Dithering and Averaging: Caveats and Limits of Application

by

István Kollár

Dept. of Measurement and Information Systems
Budapest University of Technology and Economics
H-1521 Budapest, Magyar tudósok krt. 2., HUNGARY
Tel.: +36 1 463-1774, fax: +36 1 463-4112
Home page: http://www.mit.bme.hu/~kollar/
Email: kollar@mit.bme.hu

Abstract – Dithering is an attractive way to decrease quantization bias. However, applying results obtained for ideal quantizers, especially with sampled dithers, requires special attention, since they can be wrong when they are directly extended to real analog-to-digital converters. Moreover, simulation can also be misleading since synchronization effects may show up, not necessarily occurring in realization of dithered ADC's. This paper discusses such cases, gives explanation for the cause of a few properties obtained by simulations in a recent paper (Krause, L., "Effective quantization by averaging and dithering," *Measurement*, Vol. 39, No. 8, pp. 681-694. 2006), and suggests that Gaussian dither is better than sinusoidal or uniform one.

Keywords – dither, quantization bias, analog-to-digital conversion, averaging, effective number of bits, ENOB.

I. Introduction

Quantization (analog-to-digital conversion and roundoff) is a nonlinear operation, inherent to digital computations. According to quantization theory [1], [2], [3], [4], the errors due to quantization can be effectively reduced by using dither. Applications of dither are discussed in several recent papers, like in [5]. However, the immediate use of results of simulation-based investigations for practice requires extreme care. One needs to precisely define the circumstances, otherwise the simulated – albeit beautiful – results cannot be reproduced in practical cases. Furthermore, certain settings yield good results without the true possibility of generalization. We will discuss such cases here.

II. Ideal quantizer?

The basic assumption of many papers on quantization is that the quantizer is ideal. This statement precisely defines the cases for which the subsequent investigations are valid. However, in many cases, non-idealities corrupt the strength of these statements, correctly obtained *for the case of ideal quantization only*. One needs to be aware of which statements, true for ideal quantization, have serious limitations for non-ideal ones.

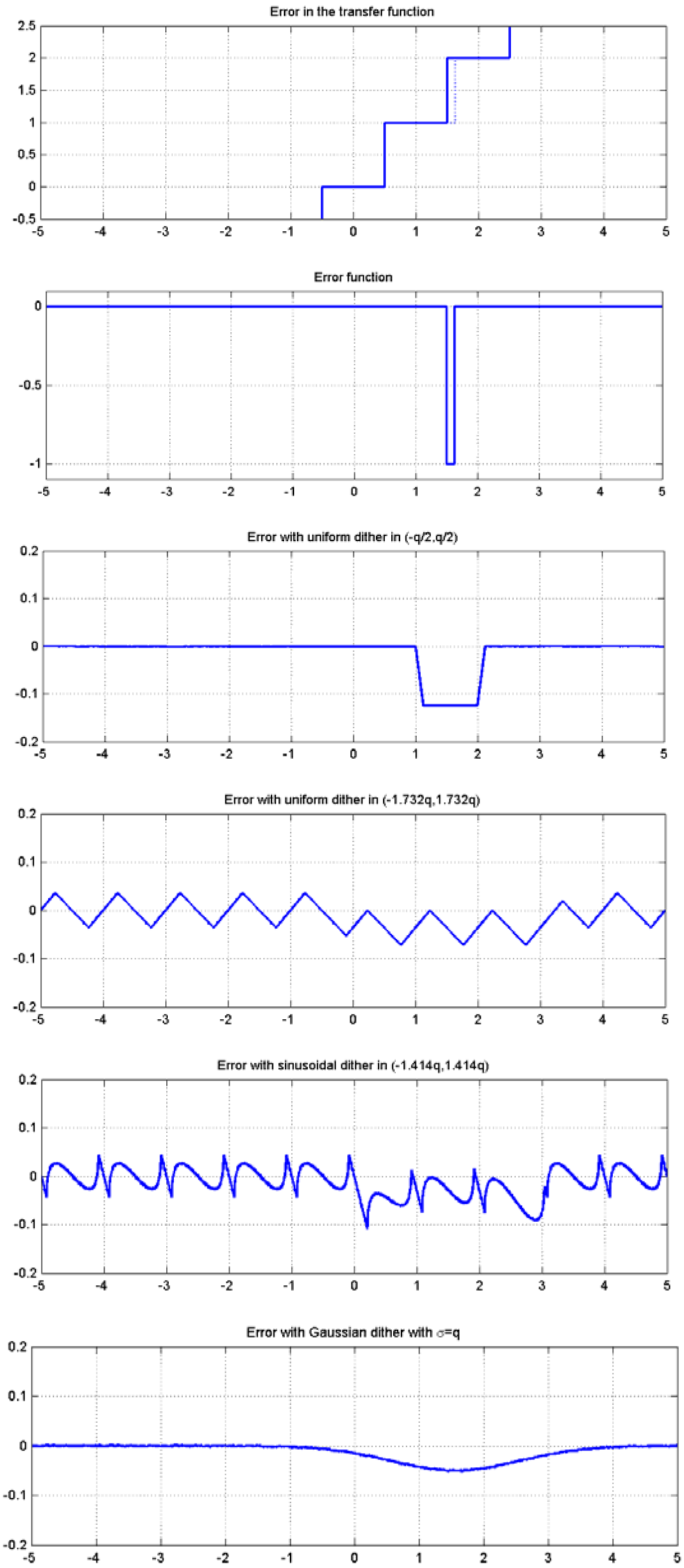
There are two main cases of quantization: analog-to-digital conversion, and arithmetic roundoff.

A. Analog-to-digital conversion

Analog-to-digital conversion is never ideal. Due to component imperfections, comparison levels deviate from the desired ideal values. Because of this, statements which are based on the uniform pattern of the comparison levels are applicable with restrictions only. Eminent examples are dithers with jumps at the sides of the probability density function (e.g. uniform, sinusoidal).

Dither signals are often used to smooth the quantization characteristic. This is possible in general with dither distributions which are wide enough (large amplitude), are smooth, and smoothly approach zero at the edges (these were the conditions for the input PDF, given by Sheppard, to validate his famous corrections [6]). If these are not true, results may be sensitive to relative settings (dither amplitude plus input signal value, with respect to comparison levels), therefore results can be applied to real ADC's with extreme care. Since one cannot avoid that the jumps at the edges of the dither PDF (nearly) coincide with true comparison levels, in A/D conversion *application of small-amplitude uniform or sinusoidal dither is generally of no use.* The following example illustrates why this is so. For this example, we discuss expected values – finite sample effects are discussed later.

Figure 1 illustrates why these dithers can act wrongly, and why normally distributed dither can be suggested rather ([4], [7], [8]).



a) b) c) c) d) e)

Fig. 1 Mean value of the quantization error of a dithered constant value, as a function of the constant value (unit: q), for a uniform quantizer with the ideal comparison level $1.5q$ being off by $+q/8$, that is at $1.625q$ for typical dithers, a) error in the transfer function; b) error withput dither; c) uniform dither in $\pm q/2$ (the common setting); d) uniform dither in $\pm 1.732q/2$; e) sinusoidal dither, amplitude $A_d = 1.414q$ f) Gaussian dither, $\sigma = q$

The dither in c) has the most common setting for uniform: (-q/2,q/2). It is observable that for ideal quantizer transfer function parts, there is no bias, but the erroneous comparison level causes significant error. Dithers d)-e)-f) have approximately the same variance. Gaussian dither more effectively smoothes the error than the others. Not only there is virtually no error further away from the erroneous comparison level, but the maximum error is also smaller, and it smoothly depends on the input value.

B. Arithmetic roundoff

The only really ideal quantization situation is arithmetic roundoff. This is however a case where artificial dither is usually not used. One would not really appreciate a computer which returns different results for different runs… moreover, this is *requantization* which is special in itself. For this, a special theorem is formulated ([7],[8]) which requires a digital dither which has as many bits at the number of bits to be eliminated, and have uniform distribution to cause no bias.

III. Sampled sinusoidal dither?

When one talks about *sinusoidal* dither, this means that one thinks in terms of a signal with the following probability density function / characteristic function (PDF/CF):

$$f(x) = \frac{1}{\pi\sqrt{A^2 - x^2}}, \qquad \Phi(u) = J_0(Au) \tag{1}$$

This is a good model of a sampled random-phase sine wave, that is, when samples are taken in an unsynchronized manner from a sine wave, or when many samples are taken which cover the phase values more or less uniformly. However, when only a few samples are taken from a fixed-phase sine wave, this has nothing to do with sinusoidal distribution. For example, the following case [5]: “AVR2: frequency of sine dither signal: $f\,/\,f_s = 0.5$; amplitude of sine dither signal: $0.85q$; phase of dither signal (referring to f_s): $\varphi = 0.3$ rad", this is just two different values: $x_1 = 0.85q\cdot\sin(0.3) \approx 0.2512$, $x_2 = -0.85q\cdot\sin(0.3\cdot 2\pi) \approx -0.2512$. These two values ±0.25 represent a *binary distribution,* at values $\pm q/2$, and thus allow doubling of the resolution. The case is similar for the value $5.9q{*}\sin(0.3) \approx 1.744$ also mentioned in [5]. None of the properties of the sinusoidal dither is used, except for the one that samples through a full period sum up to zero, a property which can be achieved also by other means. Any deterministic sequence can be used which has zero mean, and is reasonably smeared (above in the simples way, to just two values at $\pm q/2$), thus effectively decreases bias.

IV. Normally distributed dither

A special case is normal distribution. A "normally" distributed sequence with zero arithmetic mean can be generated by

- generating a uniformly distributed sequence e.g. in $[0+1/(2n), 1-1/(2n)]$,
- applying the inverse of the normal cumulative distribution function to this sequence, maybe with subsequent amplitude limitation
- correcting the sequence to have zero mean.

Applying the algorithm AV*n* in [5] to the measured sequence (averaging each group of n samples) will yield the "best" approximation of the input.

A better approximation can be achieved when the uniform distribution is generated not through independent samples, but as a randomly ordered sequence containing the values $q\cdot k/N$, $k = -(N-1), -(N-3), \ldots N-1$. In this case the "normally distributed" sequence has a histogram which has close-to-zero randomness, thus realizing the expected average behavior.

V. Bias from *n* Samples

Assume that an unknown dc value x is measured. One can take n samples, with n different dither values. The sum of the dither values is zero. What is the largest error one can observe for a given dither sequence?

n values define n different offsets before quantization. If x changes by small steps, the same values are measured for x, until the first such dithered value crosses a quantization level. That is, the output is insensitive to an increase in x until the next dithered value crosses a comparison level. Therefore, the maximum error is determined by the maximum insensitivity interval. If the dither values are uniformly distributed, the maximum interval is q/n, and the maximum error is $q/(2n)$. This is the best case, observed experimentally by [5]. This can be achieved with digital uniform dither with values $q/2 \cdot k/n$, $k = -(n-1), -(n-3), \ldots n-1$. Therefore, uniform dither is the best – at least for an ideal quantizer.

The worst case is when the dither values are “unlucky” and do not help. This extreme is also observed for certain sample sequences of a sinusoidal dither.

One may wonder if with sinusoidal dither it is also possible to achieve or at least approach the limiting error value q/n. From Tables 2-3 in [5] it is clear that looking at the mean square error,

- for small values on n, if the discrete values can be situated more or less uniformly, it is possible to approach discrete uniform distribution with samples of the sinusoidal in many settings,
- for larger values, this situation is more rare to reproduce.

One must recognize however that the above mentioned simulation in [5] measures the *error variance for an input uniformly distributed in the input range.* That is, the Q_{eff}

values are *average* values, and are not directly related to the *maximum* error value. The maximum error can be significantly worse than the error calculated from the average noise power, with uniform assumption.

It is possible to give an upper limit of the bias using sinusoidal distribution ([9], Fig. 2).

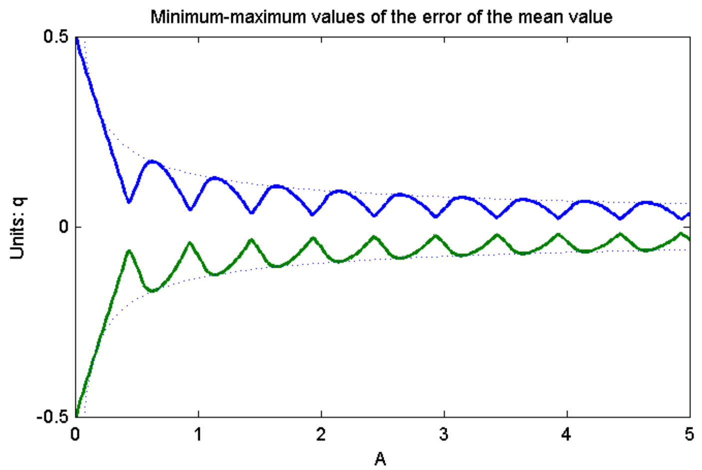


Fig. 2 Limits of bias with sinusoidal dither (minimum-maximum values are taken by changing the input mean value)

The envelope of this upper limit is approximately $\pm\frac{0.1355q}{\sqrt{A/q}}$. One can achieve rather small errors for special sine amplitudes, but this is sensitive to setting errors. This is related to the argument against sine wave as a dither: it is sensitive to non-idealities.

An example for the practical use of sinusoidal dither is given in [10]. Here, a sinusoidal dither of peak-to-peak amplitude of about $23q$ is used at the input of a signal analyzer. This seems to be in contradiction with the above statement.

In this example, a large-amplitude sinusoidal dither is used to average out ADC nonlinearities. This is good, but then subtractive dither has to be used, otherwise the signal is too much corrupted by dither. The designers suppressed the dither variance (power) from the processed signal in the frequency domain. In this respect, the sine wave is desirable, since its power is concentrated to one frequency, therefore this can be easily eliminated from the spectrum.

If averaging of several samples is done, concentrated power has no practical importance: if the arithmetic mean of the dither samples is zero, averaging of samples is enough to get rid of the dither at the quantizer output. This is, however, a special setting. Its effectiveness depends on the effect of dithering on the quantization error. By this, we are back to Fig. 1: dither in averaging AD conversion is good only if it effectively decreases quantization distortion for the non-ideal case. And this requires dither with smooth PDF edges: that is, normally distributed dither.

References

[1] B. Widrow, "Statistical analysis of amplitude-quantized sampled-data systems," *Trans. AIEE, Part II.: Applications and Industry*, vol. 79, no. 52, pp. 555-68, Jan'61 Section, 1961.

[2] B. Widrow, I. Kollár, and M.-C. Liu, "Statistical theory of quantization," *IEEE Trans. on Instrumentation and Measurement*, vol. 45, no. 6, pp. 353-61, 1995.

[3] I. Kollár, "*Quantization Noise*," Doct. Sci. Thesis, Hungarian Academy of Sciences, Budapest, 1997.

[4] B. Widrow and I. Kollár, "*Quantization Noise in Digital Computation, Signal Processing, and Control*." Cambridge University Press, in preparation. Home page: http://www.mit.bme.hu/books/quantization/ Appendix on Digital Dither: http://www.mit.bme.hu/books/quantization/digital_dither.pdf

[5] Krause, L., "Effective quantization by averaging and dithering," *Measurement*, Vol. 39, No. 8, pp. 681-694. 2006.

[6] Sheppard, W. F. "On the calculation of the most probable values of frequency-constants, for data arranged according to equidistant divisions of scale," Proc. London Math. Soc., Vol. 29, pp. 353-380. 1898.

[7] R. A. Wannamaker, S. P. Lipshitz, J. Vanderkooy, and J. Nelson Wright, "A mathematical theory of non-subtractive dither," *IEEE Trans. on Signal Processing*, vol. 48, no. 2, pp. 499-516, Feb. 2000.

[8] Kollár I., "Digital Non-Subtractive Dither: Necessary and Sufficient Condition for Unbiasedness, with Implementation Issues," in: Proceedings of the 23rd IEEE Instrumentation and Measurement Technology Conference, Sorrento, Italy, 24-27 April 2006., Piscataway: IEEE, 2006. pp. 140-145. Revised and submitted to Trans. Instrumentation and Measurement, 2006.

[9] Kollár I., "Bias of Mean Value And Mean Square Value Measurements Based on Quantized Data," IEEE Trans. Instrum. Meas. 43: 733-739 (1994).

[10] I Kollár, "Statistical Theory of Quantization: Results and Limits," Periodica Polytechnica Ser. Electrical Engineering, 28: (2-3) 173-189 (1984).

A few PDF files are readable at http://www.mit.bme.hu/~kollar/publist.html