

Dither y Noise Shaping. Un abordaje práctico.

Abstracto

Al reducir la profundidad de bits de una señal digital (por ejemplo de 24 a 16), se producirá lo que se denomina “ruido de cuantización”. Esto es producto de realizar un “truncamiento” (sacar los 8 bits menos significativos) de la señal de 24 y así llegar a la de 16 bits. A fin de reducir los efectos adversos de este proceso, es que aplicamos una técnica llamada “Dithering”. Este artículo no intenta presentar la teoría detallada del proceso (que por cierto hay toneladas de bibliografía al respecto), sino darle un abordaje más práctico, el cual es un poco más difícil de encontrar.¹

Dither. Un poco de teoría.

Ya dijimos que el “*ruido de cuantización*” es un efecto que aparece luego de haber reducido la cantidad de bits a través del *truncado* la señal. La forma de minimizar este ruido de cuantización es a través del proceso de **Dithering** el cuál es simplemente el agregado de ruido blanco (aleatorio y uniformemente distribuido) a la señal original, antes de la reducción.

¿Agrego ruido blanco para eliminar “ruido de cuantización”? Esto no parece tener mucho sentido. Y es verdad, en una primera lectura, esto no parece ser un proceso muy útil.

La realidad es que el llamado “*ruido de cuantización*” **no es un ruido** como tal, sino más bien una **distorsión**. Por definición, para que una señal sea considerada como ruido, debe ser aleatoria y, por lo tanto, no debe estar relacionada (correlacionada) con otras variables. Y resulta ser que el ruido de cuantización, si lo está, y su relación es con el nivel de la señal a cuantizar.

Por lo tanto, lo que buscamos con proceso de Dithering es **cambiar distorsión por ruido**.²

¿Y por qué eso sería un buen cambio? Bueno, resulta que la psicoacústica humana *tolera mejor el ruido que la distorsión*. Esta última resulta mucho más molesta y distractiva que el ruido en sí mismo.

¹ Importante: El presente trabajo no constituye ninguna representación, recomendación o preferencia hacia ninguna marca o modelo mencionado. No se realiza un ranking de mejor o peor funcionamiento, sino solamente se exponen los resultados.

² El agregado de ruido a la señal a recuantizar hace decorrelacionar el error de cuantización del nivel de señal, y por lo tanto lo transformamos en verdaderamente ruido en vez de distorsión.

Dither. La práctica.

A continuación se presentan los resultados de mediciones realizadas con una señal de 24-bits 997Hz a 2LSB (least significant bits) a la que se ha procesado con distintos plugins o DAWs, para convertirla a 16-bits. Como se mencionó anteriormente, con el proceso intentamos bajar la distorsión a costa de un mayor ruido, y por lo tanto podemos medir parámetros relacionados con estos fenómenos, a saber: TD (Total Distortion) y SNR (Signal to Noise Ratio).

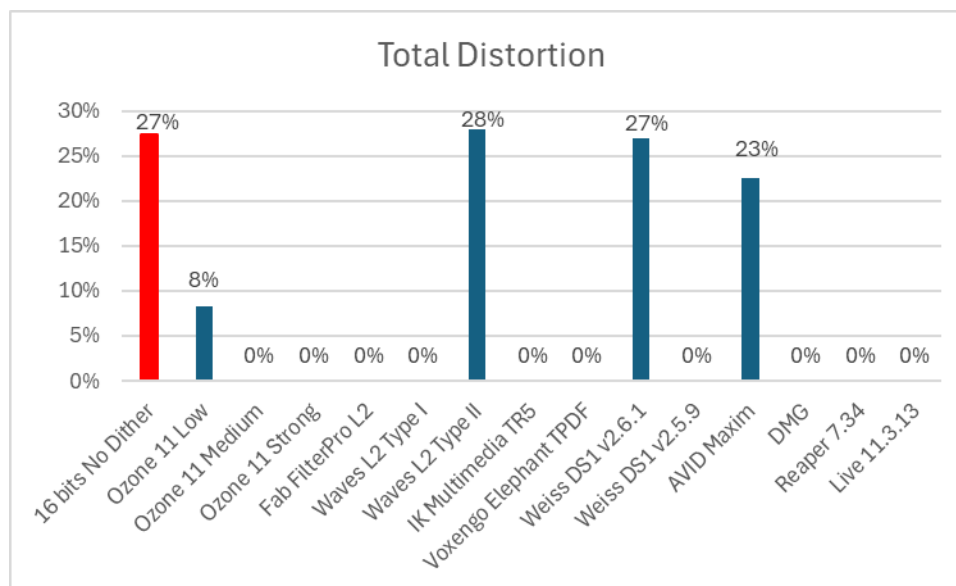


Figura 1: Distorsión Total. Señales -30dBc de la fundamental

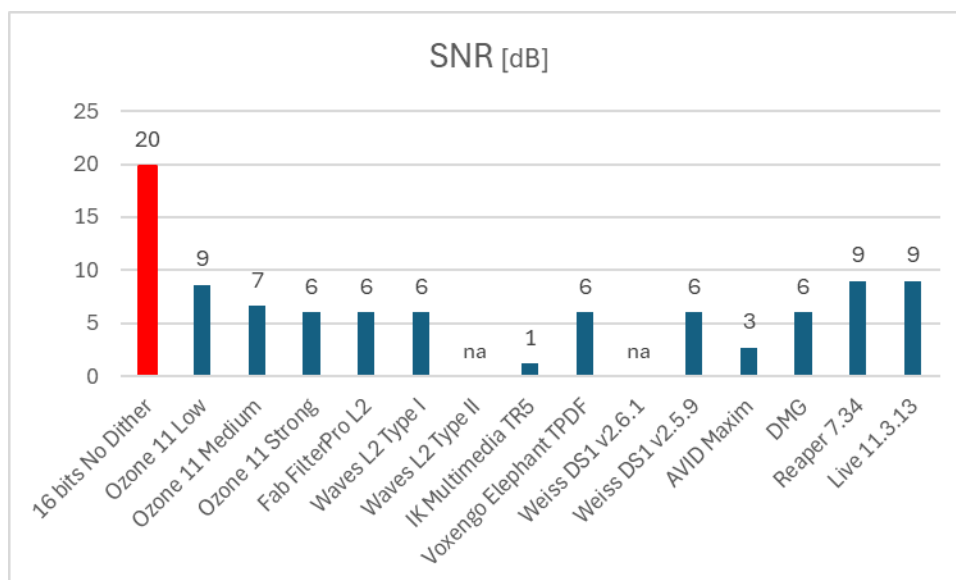


Figura 2: Relación Señal a ruido (SNR). Si el proceso no baja significativamente la distorsión, no se procede a medir este parámetro (na).

En la Figura 1 podemos ver el impacto en la distorsión en las señales cuantizadas a 16-bits con el agregado de dither. Vemos como **la distorsión se reduce drásticamente (o se anula)** con respecto a la señal original sin dither. ¡Exactamente lo que queríamos lograr!

En la Figura 2 vemos el precio a pagar por esa reducción. Se observa fácilmente la **degradación de la relación SNR**, debido al haber introducido el dither (aka ruido).

Vemos también que hay algunos procesadores que arrojan unos resultados particulares (Waves L2, Weiss DS1 v2.6.1, Maxim).

En relación a estos, me gustaría solo comentar el caso del plugin de Softube Weiss DS1, en el cuál la versión actual (v2.6.1) incorpora autoblanking mientras que la anterior (v2.5.9) no lo tenía. Esto hace que para los niveles muy bajos de señal (como los utilizados en estas mediciones), la nueva versión no aplique dithering y, por lo ende, vemos los mismos valores de distorsión y SNR que en la señal de 16-bits sin dither.³ Por tal motivo, el comportamiento de ambas versiones es diferente para éstos niveles.

La selección de la intensidad del dither (en aquellos plugin que lo permiten), afectan al porcentaje de muestras que cambian de estado, o dicho de otra manera, que modifican su nivel con respecto a la señal original. Valores típicos para dither de baja intensidad rondan el 20% , para medios un 30% y altos de 35% en adelante.⁴

Algo interesante para mencionar es que la aplicación del *dither* es *independiente en cada canal*. Por ejemplo, para el caso del procesamiento con dithering en stereo, el generado para el canal izquierdo será independiente al generado para el derecho.

Noise Shaping.

El modelado del ruido (Noise Shaping) es una técnica que me va a permitir, redistribuir el ruido a *zonas de frecuencia de menor percepción acústica*. **Si el dither agrega ruido uniformemente distribuido, el noise shaping lo “acomoda”**.

En otras palabras, se trata de reducir la potencia de ruido en zonas de mayor sensibilidad psicoacústica (medias y medias altas) y aumentarla en las zonas de menor sensibilidad (altas frecuencias). Con esto logro **mejorar la percepción de la relación señal a ruido**.

En la teoría, la potencia total de ruido se mantiene inalterada (solo existe redistribución) pero en la implementación práctica la misma aumenta luego de aplicar noise shaping. Sin

³ Ingenieros de Softube me han confirmado que se ha introducido este cambio con el fin de reducir el consumo de CPU para señales de entrada de tan bajo nivel.

⁴ La intensidad del Dither también puede relacionarse con los LSB (nivel) agregados (típicamente 1, 2 o 3 LSB). Generalmente para distribuciones triangulares, lo común son 2 LSB.

embargo, esto es transparente para la escucha, ya que el oyente efectivamente percibirá una reducción del ruido frente a la señal útil, logrando una mejora notable en la calidad total percibida.

Los distintos niveles de noise shaping (en el caso que sea seleccionable), tiene que ver con la cantidad de potencia desplazada hacia las frecuencias psicoacústicamente menos sensibles.

A diferencia del Dihter, que no tiene mucha personalización posible⁵, el Noise shaping si puede tener (y de hecho lo tiene) un trabajo de ajuste importante. Muchos de los fabricantes utilizan distribución de ruido basada en modelos desarrollados por ellos de manera de optimizar los efectos psicoacústicos. Es así que vemos siglas como: POWr, MBIT, IDR, UV22, etc.

De forma sencilla se puede insertar un analizador de espectro, y analizar las curvas de ruido de las distintas marcas y modelos de plugins.

Resultados.

En la Figura 3, vemos reflejado los valores obtenidos de Noise Power(Npow) y Noise to Mask Ratio (NMR)⁶ para distintos niveles de dithering de los plugins medidos.

La convención de etiquetado es:

Nombre plugin / Tipo de Dither / Intensidad Noise Shaping.

↓

↓

↓

Por ejemplo: Ozone / Medium / Low

indica que corresponde al plugin Ozone con un Dither Medium y una intensidad de Noise Shaping Low.

⁵ Parámetros que pueden variar en el Dither son su intensidad y su distribución (Rectangular, Triangular, Gaussian). En audio, el uso común y recomendado es la distribución triangular (aunque unos pocos plugin permiten seleccionar el tipo a aplicar)

⁶ Noise to Mask Ratio (NMR) es una medida de enmascaramiento del ruido. Valores más negativos indican que el perfil de ruido se encuentra más alejado del umbral de enmascaramiento, presentando una menor percepción psicoacústica.

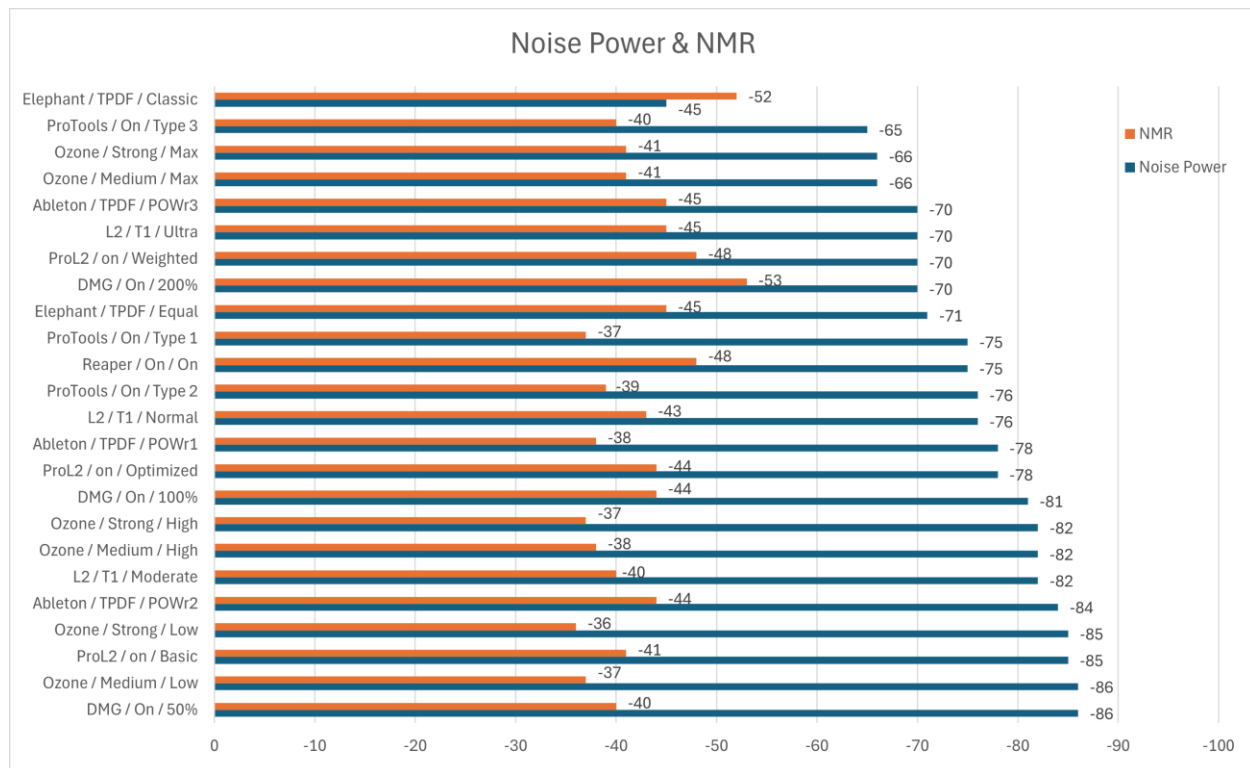


Figura 3: Niveles de Noise Power y Noise to Mask Ratio para cada combinación medida. Si el procesador posee autoblanking, no se muestra en el gráfico

Recordar que el Noise Power se relaciona con la potencia de ruido que se encuentra en la banda y el NMR a cuánto psicoacústicamente se percibe dicho ruido. Mayores niveles de nPower (menos negativos) significan mayor ruido presente mientras que menores valores de NMR (más negativos) resultan en una menor percepción del mismo.

Se puede apreciar cómo, a medida que aumento la intensidad del noise shaping, el parámetro NPower (potencia de ruido sobre toda la banda) aumenta y en NMR (que es un parámetro de enmascaramiento de ruido) en general disminuye.

A escuchar. Comparaciones

Si bien los resultados son indicadores que nos orientan hacia las características de la señal procesada, no nos dan una idea cabal de cómo se escucha cada combinación. ¿Es una distorsión de x% algo tolerable a estos niveles? ¿qué tan bueno o malo es un valor de SNR de ydB? O si cuánto es la diferencia de percepción entre los distintos valores de NMR.

Para que el lector pueda tener una representación auditiva, he creado una página en donde

se pueden escuchar cada una de las señales procesadas, con el fin de poder ayudarlo a seleccionar el Dither y/o Noise Shaping más le resulte conveniente.

Se puede validar cómo mediciones con una potencia de ruido (NPower) menor efectivamente se traduce en una menor percepción auditiva, o incluso combinaciones con el mismo valor de NPower pero menor NMR, se perciben con menor carga ruidosa, validando los resultados de las mediciones.

La realidad es que como todo esto ocurre a niveles muy bajos de señal, en la mayoría de los casos no será un tema de extrema preocupación. Pero como siempre estamos en la búsqueda de la excelencia... ¿por qué privarse de elegir lo que más nos sirva?

El siguiente link apunta a una página donde he creado una herramienta para poder comparar varias combinaciones de Dither y Noise Shaping, tanto para una señal senoidal como para una de instrumento.

Si bien se puede acceder desde cualquier dispositivo, se recomienda ingresar desde una PC ya que se accederá a la versión completa. El acceso desde un dispositivo móvil, redireccionará a la versión reducida de la herramienta.

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

Pablo Panitta