Dither & Noise Shaping: a Pragmatic Approach

Pablo F. Panitta, AES Member.

pfp2017@gmail.com

ABSTRACT

When reducing the bit depth of a digital signal (e.g., from 24 to 16), what is known as *"quantization noise"* will occur. This is a result of performing a "truncation" (removing the 8 least significant bits) of the 24-bit signal to convert it to 16-bit word length signal. In order to reduce the adverse effects of this process, a technique called **"Dithering"** is applied. This article does not attempt to address the theory of the process (of which there is plenty of literature on the subject), but rather to give it a more practical approach, which is a little more difficult to find.[1]

1. Dither. Basic theory.

As we mentioned, "quantization noise" is an artifact that appears after reducing the number of bits by truncating the signal. The way to minimize this quantization noise is through the process of Dithering, which is simply the addition of white noise (random and uniformly distributed) to the original signal before the reduction.

Adding white noise to eliminate 'quantization noise'? This does not seem to make much sense. And it is true, at first glance, this does not seem to be a very useful process.

The reality is that the so-called "quantization noise" **is not noise** as such, but rather a **distortion**. By definition, for a signal to be considered noise, it must be random and, therefore, not related (correlated) to other variables. And it turns out that quantization noise is related, and its relationship is with the level of the signal to be quantized.

Therefore, what we aim to achieve with the **Dithering process** is **replace distortion for noise**.[2]

And why would that be a good change? Well, it turns out that human psychoacoustics tolerate noise better than distortion. The latter is much more annoying and distracting than the noise itself.

2. Dithering in Practice.

The following are the results of measurements made with a 24-bit 997Hz signal at 2LSB (least significant bits) that has been processed with different plugins or DAWs to convert it to 16-bits. As mentioned above, with the process we try to lower the distortion at the cost of increased noise, and therefore we can measure parameters related to these phenomena, namely: TD (Total Distortion) and SNR (Signal to Noise Ratio).

For the purpose of considering what constitutes distortion and what constitutes noise, I have selected the following criterion: if the present signals are at a level greater than -30dBc, they are considered as distortion. For the remaining signals, their classification will be as noise.

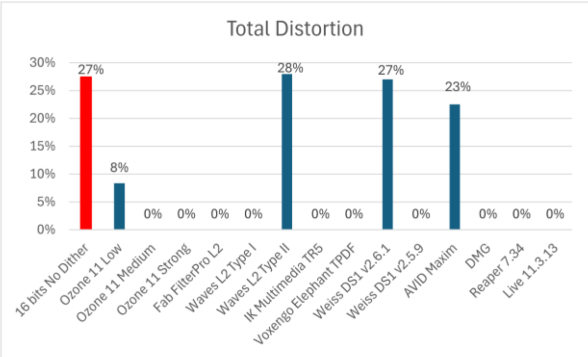

Fig.1. Total Distortion.

In Figure 1 we can see the impact on the distortion in the 16-bit quantized signals with the addition of dither. We see how the distortion is **drastically reduced** (or canceled) with respect to the original signal without dither. *This is exactly what we wanted to achieve!*

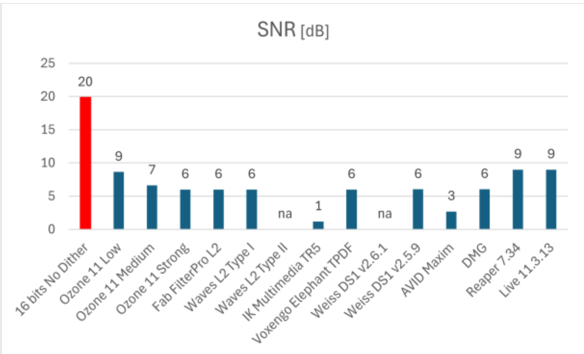

Fig.2. Signal to Noise Ratio (SNR). If the process does not significantly lower the distortion, this parameter is not measured ('na').

In Figure 2 we see the price to pay for that reduction. The **degradation of the SNR ratio**, easily observed, due to the introduction of dither (aka noise).

[1] Important disclosure: This work does not constitute any representation, recommendation, or preference towards any brand or model mentioned. No ranking of better or worse performance is made, but only the results are presented.

[2] The addition of noise to the signal being requantized decorrelates the quantization error from the signal level, and therefore, we transform it into true noise instead of distortion.

We also see that there are some processors that give peculiar results (Waves L2, Weiss DS1 v2.6.1, Maxim).

In relation to these, I would like to comment only the case of the Softube Weiss DS1 plugin, in which the current version (v2.6.1) incorporates autoblanking while the previous one (v2.5.9) did not. This means that for very low signal levels (such as those used in these measurements), the new version does not apply dithering and, therefore, we see the same distortion and SNR values as in the 16-bit signal without dither.[3] For this reason, the behavior of each version is different at these levels.

The selection of the dither intensity (in those plugins that allow it), affects the percentage of samples that change state, or in other words, that modify their level with respect to the original signal. Typical values for low intensity dither are around 20%, for medium 30% and high 35% and above.[4]

Something interesting to mention is that the application of dither *is independent in each channel*. For example, in the case of stereo dithering processing, the one generated for the left channel will be independent of the one generated for the right.

3. Noise Shaping.

Noise shaping is a technique that will allow me to redistribute noise to frequency zones of lower acoustic perception. **If dither adds uniformly distributed noise, noise shaping "molds" it**.

In other words, it is about reducing the noise power in areas of greater psychoacoustic sensitivity (mid and mid-high) and increasing it in areas of lower sensitivity (high frequencies). This makes it possible the ***improvement in the perception of the signal-to-noise ratio***.

In theory, the total noise power remains unchanged (there is only redistribution) but in practical implementation it increases after applying noise shaping. However, this is transparent to listening, since the listener will effectively perceive a reduction of noise in front of the useful signal, achieving a noticeable improvement in the overall perceived quality.

The different levels of noise shaping (if selectable), have to do with the amount of power displaced towards the psychoacoustically less sensitive frequencies.

Unlike Dither, which does not have much possible customization[5], Noise shaping can (and in fact does) have significant adjustment work. Many manufacturers use noise distribution based on models developed by them in order to optimize psychoacoustic effects. This is how we see acronyms such as: POWr, MBIT, IDR, UV22, etc.

Very easily, you can insert a spectrum analyzer to analyze the noise curves of different plugin brands and models.

4. Measurements Results.

Two indicators were measured, namely:

- **Noise Power** (NPow) [dBFS]: indicates the total noise power.
- **Noise to Mask Ratio** (NMR) [dB]: indicates the masking of the noise.

Both indicators are evaluated across the full spectrum (10Hz – 22050Hz).

Less negative values of Noise Power indicate higher power values, which signifies a greater total noise load in the spectrum. However, depending on its distribution, it may be more or less perceptible, and this is where the Noise to Mask Ratio helps us estimate this parameter. In this case, lower values of the NMR will indicate greater masking, which translates to a lower psychoacoustic perception of the noise.

The labeling convention is:

Plugin name / Dither Type / NS intensity

For instance:

Ozone / Medium / Low

Means the measurement was performed on Ozone plugin with Medium Dither and Low Noise Shaping intensity setup.

In Fig.3 (at page 4), we can see the measurement results.

It can be observed how, *as the intensity of the noise shaping increases, the NPower parameter (noise power over the entire band) increases, and the NMR (which is a noise masking parameter) generally decreases.*

[3] Softube engineers have confirmed that this change was introduced in order to reduce CPU consumption for signals with such low input levels.

[4] Dither intensity can also be related to the added LSBs (level) (typically 1, 2, or 3 LSBs). Generally, for triangular distributions, 2 LSBs is the common choice.

[5] Dither parameters that can be varied include its intensity and distribution type (Rectangular, Triangular, Gaussian). In audio, the triangular distribution (TPDF) is the most common and recommended choice (though a few plugins allow you to select the type to apply).

5. The listening. Benchmark.

Although the results are based on real measurements, they do not give us a complete idea of how each combination sounds. Is a distortion of x% something tolerable at these levels? How good or bad is an SNR value of y dB? Or... how much is the perception difference between the different NMR values?

In order to provide the reader with an auditory representation, I have created a page Search for results of the tests, which can help them choose the Dither and/or Noise Shaping that best suit their needs. It can be validated how measurements with a lower noise power (NPower) effectively translate into a lower auditory perception, or even combinations with the same NPower value but lower NMR, are perceived with less noise load, validating the measurement results.

The reality is that as all this occurs at very low signal levels, in most of the cases it will not be a matter of extreme concern. But as we are always in search of excellence... why deprive ourselves of choosing what best suits us?

The following link points to a page where I have created a tool to compare various combinations of Dither and Noise Shaping, both for a sinusoidal signal and for an instrument signal.

Although it can be accessed from any device, it is recommended to enter from a PC since you will access the full version. Access from a mobile device will redirect you to the reduced version of the tool.

https://soundins.com.ar/Publications/Dither/DitherNS_Tool_Main.html

6. A1. Brief tool description and usage guidelines.

The tool presents the processing results of two types of low-level signals:

1) Sine: sinusoidal signal at 997Hz , 2LSB (Least Significant Bit) peak level / 44.1kHz /24 bits

2) Instrument: instrument signal -86 dBFS RMS at 44.1kHz/24bits

Both signals were converted from 24-bit to 16-bit depth through dithering and/or noise shaping processing, using the plugins or DAWs indicated in each case.

Except for the gain application (after processed signal), no other processes, effects, or signal alterations were included in the measurement.

First, select the type of source you want to play (sine wave or instrument) using the "Source" buttons in the panel.

Next, choose the process applied to the selected source. To do this, select the corresponding rectangle in the main panel. Each one describes the applied process inside it (or via an associated tooltip). Pressing it will play the corresponding processed audio.

The 997Hz button enables or disables the sine tone. Disabling it is useful for listening the noise profile and distortion.

At the bottom of the page, you will find the Loop and Stop audio buttons.

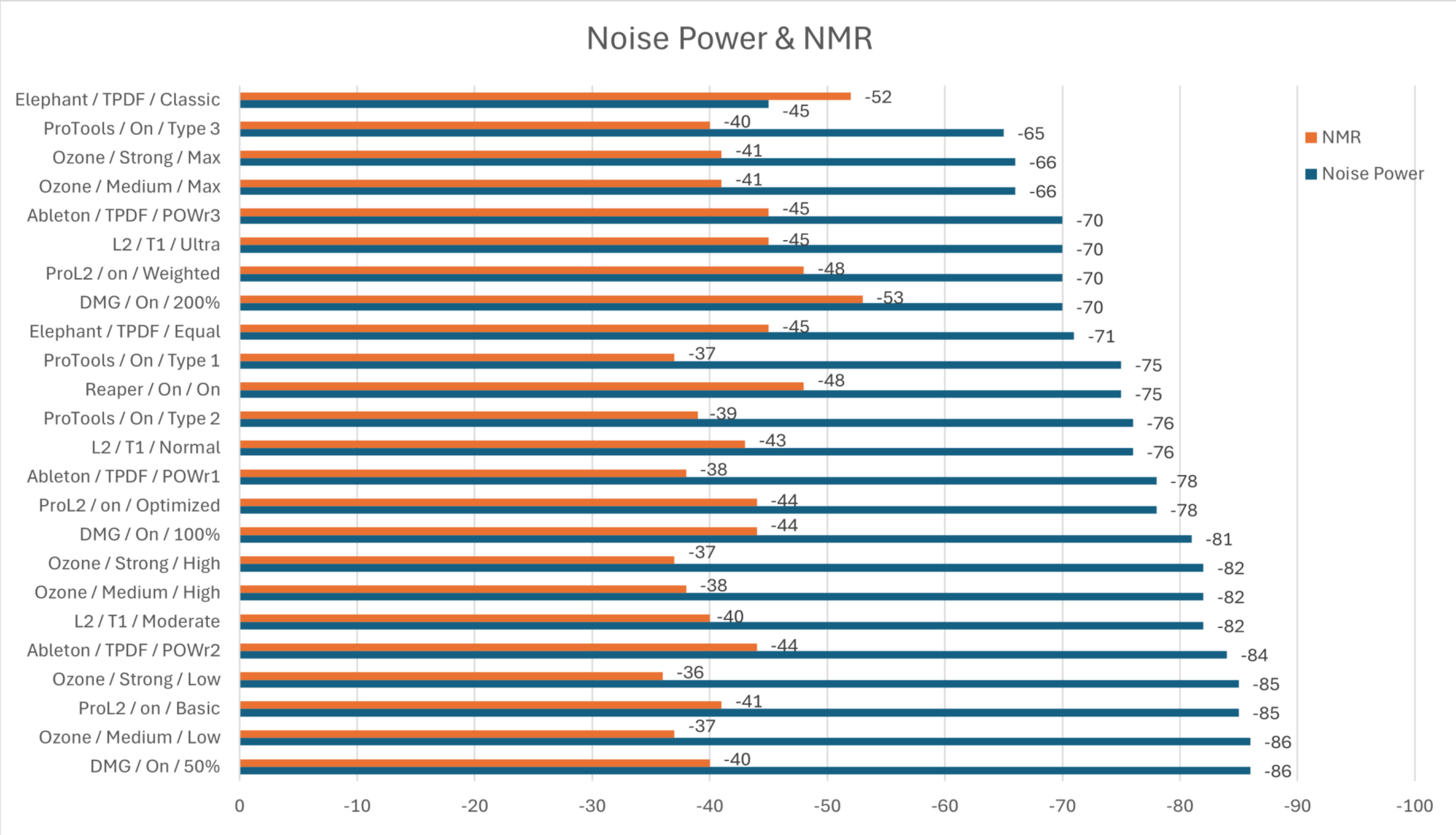

Fig.3. Noise Power y Noise to Mask Ratio for each combination shown. If the processor has auto-blanking, it is not shown on the graph.