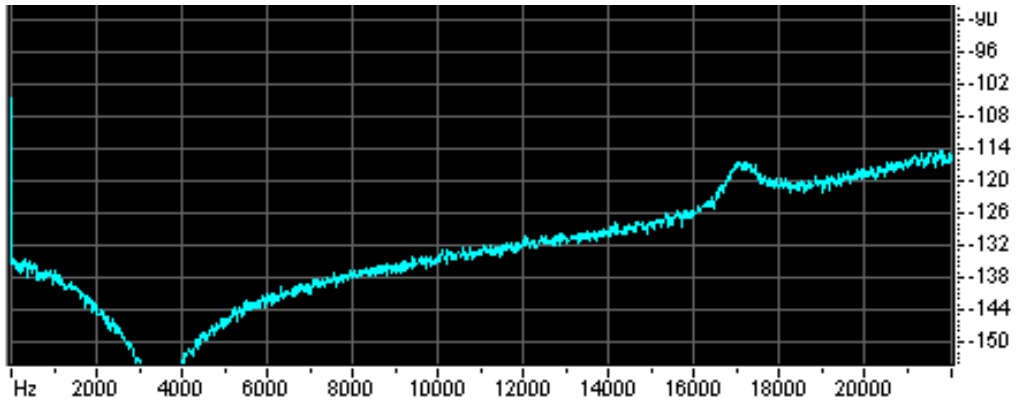
No noise shaping (TPDF dithering)



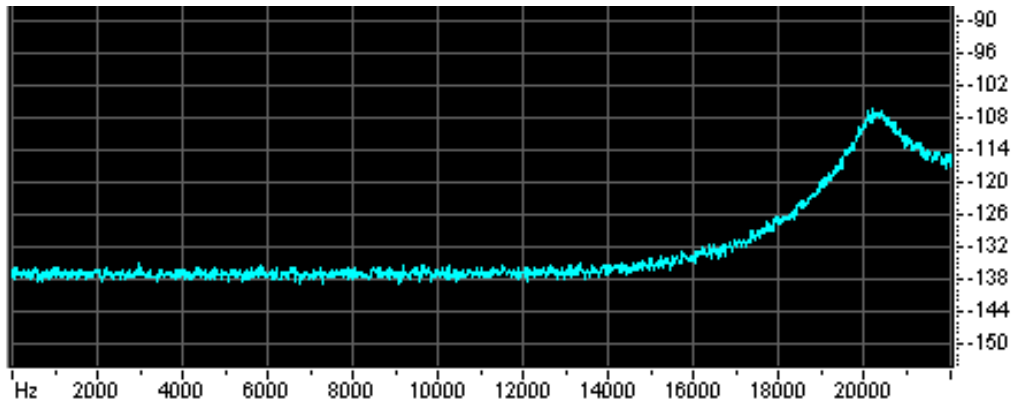
Cool Edit (C1 curve)



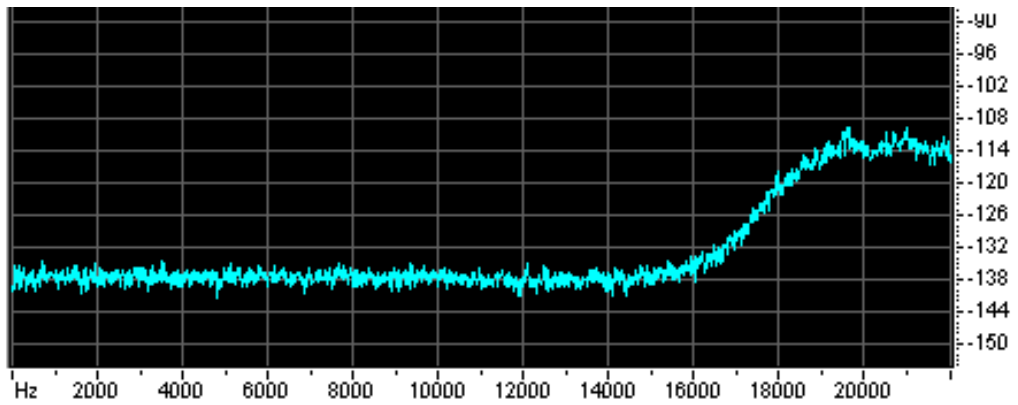
Cool Edit (C3 curve)



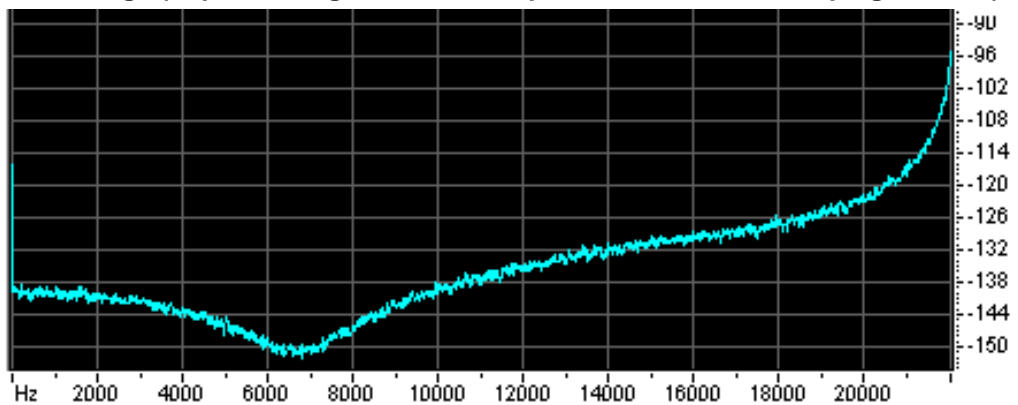
POW-R (pow-r1 mode)



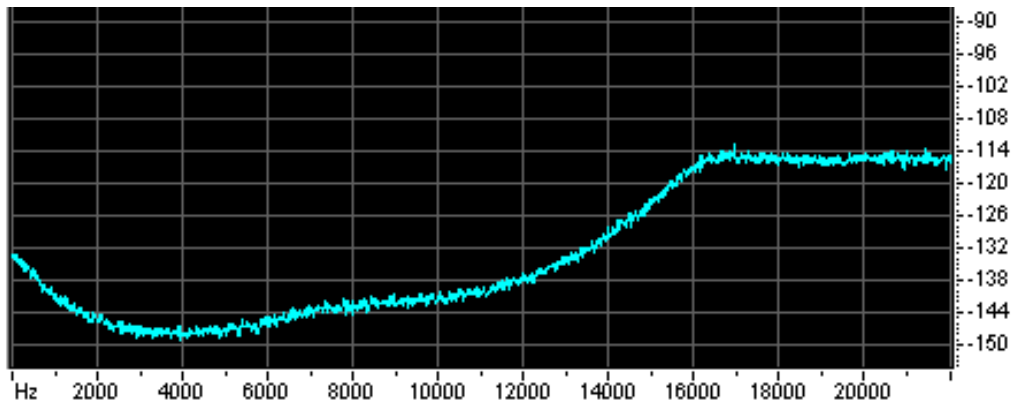
UV22



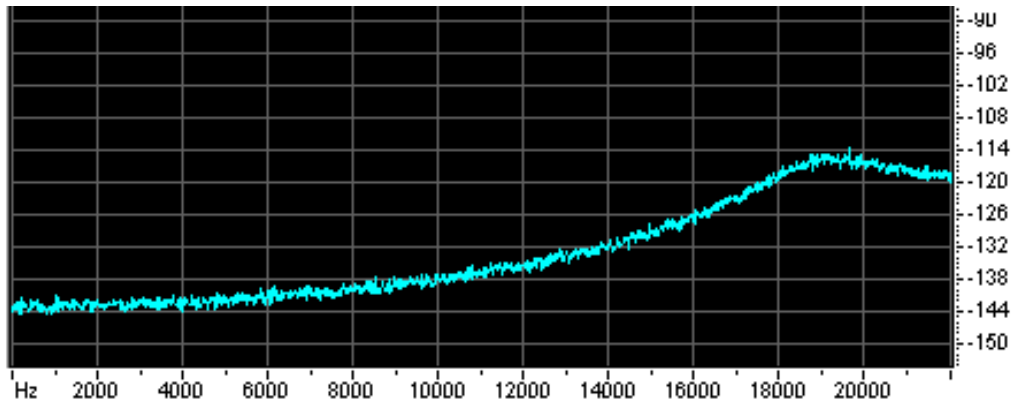
Sound Forge (hi-pass triangular dither + equal loudness noise shaping contour)



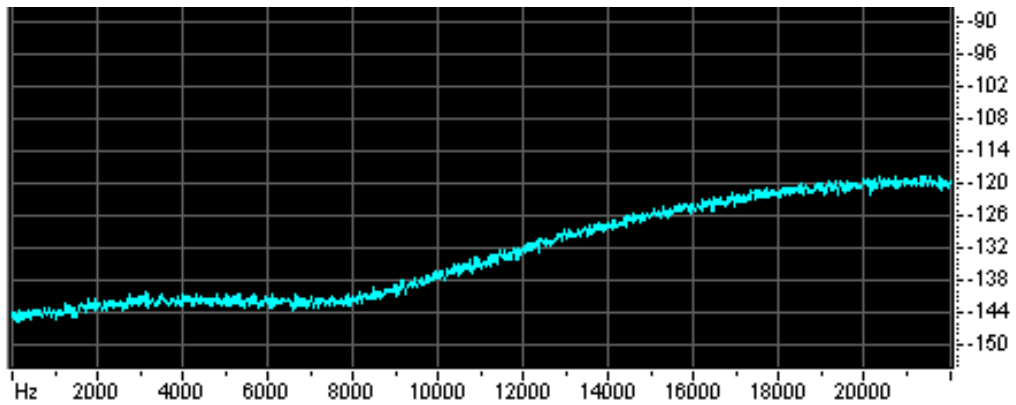
L2 (“Normal” preset)



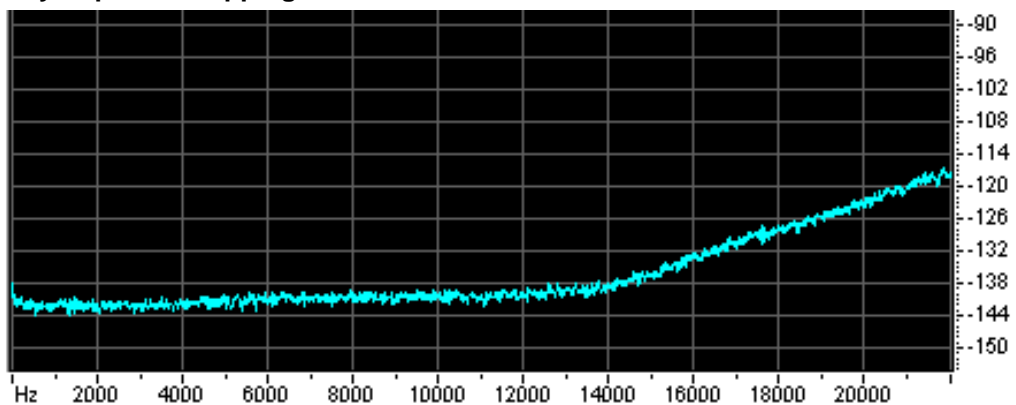
L1-Ultramaximizer (“Normal” preset)



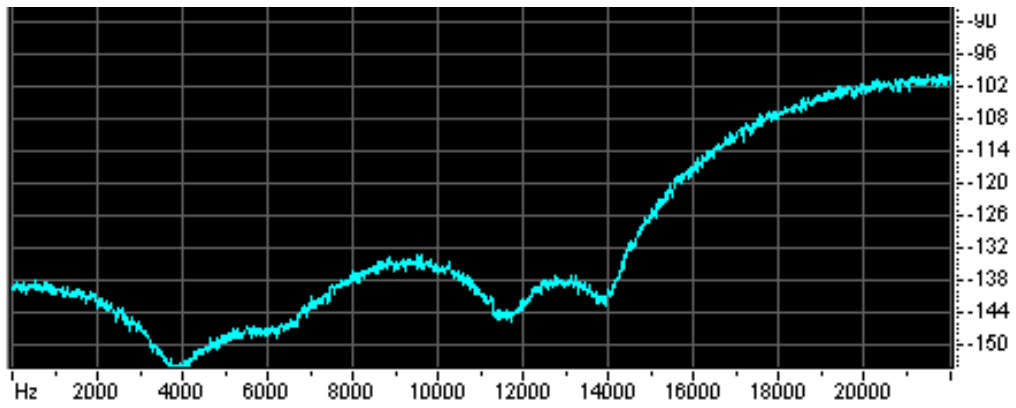
POW-R (pow-r2 mode)



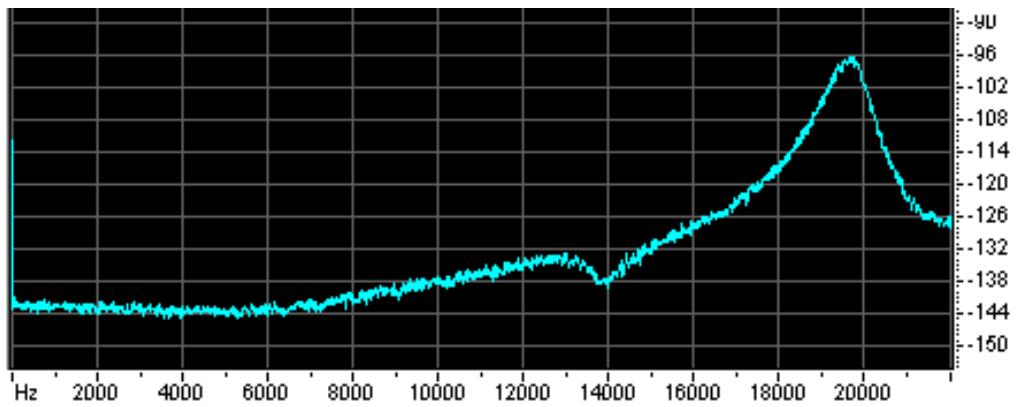
Sony Super Bit Mapping



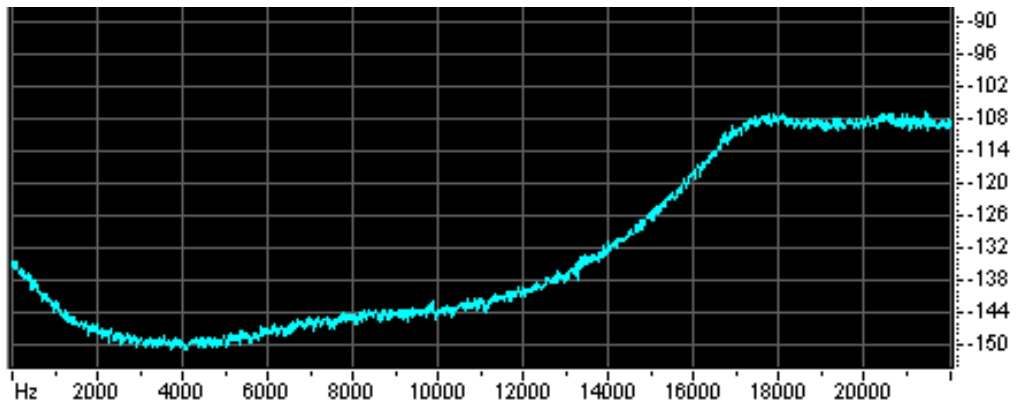
WaveLab (Type3 curve)



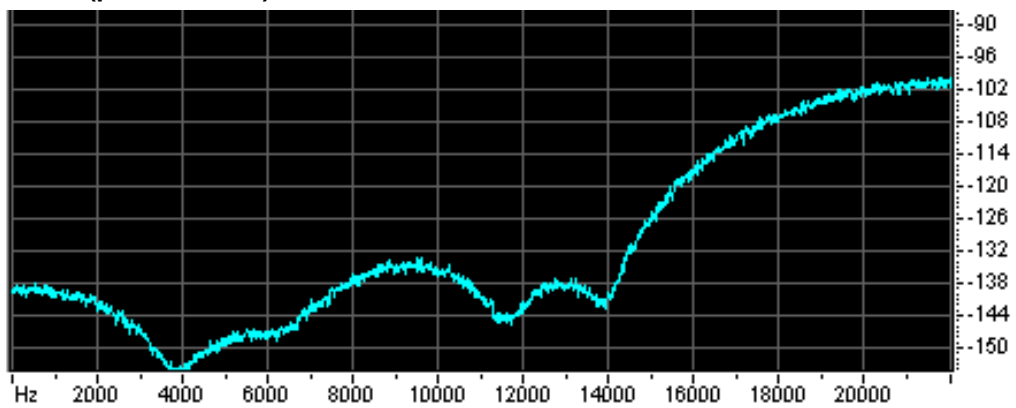
Cool Edit (44.1 KHz curve)



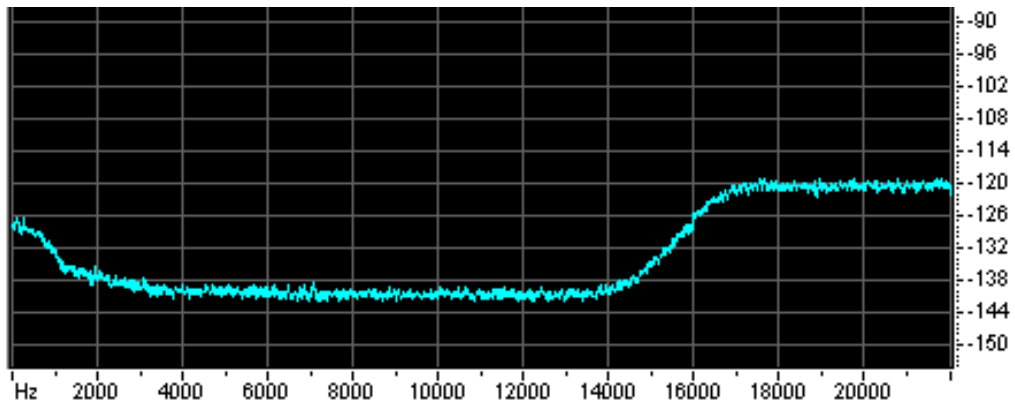
L2 ("Ultra" preset)



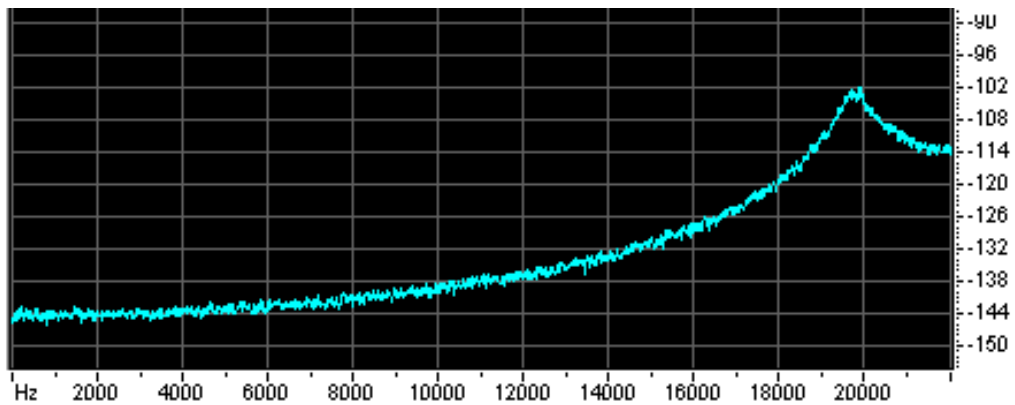
POW-R (pow-r3 mode)



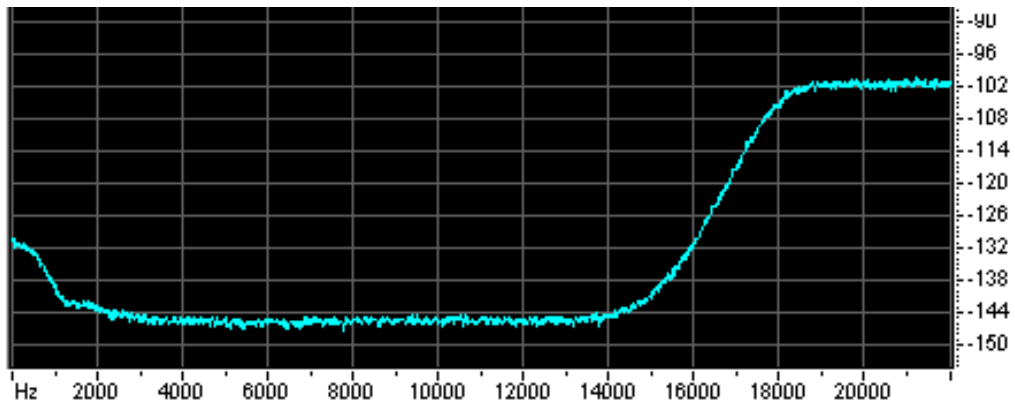
ExtraBit (“Medium+LPDN” preset)



L1-Ultramaximizer (“Ultra” preset)



ExtraBit (“Ultra+LPDN” preset)



Noise samples

From the following web pages you can download short samples of noise from different noise shaping systems or examples of processed audio:

<http://audio.rightmark.org/lukin/dither/nsamples.zip> - short samples of noise in mono 16-bit 44.1 kHz WAV files. After quantization, the noise samples were amplified by 40 dB for easy listening.

<http://audio.rightmark.org/lukin/dither/> - here you can find samples of audio processed by these noise shapers, and the latest updates for this paper, and links to other documents on noise shaping.

Result analysis

Here we try to analyze why most of commercially available word length reduction systems did not reach the noise reduction level of ExtraBit noise shaper. Also we emphasize specific properties and drawbacks of different examined noise shapers.

TPDF dithering features the most audible quantization noise. It does not shape the spectrum of noise to reduce the perceived noise loudness.

Cool Edit C1 curve features reduced amount of low-frequency noise below 9 kHz, but high-frequency noise in audible range is significantly increased. The most audible noise is concentrated around 12 kHz (around second spectral peak of human's ear sensitivity). That is why the perceived amount of noise reduction is low.

Cool Edit C3 curve tries to approximate the curve of human hearing threshold. But the approximation is very rough (most probably because of a too low order of noise shaping filter), and high peaks and pits in noise spectrum can lead to unpleasant noise coloration.

UV22 word length reduction system reduces the noise floor over the whole audible range by approximately 5 dB. The perception of noise is similar to white TPDF noise, reduced by 5 dB.

Sound Forge's noise shaping features insufficient dithering amount, which leads to noise modulation and even small harmonic distortion. The noise shaping curve tries to approximate the curve of human hearing, but the approximation is very rough again.

L2 "Normal" preset features good noise reduction around 4 kHz, at the expense of highaudible noise power around 14 kHz (clearly noticed esp. by younger listeners).

L1-Ultramaximizer's "Normal" preset features some noise reduction below 9 kHz, at the expense of moving noise from these frequencies to less audible range of 9 kHz – 16 kHz.

WaveLab's Type3 curve is another approximation of human hearing threshold, but the approximation is rough. There's an excessive amount of audible noise around 9 kHz.

Cool Edit 44.1 KHz curve is a more adequate approximation of the human hearing threshold. Nevertheless the curve is far from optimal (too low filter order again?). It could be much closer to optimal if the spectral pit at 21 kHz – 22 kHz is filled with frequencies from around 12 kHz.

L2 "Ultra" preset is much like L2 "Normal" preset, but with less noise below 15 kHz, and more noise above 15 kHz. Still high audible noise power around 15 kHz.

POW-r3 spectrum curve resembles Meridian-D and WaveLab Type3 noise shaping curves. The most audible noise of POW-r3 curve is around 8 kHz. We also suggest that there is excessive amount of noise reduction around 3.5 kHz. The noise floor of most high-resolution professional electronics is much higher than the noise floor of POW-r3 curve around 3.5 kHz. Probably, moving some noise from 7 kHz – 10 kHz to the band around 3.5 kHz would reduce the perceived noise audibility (and smooth the noise spectrum too).

L1-Ultramaximizer's "Ultra" preset is similar to "44.1 KHz" curve from Cool Edit. The most audible noise lies around 12 kHz, which could be moved to 20 kHz – 22 kHz spectral band. The overall rating of this curve is very good.

In our **ExtraBit** noise shaping system we have tried approximate the human hearing threshold as close as possible. We have performed extensive experiments to analyze the perception of broadband noise. Our results differ from classical Fletcher-Munson curve and even from its tilted (according to critical bands) "Shannon" form. Simultaneously we tried to achieve a smooth noise spectrum in audible range to reduce the noise coloration. We have selected the curve with flat spectrum top to achieve the

maximum efficiency from using high-frequency band at the constraints of maximal noise peak values of -60 dB FS. Our listening tests show that ExtraBit noise shaper has the lowest audible noise among all popular noise shaping systems.

Summary

We have evaluated different popular word length reduction systems and compared them according to specified criterions. We have concluded that several systems give results, which are close to optimal, but further optimization can be done yet. We have suggested a new ExtraBit noise shaping system, which yields best results.

Feedback

I will highly appreciate any comments or suggestions on this article. Also I am ready to evaluate and include into comparison any other word length reduction system. Want to know details? All the questions are welcome too. Please write me to: lukin@ixbt.com

From this web page you can download the most up to date demo version of ExtraBit Mastering Processor: <http://audio.rightmark.org/lukin/dither/>

Acknowledgements

I'd like to thank

Stanley P. Lipshitz and **Robert Wannamaker**, *Univ. of Waterloo*,

Ambrose Field, *Univ. of York UK*,

Sven Duwenhorst and **Ralf Schluenzen**, *TC Works*,

Andrew Nemeth, *Blue Mountains*,

Daniel Weiss, *Weiss*,

Richard Elen, *Apogee*,

Meir Shashoua, *Waves*,

and **Sergey Morozov**

for their helpful advice and support.