EXPERIMENTAL INVESTIGATION INTO THE OPTIMAL USE OF DITHER

PACS: 43.60.Cg

Preben Kvist[1], Karsten Bo Rasmussen[2], Torben Poulsen[1]
[1]Acoustic Technology, Ørsted•DTU, Technical University of Denmark
DK-2800 Lyngby, Denmark
[2]Oticon, Strandvejen 58, DK-2900 Hellerup, Denmark

ABSTRACT
The use of dither in A/D as well as D/A conversion is standard, since it is well known that dither can reduce the audibility of quantisation noise as well as converter errors. While many elaborate dither applications exist today, it is a fact that many psychoacoustic aspects of dither are still largely unknown. The work presented is from an experimental study into the psychoacoustic effects of dither, in terms of subjective preference. The purpose is to improve our understanding of what can be obtained through a specific dither application.

QUANTISATION AND DITHERING

In order to understand A/D conversion, it is helpful to divide it into two separate operations, i.e. sampling and amplitude quantisation. The sampling is simply amplitude measurements done at discrete time intervals. Quantisation is conversion from the precise amplitude values to discrete amplitude values, i.e. finite resolution. This process is necessary if the signal is to be stored or transmitted digitally. Throughout the paper, the step size of the quantisers will be referred to as one least significant bit (LSB).

In the following, quantisers of the socalled midtread type will be used. All results will also be valid for the midriser type. It is well known that for low signal levels, distortion-like components will appear in the spectrum of a quantized sinusoid. In the following, we shall see how the use of dither can remove these undesired components.

Nonsubtractive Dither

The simplest form of dithering is nonsubtractive dither. The difference compared to the undithered quantisation is the addition of random noise n(t) prior to the quantisation. The choice of Probability Density Function (PDF) and dither amplitude will be discussed below.

One of the advantages of using dither is the improvement of the low-level resolution in the quantiser. Figure 1(a) shows an undithered quantisation of a 1½LSB sinusoid. If the signal is dithered prior to quantisation the quantiser output will rapidly toggle between levels, Figure 1(b). If this output is averaged over 20 traces then the sinusoid is restored, Figure 1(c). Hence, if the signal is properly dithered prior to quantisation, resolution below the least significant bit can be obtained in an average sense.

Figure 2 shows the dither amplitude needed to linearise the average input-output characteristic of the quantiser, for commonly used dither PDFs, rectangular, triangular and Gaussian.
An amplitude of 1 LSB is needed to linearise the input-output characteristic if the rectangular distribution is used. For the triangular distribution, a peak-peak amplitude of 2 LSB is needed. Since the Gaussian distribution in theory could have values from $\pm\infty$ the Root Mean Square (RMS) value is used instead of the peak-peak amplitude. In order to linearise the input-output characteristic with Gaussian PDF dither, a RMS-value of ½LSB is required.

For both the triangular and the Gaussian PDF dither the input-output characteristic will continue to be linear when more than the minimum required amplitude is added. However, in the case of rectangular PDF dither the input-output characteristic will only be linear for n LSB dither, where n is a positive integer.

The RMS values of sufficient dither, regarding the linearisation of input-output characteristic, with the three different PDFs are:

1 LSB rectangular dither: $LSB/\sqrt{12}$
2 LSB triangular dither: $LSB/\sqrt{6}$
½LSB RMS Gaussian dither: $\frac{LSB}{2}$

Figure 3 shows the Power spectral density of a quantized 4 LSB sinusoid dithered with the three PDF mentioned above. All three PDFs removes the undesired frequency components.

Judging from the previous results, 1 LSB rectangular PDF dither seems like a good choice. However, with only 1 LSB dither there is a risk of the noise being modulated by the signal. Imagine a slow varying signal dithered with 1 LSB, i.e. [-0.5,0.5]. When the input signal is close to a quantisation step, the 1 LSB dither will not be sufficient to get the quantisers output to toggle between two values. As the signal increases, the dither will start taking effect. In this way the noisy part of the output signal becomes modulated by the input signal. Hence, a peak-peak dither amplitude of more than 1 LSB is necessary in order to remove this unwanted modulation. Since rectangular PDF dither with a peak-peak amplitude of more than 1 LSB does not linearise the average input-output characteristic, it is not suitable. 2 LSB triangular PDF dither will be the best choice judging from the simulations, but since Gaussian PDF noise is much easier to generate in the analog domain, it is frequently used instead.

<u>Subtractive dither</u>
An extension of the nonsubtractive dither is the subtractive dither. The difference compared to the nonsubtractive scheme is that the dither noise is being subtracted at the output of the DAC. This scheme gives all the benefits of the nonsubtractive dither but as figure 4 shows, the noise floor is lowered compared to nonsubtractive dither.

In practice, subtractive dithering is difficult to implement because of the need to transmit the dither to the output. One way of doing this is to use a pseudo random number generator followed by a DAC as the source of analog dither noise. This makes it possible to transmit discrete values instead of the analog noise, and hence dither subtraction becomes more realistic. It is important to note that this type of dither cannot linearise the input-output characteristic completely but it will increase the resolution according to the number of levels used in the DAC to generate the dither.

If a random number generator provides dither of an integral number of LSBs, then it will not improve the resolution below the least significant bit. It will, however, be able to alleviate problems with nonlinearity in the DAC. In the present work, the quantisers are assumed to be ideal (i.e. perfectly linear) and hence this property will not be investigated further.

NOISE SHAPING

Noise shaping is a feedback technique, which in its simplest form means that the quantisation error of the current sample is stored and then subtracted from the next sample. The idea is that at least for slowly varying signals, the average output will be a better approximation to the input.

If a subtractive dither scheme is implemented, the error should be calculated as the difference between the input and the output of the quantiser, i.e. not including the dither.

Figure 5(a) shows the Power Spectral Density (PSD) for a sinusoid if simple noise shaping is applied, i.e. subtraction of the quantisation error from the previous sample. This form is called 1st order noise shaping. The total noise power is not reduced by noise shaping but the frequency spectrum rises with frequency and this way some of the noise is moved outside the audio band.

Oversampling is sampling at a much greater frequency than required by the Nyquist theorem. The total noise introduced by the quantisation will remain the same when oversampling is applied but it will be spread over a wider frequency range and hence only a fraction of the noise will be in the audio band. The achieved improvement in Signal to Noise Ratio (SNR) will be 3 dB when the sampling ratio is increased by a factor of 2. Figure 5(b) shows the 1st order noiseshaper used in (a) but now oversampling by a factor of 4 was used. Because of the noiseshaper, the improvement of SNR in the audioband is much greater than the 6 dB predicted by oversampling alone. Hence, using oversampled conversion with noise shaping proves to be much more powerful than regular oversampled conversion.

For simple noise shaping the noise increases 6 dB/octave with frequency. With more complex filters in the feedback loop it is possible to shape the noise to more complex shapes. Of particular interest would be the equal loudness level curves. By shaping the noise to equal loudness level curves the noise could be expected to be optimised to the human hearing, i.e. as inaudible as possible for a specific overall level.

Figure 6(a) shows the PSD of a sinusoid quantised with the 9th order noiseshaper suggested by Wannamaker. The coefficient for the FIR-filter is calculated from a fitting of the filter to the inverse of the ISO standard 15-phon equal loudness data with a free-to-diffuse field correction applied. Wannamaker concludes that the perceived output noise power will be lowered by 17 dB with this scheme.

If a finite quantiser is used, there will be a risk of numerical overload. To avoid this, the contribution from the feedback loop must be limited, and hence the noise shaping will be less ideal. This will degrade the efficiency of the noise shaper. Figure 6(b) shows the PSD of a 9 LSB noiseshaper with a 4 LSB peak-peak sinusoid as input. In this case the noisefloor will be higher than if no noise shaping had been used. Due to this limitation, noise shaping will be most efficient with relatively low inputs, but the case of a weak input is also critical in the sense of SNR. Hence, the approach might still be useful.

SUBJECTIVE PREFERENCE TESTS

Despite the numerous articles on the theory of dither only very few published subjective tests are available. In the famous work by Rabiner and Johnson only the subtractive dither scheme was investigated. Since the nonsubtractive scheme is commonly used it would be interesting to have a similar investigation for nonsubtractive dither. To find out how big an advantage these two schemes provide, the following subjective comparisons has been made,

Nonsubtractive dither using 1 LSB rectangular PDF dither vs. undithered PCM.
Nonsubtractive dither using 2 LSB triangular PDF dither vs. undithered PCM.
Subtractive dither using 2 LSB triangular dither vs. undithered PCM.

In order to find out if rectangular PDF or triangular PDF nonsubtractive dither is preferred, the following comparison has been made,

Nonsubtractive dither using 2 LSB triangular PDF dither vs. 1 LSB rectangular PDF dither.

Regarding noiseshaping, Wannamaker suggests that noiseshaping can reduce the perceived output noise by 17 dB when using a 9th order noise-shaper. However, this result is obtained for infinite quantisers. In real world applications the improvements might be much smaller. To test the improvement due to noiseshaping, the following comparisons has been made,

9th order nonsubtractive psychoacoustically optimal noiseshaping vs. nonsubtractive dithered PCM (2 LSB triangular PDF dither).

The magnitude of the perceived improvement has been tested by comparing the two signals at equal bitrate, then giving a 1 bit advantage to one of the signals and then a 2 bit advantage. In this way it was tested how many more bits would be needed to obtain the same sound quality in for example undithered quantisation compared to dithered quantisation.

The comparisons were carried out using 16, 64, 256 and 1024 LSB inputs, roughly equivalent of 4, 6, 8, and 10 bits.

RESULTS

The subjective preference tests were carried out using a computerized setup with monaural headphone presentation. Ten subjects were used. A preliminary analysis shows the following key results,

<u>Nonsubtractive dither</u>

In the case of music, a dithered signal is preferred over undithered at a 4 bit level of resolution, but for 6-10 bit undithered is preferred. Rectangular PDF dither is preferred over triangular PDF dither for all resolutions.

For speech, dithered is preferred at 4-8 bits. For higher resolution undithered is preferred. At 4-6 bit triangular is preferred over rectangular. A probable cause is the pauses in a speech signal, where noise modulation can be disagreeable. For 8 bits and higher resolution, rectangular PDF is again preferred.

<u>Subtractive dither</u>

In the case of 4-8 bits subtractive dither is preferred in the case of music.

For speech, subtractive dither is preferred over no dither even when the undithered signal has 1 bit better resolution, in the range 4-6 bit.

<u>Noiseshaping</u>

Noise shaping is preferred over PCM with the same resolution, and for higher resolutions (10 bits) it may be preferable even if PCM is given a 1-bit advantage. This result could be a consequence of the chosen filtering of the noise shaper, being shaped according to an equal loudness level curve.

CONCLUSION

The subjective results regarding subtractive dither can be compared with the classical results from Rabiner and Johnson. The findings in the present work agrees well with the previous results, stating the usefulness of subtractive dither for low resolutions. Nonsubtractive dither is only preferred for resolutions less than about 8 bits. Noise shaping has been confirmed as a useful approach.

References

L.R. Rabiner and J.A. Johnson. Perceptual evaluation of the effects of dither on low bit rate pcm systems. The Bell System Technical Journal, 51(7), September 1972.

R.A. Wannamaker. Psychoacoustically optimal noise shaping. The Journal of the Audio Engineering Society, 40(7/8), July/August 1992

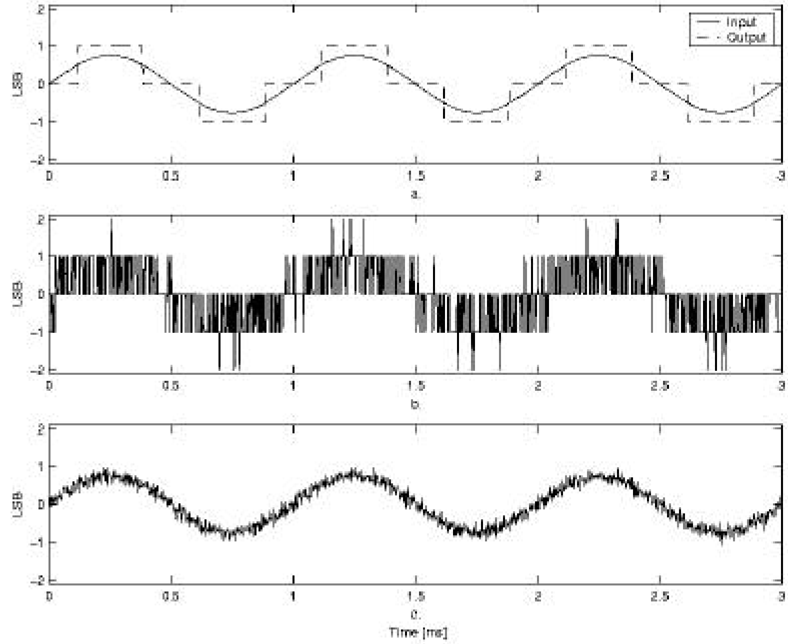


Figure 1. Undithered and dithered quantisation of 1½LSB sinusoid. (a) Input and output of an undithered quantiser. (b) Output of quantiser with a 1½LSB dithered sinusoid as input. Dithered with 2 LSB triangular PDF dither. (c) Averaged Output of quantiser. Averaged over 20 traces.

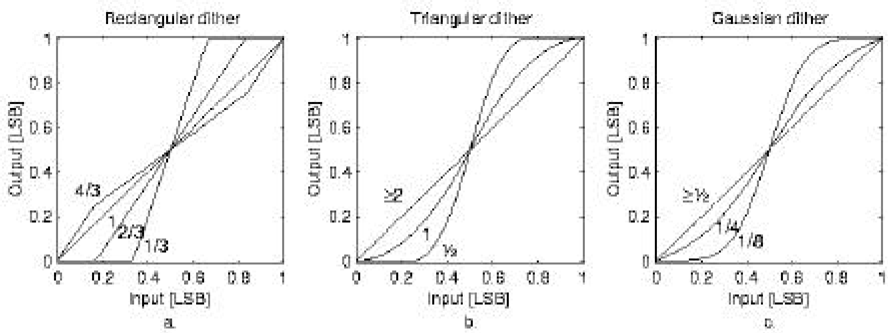
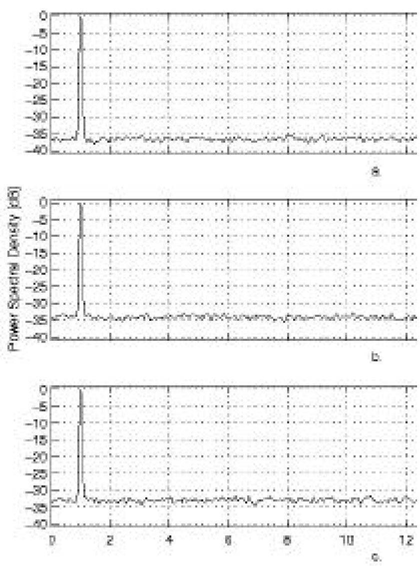


Figure 2. Averaged input-output characteristic. Averaged over 200000 traces. (a) Rectangular PDF dither with peak-peak amplitude of 1/3, 2/3, n and 4/3 LSB where n is a positive integer. (b) Triangular PDF dither with peak-peak amplitude of 1/2, 1 and ≥2 LSB. (c) Gaussian PDF dither with RMS of 1/2, 1/4, 1/8 and 1/16 LSB.



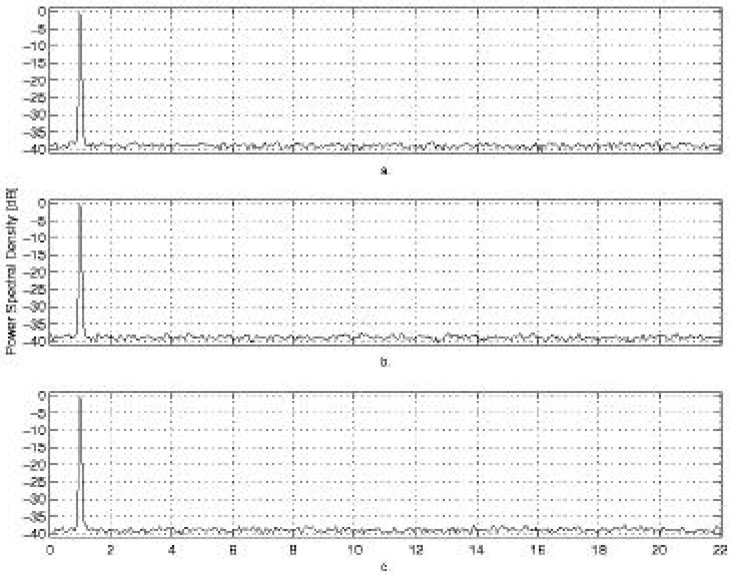
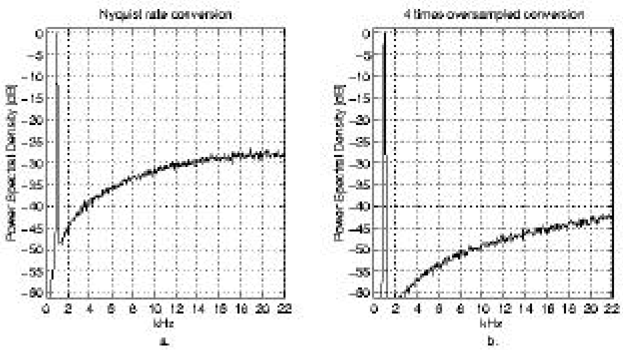


Figure 3 (Left). Power Spectral Density for a quantised 4 LSB sinusoid, nonsubtractively dithered. The PSD is normalised to give 0 dB at the input frequency. (a) 1 LSB rectangular PDF dither. (b) 2 LSB triangular PDF dither. (c) ½LSB RMS Gaussian PDF dither.

Figure 4 (Right). Power Spectral Density for a quantised 4 LSB sinusoid, subtractively dithered. The PSD is normalised to give 0 dB at the input frequency. (a) 1 LSB rectangular PDF dither. (b) 2 LSB triangular PDF dither. (c) ½LSB RMS Gaussian PDF dither.



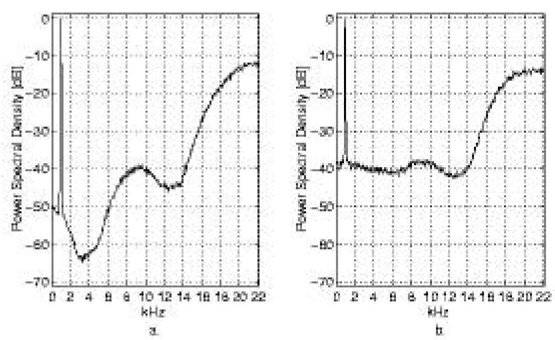


Figure 5 (Left). Power spectral density for a 4 LSB sinusoid, nonsubtractively dithered with 2 LSB triangular dither. Noise shaping is applied. The PSD is normalised to give 0 dB at the input frequency. (a) Nyquist rate conversion. (b) 4 times oversampled conversion.

Figure 6 (Right). Power Spectral Density for a psychoacoustically optimal noiseshaped 4 LSB sinusoid subtractively dithered with 2 LSB triangular dither. The PSD is normalised to give 0 dB at the input frequency. (a) Infinite quantiser. (b) 9 LSB quantiser